

**Copyright 2011 Emily Ruth Wheeler Lankau**

SITE EFFECTS ON WILDLIFE ENTERIC BACTERIAL DIVERSITY:  
YOU ARE *WHERE* YOU EAT?

BY

EMILY RUTH WHEELER LANKAU

DISSERTATION

Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Animal Sciences  
in the Graduate College of the  
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor Roderick I. Mackie, Chair and Director of Research  
Professor H. Rex Gaskins  
Professor Ken N. Paige  
Professor Patrick J. Weatherhead  
Associate Professor Isaac K. O. Cann

## ABSTRACT

The study of microbial biogeography has made vast strides in recent years due, in part, to advances in the technological capacity to document microscopic biodiversity at the community, population and genomic levels. In this dissertation, we apply a combination of molecular and cultivation approaches to the study of enteric microbial diversity in a unique system of sister iguanid species, the Galápagos land and marine iguanas (*Conolophus* spp. and *Amblyrhynchus subcristatus*, respectively). We explored the spatial diversity of enteric bacteria in these two species at multiple levels, from communities to genetic traits. The unique host population history and geography of this island chain informed hypotheses about expected biogeographical patterns and the underlying processes that shape microbial diversity in this system.

In Chapter 1, we reviewed current understanding of microbial biogeography and explored how incorporation of such ecological theory might benefit understanding of enteric microbial community structure and function across space and time.

Chapter 2 explored the spatial community diversity of marine and land iguana enteric communities using a molecular approach based on 454 pyrosequencing. Firstly, we demonstrated that while host species was the strongest force shaping these communities, within a host species, geographical proximity also determined overlap in community composition at the genus level. In addition, we found that the degree of contact among host species can produce distinct local effects in community richness and composition, especially when host species have amplified opportunities for microbial exchange such as higher population densities or limited habitat area.

Next in Chapter 3, we explored taxonomically finer-scale diversity patterns among marine iguana populations for the rich enteric genus *Clostridium*. In contrast to the expectation that each host population should have relatively distinct *Clostridium* communities, we found a surprising amount of phylogenetic conservation across all sites, despite also demonstrating evidence suggesting on-going taxonomic turnover – forces which might otherwise lead to rapid divergence of enteric communities in allopatric host populations.

We then applied a more traditional cultivation and molecular genetic approach in Chapter 4 in order to document *Salmonella enterica* strain diversity among sites. We serotyped and genotyped *Salmonella enterica* isolates carried by land and marine

iguanas across a geographical gradient and found nearly complete isolation among strain pools. However, we also found suggestions of geographically-dependent genomic similarity among sites, possibly due to long-distance transport of genetic elements by oceanic currents.

Finally, in Chapter 5 we explored genetic trait biogeography in this system by documenting phenotypic and genetic patterns of antibiotic resistance. We found that sites farther from high densities of humans (i.e. major port towns) harbored fewer resistant bacteria. We also noted that these antibiotic resistance traits may not be retained within the broader endemic bacterial community of Galápagos wildlife for any appreciable length of time, as *Salmonella enteric* isolates did not share resistance traits found in *Escherichia coli* within the same site or even within the same individual fecal sample.

Through this work as a whole we demonstrated that, within the context of a single study system, biogeographical patterns and mechanisms can vary widely across bacterial taxonomic levels, with both ecological and evolutionary forces acting in concert on enteric biodiversity. Increased understanding of the interplay among these forces for shaping microbial community form and function across taxonomic scales has potential to improve not only theoretical understanding of microbial ecology but also to advance management of pressing issues such as novel disease emergence or antibiotic resistance dissemination.

## **ACKNOWLEDGMENTS**

### **Research support**

Thanks to my adviser Dr. Roderick Mackie, my committee, Drs. Isaac Cann, Rex Gaskins, Ken Paige and Patrick Weatherhead and to GEEB Micro for their valuable feedback and guidance on this work. Thanks to the Department of Animal Sciences and to Nancy Henry for administrative support of these projects. A special thanks to Richard Lankau for statistical support and for many critical discussions throughout the development of this dissertation.

I appreciate the help of Martin Wikelski, the Charles Darwin Research Station – especially Paulina Couenberg, and Sonia Cisneros – the Parque Nacional Galápagos and PNG director Washington Tappía for facilitating field collections in the Galápagos Islands, Ecuador. All sampling and exports were approved by the Charles Darwin Research Station and Galapagos National Park (PNG Autorization de Proyecto PC-21-06 Ext 01-09). Exportation of samples was also approved by the Convention on International Trade in Endangered Species (CITES Permit No. 007-09/PNG ). Many thanks to Augusto G. Haz Beltran, Lenin Cruz Bedon and the crew of the Pirata for assistance with sample collection in the field. Mark Mitchell was instrumental in helping with preparations for field sampling and served as a valuable source of feedback and support throughout this process. I would also like to express my appreciation to Peiyong Hong who performed the laboratory data collection for two chapters of this work and to Yuejian Mao for bioinformatics assistance. Finally, thanks to Bryan White for providing access to BL2 laboratory resources for safe handling of feces and potentially pathogenic isolates and to Isaac Cann for access to other general laboratory resources.

All handling of live animals was performed under an approved Institutional Animal Care and Use protocol (IACUC Project Number # 09041) and all biological specimen handling was performed with university oversight under an approved biosafety project registration (Project title: Salmonella microdiversity in Galápagos Iguanas). Many thanks to both the IACUC and biosafety committees for ensuring the welfare, health and safety of all animals and humans involved in this study.

## **Funding**

The University of Illinois Graduate College, the United States Environmental Protection Agency, the University of Illinois College of Veterinary Medicine and a number of other extramural agencies have been most generous in providing me with personal stipend, tuition and research support throughout my graduate experience. Without these financial resources, I would not have been able to make such rapid strides towards simultaneous completion of two doctoral degree programs. Research funding specifically for the projects described in this dissertation was provided by the USDA National Research Initiative (Proposal no. AG-2008-35206-18784, R. I. M., Chapters 2 and 3) and the Conservation Medicine Center of Chicago (Chapters 4 and 5). Many thanks to all of the agencies and groups who have supported all of my research and educational activities throughout my years at the University of Illinois.

## **Personal support**

Completion of the Veterinary Medical Scholars Program has been a complicated, challenging and sometimes seemingly impossible task. This accomplishment would not have been possible without the administrative and personal support of the staff and faculty of the College of Veterinary Medicine, including but not limited to Ned Hahn, Lois Hoyer, Nikki Hausmann, Jacque Dalzell, Jon Foreman, Jerry Pijanowski, Mary Kelm, Shelley Rairden, and Karen Eichelberger. All of my gratitude to my fellow graduate students and dear friends who were with me through this processes - Cassandra Allsup, Katie Amato, Sara Paver, Ariane Peralta, Johanna Salzer, Nora Ortinau and Leslie Rye – thank you for discussions, constructive criticisms, fair and honest reflections, and many fun times outside the lab. Many thanks also to Marvin and Caroline Case for their interest in my academic progress and their continued support for veterinary students with atypical career goals. Finally, I owe an endless debt of gratitude to my husband Rick, and to Sebastian, Boots, Owen and Toby for keeping me company during the writing of this work and for holding down the fort during late nights in the lab and clinic. I would not have found the persistence to continue in this deeply frustrating experience without your constant reminders that science does have great value for advancing humanity, despite the inherently flawed nature of many of those who do this important work.

## **TABLE OF CONTENTS**

<b>CHAPTER 1: CONCEPTS AND QUESTIONS IN THE BIOGEOGRAPHY OF ENVIRONMENTAL AND HOST-ASSOCIATED MICROBES .....</b>	<b>1</b>
<b>CHAPTER 2: BIOGEOGRAPHICAL PATTERNS IN THE GASTRO-INTESTINAL BACTERIAL COMMUNITIES OF GALÁPAGOS IGUANAS SUGGEST COMPLEX METACOMMUNITY PROCESSES ....</b>	<b>30</b>
<b>CHAPTER 3: <i>CLOSTRIDIUM</i> PHYLOGENETIC DIVERSITY CONSERVED ACROSS ALLOPATRIC MARINE IGUANA POPULATIONS DESPITE EVIDENCE OF TAXA TURNOVER .....</b>	<b>60</b>
<b>CHAPTER 4: SEROTYPING AND GENOMIC FINGERPRINTING OF <i>SALMONELLA ENTERICA</i> FROM GALÁPAGOS IGUANAS REVEALS COMPLEX BIOGEOGRAPHICAL PROCESSES .....</b>	<b>89</b>
<b>CHAPTER 5: CARRIAGE OF ANTIBIOTIC RESISTANT ENTERIC BACTERIA VARIES AMONG SITES IN GALÁPAGOS REPTILES .....</b>	<b>121</b>
<b>CHAPTER 6: CONCLUSIONS .....</b>	<b>145</b>
<b>APPENDIX A: SUPPORTING MATERIALS FOR CHAPTER 2 .....</b>	<b>151</b>
<b>APPENDIX B: SUPPORTING MATERIALS FOR CHAPTER 3 .....</b>	<b>152</b>
<b>APPENDIX C: SUPPORTING MATERIALS FOR CHAPTER 4.....</b>	<b>153</b>
<b>APPENDIX D: SUPPORTING MATERIALS FOR CHAPTER 5 .....</b>	<b>155</b>

# **CHAPTER 1: CONCEPTS AND QUESTIONS IN THE BIOGEOGRAPHY OF ENVIRONMENTAL AND HOST-ASSOCIATED MICROBES**

## **Introduction**

The field of microbiology initially developed as an observational, rather than theoretical science, and as such was slow to develop a cohesive underlying conceptual framework (Prosser et al. 2007, McMeekin et al. 2010). However, recently a growing interest in how classical ecological principles can benefit understanding of microorganisms has prompted efforts to apply traditional “macrobial” theory to explaining and predicting microbial community composition and function (Prosser et al. 2007). These ecological studies of microbiological systems have demonstrated that established ecological theory can translate to minute scales. However such studies have also demonstrated that the spatial and temporal extremes of diminutive organisms can also result in intriguing deviations from the expectations of classical theories, making microbial communities a fertile area for continued investigation (Green and Bohannan 2006).

Application of classical ecological principals to a wide variety of environmental microbial systems has perhaps revealed more questions than answers regarding the rules that govern the assembly and function of microbial communities. In particular, recent empirical challenges to the notion that small organisms are universally catholic in their distributions has encouraged more fine-grained studies of biogeographical patterns and processes in microbial systems (Martiny et al. 2006, O’Malley 2007). These studies have revealed a plethora of patterns, leading to the perhaps unsurprising conclusion that both neutral and selective forces interplay to shape microbial communities. And as for many hard-fought ecological debates, the question of which driver matters most has proved to be far less fascinating than the questions of when and why a particular mechanism acts as the primary driver in a given microbial system.

While it has been suggested that studying host-associated microbes is of limited value for elucidating general principals of microbial biogeography due to the inherent geographical limitations presented by host distributions (Fenchel and Finlay 2004), we would argue that relatively well-studied host-microbial systems have much to contribute to understanding of microbial ecology, in addition to being most directly applicable to



problems solving problem in animal health and conservation. In this chapter, we will provide a short review of the current literature on bacterial population and community biogeography, both broadly and specifically as these theories relate to host-associated communities.

### **Microbial Biogeography: Patterns and Processes**

Biogeography is the study of the distribution of species over space and time and of the forces which determine those distributions (MacArthur and Wilson 1967, Martiny et al 2006). Studies of microbial biogeography pose the questions: *What biogeographical patterns do microbial communities exhibit?* and *What relative role do contemporary environmental selection and historical or ongoing stochastic processes play in determining microbial community structure and function?*

It had long been thought that microbes were largely catholic in their distribution, due to their small size, relatively large populations, metabolic flexibility and presumed high mobility (Finlay et al. 2001, Fenchel and Finlay 2003, Fenchel and Finlay 2004, Green et al. 2008). Due to this presumed lack of dispersal limitation, assembly of microbial communities was thought to be primarily driven by local habitat selection (i.e. species sorting) from among members of the global species pool (Brock 1961). This idea has been commonly cited as: *Everything is everywhere, but the environment selects* (Baas-Becking 1934, de Wit and Bouvier 2006). However, recent studies in microbial biogeography have demonstrated that both neutral (i.e. dispersal or demographic stochasticity) and niche (i.e. environmental selection or interspecies interactions) processes can contribute to bacterial distribution patterns (e.g. Cho and Tiedje 2000, Finlay 2002, Oda et al. 2003, Papke et al. 2003, Whitaker et al. 2003, Finlay and Fenchel 2004, Fontaneto et al. 2008, Takacs-Vesbach et al. 2008, Vos and Velicer 2008), suggesting that the assemblage of microbial communities is more complex and varied than initially presumed.

#### *Ecological patterns in microbial communities*

Understanding microbial dispersal rates has been identified as a key to understanding how and when microbes present specific biogeographical patterns and for tackling the question: “Is dispersal so rapid that regional biota are homogenized into

a single global metacommunity, or is it slow enough to allow signatures of regional-scale processes to develop?” (Telford et al. 2008). Dispersal rates may vary widely for microorganisms, despite their small size and large population sizes, because the ability to move among and successfully colonize suitable habitat may vary with a variety of ecological characteristics (e.g. habitat specialization, presence of resistant life stages such as spores, ability to defend against oxidative damage, etc.) that may affect dispersal capacities. Measuring dispersal rates in such small organisms may be a challenge that current technology is yet to fully meet, leading to a reliance on pattern detection to determine which mechanisms are dominant in a given system. Dispersal limitations can produce characteristic population genetic patterns, such as distance-decay (i.e. correlation between genetic and geographical distance among populations) and species-area relationships (i.e. correlation between the number of species and patch size) which can suggest that dispersal limitation are an important force for determining species distributions (Green and Bohannan 2006). In addition, communities may show characteristic diversity and variability patterns which suggest stochastic dispersal and demography at play in the formation and maintenance of microbial communities.

### *The evidence for microbial cosmopolitanism*

It is clear that environmental conditions (i.e. abiotic factors or interspecies interactions) can be a strong force for structuring the observed distributions of highly mobile microorganisms. A number of studies have demonstrated strong dependence of community structure on environmental factors, such as pH or nutrient availability, for explaining population or community variation in the absence of concurrent geographical relationships (Franklin et al. 2000, Finlay et al. 2006, Fierer et al. 2007). Similarly at the community level, a study of phyllosphere bacterial communities on tree leaves revealed community similarities driven by tree taxonomy across continents with minimal geographical patterning detected among host species across the sampling range (Redford et al. 2010). Similarly, composition of epilimnetic bacterial communities in several lakes in Minnesota spanning a range of abiotic factors and geographic locations found the lakes with similar physical conditions had more similar bacterial communities (Yannarell and Triplett 2004).

Further, when niche selection is the primary force shaping microbial distributions, a single strain or species may be widely or globally detected in patches of appropriate habitat. For example, a comparison of aquatic microbes from a Danish marine inlet and an English fresh water lake found striking similarities between the two habitats in terms of both species size and identity (Fenchel and Finlay 2004). Such patterns have also been demonstrated on very broad scales for some eukaryotic taxa (Finlay et al. 2006, Bass et al. 2007, Fontaneto et al. 2008).

Finally if microorganisms are globally distributed, then species present in any local sample should represent a large portion of the cumulative regional species pool from all similar habitat samples (Green and Bohannan 2006). One study of the flagellate genus *Paraphysomonas* from a fresh-water pond in England showed that this single site was host to 80% of the known species for this group (Finlay 2002). Further evaluation across a wide range of eukaryotic taxa at two separate locations in Europe suggested a more general relationship between body size and global species distributions. The ratio of locally to globally detected species decreased significantly with mean body size (Fenchel and Finlay 2004).

### *The evidence for microbial endemism*

There is certainly some evidence to suggest that microbial dispersal can be quite rapid, obscuring any patterns of allopatric population or community divergence through homogenizing gene or species flow among locations. In contrast, detection of characteristic ecological patterns, either genetic or community structure, associated with geographical location can suggest the presence of ecologically meaningful dispersal limitation (Green and Bohannan 2006). For example, a distance-decay pattern (or isolation-by-distance relationship), where genetic similarity is highest in populations that are more geographically proximate, suggests that dispersal and genetic mixing occurs more often between sites that are closer together because of limitations in how far the organism is able to move (Wright 1943, Nekola and White 1999).

Detection of population genetic patterns associated with geographical location has been used to explore the possibility that at least some types of very small organisms are dispersal limited. For example, population genetic analysis of a hyperthermophilic archeon (*Sulfolobus*) which grows in isolated hot springs across the globe showed very

little gene flow among locations and no relationship between populations structure and environmental variables like pH or water chemistry. *Sulfolobus* grow optimally at 80°C and pH 3, experiences cell cycle arrest with abrupt changes in temperature and has no known spore state (Whitaker et al. 2003). Similar studies have found comparable evidence for dispersal limitation and endemism in a variety of habitats at varying spatial scales (Cho and Tiedje 2000, Hillebrand et al. 2001, Franklin and Mills 2003, Papke et al. 2003, Vos and Velicer 2008).

### **Coming to Terms with Complexity**

The processes that determine microbial geographical distribution patterns are not necessarily mutually exclusive nor are either extreme ubiquity or endemism the only possible patterns. While some studies do find strong evidence for environmental selection or for isolation-by-distance, other studies have found either no evidence for either effect (Beisner et al. 2006, Nabout et al. 2009) or evidence of both acting simultaneously, albeit with differing relative strengths (Pangaling et al. 2009, Soininen et al. 2009). It is likely as these patterns continue to be explored on many spatial scales, that increasingly complex biogeographical patterns may be revealed.

#### *Alternative patterns and processes*

Biogeography is inherently an observational field where application of pattern discovery and causal reasoning can allow for strong support for mechanistic models. However, a reduced capacity for experimental approaches must also be accompanied with a cautious interpretation of patterns as process.

If dispersal limitations primarily drive microbial similarity patterns among habitat patches then a significant relationship between similarity and geographic distance might reasonably be expected (Figure 1.1a). In contrast when niche selection processes are strongest, then similarity should be correlated to differences in environmental measures (Figure 1.1b). However, if temporal history of population or community establishment is a primary driver of diversity, then population or community similarity may not correlate to either of the previous variables, instead structuring based on temporal variables, such as geological history or time since last disturbance (Figure 1.1c). Historical differences in regional pools at the time of initial

colonization may result in priority effects that shape alternative stable community structures. The effects of such historical processes has received relatively little attention in microbial studies to date (but see, for example, Fraterrigo et al. 2006, Takacs-Vesbach et al. 2008). Finally, if the environmental driver is located along a directional gradient, then both geographic distance and environmental measures may be associated with microbial population or community similarities, although only one of these variables is causal (Figure 1.1d). While some microbial types may be responsive to the detected gradient, others may be less reliant on that same variable and may instead be structured by other processes along this same gradient (Rousk 2010). Such correlations among potential drivers of microbial distributions may make interpretation of patterns much more challenging.

Another potential pitfall for drawing conclusions of process from observed patterns is a failure to fully incorporate necessary biological realism. For example, a lack of significant distance-decay patterning using a model which assumes equal dispersal potential in all directions does not necessarily imply ubiquity as the only alternative. Yet in studies which do not detect the expected dispersal pattern, the common conclusion has often been that dispersal limitations do not play a role in the system of interest. However, alternative explanations may be viable in some systems that would possibly better fit the unique ecological conditions. While simple isolation-by-distance models that assume a capacity for random dispersal in all directions remain a reasonable initial model for consideration (see review by Jenkins et al. 2010), if dispersal processes are asymmetrical in space or time, the resulting genetic or community patterns may not be evident unless this asymmetry is incorporated into analyses (Cook and Crisp 2005). Only a few studies to date have begun to incorporate potential asymmetries in dispersal potential into exploring microbial biogeography (Boyer 2008, Fierer et al. 2007).

### *Methodological challenges*

A variety of approaches have been applied to identifying microbial types. This diversity has perhaps lead to disagreement and confusion in comparing and interpreting findings among studies. On the one hand, many of the described studies demonstrating cosmopolitanism have been based on morphologic, rather than molecular identification. Such an approach could have resulted in failure to recognize cryptic species within

morphotypes that might reflect dispersal-limitation processes (Green and Bohannan 2006). As even defining what a “species” signifies for microbial species is still under active debate (Achtman and Wagner 2008, Gevers et al. 2005), this raises the question as to what functional or morphological level of microbial variation is appropriate for meaningful pattern detection. On the other hand, it has been argued that studies which show regional population uniqueness based on neutral or nearly neutral genetic markers may not reflect meaningful phenotypic or functional differences (Fenchel 2003). Both of these arguments may have some merit. The choice of genetic or phenotypic measures by which microbial species and strains are identified may strongly influence the types of patterns that are detected, as different elements of the genome may respond very differently to different evolutionary pressures. In the extreme case, horizontal gene transfer may result in very local patterns in functional traits due to selective pressures which may not reflect deeper strain history but may reflect important patterns in trait distribution (Parnell et al. 2010, Hall et al. 2010).

Traditionally, biogeographical studies of plants and animals have accounted for this possibility by combining neutral and trait-based genetic elements to fully explore how migration and selection combine to shape geographic patterns (stochastic versus deterministic, Hendry 2002, McKay and Latta 2002, Porcher et al. 2006, Steane et al. 2006; Gandour et al. 2008). Given the growing understanding of how horizontal gene transfer can contribute to patterns in bacterial genetics, it is a reasonable parallel to suggest that bacterial biogeography might also benefit from simultaneous measures of diversity at different loci or levels to better understand the complexity of these questions.

While one area of the genome might be highly conserved and show a homogeneity consistent with ubiquity, strong selection on individual traits may result in rapid divergence of other sites within the same genome (Coleman and Chisholm 2010). For example, a study of *Salinibacter ruber* isolates from three regions of the world showed no geographical patterns when using a more traditional genotyping approaches, but saw strong geographical isolation patterns among regional isolates with mass spectrometric analysis of cell envelope sulfonolipids (Roselló-Mora et al. 2008). The authors suggested that these phenotypic differences might reflect genetic differences in transcriptional regulation, a genetic element not normally targeted for population

studies. Similarly, a study of genomic patterns within and among populations of marine bacteria (*Prochlorococcus* and *Pelagibacter*) showed that housekeeping genes were without geographical patterns but phosphorus acquisition genes showed strong site patterns which paralleled major biogeochemical differences among sampling sites. A combined approach that incorporates both selectively functional traits and neutral markers may reveal illuminating conflicts that could more clearly reveal the complexity of microbial biogeography, especially when multiple processes are acting in concert.

### *Approaching conceptual resolution*

Improved understanding of microbial biogeography will require increased comfort with increasingly complex patterns, mechanisms, and technological approaches to linking the two. Incorporation of this complexity into biogeographical models may lead to resolution of the current theoretical debates in this field. The tension between the effects of stochastic processes like dispersal and environmental selection is certainly not new. Early in the development of classical ecological theory there was a similar emphasis on deterministic environmental and interspecies effects on species distributions (Tilman 1982, Chase and Leibold 2003, Ricklefs 2004). More recent challenges to this view have illuminated the contribution of stochastic variation in demographic processes and dispersal on ecological communities (Hubble 2001). However, in such classical ecological theory, recent synthesis increasingly supports a shift in emphasis from dichotomous outcomes to exploring how specific ecosystem traits might result in differences in the relative contribution of neutral or selective forces on population and community processes (Chase 2007).

This is not to suggest that amid this complexity general principles are not waiting to emerge, based for example, on microbial species traits (e.g. generalist versus specialist, growth rate, sporulation etc.) or microbial community traits (e.g. diversity, stability, environmental stresses etc.). For example, dispersal ability and colonization success in microbes may be linked to traits affecting competitive abilities, where good competitors are often habitat specialists which are poor dispersers and vice versa (i.e. the competition-colonization trade-off; Levins and Culver 1971, Slatkin 1974, Hastings 1980, Tilman 1994). Similarly, different types of “macrobial” taxa have been shown to have different maximum dispersal extents due to resistance to intervening conditions,

resulting in different degrees of dispersal limitation in different taxa based on various dispersal and establishment traits (e.g. mode of transportation, resistance to desiccation, fecundity etc.; Brown and Lomolino, 1998).

### **The Biogeography of Host-Associated Communities**

Host-associated microbial systems have long been studied for their role in host health and physiology, but only recently has ecological theory been applied to understanding these communities. Such an approach may reveal not only deeper insight into how these systems operate, but may also more broadly contribute to understanding the complexity of microbial ecological processes. Host-associated systems represent a relatively unique type of microbial community in terms of potential specialization of microbes to specific host species, in terms of geographical restrictions based on host distributions, and in terms of the dynamic feedbacks which may occur between host and microbes (Camp et al. 2009, Robinson et al. 2010). However, while these traits may impact the outcomes of ecological processes operating in these systems, these characteristics certainly do not prohibit the application of ecological thinking to understanding the composition and function of these systems.

#### *Scale in host-associated systems*

The biogeography of host-associated microbial communities can be considered at a variety of scales from within a single host to among communities of hosts (Figure 2.1). These difference scales may contribute to the complexity of understanding the biogeography of host-associated microbes. However, these different scales do not preclude drawing parallels to established theory in environmental systems. Within and across these scales, habitat selection, dispersal or exposure differences, and exposure history may all play a role in community formation and maintenance.

Observations of within host microbial biogeography have demonstrated that microbial communities differ in diversity and composition among body regions (e.g. skin versus digestive tract and mucus membrane in human subjects; Costello et al. 2009; fore-stomach versus feces in bovine subjects, Michell and et al. 2009). While these studies demonstrate that between body location differences are very strong, inter-individual variability at any specific body location is quite high (Costello et al. 2009,



Fierer et al. 2010, Lazarevic et al. 2010). In contrast, within a single individual, community composition can be very stable at a given body region over time (Costello et al. 2009). These within individual patterns can be very fine-scale with predictable differences in biota detected even between the right and left hands of the same individual (Fierer et al. 2008). As studies at this scale are still relatively novel, it remains to be seen what forces structure differences among such proximate body locations, but differences in host-microbial interactions at different body sites appears to play a key role in some circumstances (Costello et al. 2009).

Host-associated community patterns can also be studied among allopatric host populations, in parallel to established environmental microbial studies which compare distinct average community traits of local habitats across geographical locations. At the host population or community level, differences in microbiota from the same body region among locations might reflect host genetic or behavior differences, differences in regional microbial pools, or historical differences in the establishment of communities.

### *Theoretical expectations*

Dispersal limitation has been proposed to dominate biogeographical distributions of host-associated microbes because of the inherent confines on microbial dispersal imposed by host species distributions (Zhang and Blackwell 2002, Bala et al. 2003, Falush et al. 2003, Hedlund et al. 2003, Fenchel and Finlay 2004, Papke and Ward 2004). A tendency towards strong dispersal limitation may, in fact, more broadly reflect a characteristic of extremely specialized or geographically restrictive habitats and very likely is not unique to host-associated system. Certainly host-associated microbes can be so intimately associated with their hosts that microbial population genetics may reflect deep evolutionary or ecological history of the host. For example, the biogeography of the human stomach bacterium *Helicobacter pylori* reflects the deep evolutionary history of human populations from origins in Africa to our modern global distribution (Falush et al 2003).

While dispersal limitations may be a rational expectation for obligatory or highly host-specific associations such as pathogens or co-evolved mutualists, this restriction may possibly be relaxed for host-associated microbes which have the ability to survive, even temporarily, in non-host environments or to associate with many different host

species. For example, gastrointestinal bacteria responsible for food-borne outbreaks can persist for days to months in a variety of environmental sources (e.g. brackish water - Rhodes and Kator 1988, seawater - Rozen and Belkin 2001, swine manure - Guan and Holley 2003, produce- Harris et al. 2003, plants - Brandl 2006, fruit - Alegre et al. 2010, cattle manure - Semenov et al. 2010), resulting in wide and rapid dispersal by agricultural distribution systems (e.g. CDC 2006, 2007, 2009, 2010). For at least some host-associated bacteria, a single strain may quickly travel long distances, as during a food-borne outbreak, and thus would not necessarily be expected to demonstrate clear genetic-by-distance relationships.

As host-associated microbial communities are our most intimate associates, understanding these communities is vital to understanding health and disease. But in addition to these applied goals, studies of these systems have much to contribute to such theoretical discussions as well.

### *Biogeographic patterns in enteric communities*

A number of recent studies have focused specifically on the roles of environmental selection, regional differences in exposures and colonization history on intestinal microbial communities. Environmental selection appears to be a key factor for determining differences in specific host-associated communities (e.g. biota of the intestinal tract) among individuals or among host species. Strong selective forces, such as diet or exposure to antibiotic medications for example, can have notable impacts on gastrointestinal microbial communities among individuals (e.g. Dietary effects: Tannock 1995, Gil and Reuda 2000, Rabiou and Gibson 2002, Macfarlane and Macfarlane 2003, Turnbaugh et al. 2009b; Antibiotic effects: Tannock 1995, Sullivan and Nord 2006, Antonopoulos et al. 2009). Diet in particular can act as a strong selective force on gastrointestinal microbial communities, contributing to differences in microbial community form and function, both within and among host species (Ley et al. 2008a, Ley et al. 2008b, Turnbaugh et al. 2009a, Turnbaugh et al. 2009b, Hong et al. In press). Within mammalian taxa, for example, dietary differences have been found to cross phylogenetic relationships, with microbial community composition determined more strongly by dietary groups (i.e. herbivores versus omnivores versus carnivores) than by host taxonomic relationships or geographical location (Ley et al. 2008a).

These selective differences between host species can be quite strong, having the ability to produce major shifts from initial inoculum composition. When germ-free mice and zebra fish were colonized with the intestinal biotas of conventional zebrafish and mice, respectively, the resulting microbial communities were able to shift over time to more closely resemble the normal host biota compared to the donor biota (Rawls et al. 2006). But what this study also suggested is that different host species may be able to accommodate a wide variety of microbial species or strains to which they may not normally be exposed. While gastrointestinal communities were drastically shifted from their initial inocula, cross-inoculated individuals still more strongly resembled the donor biota in overall composition.

The potential for the host environment to drive community convergence should be substantially limited by any existing differences in microbial exposures (i.e. differences in the regional or local species pool) either when communities are first developing or by the on-going introductions of novel species. However, the manner by which such environmental or historical forces shape host-associated microbial communities is less well-described (Robinson et al. 2010). Such patterns may be especially striking among allopatric, conspecific host populations, where holding host selection relatively constant may allow for exploring such dispersal or temporal effect on microbial communities.

A number of studies have demonstrated geographic patterns among gastrointestinal microbial communities (Mueller et al. 2006, Dethlefsen et al. 2007, Frank et al. 2007, Ley et al. 2008a, Costello et al. 2009, Turnbaugh et al. 2009a, Benson et al. 2010, Fallani et al. 2010, Hong et al. In press). Such effects can even supersede dietary influences within a host species. For example, a geographical gradient in gut microbiota composition and similarity was detected in a study of European infant enteric microbial communities, an effect that was stronger than other explanatory variables including feeding method (breast milk or formula), antibiotic exposures, and differing mode of delivery (cesarean or vaginal; Fallani et al. 2010). This suggests that differences in regional strain or species pools may contribute strongly to gastrointestinal community structure, despite or in addition to strong environmental selective pressures. Mouse mothers and their pups share phylogenetically similar cecal microbiota, suggesting that pups acquire the structure of their microbiota from local exposures (Ley

et al. 2005). Further, these exposure effects can over-ride any contribution of host genetics, with transplanted pups resembling gestating mother despite the genetic line of the mice (Friswell et al. 2010).

Priority effects through differences in exposure history may also have strong influences in the initial development of host-associated microbial communities that may persist in shaping divergent communities. For example, newborn infants acquire their microbiomes during or soon after birth. The composition of this initial biota may be strongly influenced by differences in initial exposures during the birthing process. While infants that experience a vaginal birth have early skin and oro-nasopharyngeal microbiomes which are more similar to maternal vaginal communities, infants delivered by cesarean section have a microbiomes which are more similar to maternal skin communities (Dominguez-Bellow et al. 2010).

While enteric microbes have been explored at the community level in terms that can translate to biogeography, population genetic exploration of ecological patterns across host populations outside of the disease context are not well established. Certainly it has been long established that local exposures can determine disease risk and that genetic evaluation of specific pathogens can help with tracking the spread outbreaks or virulent strains (Fan et al. 2009, Girones et al. 2010, Harris et al. 2010). However, how location affects strain carriage of non-pathogenic bacteria has not been fully explored (but see Wheeler et al. 2011). Thus, understanding of the biogeography of host-associated microbes is but in the nascent stages and a more rigorous application of the theoretical structure established for environmental microbes may strongly contribute to both furthering microbial ecology theory and to bettering our understanding of the role these communities play in health and disease.

There are many questions left to be explored in the area of host-associated microbial biogeography: *At the broad geographical scale, how do host differences in diet or exposures shape differences in microbial community form and function? Does geographic overlap among host populations contribute to increased similarity in individual microbial communities? How do geographic separations, either historical or current, shape host-associated communities among conspecific host populations? Finally, do local exposures drive microbial species or strain distributions within geographically structured host populations?* Such questions not only have theoretical

implications for understanding of the ecology of host-associated microbes, but may also aid with understanding how contacts among host populations or host species may contribute to disease risk.

## **Dissertation Study Context and Description**

Island systems have long contributed to the study of how geographic isolation contributes to ecological and evolutionary processes in both macrobial and microbial systems (Darwin 1859, Mayr 1963, MacArthur and Wilson 1967, Diamond and May 1976, Krebs 1994, van der Gast 2008, Losos and Ricklefs 2009). The Galápagos Islands located approximately 1000 km off the coast of Ecuador are particularly famous for their unique ecology and contribution to current understanding of island biogeography and evolution (Darwin 1859, Boag and Grant 1984, Grehan 2001, Parent et al. 2008).

### *Study system*

The Galápagos Islands are host to a number of endemic wildlife species including 27 reptile species that are distributed throughout the island chain. This unique biodiversity includes two ecologically distinct genera of large herbivorous lizards, the marine and land iguanas (*Amblyrhynchus cristatus* and *Conolophus* species; Jackson 2007). Land and marine iguanas are phylogenetically quite closely related, having diverged from a common Ctenosaur ancestor approximately 10 to 20 million years ago (Wyles and Sarich 1983, Sites et al. 1996, Rassmann 1997, Rassmann et al. 1997a, Cogger and Zweifel 1998).

Populations of both genera are both found on a number of sites throughout the island chain and their distributions overlap at some locations (Jackson 2007). Marine iguanas are locally abundant and widely distributed throughout the islands, living in dense colonies along the rocky coasts and feeding on marine algae (Trillmich and Trillmich 1986, Wikelski et al. 1993, Shepherd and Hawkes 2005). Land iguanas are less commonly found in dense groups except on smaller islands, where space is limited (Kritcher 2006). *Conolophus subcristatus* is found on a number of islands while *C. pallidus* is entirely endemic to the island of Santa Fe. The rarest species, *C. marthae*, is a newly discovered pink species endemic to a single volcano on Isla Isabela (Gentile et al. 2009). Land iguanas prefer inland habitat (Snell et al. 1984) and primarily eat opuntia

cactus pads and leaves and flowers of lantanas and cordia plants (Christian et al. 1984). The ecological distinctions between marine and land iguanas generally results in relatively low overlap in habitat use on all but the smallest and most densely occupied islands on which these two forms cohabit, such as Isla Plaza Sur, where the two ecotypes have been shown to hybridize (Rassmann et al. 1997b).

The unique biology, phylogeny and biogeography of the Galápagos iguanas provide a distinctive system for studying the effects of host ecology and host population isolation on enteric bacterial communities. In Chapter 2, we investigated enteric bacterial community patterns in allopatric and sympatric populations of Galápagos marine and land iguanas using 16S rRNA gene pyrosequencing of fecal community genomic DNA. We hypothesized that gastrointestinal bacterial communities at this level of taxonomic organization would be most strongly influenced by the different host environments; however, we also expected that among population of each host, there would be strong geographical patterns in the overall similarity of these communities, possibly reflecting either contemporary variability in local exposures or historical patterns of divergence among host populations.

Chapter 3 investigated the evolutionary relationships among OTUs of the diverse genus *Clostridium* within and between sampling sites. We hypothesized that if host colonization history explained the geographic distribution and site-overlap of *Clostridium* OTUs, then evolutionary relationships among those OTUs would show characteristic phylogenetic signals that paralleled the geological history of the sites.

In Chapter 4, we explore the biogeographic patterns of a single enteric bacterial species, *Salmonella enterica*, using two complementary typing methods – phenotypic serotyping and genomic fingerprinting. We hypothesized that *S. enterica* serovars would vary substantially among sampling locations but not between host species within sampling sites.

Finally, Chapter 5 explores the biogeography of a mobile genetic trait, antibiotic resistance, to explore connectivity among human influences on the island chain and wildlife species. We hypothesized that proximity to major port towns would result in higher levels of antibiotic resistance in environmental and reptile isolates of *Escherichia coli* and *S. enterica*.

Together these four studies present a novel contribution to studies of bacterial biogeography by characterizing patterns and processes in this unique system from the genus to trait level in order to understand how environmental selection, history and local differences in microbial exposures can interplay to shape enteric communities among host populations.

## References

- Achtman M, Wagner M. 2008. Microbial diversity and the genetic nature of microbial species. *Nature Reviews Microbiology* 6: 431-440.
- Alegre I, Abadias M, Anguera M, Usall J, Viñas I. 2010. Fate of *Escherichia coli* O157:H7, *Salmonella*, and *Listeria innocua* on minimally processed peaches under different storage conditions. *Food Microbiology* 27: 862-868.
- Antonopoulos DA, Huse SM, Morrison HG, Schmidt TM, Sogin ML, Young VB. 2009. Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. *Infection and Immunity* 77: 2367-2375.
- Baas-Becking LGM. 1934. *Geobiologie of inleiding tot de milieukunde*. Van Stockkum & Zoon: The Hague, Netherlands.
- Bala A, Murphy P, Giller KE. 2003. Distribution and diversity of rhizobia nodulating agroforestry legumes in soils from three continents in the tropics. *Molecular Ecology* 12: 917–929.
- Bass D, Richards TA, Matthai L, Marsh V, Cavalier-Smith T. 2007. DNA evidence for global dispersal and probable endemism of protozoa. *BMC Evolutionary Biology* 7: 162.
- Beisner BE, Peres PR, Lindstrom ES, Barnett A, Longhi ML. 2006. The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology* 87: 2985-2991.
- Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, Zhang M, Oh PL, Nehrenberg D, Hua K, Kachman SD, Moriyama EN, Walter J, Peterson SA, Pomp D. 2010. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proceedings of the National Academy of Sciences* 107: 18933-18938.

- Boag PT, Grant PR. 1984. The classical case of character release: Darwin's finches (Geospiza) on Isla Daphne Major, Galápagos. *Biological Journal of the Linnean Society* 22: 243-287.
- Boyer DG. 2008. Fecal coliform dispersal by rain splash on slopes. *Agricultural and Forrester Meteorology* 148: 1395-1400.
- Brandl MT. 2006. Fitness of human enteric pathogens on plants and implications for food safety. *Annual Review of Phytopathology* 44: 367-392.
- Brock TD. 1961. *Milestones in Microbiology*. Prentice-Hall: Englewood Cliffs, NJ.
- Brown JH, Lomolino MV. 1998. *Island biogeography: Assembly and evolution*. In *Biogeography* (2<sup>nd</sup> ed.). pp. 407-447. Sinauer Associates, Inc.: Sunderland, MA.
- Camp JG, Kanther M, Semova I, Rawls JF. 2009. Patterns and scales in gastrointestinal microbial ecology. *Gastroenterology* 136: 1989-2002.
- CDC. 2006. Ongoing Multistate Outbreak of *Escherichia coli* serotype O157:H7 Infections Associated with Consumption of Fresh Spinach – United States, September 2006. *Morbidity and Mortality Weekly Reports* 55: 1-2.
- CDC. 2007. Multistate outbreak of *Salmonella* serotype Tennessee infections associated with peanut butter – United States, 2006-2007. *Morbidity and Mortality Weekly Reports* 57: 521-524.
- CDC. 2009. Outbreak of *Salmonella* Serotype Saintpaul infections associated with Eating alfalfa sprouts – United States, 2009. *Morbidity and Mortality Weekly Reports* 58: 500-503.
- CDC. 2010. Two multistate outbreaks of shiga toxin–producing *Escherichia coli* infections linked to beef from a single slaughter facility – United States, 2008. *Morbidity and Mortality Weekly Reports* 59: 557-560.
- Chase JM, Leibold MA. 2003. *Ecological niches: Linking classical and contemporary approaches*. University of Chicago Press: Chicago, IL.
- Chase JM. 2007. Drought mediates the importance of stochastic community assembly. *Proceedings of the National Academy of Sciences* 104: 17430-17434.
- Cho JC, Tiedje JM. 2000. Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Applied and Environmental Microbiology*. 66: 5448-5456.



- Christian KA, Tracy CR, Porter WP. 1984. Diet, digestion, and food preferences of Galapagos land iguanas. *Herpetologica* 40: 205-212.
- Cogger HG, Zweifel RG (eds.). 1998. *Encyclopedia of reptiles and amphibians* (2<sup>nd</sup> ed.). Academic Press: San Diego.
- Coleman ML, Chisholm SW. 2010. Ecosystem-specific selection pressures revealed through comparative population genomics. *Proceedings of the National Academy of Sciences* 107: 18634-18639.
- Cook LG, Crisp MD. 2005. Directional asymmetry of long-distance dispersal and colonization could mislead reconstructions of biogeography. *Journal of Biogeography* 32: 741-754.
- Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. 2009. Bacterial community variation in human body habitats across space and time. *Science* 326: 1694-1697.
- Darwin C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray: London, United Kingdom.
- de Wit R, Bouvier T. 2006. “Everything is everywhere, but, the environment selects”; what did Baas Becking and Beijerinck really say? *Environmental Microbiology* 8: 755–758.
- Dethlefsen L, McFall-Ngai M, Relman DA. 2007. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* 449: 811–818.
- Diamond JM, May RM. 1976. Island biogeography and the design of natural reserves. In *Theoretical Ecology: Principles and Applications* (May RM, ed.). pp. 163–186. Blackwell Scientific: Oxford, United Kingdom.
- Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R. 2010. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences* 107: 11971-11975.

- Fallani M, Young D, Scott J, Norin E, Amarri S, Adam R, Aguilera M, Khanna S, Gil A, Edwards CA, Doré J, and other members of the INFABIO Team. 2010. Intestinal microbiota of 6-week-old infants across Europe: Geographic influence beyond delivery mode, breast-feeding and antibiotics. *Journal of Pediatric Gastroenterology and Nutrition* 51: 77-84.
- Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, Blaser MJ, Graham DY, Vacher S, Perez-Perez GI, Yamaoka Y, Mégraud F, Otto Kristina, Reichard U, Katzowitsch E, Wang X, Achtman M, and Suerbaum S. 2003. Traces of Human migrations in *Helicobacter pylori* populations. *Science* 299: 1582-1585.
- Fan J, Shu M, Zhang G, Zhou W, Jiange Y, Zhu Y, Chen G, Peacock SJ, Wan C, Pan W, Feil EJ. 2009. Biogeography and virulence of *Staphylococcus aureus*. *PLoS One* 4: e6216. doi: 10.1371/journal.pone.0006216.
- Fenchel T, Finlay B. 2004. The ubiquity of small species: Patterns of local and global diversity. *BioScience* 54: 777-784.
- Fenchel T, Finlay BJ. 2003. Is microbial diversity fundamentally different from biodiversity of larger animals and plants. *European Journal of Protistology* 39: 486-490.
- Fenchel T. 2003. Biogeography for bacteria. *Science* 301: 925-926.
- Fierer N, Hamady M, Lauber CL, Knight R. 2008. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proceedings of the National Academy of Sciences* 105: 17994-17999.
- Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R. 2010. Forensic identification using skin bacterial communities. *Proceedings of the National Academy of Sciences* 107: 6477-6488.
- Fierer N, Morse JL, Berthrong ST, Bernhardt ES, and Jackson RB. 2007. Environmental controls on the landscape-scale: Biogeography of stream bacterial communities. *Ecology* 88: 2162-2173.
- Finlay BJ. 2002. Global dispersal of free-living microbial eukaryote species. *Science* 296: 1061-1063.
- Finlay BJ, Fenchel T. 2004. Cosmopolitan metapopulations of free-living microbial eukaryotes. *Protist* 155: 237-244.

- Finlay BJ, Estaban GF, Clarke KJ, Olmo JL. 2001. Biodiversity of terrestrial protozoa appears homogenous across local and global spatial scales. *Protist* 152: 355-366.
- Finlay BJ, Estaban GF, Brown S, Fenchel T, Hoef-Emden K. 2006. Multiple cosmopolitan ecotypes within a microbial eukaryote morphospecies. *Protist* 157: 377-390.
- Fontaneto D, Barraclough TG, Chen K, Ricci C, Herniou EA. 2008. Molecular evidence for broad-scale distributions in bdelloid rotifers: everything is not everywhere but most things are very widespread. *Molecular Ecology* 17: 3136-3146.
- Frank DN, St. Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proceedings of the National Academy of Sciences* 104: 13780–13785.
- Franklin RB, Mills AL. 2003. Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. *FEMS Microbial Ecology* 44: 335-346.
- Franklin RB, Taylor DR, Mills AL. 2000. The distribution of microbial communities in anaerobic and aerobic zones of a shallow coastal plain aquifer. *Microbial Ecology* 38: 377-386.
- Fraterrigo JM, Balsler TC, Turner MG. 2006. Microbial community variation and its relationship with nitrogen mineralization in historically altered forests. *Ecology* 87: 570-579.
- Friswell MK, Gika H, Stratford IJ, Theodoridis G, Telfer B, Wilson ID, McBain AJ. 2010. Site and Strain-Specific Variation in Gut Microbiota Profiles and Metabolism in Experimental Mice. *PLoS ONE* 5: e8584. doi: 10.1371/journal.pone.0008584.
- Gandour M, Hessini K, Abdelly C. 2008. Understanding the population genetic structure of coastal species (*Cakile maritima*): seed dispersal and the role of sea currents in determining population structure. *Genetical Research* 90: 167-178.
- Gentile G, Fabiani A, Marquez C, Snell HL, Snell HM, Tapia W, Sbordoni V. 2009. An overlooked pink species of land iguana in the Galápagos. *Proceedings of the National Academy of Sciences* 106: 507-511.

- Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, Stackebrandt E, Van de Peer Y, Vandamme P, Thompson FL, Swings J. 2005. Re-evaluating prokaryotic species. *Nature Reviews Microbiology* 3: 733-739.
- Gil A, Reuda R. 2000. Modulation of intestinal microflora by specific dietary components. *Microbial Ecology in Health and Disease* 12: 31–39.
- Girones R, Ferrus MA, Alonso FL, Rodriguez-Manzano J, Calgua B, Correa AS, Hundesa A, Carratala A, Bofill-Mas S. 2010. Molecular detection of pathogens in water- The pros and cons of molecular techniques. *Water research* 44: 4325-4339.
- Green J, Bohannan BJM. 2006. Spatial scaling of microbial biodiversity. *Trends in Ecology and Evolution* 21: 501-507.
- Green JL, Bohannan BJM, Whitaker RJ. 2008. Microbial biogeography: From taxonomy to traits. *Science* 320: 1039-1042.
- Grehan J. 2001. Biogeography and evolution of the Galápagos: integration of the biological and geological evidence. *Biological Journal of the Linnean Society*. 74: 267-287.
- Guan TY, Holley RA. 2003. Pathogen survival in swine manure environments and transmission of human enteric illness – a review. *Journal of Environmental Quality* 32: 383-392.
- Hall BG, Ehrlich GD, Hu FZ, Fen Z. 2010. Pan-genome analysis provides much higher strain typing resolution than multi-locus sequence typing. *Microbiology* 156: 1060-1068.
- Harris LJ, Farber JN, Beuchat LR, Parish ME, Suslow TV, Garret EH, Busta FF. 2003. Outbreaks associated with fresh produce: incidence, growth, and survival of pathogens in fresh and fresh-cut produce. *Comprehensive Reviews in Food Science and Food Safety* 2: 78-141.
- Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK, Chantratita N, Gardte S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD. 2010. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327: 469-474.
- Hastings A. 1980. Disturbance, coexistence, history and competition for space. *Theoretical Population Biology* 18: 363-373.

- Hedlund BP, Staley JT. 2003. Microbial endemism and biogeography. In *Microbial Diversity and Bioprospecting* (Bull, AT ed.). American Society of Microbiology: Washington, DC.
- Hendry AP. 2002.  $Q(ST) \geq \text{not equal} < F\text{-}ST$ ? *Trends in Ecology & Evolution* 17: 502.
- Hillebrand H, Watermann F, Karez R, Berninger UG. 2001. Differences in species richness patterns between unicellular and multicellular organisms. *Oecologia* 126: 114-124.
- Hong PY, Wheeler E, Cann IKO, Mackie RI. In press. Phylogenetic analysis of the intestinal microbial community in herbivorous land and marine iguanas of the Galapagos Islands using 16S rRNA-based pyrosequencing. *ISME Journal*.
- Hubble SP. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press: NJ.
- Jackson MH. 2007. *Galapagos: A Natural History* (11<sup>th</sup> ed.). University of Calgary Press: Calgary, Alberta, Canada.
- Jenkins DG, Carey M, Czerniewska J, Fletcher J, Hether Y, Jones A, Knight S, Knox J, Long T, Mannino M, McGuire M, Riffle A, Segelsky S, Shappell L, Sterner A, Strickler T, Tursi R. 2010. A meta-analysis of isolation by distance: relic or reference standard for landscape genetics? *Ecography* 33: 315-320.
- Krebs CJ. 1994. *Ecology: The Experimental Analysis of Distribution and Abundance*. HarperCollins, New York, NY.
- Kritcher JC. 2006. Hideous Looking Creatures. *Galapagos: A Natural History*. pp. 81-97. Princeton University Press: Princeton, NJ.
- Lazarevic V, Whiteson K, Hernandez D, Francois P, Schrenzel J. 2010. Study of inter and intra-individual variation in the salivary microbiota. *BMC Genomics* 11: 523.
- Levins R, Culver D. 1971. Regional coexistence of species and competition between rare species. *Proceedings of the National Academy of Sciences* 68: 1246-1248.
- Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. 2005. Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences* 102: 11070–11075.
- Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI. 2008a. Evolution of mammals and their gut microbes. *Science* 320: 1647–1651.

- Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. 2008b. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nature* 6: 776-788.
- Losos JB, Ricklefs RE. 2009. Adaptation and diversification on islands. *Nature* 457: 830-836.
- MacArthur R, Wilson E. 1967. *The Theory of Island Biogeography*. Princeton University Press: Princeton, NJ.
- Macfarlane S, Macfarlane GT 2003. Food and large intestine. In *Gut Flora, Nutrition, Immunity and Health* (Fuller R, Perdigon G, eds.). pp. 24-51. Blackwell Publishing: Oxford, United Kingdom.
- Martiny JB, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR, Morin PJ, Naeem S, Øvreås L, Reysenbach A, Smith VH, Staley JT. 2006. Microbial biogeography: putting microorganisms on the map. *Nature Reviews* 4: 102-112.
- Mayr E. 1963. *Animal Species and Evolution*. Harvard University Press: Cambridge, MA.
- McKay JK, Latta RG. 2002. Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evolution* 17: 285-291.
- McMeekin TA, Hill C, Wagner M, Dahl A, Ross T. 2010. Ecophysiology of food-borne pathogens: Essential knowledge to improve food safety. *International Journal of Food Microbiology* 139: S64-S78.
- Michell and RJ, Monteils V, Zened A, Combes S, Cauquil L, Gidenne T, Hamelin J, Fortun-Lamothe L. 2009. Spatial and temporal variations of the bacterial community in the bovine digestive tract. *Journal of Applied Microbiology* 107: 1642-1650.
- Mueller S, Saunier K, Hanisch C, Norin E, Alm L, Midtvedt T, Cresci A, Silvi S, Orpianesi C, Verdenelli MC, Clavel T, Koebnick C, Zunft H-JF, Doré J, Blaut M. 2006. Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study. *Applied and Environmental Microbiology* 72: 1027-1033.
- Nabout JC, Siqueira T, Bini LM, Nogueira I de S. 2009. No evidence for environmental and spatial processes in structuring phytoplankton communities. *Acta Oecologia* 35: 720-726.

- Nekola JC, White PS. 1999. The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* 26: 867–878.
- O'Malley MA. 2007. The nineteenth century roots of 'everything is everywhere'. *Nature Reviews* 5: 647-651.
- Oda Y, Star B, Huisman LA, Gottschal JC, Forney LJ. 2003. Biogeography of the purple nonsulfur bacterium *Rhodospirillum rubrum*. *Applied and Environmental Microbiology* 69: 5186-5191.
- Pangaling E, Wang H, Venables M, Wallace A, Grant WD, Cowan DA, Jones BE, Ma Y, Ventosa A, Heaphy S. 2009. Microbial biogeography of six salt lakes in Inner Mongolia, China, and a salt lake in Argentina. *75*: 5750-5760.
- Papke RT, Ramsing NB, Bateson MM, Ward DM. 2003. Geographical isolation in hot springs cyanobacteria. *Environmental Microbiology* 5: 650-659.
- Papke RT, Ward DM. 2004. The importance of physical isolation to microbial diversification. *FEMS Microbiology Ecology* 48: 293–303.
- Parent CE, Caccone A, Petren K. 2008. Colonization and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical Transactions of the Royal Society B*. 363: 337-3361.
- Parnell JJ, Rompato G, Latta LC IV, Pfrender ME, Van Nostrand JD, Andersen G, Champine P, Ganesan B, Weimer BC. 2010. Functional Biogeography as Evidence of Gene Transfer in Hypersaline Microbial Communities. *PLoS ONE* 5: e12919. doi:10.1371/journal.pone.0012919
- Porcher E, Giraud T, Lavigne C. 2006. Genetic differentiation of neutral markers and quantitative traits in predominantly selfing metapopulations: confronting theory and experiments with *Arabidopsis thaliana*. *Genetical Research* 87: 1-12.
- Prosser JI, Bohannan BJM, Curtis TP, Ellis RJ, Firestone MK, Freckleton RP, Green JL, Green LE, Killham K, Lennon JJ, Osborn MA, Solan M, van der Gast CJ, Young JPW. 2007. The role of ecological theory in microbial ecology. *Nature Reviews Microbiology* 5: 384-392.
- Rabiu BA, Gibson GR. 2002. Carbohydrates: a limit on bacterial diversity within the colon. *Biol Rev* 2002; 77: 443–453.
- Rassmann K. 1997. Evolutionary age of the Galápagos iguanas predates the age of the present Galápagos Islands. *Molecular phylogenetics and evolution* 7: 158-172.

- Rassmann K, Tautz D, Trillmich F, Gliddon C. 1997a. The microevolution of the Galápagos marine iguana *Amblyrhynchus cristatus* assessed by nuclear and mitochondrial genetic analyses. *Molecular Ecology* 6: 437-452.
- Rassmann K, Trillmich F, and Tautz D. 1997b. Hybridization between the Galapagos land and marine iguana (*Conolophus subcristatus* and *Amblyrhynchus cristatus*) on Plaza Sur. *Journal of Zoology* 242: 729-739.
- Rawls JF, Mahowald MA, Ley RE, Gordon JI. 2006. Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell* 127: 423-433.
- Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N. 2010. The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environmental Microbiology* 12: 2885-2893.
- Rhodes MW, Kator H. 1988. Survival of *Escherichia coli* and *Salmonella* spp. in estuarine environments. *Applied and Environmental Microbiology* 54: 2902-2907.
- Ricklefs RE. 2004. A comprehensive framework for global patterns in biodiversity. *Ecological Letters* 7: 1-15.
- Robinson CJ, Bohannan BJM, Young VB. 2010. From structure to function: the ecology of host-associated microbial communities. *Microbiology and Molecular Biology Reviews* 74: 452-476.
- Rosselló-Mora R, Lucio M, Peña A, Brito-Echeverría J, López-López Aránzazu, Valens-Vadell Maria, Frommberger M, Antón J, and Schmitt-Kopplin P. 2008. Metabolic evidence for biogeographic isolation of the extremophile bacteria *Salinibacter ruber*. *Journal of the International Society of Microbial Ecology* 3: 242-253.
- Rousk J. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME Journal* 4: 1340-1351.
- Rozen Y, Belkin S. 2001 Survival of enteric bacteria in seawater. *FEMS Microbiology Reviews* 25: 513-529.
- Semenov AM, Kuprianov AA, van Bruggen AHC. 2010. Transfer of enteric pathogens to successive habitats as part of microbial cycles. *Microbial ecology* 60: 239-249.
- Shepherd SA, Hawkes MW. 2005. Algal food preferences and seasonal foraging strategy of the marine iguana, *Amblyrhynchus cristatus*, on Santa Cruz, Galápagos. *Bulletin of Marine Science* 77: 51-72.



- Sites JW, Scott KD, Guerra T, Iverson JB, Snell HL. 1996. Character congruence and phylogenetic signal in molecular and morphological data sets: A case study of the living iguanas. *Molecular Biology and Evolution* 13: 1087-1105.
- Slatkin M. 1974. Competition and regional coexistence. *Ecology* 55: 128-134.
- Snell HL, Snell HM, Tracy CR. 1984. Variation among populations of Galapagos land iguanas (*Conolophus*): contrasts of phylogeny and ecology. *Biological Journal, Linnean Society* 21: 185-208.
- Soininen J, McDonald R, Hillebrand H. 2007. The distance decay of similarity in ecological communities. *Ecography* 30: 3-12.
- Soininen J, Weckstrom J. 2009. Diatom community structure along environmental and spatial gradients in lakes and streams. *Fundamental and Applied Limnology* 174: 205-213.
- Steane DA, Conod N, Jones RC, Vaillancourt RE, Potts BM. 2006. A comparative analysis of population structure of a forest tree, *Eucalyptus globulus* (Myrtaceae), using microsatellite markers and quantitative traits. *Tree Genetics and Genomes* 2: 30-38.
- Sullivan A, Nord CE. 2006. Modifying the intestinal microbiota with antibiotics. In *Gastrointestinal Microbiology* (Ouweland AC, Vaughan EE, eds.). pp. 335-369. Taylor and Francis: New York, NY.
- Takacs-Vesbach C, Mitchell K, Jackson-Weaver O, Reysenbach A-L. 2008. Volcanic calderas delineate biogeographic provinces among Yellowstone thermophiles. *Environmental Microbiology* 10: 1681-1689.
- Tannock GW. 1995. Modification of the normal microbiota by diet, stress, antimicrobial agents, and probiotics. In *Gastrointestinal Microbiology, Vol. 2* (Mackie RL, White BA, Isaacson RE, eds.). pp. 434-456. Chapman and Hall: New York, NY.
- Telford RJ, Vandvik V, and Birks HJB. 2008. Dispersal limitations matter for microbial morphospecies. *Science* 312: 1015.
- Tilman D. 1982. Resource competition and community structure. Princeton University Press: Princeton, NJ.
- Tilman D. 1994. Competition and Biodiversity in spatially structured habitats. *Ecology* 75: 2-16

- Trillmich KGK, Trillmich F. 1986. Foraging strategies of the marine iguana, *Amblyrhynchus cristatus*. *Behavioral Ecology and Sociobiology* 18: 259-266.
- Tuomisto H, Ruokolainen R, Yli-Halla M. 2003. Dispersal, environment, and floristic variation of western Amazonian forests. *Science* 299: 241–244.
- Turnbaugh PJ, Hamady M, Yatsunencko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI. 2009a. A core gut microbiome in obese and lean twins. *Nature* 457: 480-484.
- Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. 2009b. The effect of diet on the human gut microbiome: A metagenomic analysis in humanized gnotobiotic mice. *Science Translational Medicine* 1: 6ra14.
- van der Gast CJ. 2008. Islands shaping thought in microbial ecology. *Advances in Applied Microbiology* 64: 167-182.
- Vos M, Velicer GJ. 2008. Isolation by distance in the spore-forming soil bacterium *Myxococcus xanthus*. *Current Biology* 18: 386-391.
- Wheeler E, IKO Cann, RI Mackie. 2011. Genomic fingerprinting and serotyping of *Salmonella* from Galápagos iguanas demonstrates island differences in strain carriage. *Environmental Microbiology Reports* 3: 166-173.
- Whitaker RJ, Grogan DW, Taylor JW. 2003. Geographic barriers isolate endemic populations of hyperthermophilic Archaea. *Science* 301: 976-978.
- Wikelski M, Gall B, Trillmich F. 1993. Ontogenetic changes in food intake and digestion rate of the herbivorous marine iguana (*Amblyrhynchus cristatus*, Bell). *Oecologia* 94: 373-379.
- Wright S. 1943 Isolation by distance. *Genetics* 28: 114-138
- Wyles JS, Sarich VM. 1983. Are the Galápagos iguanas older than the Galápagos? Molecular evolution and colonization of the archipelago. *Patterns of Evolution in Galápagos Organisms*. (Bowman RI, Berson M, Leviton AE, eds.). pp. 177-186. American Association for the Advancement of Science: San Francisco, CA.
- Yannarell AC, EW Triplett. 2004. Within- and between-lake variability in the composition of bacterioplankton communities: investigations using multiple spatial scales. *Applied and Environmental Microbiology* 70: P214-P223.
- Zhang N, Blackwell M. 2002. Population structure of dogwood anthracnose fungus. *Phytopathology* 92: 1276–1283.

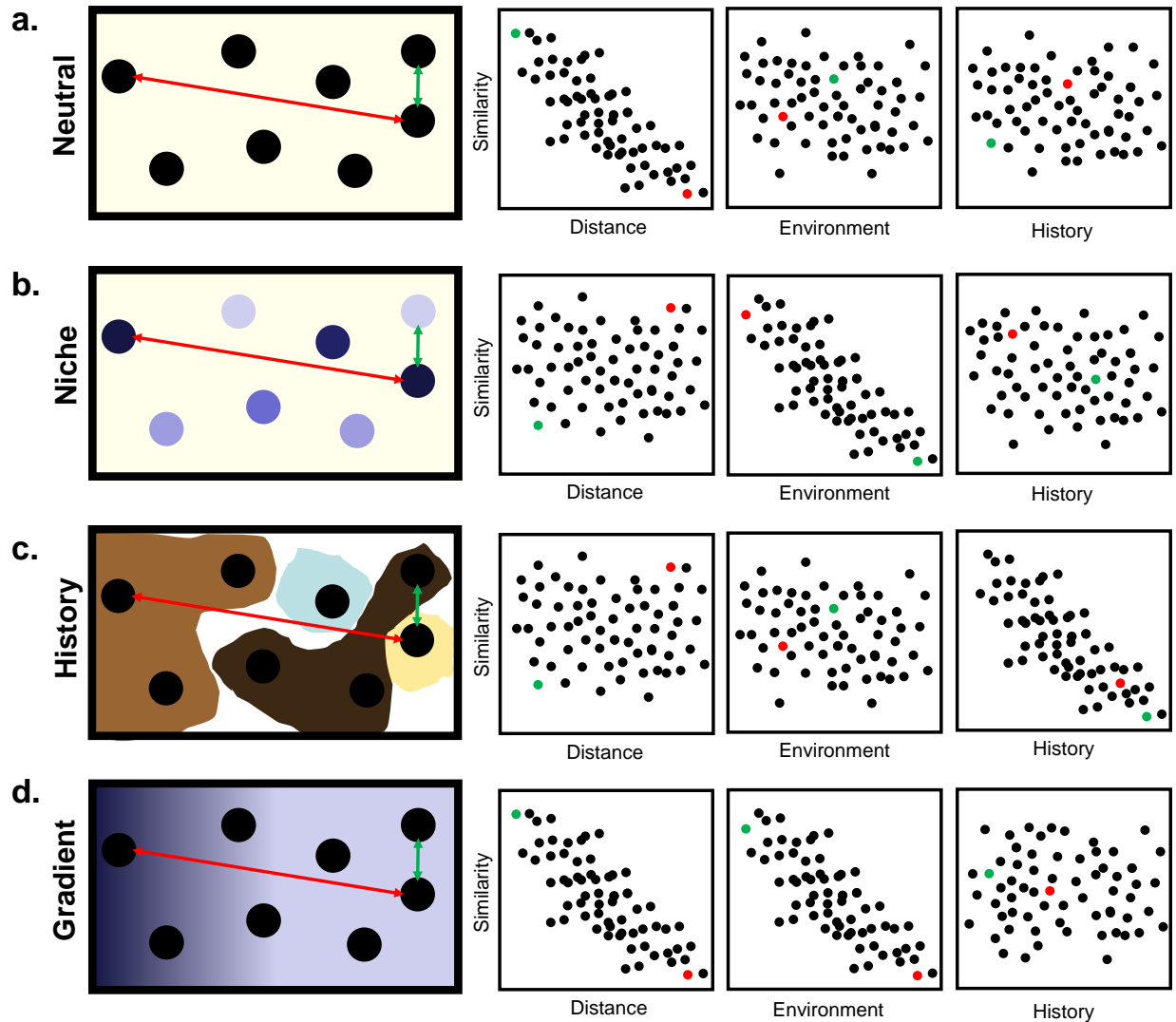


Figure 1.1. Hypothetical habitat patches and expected genetic (or community) similarity patterns. Expected correlations of similarity are shown plotted against geographic distance (“Distance”) between patches (black or blue dots), difference in measure of an environmental condition (“Environment”) and differences in time of colonization (“History”) for each of four possible drivers: a.) neutral effects (i.e. dispersal limitation), b.) niche effects (i.e. environmental selection, shades of blue represent different values of an environmental variable c.) temporal history of colonization (indicated by different color patches representing different colonization histories), and d.) environmental selection gradient (an environmental factor which changes over distance). Throughout, the comparison of two habitat patches to the same patch are provided (red arrows and dots and green arrows and dots) to aid with visual comparison between the hypothetical habitat situation and associated correlation graphs.

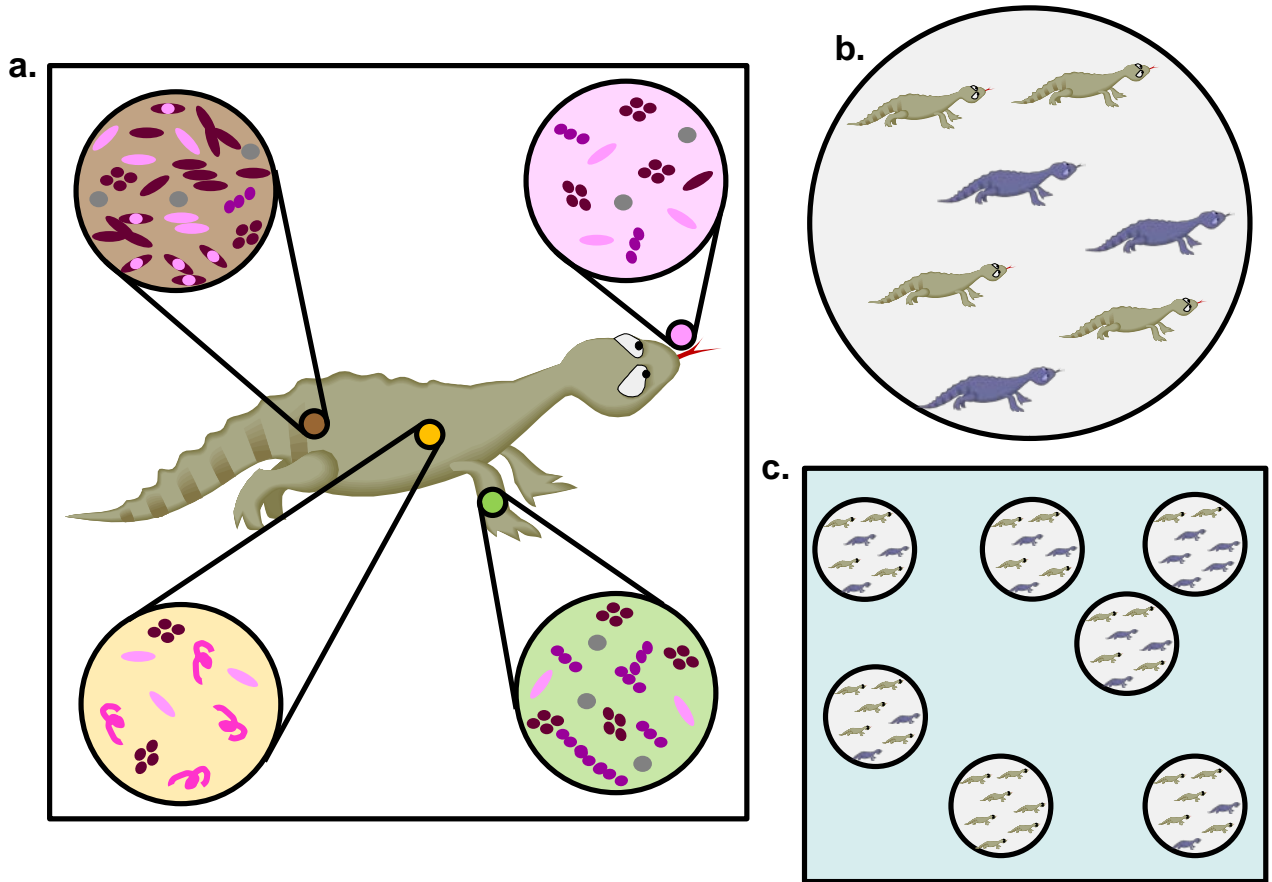


Figure 1.2. Biogeographical scales in host-associated microbial communities. Host-associated microbial communities may express biogeography within an individual host due to differential selection or dispersal to different locations within the host body (a.), among conspecific or heterospecific individuals within a single habitat patch (b.) or among allopatric host populations or communities (c.).

## **CHAPTER 2: BIOGEOGRAPHICAL PATTERNS IN THE GASTROINTESTINAL BACTERIAL COMMUNITIES OF GALÁPAGOS IGUANAS SUGGEST COMPLEX METACOMMUNITY PROCESSES**

### **Chapter summary**

Diet can strongly influence gastrointestinal microbial communities through species sorting. Ecological drift can also produce stochastic differences in community composition via variation in local microbial exposures and chance losses of gastrointestinal taxa in allopatric host populations. In this chapter, we investigated biogeographical patterning of enteric bacterial communities of Galápagos marine (*Amblyrhynchus cristatus*) and land iguanas (*Conolophus pallidus* and *C. subcristatus*) using 16S rRNA gene pyrosequencing of fecal community genomic DNA. Host ecotype (marine versus land iguana) was the strongest determinant of bacterial community structure but within each host the geographical location of sampling sites also exhibited strong effects on community similarity and structure. Microbial community dissimilarity was strongly associated with geographic distance among sampling sites, with more proximate sites having more similar communities. However, within site dynamics disrupted this trend on Isla Plaza Sur, where marine iguana bacterial communities were uniquely more diverse (both within and between host individuals) than observed in other marine iguana populations. Within this site, marine and land iguana bacterial communities were also more similar to each other than other sympatric marine and land iguana population pairs, possibly due to higher overlap in habitat use on this extremely small island. Finally, a subset of genera were strongly associated with one or a few sampling sites while they were often absent or found at low abundances at other sites. These patterns together suggest that host-bacterial interactions and host dietary differences likely are dominant drivers of gastrointestinal community composition at the broad functional level, such as might be represented by genus-level diversity patterns. However, site differences in bacterial composition are also likely driven by ecological drift, potentially leading to geographic variation in the ecological function of these microbial communities.

### **Introduction**

Strong selective forces, such as diet or exposure to antibiotic medications for example, can have notable impacts on gastrointestinal microbial communities (e.g. Dietary effects: Tannock 1995, Gil and Reuda 2000, Rabiú and Gibson 2002, Macfarlane and Macfarlane 2003, Turnbaugh et al. 2009b; Antibiotic effects: Tannock 1995, Sullivan and Nord 2006, Antonopoulos et al. 2009 - note that throughout we use the word “selection” in reference to environmental forces leading to species sorting within a community, rather than natural selection of fitter genotypes within a populations). Diet in particular can act as a strong selective force on gastrointestinal microbial communities, contributing to differences in microbial community form and function, both within and among host species (Ley et al. 2008a, Ley et al. 2008b,

Turnbaugh et al. 2009a, Turnbaugh et al. 2009b). Within mammalian taxa, for example, dietary differences have been found to cross phylogenetic relationships, with microbial community composition determined more strongly by dietary groups (i.e. herbivores versus omnivores versus carnivores) than by host taxonomic relationships or geographical location (Ley et al. 2008a).

The potential for such host-associated selection to result in convergent enteric communities should be substantially limited by any existing differences among populations in environmental exposures (i.e. differences in the regional or local species pool). However, the manner by which such environmental or historical forces shape host-associated microbial communities is less well-described (Robinson et al. 2010). A number of studies have demonstrated strong differences among host-associated microbial populations and communities, including gastrointestinal microbes, due to environmental influences (Dethlefsen et al. 2007, Frank et al. 2007, Ley et al. 2008a, Costello et al. 2009, Turnbaugh et al. 2009a, Benson et al. 2010) and due to the geographical distribution of host populations (Benson et al. 2010, Fallani et al. 2010, Mueller et al. 2006, Hong et al. In press). Such effects can even supersede dietary influences in some cases. For example, a geographical gradient in gut microbiota composition and similarity was detected in a study of European infant enteric microbial communities, an effect that was stronger than other explanatory variables including feeding method (breast milk or formula), antibiotic exposures, and differing mode of delivery (cesarean or vaginal; Fallani et al. 2010). This suggests that differences in regional strain or species pools may contribute strongly to gastrointestinal community structure, despite or in addition to strong environmental selective pressures.

These local or regional metacommunity dynamics may shape gastrointestinal communities on both contemporary and historical time scales. Environmental sources such as water, food or soils may provide local sources of exposure for gastrointestinal microbes (e.g. Lee et al. 2010, Al-Ahmad et al. 2010, Troyer 1983) that may vary considerably among sites on a contemporary time scale. Additionally, microbial communities may also drift apart, as a host species sequentially colonizes new geographical areas; conceptually parallel to genetic drift, taxa may be lost by chance during the host founding events or new taxa may be gained over time from either local exposures or differential adaptive radiations of existing taxa in different populations

resulting in divergence from ecological drift (Hubbell 2001). While patterns of contemporary versus historical losses of taxa may be difficult to distinguish in many cases, the gain of bacteria via processes at these two time scales should create recognizable patterns in community similarities. When local microbial species pools vary among host populations, these differences in exposures could result in increased similarity of gastrointestinal community composition to local sources of enteric bacterial diversity. One potential source of new bacterial exposure is heterospecific host taxa, which are likely due to host differentiated microbial communities. If contact rates between heterospecific host species varies geographically, for example through varying overlap in habitat use, this could lead to decreased differentiation in gut communities between host species, while simultaneously creating increased variation between host populations of the same species. In contrast, historical ecological drift could lead to geographically distinct gastrointestinal microbial communities in allopatric host populations which parallels host population colonization, degree of contact, or population genetic patterns.

Island systems have long contributed to the study of how geographic isolation contributes to ecological and evolutionary processes in both microbial and microbial systems (Darwin 1859, Mayr 1963, MacArthur and Wilson 1967, Diamond and May 1976, Krebs 1994, van der Gast 2008, Losos and Ricklefs 2009). The Galápagos Islands off the coast of Ecuador are particularly famous for these contributions and are host to 27 reptile species that are distributed throughout the island chain, including two ecologically distinct “ecotypes” of large herbivorous lizards, the marine and land iguanas.

Land and marine iguana populations are both found on a number of sites throughout the island chain and their distributions overlap at some locations (Jackson 2007). Marine iguanas (*Amblyrhynchus cristatus*) are locally abundant and widely distributed throughout the islands, living in dense colonies along the rocky coasts and feeding on marine algae (Trillmich and Trillmich 1986, Wikelski et al. 1993, Shepherd and Hawkes 2005). Land iguanas (species *Conolophus subcristatus* and *C. pallidus* included in this study) are less commonly found in dense groups (Kritcher 2006). *C. subcristatus* is more widely distributed among the islands than *C. pallidus*, which is entirely endemic to the island of Santa Fe. Land iguanas rarely frequent the coastlines

(Snell et al. 1984) and primarily eat *Opuntia* cactus pads and leaves and flowers of lantanas and cordia plants (Christian et al. 1984). The ecological distinctions between marine and land iguanas generally result in relatively low overlap in habitat use on all but the smallest and most densely occupied sympatric islands, such as Isla Plaza Sur, where the two ecotypes have been shown to hybridize (Rassmann et al. 1997).

Despite these ecological distinctions, land and marine iguanas are phylogenetically closely related, having diverged from a common Ctenosaur ancestor approximately 10 to 20 million years ago (Wyles & Sarich 1983, Sites *et al.* 1996, Rassmann 1997, Cogger and Zweifel 1998). Yet despite this relatively close genetic kinship, previous work in this system has demonstrated that extreme dietary divergence over the divergent evolution of these two groups has resulted in a significant departure in the structure and composition of their gastrointestinal microbial communities. In fact, land iguana gastrointestinal communities are more similar in composition to those of green iguanas (*Iguana iguana*) from El Salvador and Galápagos giant tortoises (*Geochelone gigantea*) than to marine iguanas (Hong et al. In press). It has been proposed that this pattern may result from the dietary preferences of these species: land iguanas, green iguanas, and giant tortoises all share a similar diet of fibrous terrestrial plant material (Christian et al. 1984, Troyer 1982, Hatt et al. 2005) while marine iguanas are unique in their preference for macrophytic algae (Shepherd and Hawkes 2005).

The unique biology, phylogeny, and biogeography of the Galápagos iguanas provides a valuable system for studying the effects of differing host gastrointestinal environments and host population isolation on enteric bacterial community composition and structure. In this chapter, we investigated enteric bacterial community patterns in allopatric and sympatric populations of Galápagos marine and land iguanas using 16S rRNA gene pyrosequencing of fecal community genomic DNA. We evaluated microbial community diversity differences at the broad taxonomic level of the bacterial genus to focus this study on biogeographical patterns which might directly relate to differences in gastrointestinal function. We first hypothesized that gastrointestinal bacterial communities at this level of taxonomic organization would be most strongly influenced by the selective force of host ecotype (land or marine iguana), presumptively driven by factors such as diet and host-bacterial interactions. However, we also expected



that among populations of each host there would be strong geographical patterns in the overall similarity of these communities, possibly reflecting either contemporary variability in local exposures or historical patterns of divergence among host populations.

## **Methods**

### *Study design and sampling sites*

The Galápagos Islands off the coast of Ecuador are host to two unique ecotypes of large iguanas - land iguanas (*C. pallidus* and *C. subcristatus*, included in this study, and a third species *C. marthae*) and marine iguanas (*Amblyrhynchus cristatus*). Fecal specimens were obtained from marine iguanas at five locations on four islands and from land iguanas at three sites on three islands (*C. pallidus* on Isla Santa Fe and *C. subcristatus* at the other sites; Figure 2. 1). Free-living marine and land iguanas were observed from a distance for defecation and then fecal specimens were collected from the ground using a sterile wooden applicator. Individual animals were intentionally selected from different areas of each site and samples at each site were collected over a short period of time (24-48 hours) to avoid repeated sampling of the same individual. The slow gastrointestinal transit rates of herbivorous reptiles (e.g. estimated to be 5-10 days for marine iguanas, Wikelski et al. 1993), also reduces the likelihood of observing a single individual defecating twice within a short sampling period. Fecal specimens were obtained from both iguana types on Isla Plaza Sur ( $N_{\text{land}}=5$ ,  $N_{\text{marine}}=5$ ), Isla Fernandina ( $N_{\text{land}}=5$ ,  $N_{\text{marine}}=5$ ), and Isla Santa Fe ( $N_{\text{land}}=6$ ,  $N_{\text{marine}}=5$ ), where both species are present and from only marine iguanas from the sampling sites on Isla San Cristobal (La Loberia  $N=5$ , Punta Carola  $N=5$ ), where land iguanas are not present (Figure 2.1). Samples were placed in sterile plastic tubes, stored at 4°C during travel and transported to Urbana, Illinois for storage at -20°C until analysis could be completed. The majority of samples included in this study were collected in September, 2009. Due to the small number of samples obtained for marine iguanas on Isla Santa Fe in this 2009 sampling, 3 land iguana and 3 marine iguana fecal samples collected in 2005 which had been stored at -70°C were also included in these analyses.

### *DNA extraction*

Community genomic DNA was extracted from frozen fecal samples using the Ultraclean Soil DNA isolation kit (MoBio, Carlsbad, CA) using the manufacturer's protocol, with minor modifications. Briefly, 0.2g of fecal material was placed in the manufacturer provided lysis tube and then pre-incubated with 12  $\mu$ l each of 100mg/ml lysozyme and 1mg/ml achromopeptidase at 37°C for 1 hour before proceeding with chemical lysis and extraction according to the manufacturer's directions. We evaluated the purity and yield of DNA using a Qubit fluorometer (Invitrogen, CA) and DNA concentrations were then standardized to 10ng/ $\mu$ l before use as template for polymerase chain reaction (PCR).

### *Barcoded PCR and 454 pyrosequencing*

Pyrosequencing allows for the simultaneous processing of many thousands of heterogeneous PCR amplicons derived from a one or more mixed microbial community samples. The resulting sequence reads, or pyrotags, can then be used for detailed assessment of microbial community composition, structure and hypothesized functional capacities (Andersson et al. 2008, Hugenholtz and Tyson 2008, McKenna et al. 2008, Roesch et al. 2007). The 16S rRNA gene was amplified from community genomic DNA using bacterial-specific forward 519F (5'-Fusion A-Barcode -CAGCMGCCGCGTAATWC-3') and reverse 926R (5'-Fusion B-Barcode- CCGTCAATTCMTTTRAGTT-3') primer pairs with a unique sequence barcodes to identify each sample (Parameswaran et al. 2007, Hamady et al. 2008, [www.roche.com](http://www.roche.com)). PCR amplification mixtures comprised 100 ng of genomic DNA, 25  $\mu$ l of Premix F (Epicentre Biotechnologies, WI), 200 nM (each) of forward and reverse primers, 0.5 U of Ex *Taq* DNA polymerase (Takara Bio Inc., Japan), and the volume added up to 50  $\mu$ l with molecular-biology grade water. PCR with 30 cycles of thermal program (denaturation, 95 °C for 30 s; annealing, 55 °C for 45 s; and extension, 72 °C for 60 s) was performed. Amplicons were gel purified using the Wizard DNA purification kit (Promega, WI). DNA concentrations of the final concentrated PCR products were quantified by Qubit fluorometer (Invitrogen, CA).

### *Pyrotags handling and analysis*

PCR amplicon pyrosequencing was performed on a 454 FLX Titanium sequencer (Roche Applied Biosciences, Switzerland) by the Roy J. Carver Biotechnology Center, University of Illinois at Urbana Champaign. A total of 245,685 16S DNA pyrotags were obtained, which were sorted by primer barcodes to form independent pyrotag libraries for each sample. Each pyrotag library was processed to remove low-quality reads as described elsewhere for these data (Hong et al. In press). Primer, barcode, and adaptor sequences were removed and the pyrotags were processed using the RDP Pipeline Initial Process (Cole et al. 2009). Secondary structure alignments of processed pyrotags were produced using RDP Infernal (Cole et al., 2009). The aligned pyrotags were visually checked with Jalview (<http://www.jalview.org>), and manual adjustments were performed as necessary to ensure alignment quality. After processing and trimming the pyrotags had an average read length of 370 nt. No significant differences in the average number of sequences were detected for host ecotypes or sampling sites (overall average number of sequences/sample: 5177 +/-1218). A complete description of these pyrosequencing data and of diet-linked differences among host species in these bacterial communities is published elsewhere (Hong et al. In press). RDP Classifier was used for taxonomical assignments of the aligned 16S pyrotags at 95% confidence level (Cole et al. 2009). Primer-E was used to compile the relative abundance of sequences within each sample at the genus level for statistical analysis of community patterns. This genus-level relative abundance matrix was used for all analyses and will be referred to as genus-level OTUs or genera throughout. A more detailed taxonomic analysis of these data are described elsewhere (Hong et al. In press, Chapter 3).

### *Statistical analysis*

Rarified genus richness measures were estimated for each sample at a standardized sampling effort of 6000 sequences per sample using the rarify function in the vegan package in R (Oksanen et al. 2010). Differences among these estimated richness values were evaluated between sites and host ecotypes using the model fit function in JMP 8 (SAS Institute, Cary, NC).

Using the bacterial genera relative abundance matrix, differences among groups were first visualized using ordination by non-metric multidimensional scaling (NMDS) using the metaMDS function in the vegan package of R (Oksanen et al. 2010, R Development Team 2010). NMDS plots were made for the full dataset and for each host ecotype (land or marine) independently. To explore the effects of explanatory variables (host ecotype and sampling site) on patterns detected in these ordinations, we used the adonis function in the vegan package in R (Oksanen et al. 2010) to perform a non-parametric multivariate analysis of variance (npMANOVA). This method (also referred to as a permutation MANOVA) partitions the sums of squares of distance matrices among treatments and has relaxed assumptions relative to traditional MANOVA. Significance in the permutation tests was determined by comparing the observed effects against 5000 random permutations of the data for each model run independently. A number of independent models were run to explore the main effects of host ecotype and sampling site and their interaction (only for models including the three sympatric islands), sampling site geography (latitude and longitude) and sampling site and geography within each host ecotype.

A complete Bray-Curtis dissimilarity matrix was calculated using the vegdist function in the vegan package in R (Oksanen et al. 2010). This matrix was then used to calculate average dissimilarities within and between groups of interest (e.g. comparing differences in average dissimilarity within host ecotypes versus between sympatric host ecotypes) using the general linear model fitting function and when warranted, multiple comparisons with a Tukey's HSD adjustment in JMP8 (SAS Institute Inc, Cary, NC).

Geographic distances between all pairs of sampling sites were estimated as the Euclidean distance in kilometers using an approximate latitude and longitude of the sampling points. A linear correlation of the relationship between this geographic distance and the average dissimilarity of marine iguana gastrointestinal communities was calculated using JMP8. Only samples from marine iguanas were included in this geographic analysis due to the small number of sampling sites available for land iguanas. This analysis was performed for all marine iguana sampling sites and again excluding Isla Plaza Sur, due to the significantly different microbial community patterns and variation detected in marine iguana populations at this site.

The relationship of relative abundance of each genus-level OTU to host ecotype, sampling site and host populations was evaluated by univariate ANOVA models for each genus-level OTU independently. Two models were explored: first we evaluated a fully crossed factorial model for host type and sampling site including the three sites containing both land and marine iguanas, and secondly we evaluated a model of site effects for marine iguanas samples only. For each effect within each model, we adjusted the maximum  $\alpha$  value of 0.1 with a false discovery rate adjustment (Benjamini and Hochberg 1995). A heat-map was produced to demonstrate the patterns for taxa with significant relationships to the variables tested. The abundance (number of reads) of each genus was scaled relative to the total abundance of that genus across samples to allow for visualization of both very abundant and relatively rare taxa in the same figure.

All other statistical and graphing procedures were performed in either JMP 8 (SAS Institute, Cary, NC) or R statistical language (R Development Core Team 2008).

## **Results**

### *Richness patterns*

The number of genus-level OTUs in land iguanas (mean OTUs +/- standard error: LI 37.9 +/- 1.9) was higher than in marine iguanas (MI 22.8 +/- 0.8; t-test, 2-tailed  $p < 0.001$ ). Land iguanas did not differ among islands in genus-level OTU richness ( $F = 33.4 \pm 4.4$ ,  $P = 42.8 \pm 3.4$ ,  $S = 37.5 \pm 1.2$ : ANOVA  $p = 0.157$ ,  $R^2 = 0.248$ ), but marine iguanas did differ, with Plaza Sur (28.6 +/- 2.2 OTUs) having higher genus-level richness relative to the other sites ( $C = 21.2 \pm 1.2$ ,  $F = 21.4 \pm 1.2$ ,  $L = 20.2 \pm 1.2$ ,  $S = 22.6 \pm 1.5$ : ANOVA  $p = 0.003$ ,  $R^2 = 0.53$ ). Due to concerns over small sample sizes for some of these comparisons both parametric ANOVA or t-tests and comparable non-parametric tests (i.e. Kruskal-Wallis or Mann-Whitney) were also performed; non-parametric analyses supported the parametric findings presented here.

### *Host and site effects on bacterial community composition*

Both host and sampling site demonstrated significant effects on gut bacterial community composition, but host effects were stronger than site effects (Table 2.1; Host  $R^2 = 0.315$ , Site  $R^2 = 0.123$ ). While interaction effects could not be tested for the full dataset due to an imbalanced design, a subset of this dataset including both host

ecotypes sampled on three sites (Isla Santa Fe, Isla Plaza Sur, and Isla Fernandina) demonstrated similar host and site patterns with a significant interaction term (Table 2.1). Gastrointestinal bacterial communities also demonstrated geographical patterning, with both latitude and longitude of sampling sites explaining portions of the variation unexplained by host species. When host species were evaluated separately, land iguana microbial communities varied by site and by longitude but not latitude, while marine iguanas had strong site, latitude, and longitude effects (Table 2. 1). These patterns are demonstrated graphically by NMDS plots of all samples (Figure 2.2 a) and for each species individually (Figure 2.2 b and c, land and marine iguanas, respectively).

The geographical nature of the site differences in gut bacterial communities was also detected in the significant correlation between the dissimilarity of marine iguana gastrointestinal communities and geographical distance among sites (Figure 2.3). When the five marine iguana sampling sites were included in this analysis, the correlation was non-significant ( $R^2=0.005$ ,  $p=0.652$ ). In contrast, when site comparisons including Plaza Sur were removed from this analysis, community dissimilarity and geographic distance were strongly correlated ( $R^2=0.493$ ,  $p<0.001$ ).

#### *Within and among island patterns*

The community dissimilarity within a host species at each sampling site was much lower than dissimilarity between marine and land iguanas where these species cohabitate (either MI versus MI or LI versus LI as compared to MI versus LI; Within Species: average dissimilarity= $0.273 \pm 0.086$ ,  $N_{\text{comparisons}}=41$ ; Between Species: average dissimilarity= $0.514$ ,  $N_{\text{comparisons}}=16$ ; T-test  $p<0.001$ ; see also Figure 2.4). Across sampling sites, there were no significant differences in the between-host species dissimilarity among islands; similarly, within host dissimilarity comparisons for land iguanas did not differ by population (Figure 2.4, right-hand and middle columns, respectively). However, within-host comparisons for marine iguanas demonstrated that Isla Plaza Sur gastrointestinal communities were much more variable (i.e. had higher within-site dissimilarity) than gastrointestinal communities of the other marine iguana populations sampled (Figure 2.4 panel a, left-hand column).

For marine iguanas, the dissimilarity of gastrointestinal communities within a site was lower than the dissimilarity between sites (Within site: average

dissimilarity=0.280+/-0.108,  $N_{\text{comparisons}}=25$ ; Between site: average=0.337+/-0.116,  $N_{\text{comparisons}}=100$ ; T-test  $p=0.0247$ ). The unique variability of marine iguana gastrointestinal communities on Plaza Sur was also evident in comparison of community dissimilarity among sites. While within-site community dissimilarity was lowest of each site comparison for Isla Fernandina, Isla Santa Fe, La Lobería and Punta Carola (Figure A.1, i.e. communities within sites were more similar than the comparison of those communities to those from other sites), the within-site dissimilarity of Isla Plaza Sur was higher than these other within site comparisons and was also higher than dissimilarities among Isla Plaza Sur and the other sites (Figure A.1).

The average relative abundance of each genus was calculated for each host-by-site combination and these values were analyzed by linear correlation between sympatric marine and land iguana populations within each site containing both of these host ecotypes. Average relative genus abundances between sympatric host pairs were significantly correlated at all three sites (138 genera-level groups;  $p < 0.001$  for all three cases; correlation coefficients:  $r_F=0.648$ ,  $r_P=0.867$ ,  $r_S=0.642$ ). In addition, the correlation on Isla Plaza Sur was significantly higher than for either Isla Fernandina or Isla Santa Fe (tested as the interaction between island and marine iguana relative abundance in an ANCOVA on land iguana relative abundance,  $p < 0.001$ ).

#### *Microbial taxa associated with host and site differences*

Univariate models, controlled for the false discovery rate, demonstrated 16 genera, mostly Firmicutes in the order Clostridiales, that were significantly related to host ecotype identity (Figure 2.5 panel a). The majority of these genera-level groupings were highly abundant in land iguana communities but absent or low abundance in marine iguana samples. However, three groups (*Anaerosporobacter*, an unclassified Clostridiaceae and *Desulfosporosinus*) showed the opposite pattern, with higher abundance in marine versus land iguanas. Site effects revealed two taxa that were uniquely abundant in fecal samples from Isla Fernandina and were rarer in samples from Isla Santa Fe and Isla Plaza Sur (*Anaerosporobacter* and an unclassified Planctomycetaceae, Figure 2.5 panel b). In contrast, an unclassified group of Lachnospiraceae was more abundant on Isla Santa Fe and Isla Plaza Sur, relative to Isla Fernandina (Figure 2.5 panel b). The interaction between host ecotype and site tested

whether the difference in abundance between ecotypes itself varied among sites (i.e. genera that showed large differences between ecotypes at some sites but weak differences at others). Five groups showed such a pattern, generally caused by taxa that had higher abundance in marine versus land iguanas at Isla Fernandina and Isla Santa Cruz, with weaker differences on Plaza Sur (Figure 2.5 panel c). A number of taxa showed strong variation in their relative abundance across sites when considering all five marine iguana populations (Figure 2.5 panel d). Notably, a number of genera from within the order Lactobacillales were much more abundant in marine iguana samples from Punta Carola on Isla San Cristóbal, while two groups of Ruminococcaceae and two groups of Enterobacteriaceae were at higher abundance on Isla Plaza Sur and Isla Santa Fe, respectively.

## **Discussion**

### *Geographical patterning of enteric communities*

Acquisition of a new diet can be a fundamental driver of animal speciation and can mediate major changes in gut microbiota as hosts make these transitions (Ley et al. 2008a, Ley et al. 2008b). As has been previously reported in this system, land iguanas fecal samples demonstrated significantly higher genus richness than marine iguana samples, possibly reflecting the increased complexity of microbial functions required for the digestion of fibrous vegetation versus macrophytic algae (Hong et al. In press). Further, host ecotype (as a likely proxy for dietary selection) was the most significant explanatory factor for microbial community composition, with obvious clustering of samples by host ecotype on ordination. Host species differences have been demonstrated to be stronger than spatial effects at the gastrointestinal bacterial community level in other systems (Ley et al. 2008a).

However, despite these strong host ecotype effects, a significant portion of the variation in genera relative abundances was also attributable to site differences, especially when analyzing each host ecotype independently. These site differences were also specifically geographical, with community composition demonstrating significant relationships to site longitude for both land and marine iguanas and also to latitude for marine iguanas. Qualitatively, Isla Santa Fe and Isla Plaza Sur land iguana biota were more closely associated to each other in ordination space than to Isla Fernandina, which



corresponded to a significant longitude effect in the permutation model. The geographical patterns for marine iguana samples were stronger, in part due to the larger number of sampling sites and wider geographical variation in the placement of these populations. A specifically geographical site pattern was present for four of five sampling sites for marine iguana populations, as demonstrated by a significant correlation between geographic and community distances.

While these patterns are striking the underlying mechanism for these observations is not entirely clear. There are at least two possible explanations for the noted distance-decay in community similarity of marine iguana enteric biota. Typically, this type of distance-decay relationship for microbial population or community patterns is attributed to dispersal limitations resulting in proximate sites sharing strains or taxa more commonly than more distant sites (e.g. Cho and Tiedje 2000, Whitaker et al. 2003, Pommier et al. 2007, Pagaling et al. 2009). These dispersal limitations may reflect contemporary processes of bacterial exchange among more proximate sites. Alternatively, these geographical patterns could potentially reflect a deeper history of divergence among gastrointestinal communities in parallel to host divergence during colonization. In the Galápagos Islands, population genetic patterns of many endemic organisms suggests a pattern of colonization which may parallel the geological history of these unique islands (i.e. the “progression hypothesis” sensu Wagner and Funk 1995, Parent et al. 2008). Population genetics in many plant and animal species in the Galápagos, including marine iguanas and other endemic reptile taxa, appear to follow a directional geographical gradient. Ancestral populations are generally located on the geologically older eastern and central islands, and younger, more recently derived populations on geologically younger islands to the north-west (as reviewed in Parent et al. 2008; Beheregaray et al. 2004, Kizirian et al. 2004, Ciofi et al. 2006, Benavides et al. 2007, Tzika et al. 2008, Gentile et al. 2009, Steinfartz et al. 2009).

#### *Local exposures can strongly influence composition*

One sampling site, Isla Plaza Sur, deviated strongly from the general geographical trends among other sampling sites. Marine iguanas from Isla Plaza Sur had much higher variability in community composition, with relatively high average dissimilarities both within site and among sites. In fact, the variability among any two marine iguanas on

Isla Plaza Sur was more similar to comparisons between host ecotypes than for within host ecotype comparisons. In addition, marine iguanas on Isla Plaza Sur had higher within-individual genus richness than the other marine iguana populations. These findings suggest that marine iguanas on Isla Plaza Sur have exposure to a more diverse microbial source pool and that incorporation of this increased diversity is occurring stochastically among individual animals, resulting in higher population level variability in composition. If assembly of diversity is primarily driven by niche selection, then community composition should be similar among similar selective environments (Chase 2003, Tuomisto et al. 20003, Dornelas et al. 2006). In contrast, if stochastic dispersal is a primary driver of community composition, there should be considerable site-to-site variation as bacterial groups arrive and establish by chance at each site independently (Condit et al. 2002, Chave and Leigh 2002). Thus, the high variability in community composition among marine iguana individuals at Isla Plaza Sur (analogous to "sites") suggests a strong role for stochastic community assembly on this island. High variability among individuals suggests that while marine iguanas can acquire genera that are potentially not part of their typical biota, these acquired genera may be transient or relatively unimportant to digestive function, resulting in wide variability between individuals in gut community composition.

Further evidence in these data suggests that the source of this novel exposure for marine iguanas on Isla Plaza Sur may be increased ecological overlap with cohabitating land iguanas. Specifically, the average relative abundances of each bacterial genus between sympatric marine and land iguanas were significantly more strongly correlated on this island compared to Isla Santa Fe or Isla Fernandina. This suggests that there is a large overlap in the bacterial groups shared by sympatric host ecotypes on this island relative to the two other island pairs.

The ecological distinctions between marine and land iguanas generally result in relatively low overlap in habitat use. Marine iguanas (*Amblyrhynchus cristatus*) are locally abundant and widely distributed throughout the islands, living in dense colonies along the rocky coasts to feed (Trillmich and Trillmich 1986, Wikelski et al. 1993 Shepherd and Hawkes 2005). Land iguanas are less commonly found in dense groups and rarely frequent the shorelines (Snell et al. 1984, Kritcher 2006). However, Isla Plaza Sur's small land area represents an extreme case where land (*C. subcristatus*) and

marine iguanas are both found at high population densities and overlap considerably due to the dearth of truly interior land (island areas in km<sup>2</sup>: Isla Plaza Sur=0.13, Isla Santa Fe=24, and Isla Fernandina=642; Jackson 2007). Alternatively, on the larger islands, land iguanas (*C. pallidus* on Santa Fe and *C. subcristatus* on Fernandina) are found at relatively low densities and experience far less overlap in habitat use with the dense populations of marine iguanas at these sites due to a general preference for upland interior habitat (personal observation E. Wheeler, see also Snell et al. 1984, Kritcher 2006).

While marine iguana gastrointestinal communities on Isla Plaza Sur were uniquely variable and rich at the genus level, the land iguana gastrointestinal communities on Isla Plaza Sur were not notably different from other con-specific populations at the genus level. The large differences in community characteristics between marine and land iguanas, in particular the latter's notably higher genus richness, may be key for understanding this asymmetry. Both empirical and theoretical work have suggested that more diverse communities have an increased capacity to resist or recover from perturbations and resist invasions by novel species (i.e. have increased community stability, MacArthur 1955, McNaughton 1977, Pimm 1984, Tilman et al. 1998, Levine and D'Antonio 1999, Lehman and Tilman 2000, Naeem et al. 2000, Ives and Carpenter 2007). Further, the phylogenetic or functional diversity of a community, rather than the total number of species, may be a key factor for determining resistance to invasion (Wardle et al. 2008, Hooper and Dukes 2010).

These proposed ecological mechanisms are also qualitatively supported by the univariate analysis of genera which strongly contribute to species versus site effects. The majority of genera which were identified as strongly associated with host ecotype difference were abundant in land iguanas and were rare to absent in marine iguanas, suggesting that marine iguana communities may represent a nested subset of diversity which is broadly common to both ecotypes. The analysis of genera associated with site effects also suggested stochastic contributions to differences in community composition at other sites. While Isla Plaza Sur showed the strongest suggestions of local stochastic site effects on gastrointestinal communities, a number of additional genera were identified as uniquely present at one or more sites. For example members of the order Lactobacillales were more abundant at Punta Carola on Isla San Cristóbal, members of

the genus Ruminococcaceae were more abundant on Isla Plaza Sur, and Isla Fernandina, and members of the genus Enterobacteriaceae were more abundant on Isla Santa Fe. In the case of Isla Plaza Sur, patterns in these data suggest that increased interaction among hosts may contribute to differential carriage of Ruminococcaceae in marine iguanas at this site. Ruminococcaceae were rarely detected in other marine iguanas samples, but were quite common and relatively abundant in land iguana communities (Hong et al. In press). Determining if other local sources of exposure can contribute to differences in bacterial carriage among host populations will require additional study and concurrent evaluation of potential sources of bacterial exchange. What also cannot be ascertained by this study is whether taxa uniquely shared between host species within a site, as on Isla Plaza Sur, are able to establish permanent, self-sustaining populations or if these taxa are maintained in the atypical host by source-sink dynamics (sensu Leibold et al 2004). Such subtleties are worthy of further exploration.

## **Conclusions**

These findings suggest that gastrointestinal communities in wild Galápagos iguanas may experience complex metacommunity dynamics across both spatial and temporal scales. A metacommunity is defined as a set of local communities which are linked by dispersal of interacting species (Gilpin and Hanski 1991; Wilson 1992). A number of extant theoretical metacommunity models, including species sorting, patch-dynamics, mass effects and neutral models (Leibold et al. 2004), may apply to aspects of the community patterns seen in this study. For example, while environmental species sorting may be a driving force in selecting for community members critical for digestive functions in the face of dietary variation, either a patch-dynamics model or neutral ecological drift model may better explain the geographical patterns detected within each host species. Additionally, local site effects, especially at Isla Plaza Sur, demonstrate that sufficient exposure to novel sources of bacterial diversity can dramatically shift enteric gastrointestinal communities, although these new community members may only persist in some enteric habitats by on-going source-sink dynamics. Empirical work has demonstrated that these different metacommunity models may act simultaneously on

different subsets of a single community, resulting in complex models of community assembly and turnover (Driscoll and Lindenmayer, 2009).

Gastrointestinal systems are highly complex and may provide an interesting system for exploring metacommunity dynamics, as they can provide multiple scales at which to explore these issues – from the individual animal to the population and then to host metapopulations with unique evolutionary and ecological histories. What may prove the most valuable outcomes of such studies is not determining which models apply, but rather what types of taxa are affected by each mechanism and how the resulting changes to enteric communities affect digestive efficiencies, health and fitness of the hosts.

### **Chapter Acknowledgements**

We would like to express appreciation to the Galápagos National Park and Washington Tapía, and the Charles Darwin Research Station and Sonja Cisneros and Paulina Couenberg for facilitating sample collection. Many thanks to Augusto G. Haz Beltran, Lenin Cruz Beldon and the crew of the Pirata for help with sample collection. We also wish to thank the Department of Animal Sciences and Nancy Henry for administrative support. Many thanks to Peiyong Hong for her laboratory work and expertise in acquiring the 454 pyrosequencing data. Thanks to Richard Lankau for statistical support and constructive comments on early drafts of this chapter. Thanks to Isaac Cann and Bryan White for access to laboratory resources.

Funding for this work was provided by the USDA National Research Initiative (Proposal No. AG-2008-35206-18784 and by the US Environmental Protection Agency Science to Achieve Results Fellowship program. This chapter was developed with support from a STAR Research Assistance Agreement No. 91684301-1 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this chapter.

## References

- Al-Ahmad A, Maier J, Follo M, Spitzmüller B, Wittmer A, Hellwig E, Hübner J, Jonas D. 2010. Food-borne Enterococci integrate into oral biofilm: An *in vivo* study. *Journal of Endodontics* 36: 1813-1819
- Andersson AF, Lindberg M, Jakobsson H, Backhed F, Nyren P, Engstrand L. 2008. Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLoS ONE* 3: e2836. doi: 10.1371/journal.pone.0002836.
- Antonopoulos DA, Huse SM, Morrison HG, Schmidt TM, Sogin ML, Young VB. 2009. Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. *Infection and Immunity* 77: 2367-2375.
- Beheregaray LB, Gibbs JP, Havill N, Fritts TH, Powell JR, Caccone A. 2004. Giant tortoises are not so slow: rapid diversification and biogeographic consensus in the Galápagos. *Proceedings of the National Academy of Sciences* 101: 6514-6519.
- Benavides E, Baum R, McClellan D, Sites JW Jr. 2007. Molecular phylogenetics of the lizard genus *microlophus* (Squamata: Tropicuridae): Aligning and retrieving indel signal from nuclear Introns. *Systems Biology* 56: 776-797.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)* 57: 289–300.
- Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, Zhang M, Oh PL, Nehrenberg D, Hua K, Kachman SD, Moriyama EN, Walter J, Peterson SA, Pomp D. 2010. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *PNAS* 107: 18933-18938.
- Chase JM. 2003. Community assembly: When should history matter? *Oecologia* 236: 489-498.
- Chave J, Leigh EG, Jr. 2002. A spatially explicit neutral model of beta-diversity in tropical forests. *Theoretical Population Biology* 62: 153–168.
- Cho JC, Tiedje JM. 2000. Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Applied and Environmental Microbiology* 66: 5448-5456.
- Christian KA, Tracy CR, Porter WP. 1984. Diet, digestion and food preferences of Galápagos land iguanas. *Herpetologica* 40: 205-212.

- Ciofi C, Wilson GA, Beheregaray LB, Marquez C, Gibbs JP, Tapia W, Snell HL, Caccone A, Powell JR. 2006. Phylogeographic history and gene flow among giant Galápagos tortoises on southern Isabela Island. *Genetics* 172: 1727-1744.
- Cogger HG, Zweifel RG (eds.). 1998. *Encyclopedia of reptiles and amphibians*. 2<sup>nd</sup> ed. Academic Press: San Diego.
- Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syex-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM. 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* 37: D141-D145.
- Condit R, Pitman N, Leigh EG, Jr, Chave J, Terborgh J, Foster RB, Núñez P, Aguilar S, Valencia R, Villa G, Muller-Landau HC, Losos E, Hubbell SP. 2002. Beta-diversity in tropical forest trees. *Science* 295: 666–669.
- Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. 2009. Bacterial community variation in human body habitats across space and time. *Science* 326: 1694–1697.
- Darwin C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray: London.
- Dethlefsen L, McFall-Ngai M, Relman DA. 2007. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* 449: 811–818.
- Diamond JM, May RM. 1976. Island biogeography and the design of natural reserves. In *Theoretical Ecology: Principles and Applications* (RM May, ed.), pp. 163–186. Blackwell Scientific: Oxford, UK.
- Dornelas M, Connolly SR, Hughes TP. 2006. Coral reef diversity refutes the neutral theory of biodiversity. *Nature* 440: 80–82.
- Driscoll DA, Lindenmayer DB. 2009. Empirical tests of metacommunity theory using an isolation gradient. *Ecological Monographs* 79: 485-501.
- Fallani M, Young D, Scott J, Norin E, Amarri S, Adam R, Aguilera M, Khanna S, Gil A, Edwards CA, Doré J, and other members of the INFABIO Team. 2010. Intestinal microbiota of 6-week-old infants across Europe: Geographic influence beyond delivery mode, breast-feeding and antibiotics. *Journal of Pediatric Gastroenterology and Nutrition* 51: 77-84.

- Frank DN, St. Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proceedings of the National Academy of Sciences* 104: 13780–13785.
- Gentile G, Fabiani A, Marquez C, Snell HL, Snell HM, Tapia W, Sbordoni V. 2009. An overlooked pink species of land iguana in the Galápagos. *Proceedings of the National Academy of Sciences* 106: 507-511.
- Gil A, Reuda R. 2000. Modulation of intestinal microflora by specific dietary components. *Microbial Ecology in Health and Disease* 12: 31–39.
- Gilpin ME, Hanski IA. 1991. *Metapopulation dynamics: Empirical and theoretical investigations*. Academic Press: London, United Kingdom.
- Hamady MJ, Walker J, Harris JK, Gold NJ, Knight R. 2008. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nature Methods* 5: 235-237.
- Hatt JM, Clauss M, Gisler R, Liesegang A, Wanner M. 2005. Fiber digestibility in juvenile Galápagos tortoises (*Geochelone nigra*) and implications for the development of captive animals. *Zoo Biology* 24: 185-191.
- Hong PY, Wheeler E, Cann IKO, Mackie RI. In press. Phylogenetic analysis of the intestinal microbial community in herbivorous land and marine iguanas of the Galapagos Islands using 16S rRNA-based pyrosequencing. *ISME Journal*.
- Hooper DU, Dukes JS. 2010. Functional composition controls invasion success in a California serpentine grassland. *Journal of Ecology* 98: 764-777.
- Hubbell SP. 2001. *The unified neutral theory of biodiversity and biogeography*. – Princeton University Press: Princeton, NJ.
- Hugenholtz P, Tyson GW. 2008. Microbiology: metagenomics. *Nature* 455: 481-483.
- Ives AR, Carpenter SR. 2007. Stability and diversity of ecosystems. *Science* 317: 58–62.
- Jackson MH. 2007. *Galápagos: A Natural History* (11<sup>th</sup> ed.). University of Calgary Press: Calgary, Alberta, Canada.
- Kizirian D, Trager A, Donnelly MA, Wright JW. 2004. Evolution of Galápagos Island Lava Lizards (Iguania: Tropiduridae: Microlophus). In *Molecular Phylogenetics and Evolution* 32: 761-769.



- Krebs CJ. 1994. *Ecology: The Experimental Analysis of Distribution and Abundance*. HarperCollins: New York, NY.
- Kritcher JC. 2006. Hideous Looking Creatures. In *Galapagos: A Natural History*. pp. 81-97. Princeton University Press: Princeton, NJ.
- Lee J, Lee CS, Hugunin KM, Maute CJ, Dysko RC. 2010. Bacteria from drinking water supply and their fate in gastrointestinal tracts of germ-free mice: A phylogenetic comparison study. *Water research* 44: 5050-5058.
- Lehman CL, Tilman D. 2000. Biodiversity, stability, and productivity in competitive communities. *American Naturalist* 156: 534–552.
- Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, Holt RD, Shurin JB, Law R, Tilman D, Loreau M, Gonzalez A. 2004. The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* 7: 601-613.
- Levine JM, D'Antonio CM. 1999. Elton revisited: a review of evidence linking diversity and invisibility. *Oikos* 87: 15-26.
- Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI. 2008a. Evolution of mammals and their gut microbes. *Science* 320: 1647–1651.
- Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. 2008b. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nature* 6: 776-788.
- Losos JB, Ricklefs RE. 2009. Adaptation and diversification on islands. *Nature* 457: 830-836.
- MacArthur R. 1955. Fluctuations of animal populations and a measure of community stability. *Ecology* 36: 533–536.
- MacArthur RH, Wilson EO. 1967. *The Theory of Island Biogeography*. Princeton University Press: Princeton, NJ.
- Macfarlane GT, Macfarlane LE. 2009. Acquisition, evolution and maintenance of the normal gut microbiota. *Digestive Diseases* 27: 90-98.
- Macfarlane S, Macfarlane GT 2003. Food and large intestine. In *Gut Flora, Nutrition, Immunity and Health* (Fuller R, Perdigon G, eds.). pp 24-51. Blackwell Publishing: Oxford: United Kingdom.
- Mayr E. 1963. *Animal Species and Evolution*. Harvard University Press: Cambridge, MA.

- McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, Liu Z, Lozupone CA, Hamady M, Knight R, Bushman FD. 2008. The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathogens* 4: e20. doi: 10.1371/journal.ppat.0040020.
- McNaughton SJ. 1977. Diversity and stability of ecological communities: a comment on the role of empiricism in ecology. *American Naturalist* 111: 515–525.
- Mueller S, Saunier K, Hanisch C, Norin E, Alm L, Midtvedt T, Cresci A, Silvi S, Orpianesi C, Verdenelli MC, Clavel T, Koebnick C, Zunft H-JF, Doré J, Blaut M. 2006. Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study. *Applied and Environmental Microbiology* 72: 1027–1033.
- Naeem S, Knops JMH, Tilman D, Howe KM, Kennedy T, Gale S. 2000. Plant diversity increases resistance to invasion in the absence of covarying extrinsic factors. *Oikos* 91: 97-108.
- Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2010. *vegan: Community Ecology Package*. R package version 1.17-4. <http://cran.r-project.org/>, <http://vegan.r-forge.r-project.org/>.
- Pagaling E, Wang HZ, Venables M, Wallace A, Grant WD, Cowan DA, Jones BE, Ma YH, Ventosa A, Heaphy S. 2009. Microbial biogeography of six salt lakes in inner Mongolia, China and a salt lake in Argentina. *Applied and Environmental Microbiology* 75: 5750-5760.
- Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M, Fire AZ. 2007. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Research* 35: e130.
- Parent CE, Caccone A, Petren K. 2008. Colonisation and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical Transactions of the Royal Society B* 363: 3347-3361.
- Pimm SL 1984. The complexity and stability of ecosystems. *Nature* 307: 321–326.
- Pommier T, Canback B, Riemann L, Bostrom KH, Simu K, Lundberg P, Tunlid A, Hagstrom A. 2007. Global patterns of diversity and community structure in marine bacterioplankton. *Molecular Ecology* 16: 867-880.

- R Development Core Team. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Rabiu BA, Gibson GR. 2002. Carbohydrates: a limit on bacterial diversity within the colon. *Biol Rev* 2002; 77: 443–453.
- Rassmann K. 1997. Evolutionary age of the Galápagos iguanas predates the age of the present Galápagos Islands. *Molecular Phylogenetics and Evolution* 7: 158-172.
- Rassmann K, Trillmich F, Tautz D. 1997. Hybridization between the Galapagos land and marine iguana (*Conolophus subcristatus* and *Amblyrhynchus cristatus*) on Plaza Sur. *Journal of Zoology* 242: 729-739.
- Robinson CJ, Bohannan BJM, Young VB. 2010. From structure to function: the ecology of host-associated microbial communities. *Microbiology and Molecular Biology Reviews* 74: 452-476.
- Roesch LF, Fulthorpe RR, Riva A, Casella G, Hadwin AK, Kent AD, Daroub SH, Camargo FA, Farmerie WG, Triplett EW. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME Journal* 1: 283-290.
- Shepherd SA, Hawkes MW. 2005. Algal food preferences and seasonal foraging strategy of the marine iguana, *Amblyrhynchus cristatus*, on Santa Cruz, Galápagos. *Bulletin of Marine Science* 77: 51-72.
- Sites JW, Scott KD, Guerra T, Iverson JB, Snell HL. 1996. Character congruence and phylogenetic signal in molecular and morphological data sets: A case study of the living iguanas. *Molecular Biology and Evolution* 13: 1087-1105.
- Snell HL, Snell HM, Tracy CR. 1984. Variation among populations of Galapagos land iguanas (*Conolophus*): contrasts of phylogeny and ecology. *Biological Journal, Linnean Society* 21: 185-208.
- Steinfartz S, Glaberman S, Lanterbecq D, Russello MA, Rosa S, Hanley TC, Marquez C, Snell HL, Snell HM, Gentile G, Dell’Olmo G, Powell AM, Caccone A. 2009. Progressive colonization and restricted gene flow shape island-dependent population structure in Galápagos marine iguana (*Amblyrhynchus cristatus*). *BMC Evolutionary Biology* 9: 297.

- Sullivan A, Nord CE. 2006. Modifying the intestinal microbiota with antibiotics. In *Gastrointestinal Microbiology* (Ouweland AC, Vaughan EE, eds.). pp 335-369. Taylor and Francis: New York, NY.
- Tannock GW. 1995. Modification of the normal microbiota by diet, stress, antimicrobial agents, and probiotics. In *Gastrointestinal Microbiology, Vol 2* (Mackie RI, White BA, Isaacson RE, eds.). pp 434-456. Chapman and Hall, New York, NY.
- Tilman D, Lehman CL, Bristow CE. 1998. Diversity-stability relationships: Statistical inevitability or ecological consequence? *American Naturalist* 151: 277-282.
- Trillmich KGK, Trillmich F. 1986. Foraging strategies of the marine iguana, *Amblyrhynchus cristatus*. *Behavioral Ecology and Sociobiology* 18: 259-266.
- Troyer K. 1982. Diet selection and digestion in *Iguana iguana*: the importance of age and nutrient requirements. *Oecologia* 61: 201-207.
- Troyer K. 1983. Behavioral acquisition of the hindgut fermentation system by hatchling *Iguana iguana*. *Behavioral Ecology and Sociobiology* 14: 189-193.
- Tuomisto H, Ruokolainen R, Yli-Halla M. 2003. Dispersal, environment, and floristic variation of western Amazonian forests. *Science* 299: 241-244.
- Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI. 2009a. A core gut microbiome in obese and lean twins. *Nature* 457: 480-484.
- Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. 2009b. The effect of diet on the human gut microbiome: A metagenomic analysis in humanized gnotobiotic mice. *Science Translational Medicine* 1: 6ra14.
- Tzika AC, Rosa SFP, Fabiani A, Snell HL, Snell HM, Marquez C, Tapia W, Rassmann K, Gentile G, Milinkovitch MC: Population genetics of Galápagos land iguana (genus *Conolophus*) remnant populations. *Molecular Ecology* 2008. 17: 4943-4952.
- Wagner WL, Funk VA. 1995. *Hawaiian biogeography: evolution on a hot spot archipelago*. Smithsonian Institution Press: Washington, DC.
- Wardle DA, Lagerström A, Nilsson MC. 2008. Context dependent effects of plant species and functional groups loss on vegetation invisibility across an island area-gradient. *Journal of Ecology* 96: 1174-1186.
- Whitaker RJ, Grogan DW, Taylor JW. 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* 301: 976-978.

Wikelski M, Gall B, Trillmich F. 1993. Ontogenetic changes in food intake and digestion rate of the herbivorous marine iguana (*Amblyrhynchus cristatus*, Bell). *Oecologia* 94: 373-379.

Wilson DS. 1992. Complex interactions in metacommunities with implications for biodiversity and higher levels of selection. *Ecology* 73: 1984-2000.

Wyles JS, Sarich VM. 1983. Are the Galápagos iguanas older than the Galápagos? Molecular evolution and colonization of the archipelago. In *Patterns of Evolution in Galápagos Organisms* (Bowman RI, Berson M, and Leviton AE, eds.). pp 177-186. American Association for the Advancement of Science: San Francisco, CA.

## Tables and Figures

Table 2.1. Permutation MANOVA analyses for genus distributions

Dataset	Source	Permutation MANOVA		
		F	R <sup>2</sup>	p
Galápagos Iguanas	Host	23.956	0.315	<0.001**
	Site	2.339	0.123	0.014**
Galápagos Iguanas	Host	26.254	0.358	<0.001**
	Latitude	2.754	0.038	0.039**
	Longitude	2.100	0.029	0.084*
Galápagos Iguanas, SPF <sup>§</sup>	Host	25.721	0.414	<0.001**
	Site	2.175	0.070	0.051*
	Host*Site	3.425	0.110	0.008**
Land Iguanas, All	Site	2.962	0.313	<0.001**
Land Iguanas, All	Latitude	1.382	0.073	0.222
	Longitude	1.950	0.103	0.050**
Marine Iguanas, All	Site	2.948	0.3709	<0.001**
Marine Iguanas, All	Latitude	3.190	0.107	0.015**
	Longitude	2.389	0.080	0.039**

\* Significant at  $\alpha=0.10$ ; \*\* Significant at  $\alpha=0.05$

<sup>§</sup> Sample subset including marine and land iguanas from the islands of Santa Fe, Plaza Sur and Fernandina.

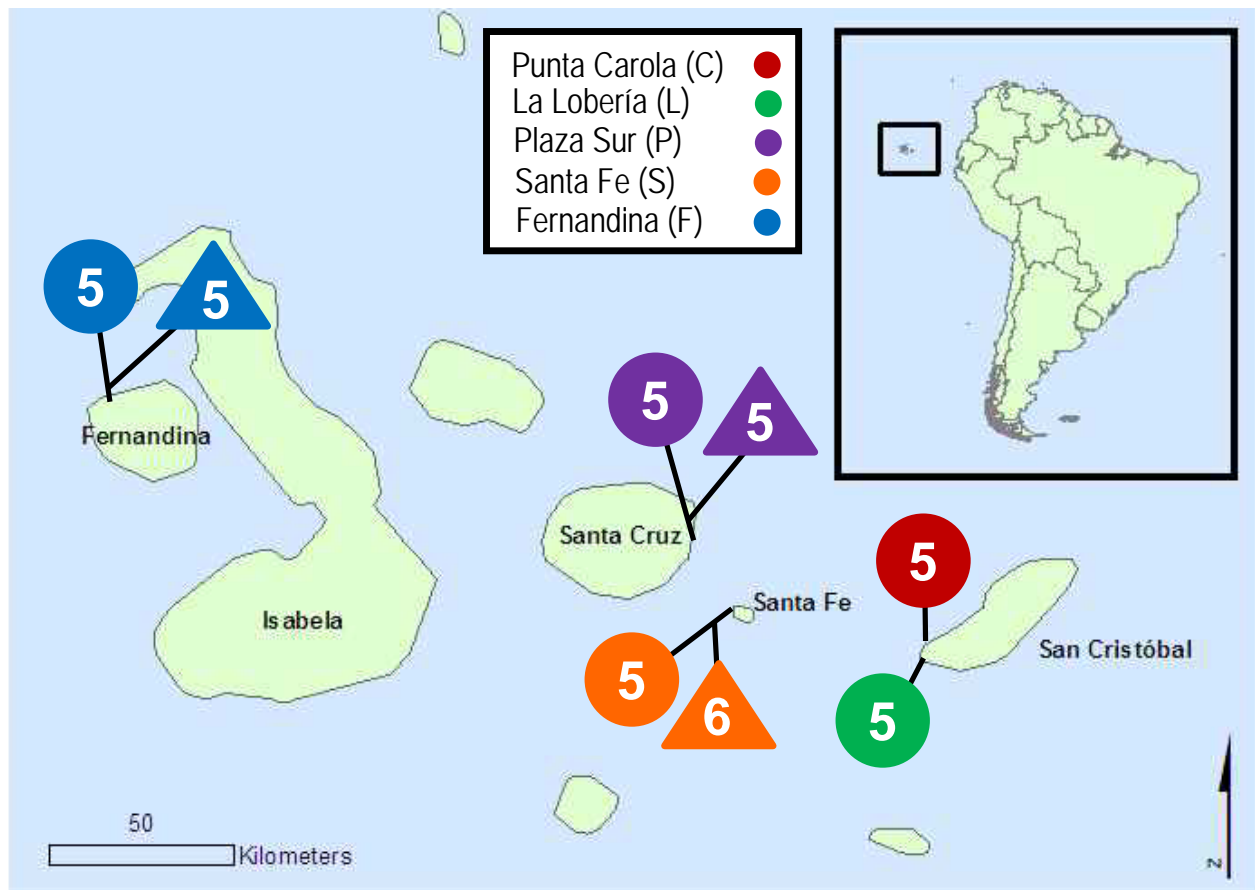


Figure 2.1 shows sampling locations for feces from marine iguanas (MI: circles; from sites C, L, S, P, and F) and land iguanas (LI: triangles; from sites S, P, and F). The number of samples analyzed for each species at each site is indicated inside the symbols. Samples from Santa Fe (S) were pooled from two sampling trips (MI: 3 from 2005 and 3 from 2009; LI: 3 from 2005 and 3 from 2009).

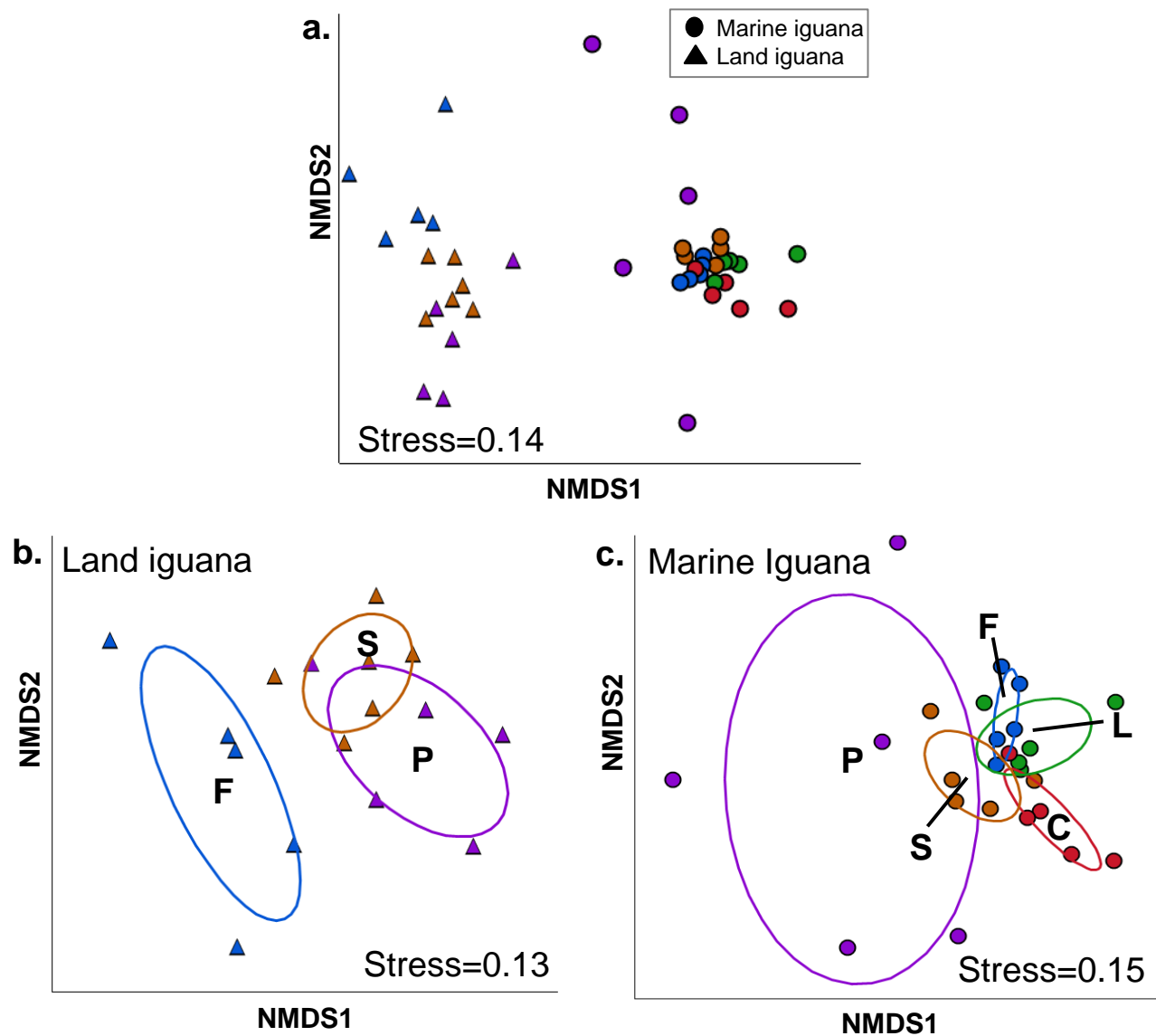


Figure 2.2. NMDS plots for all samples (a), for land iguanas only (b, symbolized by filled triangles) and for marine iguanas only (c, symbolized by filled circles). Sampling sites are represented by different colors and letter codes (Red/C=Punta Carola on Isla San Cristóbal, Green/L=La Lobería on Isla San Cristóbal, Purple/P=Isla Plaza Sur, Orange/S=Isla Santa Fe, Blue/F=Isla Fernandina). The ovals represent 50% contour ellipses by sampling site with colors corresponding to the points.

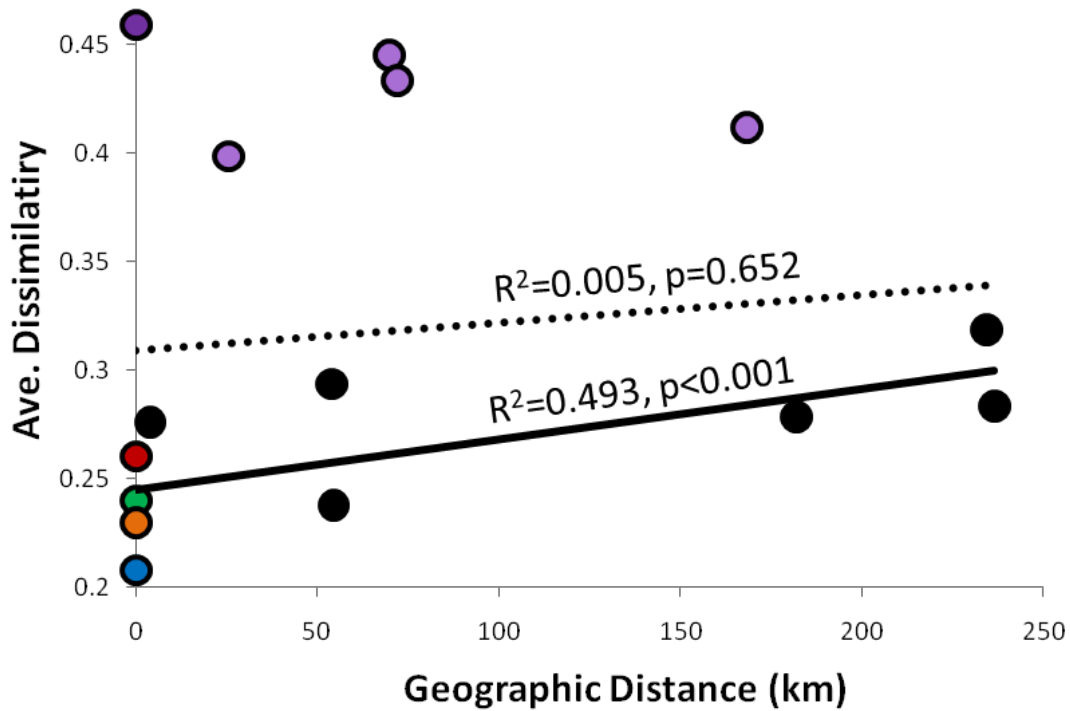


Figure 2.3. Correlation between Euclidean geographic distance and marine iguana microbial community dissimilarity. The colored points along the y-axis represent the average dissimilarity of microbial communities within each sampling site (colors corresponding to Figure 2.1), black points represent between site comparisons for all sites except Plaza Sur, and light purple points represent comparisons between Plaza Sur and all other sampling sites. The correlation among all site comparisons was non-significant (dotted line). However, removal of Plaza Sur comparisons (dark and light purple points) resulted in a significant relationship between geographic distance and microbial community dissimilarity for the remaining sites (solid line).



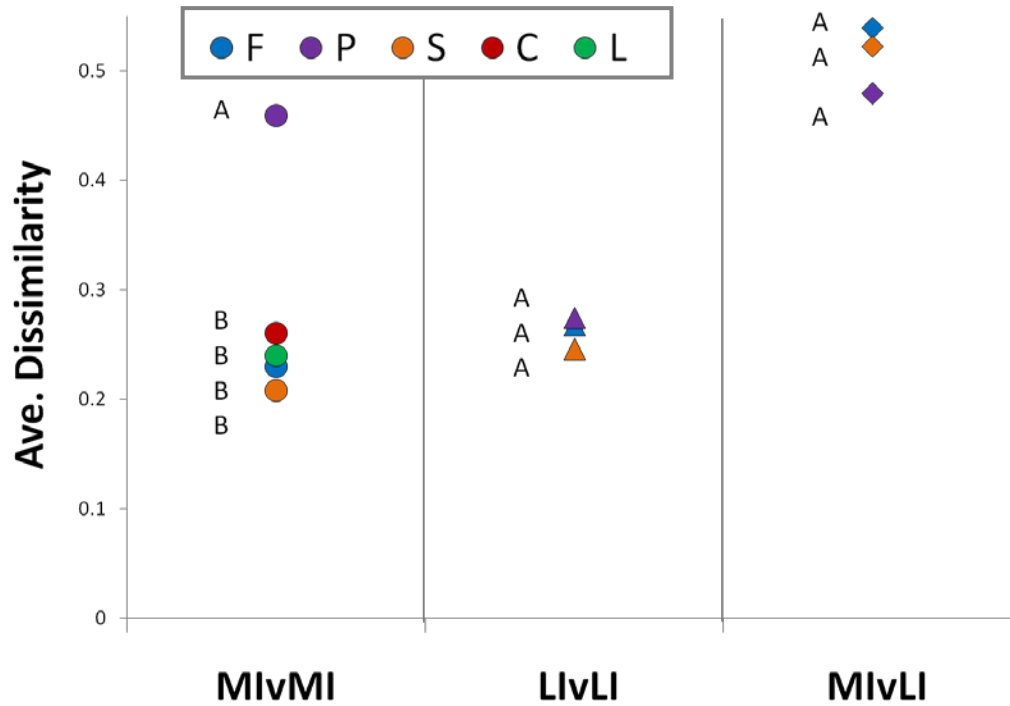


Figure 2.4. Patterns of microbial community dissimilarity within and between host species at each sampling site. Columns represent within host species - within marine iguanas (MivMI) and within land iguanas (LivLI) – or between marine and land iguanas (MivLI). Small capital letters next to each point represents significant differences among groups within each column (Tukey’s HSD adjustment for multiple comparisons).

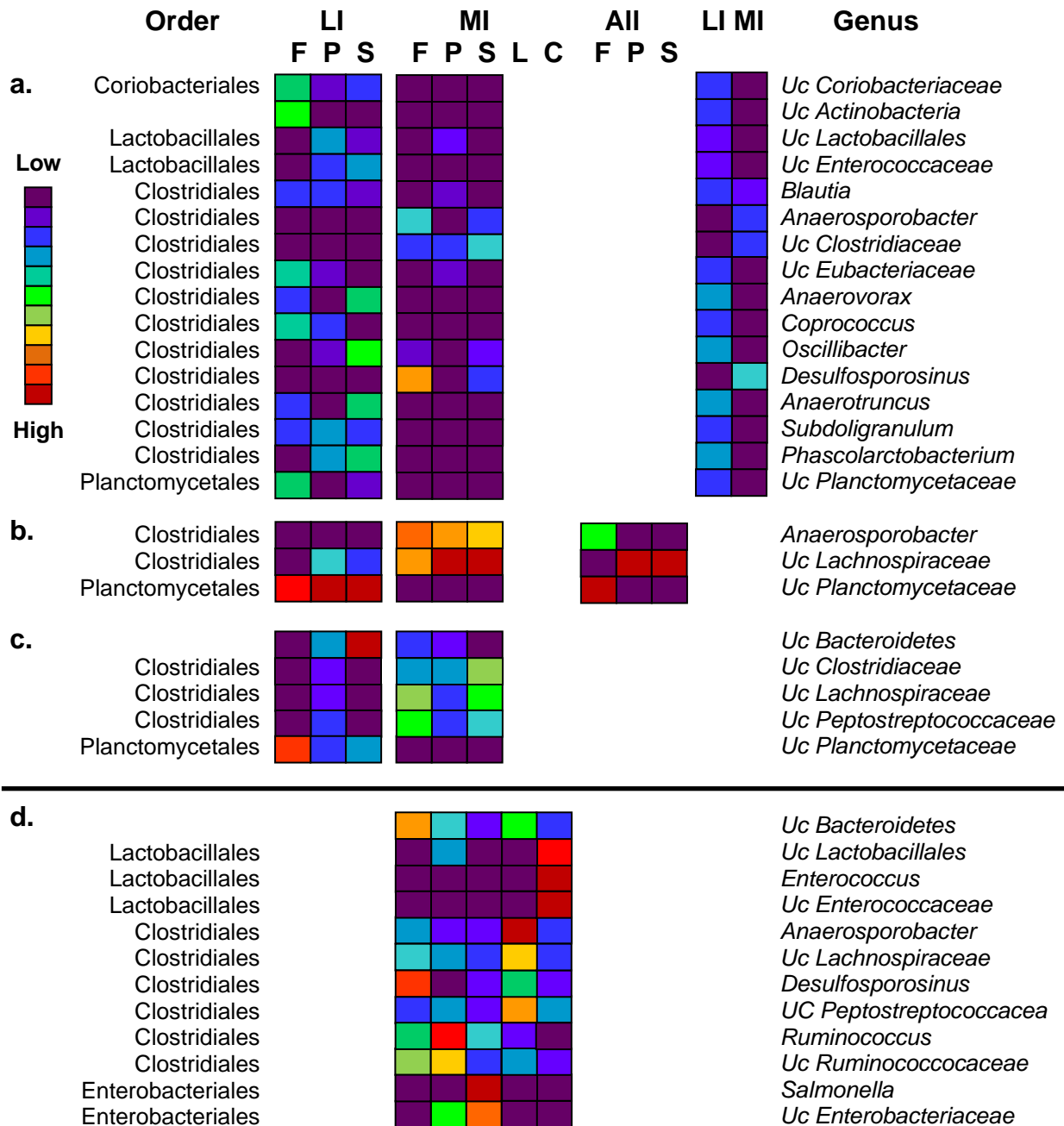


Figure 2.5. Microbial taxa contributing to detected host and site differences in microbial communities. The heat maps demonstrate the average relative abundance of each genus within each group (with sites and hosts identified by letters: C=Punta Carola, L=La Lobería, S=Santa Fe, P=Plaza Sur and F=Fernandina, LI=land iguana, and MI=marine iguana). Panels show genera significantly associated with host ecotype (a.) sites (b.), and the interaction of these terms (c.) and with sites for marine iguanas only (d.). Each panel is scaled independently to show the widest range of differences, such that the colors demonstrating the relative abundance patterns of any single taxa may differ from panel to panel.

## **CHAPTER 3: *CLOSTRIDIUM* PHYLOGENETIC DIVERSITY CONSERVED ACROSS ALLOPATRIC MARINE IGUANA POPULATIONS DESPITE EVIDENCE OF TAXA TURNOVER**

### **Chapter Summary**

Genetic studies of island taxa often demonstrate distinct population patterns which parallel colonization patterns or island geological histories. As gut associated communities are intimately associated with their hosts, it might be predicted that these communities could also demonstrate diversity patterns which are dependent on colonization history or which correspond to genetic relationships of the host. To study this question, we explored phylogenetic diversity patterns of OTUs within the genus *Clostridium* in five marine iguana populations using sequences acquired from 16S rRNA gene pyrosequencing. *Clostridium* species are known to be highly abundant and diverse in this host, but whether this diversity varies systematically among populations had not been explored. We hypothesized that if the evolutionary history of island colonization explained the geographic distribution of *Clostridium* OTUs, then evolutionary relationships among those OTUs would show characteristic phylogenetic relationships that paralleled marine iguana history of colonization, with increasing divergence of these communities from geologically old to young sites. In contrast to this expectation of strong, systematic site divergence in *Clostridium* diversity, we found that the phylogenetic structure of *Clostridium* communities was highly conserved across sites despite evidence of on-going gains and losses – processes which might be expected to lead to community divergences in allopatric host populations. These findings suggest that *Clostridium* diversity may be strongly conserved, potentially due to functional significance of these taxa to host digestion.

### **Introduction**

How communities assemble and are maintained is a fundamental question at the convergence of ecological and evolutionary theory. Forces which can shape communities on ecological time scales include abiotic filtering, competition, dispersal and demographic stochasticity. In evolutionary time, speciation and extinction can lead to further divergence among initially similar communities. The structure of any given extant community may then reflect processes at both of these time scales. Of particular interest for ecologists is then determining to what degree communities are formed by sorting among a pool of available species versus community assembly via the creation of new species *in situ* (Emerson and Gillespie 2008).

Both species sorting and *in situ* speciation can be important in community assembly, and their relative roles may differ across biogeographical conditions and taxa. In continental environments, species sorting appears to play the dominant role, with local communities resulting largely from the immigration of pre-adapted taxa. For

instance, plant communities in California's Mediterranean climates appear to be the result of immigration of taxa pre-adapted to summer droughts, with relatively little signal of adaptation to the new environmental conditions by taxa that were present prior to the shift in climate (Ackerly 2004). When barriers to dispersal are weak, new habitats are more likely to be filled by immigration of pre-adapted taxa than for the locally present taxa to adapt to the new conditions (Emerson and Gillespie 2008). This can be true even at very large scales; for instance, lupine plant endemic to the Andes are more closely related to alpine species from North America, and distantly related to low-land lupine species on the eastern half of South America (Hughes and Eastwood 2006).

Island systems may generally present a more extreme case, where species diversity and population genetic patterns are more strongly shaped by in situ evolution and geographical history due to the tendency for island species to be colonized rarely via founder events and for subsequent community developments to occur under strong isolation (Losos and Ricklefs 2009). Such patterns have been observed in a variety of taxa, including reptiles (Losos et al. 1998, Losos et al. 2003), birds (Lack 1947, Petren et al. 2005), and arthropods (Emerson and Oromí 2005). The increased barriers to dispersal on islands results in greater reproductive isolation among populations on different islands; in turn, this can lead to divergence to the point of speciation (resulting in anagenic evolution). Additionally, if new islands have many open niches, and immigration of pre-adapted taxa is restricted, cladogenesis may occur within islands as a colonizing species radiates to fill the empty niches (Emerson and Oromí 2005, Emerson and Gillespie 2008).

Incorporation of both ecological and evolutionary mechanisms into understanding of microbial community form and function are less advanced, potentially due in part to the historical expectations of microbial ubiquity resulting in large-scale species pool admixture. However, recent evidence suggests that many microbial species can be variably dispersal limited (e.g. Cho and Tiedje 2000, Franklin et al. 2000, Bala et al. 2003, Papke et al. 2003, Whitaker et al. 2003, Takacs-Vesbach et al. 2008, Vos and Velicer 2008), providing opportunity for long-term ecological and evolutionary patterning of allopatric communities. Host-associated microbial communities may be particularly affected by the present distributions and colonization histories of their hosts, as these organisms may have more limited dispersal potential relative to

environmental microbes (Zhang and Blackwell 2002, Bala et al. 2003, Falush et al. 2003, Hedlund et al. 2003, Fenchel and Finlay 2004, Papke and Ward 2004). Further, the host environment can be a strong selective filter, making differences between individuals or host populations of the same species more likely to reflect either historical or contemporary ecological processes within geographically separated communities. Enteric microbes that are quite intimately and obligatorily associated with their hosts can certainly demonstrate population genetic patterns that reflect host colonization over broad temporal and spatial scales (Falush et al. 2003). Whether this is true of microbial diversity at the community level has not been clearly documented.

The Galápagos Islands off the coast of Ecuador are a volcanic island chain where population genetic patterns of many endemic organisms suggests a pattern of colonization that runs parallel to the geological history of these unique islands (i.e. the “progression hypothesis” sensu Wagner and Funk 1995, Parent et al. 2008). Ancestral populations are generally located on the geologically older eastern and central islands, and younger, more recently derived populations are located on geologically younger islands to the north-west (as reviewed in Parent et al. 2008; Beheregaray et al. 2004, Kizirian et al. 2004, Ciofi et al. 2006, Benavides et al. 2007, Tzika et al. 2008, Gentile et al. 2009, Steinfartz et al. 2009).

The Galápagos islands are host to a unique lizard species, the marine iguana (*Amblyrhynchus cristatus*), which has demonstrated potential as a model system for exploring biogeographical patterns and mechanisms in intestinal microbial communities (Wheeler et al. 2011, Hong et al. In press, Chapter 2). Marine iguanas are large herbivorous lizards that have a complex intestinal biota, which contributes to fermentative digestion of their distinctive diet of intertidal macrophytic algae (Trillmich and Trillmich 1986, Mackie et al. 2004, Shepherd and Hawkes 2005, Mackie et al. 2008, Hong et al. In press). Marine iguanas are widely distributed throughout the Galápagos Island chain with ancestral populations distributed on the older central-eastern islands and more recent colonists located on younger sites to the northwest (Jackson 2007, Steinfartz et al. 2009). This unique population history presents the opportunity to explore the biogeographical structure of enteric communities across populations with a distinct colonization chronology.

Within the marine iguana's complex intestinal biota, the genus *Clostridium* is notably both abundant and diverse. Members of the genus *Clostridium* are gram-positive, anaerobic bacilli which are widely found in soils, water and the intestinal tracts of animals (Holt 1994). These organisms are thought to be functionally important for algal fermentation (Mackie et al. 2008, Nelson et al. 2010, Hong et al. In press), but whether these communities vary systematically among populations across the island chain has not been explored. We used 454 pyrosequencing derived sequences to explore both spatial patterning of *Clostridium* diversity and to elucidate ecological or evolutionary mechanisms which shape their distribution patterns.

First we evaluated the distribution of *Clostridium* OTUs among sites for evidence of any host colonization impacts on diversity patterns. We hypothesized that if host population history impacted the geographic distribution and site-overlap of *Clostridium* diversity, then younger, more recently colonized sites should contain a nested subset of *Clostridium* types relative to older ancestral sites. Such a pattern might arise due to stochastic loss of taxa before or during colonization of the subsequent island, for example, if each of one or a few colonists were carrying only a subset of the total microbial diversity present in the parent population. Such younger derived sites might also demonstrate decreased diversity due to community bottlenecks during which types were lost, although this pattern might be obscured if either subsequent adaptive radiations or acquisition of local types occurred, filling resulting empty niches.

We then compared the composition and phylogenetic structure of *Clostridium* communities among sites in order to better understand the community development. First, we asked whether the pattern of taxa richness suggested a role for extinction or founder effects during colonization of new islands. Secondly, we asked whether these founder effects, if present, resulted in decreased phylogenetic, in addition to or in lieu of decreased taxonomic, diversity in more recently colonized islands. Finally, community assembly may be driven by the acquisition of new taxa, in addition to the loss of taxa during colonization. Therefore, for *Clostridium* taxa unique to single islands, we evaluated sequence similarity to the other taxa to ask whether the most likely source of these novel taxa was radiation from a current member of the iguana intestinal community (cladogenesis), isolation from a population present on other islands (anagenesis), or acquisition from unique environmental sources on each island. Taken

together, these questions address the relative roles of ecological and evolutionary processes in the assemblage of enteric *Clostridium* communities across an island system.

## **Methods**

### *Study design and sampling sites*

Fecal specimens were obtained from marine iguanas at five locations on four islands in the Galápagos archipelago (Figure 3. 1). The sites selected for this study ranged in age from an estimated 3-4 million years old to <1 million years old (Cox 1983, Christie et al. 1992, White et al. 1993). Sites were distributed across this age range in approximate rank order from youngest to oldest: Isla Fernandina (<1 myo), Isla Plaza Sur (~1.5 myo), Isla Santa Fe (~2.5 myo), and Isla San Cristóbal (~3 myo). Exact timing of colonization for marine iguanas at these sites is not known, but genetic data suggest that the colonization of the island chain proceeded from older southeastern sites to geologically younger northwestern sites (Steinfartz et al. 2009).

At each site free-living marine iguanas were observed to defecate from a distance and then fecal specimens were collected from the ground using a sterile wooden applicator. Individual animals were intentionally selected from different areas of each site and samples at each site were collected over a short period of time (24-48 hours) to avoid repeated sampling of the same individual. The slow gastrointestinal transit rates of herbivorous reptiles (e.g. estimated to be 5-10 days for marine iguanas, Wikelski et al. 1993), also reduces the likelihood of observing a single individual defecating twice within a short sampling period. Five fecal specimens were obtained from a single sampling site on each of Isla Plaza Sur, Isla Fernandina, Isla Santa Fe and an additional five samples from each of two sampling sites on Isla San Cristobal (La Loberia and Punta Carola, see Figure 3.1). Samples were placed in sterile plastic tubes, stored at 4°C during travel and transported to Urbana, Illinois for storage at -20°C until analysis could be completed. The majority of samples included in this study were collected in September, 2009. Due to the small number of obtained for marine iguanas on Isla Santa Fe in this 2009 sampling, 3 land iguana and 3 marine iguana fecal samples collected in 2005 which had been stored frozen at -70°C were also included in these analyses to increase sample size for this site.

### *DNA extraction*

Community genomic DNA was extracted from frozen fecal samples using the Ultraclean Soil DNA isolation kit (MoBio, Carlsbad, CA) using the manufacturer's protocol, with minor modifications. Briefly, 0.2g of fecal material was placed in the manufacturer provided lysis tube and then pre-incubated with 12 µl each of 100mg/ml lysozyme and 1mg/ml achromopeptidase at 37°C for 1 hour before proceeding with chemical lysis and extraction according to the manufacturer's directions. We evaluated the purity and yield of DNA using a Qubit fluorometer (Invitrogen, CA) and DNA concentrations were then standardized to 10ng/ul before use as template for polymerase chain reaction (PCR).

### *Barcoded PCR and 454 pyrosequencing*

Pyrosequencing is sequencing technology which allows for the simultaneous processing of many thousands of heterogeneous amplicons derived from a one or more mixed microbial community samples. The resulting sequence reads, or pyrotags, can then be used for detailed assessment of microbial community composition, structure and hypothesized functional capacities (Andersson et al. 2008, Hugenholtz and Tyson 2008, McKenna et al. 2008, Roesch et al. 2007). The 16S rRNA gene was amplified from community genomic DNA using bacterial-specific forward 519F (5'-Fusion A-Barcode - CAGCMGCCGCGTAATWC-3') and reverse 926R (5'-Fusion B-Barcode- CCGTCAATTCMTTTRAGTT-3') primer pairs with a unique sequence barcodes to identify each sample (Parameswaran et al. 2007, Hamady et al. 2008, www.roche.com). PCR reaction mixtures comprised 100 ng of genomic DNA, 25 µl of Premix F (Epicentre Biotechnologies, WI), 200 nM (each) of forward and reverse primers, 0.5 U of Ex *Taq* DNA polymerase (Takara Bio Inc., Japan), and the volume added up to 50 µl with molecular-biology grade water. PCR with 30 cycles of thermal program (denaturation, 95 °C for 30 s; annealing, 55 °C for 45 s; and extension, 72 °C for 60 s) was performed. Amplicons were gel purified using the Wizard DNA purification kit (Promega, WI). DNA concentrations of the final concentrated PCR products were quantified by Qubit fluorometer (Invitrogen, CA).



### *Pyrotags handling and analysis*

Paired-end pyrosequencing was performed on a 454 FLX Titanium sequencer (Roche Applied Biosciences, Switzerland) by the Roy J. Carver Biotechnology Center, University of Illinois at Urbana Champaign. A total of 245,685 16S rRNA pyrotags were obtained, which were sorted by primer barcodes to form independent pyrotag libraries for each sample. Each pyrotag library was processed to remove low-quality reads as described elsewhere for these data (Hong et al. In press). Primer, barcode, and adaptor sequences were removed and the pyrotags were processed using the RDP Pipeline Initial Process (Cole et al. 2009). Secondary structure alignments of processed pyrotags were produced using RDP Infernal (Cole et al. 2009). The aligned pyrotags were visually checked with Jalview (<http://www.jalview.org>), and manual adjustments were performed as necessary to ensure alignment quality. After processing and trimming the pyrotags had an average read length of 370 nt. No significant differences in the average number of sequences were detected for host ecotypes or sampling sites (overall average number of sequences/sample: 5177 +/-1218). A complete description of these pyrosequencing data and of diet-linked differences among host species in these bacterial communities is published elsewhere (Hong et al. In press).

Pyrotags were aligned using the greengenes online alignment tool (<http://greengenes.lbl.gov>) and were then clustered using cd-hit to remove redundant sequences. Phylogenetic distances were calculated using Mothur to obtain a 95% similarity OTU relative abundance data matrix. Representative sequences for each OTU were classified to the RDP database using the greengenes classify tool. OTUs which were classified to the genus *Clostridium* were compiled into both an abundance matrix by individual iguana and an aligned sequence subset for subsequent analyses.

### *Statistical analysis*

Differences among sampling sites were first visualized using ordination by non-metric multidimensional scaling (NMDS) using the metaMDS function in the vegan package of R (Oksanen et al. 2010, R Development Team 2010). To explore the effects of sampling site and geographic location on patterns detected in these ordinations, we used the adonis function in the vegan package in R (Oksanen et al. 2010) to perform a

permutation multiple analysis of variance (MANOVA). This method (also referred to as a permutation or non-parametric MANOVA) partitions the sums of squares of distance matrices among treatments and has relaxed assumptions relative to traditional MANOVA. Significance in the permutation tests was determined by comparing the observed effects against 5000 random permutations of the data for each model run independently. The degree of *Clostridium* OTU overlap among sites was visualized using the Venn function in the gplots package of R (Warnes et al. 2010).

The distribution of taxa among sampling sites was then explored by counting the number of sites at which each OTU was detected. The distribution each OTU across sites was counted from 1 to 5. For the number of endemic (present at 1 site) through ubiquitous (present at 5 sites) and intervening were summarized by the OTU count in each category for each site. In order to test whether this observed pattern of OTU distribution among sites differed significantly from what would be expected from random sampling of a common community located at all sites, we compared the observed pattern to 100 simulations of a null model. The null model was parameterized using the observed frequency of OTUs among individual iguanas for the abundance distribution of taxa (Figure 3.2). To populate the artificial communities, each OTU was assigned to be present or absent according to one random draw from a binomial distribution where the probability of success ( $p$ ) was set equal to the proportional frequency of that particular OTU in the observed data set. This process was repeated 25 times for each simulation run, and these 25 simulated iguanas were then randomly assigned to sampling sites. The site distribution of each OTU was then calculated for each simulated run, which was used to determine the number of OTUs present on 1, 2, 3, 4, or 5 sites. Using 100 simulations, we then calculated the average and 95% confidence intervals on these distributions. The observed distribution was considered to be significantly non-random if it did not fall within the simulated 95% confidence limits.

A single representative sequence for each OTU was used to calculate a Jukes-Cantor distance matrix using MEGA version 5 (Tamura et al. 2011). The average distance was calculated within each group of interest to serve as a measure of phylogenetic diversity. The nearest neighbor sequence for each OTU was determined as the OTU with the smallest Jukes-Cantor distance and these nearest neighbors and their characteristics were recorded for analysis of their site distribution characteristics.

All statistical and graphing procedures were performed in either JMP 8 (SAS Institute, Cary, NC) or R statistical language (R Development Core Team 2010).

## **Results**

### *Clostridium diversity*

A total of 212 *Clostridium* OTUs were identified at the 95% level of sequence similarity, the largest number of OTUs within a single classified genus in this dataset (14% of 1521 total OTUs). The number of *Clostridium* OTUs detected within each site ranged from 103 to 137 OTUs, with no significant difference in the proportion of total OTUs present at each site (Figure 3.1). Individual iguanas were host to an average of 53.2 +/- 18.8 *Clostridium* OTUs with no significant difference in within individual richness among sites (ANOVA: F=1.410, p=0.267).

### *Site patterns of Clostridium diversity*

*Clostridium* OTU composition differed among sites (Permutation MANOVA: p=0.004, R<sup>2</sup>=0.327; see also Figure B.1) and 68 taxa were locally endemic to a single site out of the five and were often only present in 1 or 2 individual iguanas at that site. However, a relatively large number of taxa (47 OTUs) were common to all sites and the remaining OTUs were variably distributed across 2,3 or 4 sampling locations (Figure 3.4).

In order to explore the hypothesis that these sites share a similar community of *Clostridium* types and differed primarily due to variation in sampling of the rarer portion of the community, we ran a simulation model of this sampling event using the actual distribution of taxon abundances to parameterize the model (Figure 3.2). We ran 100 simulated sampling events of 212 OTUs across 5 locations and compiled the site distribution of both the simulation and the observed data for comparison (Figure 3.3). Notably, while the observed and simulation model distributions among sites were very similar for the endemic and variably present taxa, there was a significant difference in the number of OTUs which were present on all sites, with the observed distributions having fewer than expected ubiquitous taxa (Figure 3.4). The observed site distribution of OTUs did not differ among sites (Figure 3.4, Fisher's Exact: p>0.05 for all pair-wise comparisons). It was also notable that the majority of endemic OTUs (60 of 68) were

present in only a single individual on their respective island, while ubiquitous OTUs were of variable abundance from 5 to 25 individuals.

### *Phylogenetic diversity patterns*

As OTU distribution patterns suggested that endemic taxa (i.e. those OTUs present on a single site) are likely simply rare members of a common community, we tested the phylogenetic variability of these OTUs within each site and the relationship of these endemic OTUs to the rest of the *Clostridium* community. On four of five sampling sites the latter pattern was detected, with higher phylogenetic diversity within endemic taxa relative to more widespread taxa (i.e. those found at two or more sites; Figure 3.5). In contrast, there was no significant difference in phylogenetic diversity at Punta Carola on Isla San Cristóbal (Figure 3.5). Phylogenetic diversity did not vary when comparing the average phylogenetic distance pairwise among sites (ANOVA by site: all  $p > 0.05$ ).

We then evaluated the site distribution of nearest neighbors to evaluate the evolutionary patterns of endemic versus more widespread taxa. We identified each OTU's nearest neighbor by genetic distance and then compared the site distributions of these nearest neighbor OTUs to the observed distribution of all OTUs as a null model for random association. A total of 112 OTUs were placed as nearest neighbor to one or more of the OTUs within the dataset. The distribution of these nearest neighbors across sites was non-random, with a significant increase in the number of ubiquitous nearest neighbors (i.e. found on all five sites) and a significant decrease in the number of nearest neighbors that were endemic to one site (Figure 3.6). However, the distribution of nearest neighbors did not differ between OTUs found on a single site and all other OTUs in the dataset (Figure 3.6).

The average distance between each OTU and their nearest neighbor was calculated for endemic taxa within island (with sites Punta Carola and La Lobería averaged because they derived from a single island). These average distances from nearest neighbors within sites were strongly correlated to rank age of the island (from youngest to oldest: Isla San Cristóbal > Isla Santa Fe > Isla Plaza Sur > Isla Fernandina; Pearson's correlation:  $R^2 = 0.93$ ,  $p = 0.038$ ).

## Discussion

### *Clostridium* diversity conserved across sites despite extinctions

At 14% of all OTUs, members of the genus *Clostridium* represented a large portion of the total bacterial diversity found in marine iguana feces. Multivariate analysis of community composition indicated significant differences among sites. However, compared to a parallel analysis of these samples incorporating all OTUs obtained from complete community pyrosequencing, these site differences were more subtle, with less differentiation among sites (Figure B.1).

We originally hypothesized that founder effects during host colonization of the island chain would lead to reductions in the taxa richness and phylogenetic diversity of their associated *Clostridium* communities. Contrary to these predictions, a relatively large proportion of OTUs were present on all sites and many OTUs were shared across multiple sites with neither notable differences in the degree of overlap among specific sites nor in the degree of endemism or ubiquity across sites.

Although much of the *Clostridium* diversity was conserved across multiple or all host-populations, each site was also host to a number of endemic OTUs. Notably, the majority of these endemic isolates were present in only a single host individual and were, thus, were rare community members even within a sampling site. These presumptively endemic taxa could represent true divergence in community composition among sites, but could equally represent expected sampling error for rare community members present across all sites but were simply rarely detected. To test this hypothesis, we performed a simulation model to explore the hypothesis that all sites contain a common *Clostridium* community with the observed abundance distribution across all individuals. Through this sampling simulation, we were able to compare the observed distribution of OTUs among sites to this simulated distribution. Since the number of endemic taxa did not differ between the simulation and our observed data, we cannot rule out the possibility that these are simply rare taxa common to all islands but infrequently observed due to shallow sampling depth. However, this simulation also suggested that extinction or restricted colonization (i.e. stochastic loss of taxa in specific host populations during or subsequent to colonization processes) may have played a role in shaping the communities present at these sites.

*Rare OTUs primarily radiated from internal diversity*

While the number of endemic OTUs did not differ from what would be expected due to resampling a common distribution, this does not speak to the origin of these rare taxa or their relationship to the broader community. We found evidence of taxa losses, but no associated differences in community richness across sites. This finding suggests that communities must be acquiring new OTUs to compensate for those lost during colonization or to subsequent extinction.

To investigate the potential sources of new taxa, we then tested the phylogenetic variability of the endemic isolates relative to more widespread community members. New taxa could be acquired from environmental sources on each island (soils, water, or other host species), they could be the result of divergence from an ancestral taxa due to isolation (anagenesis), or they could be the result of adaptive radiations from extant members of the community to fill empty niches (cladeogenesis; Emerson and Gillespie 2008). First, we evaluated the within-site variance in phylogenetic relationships among the endemic taxa and all other taxa using a genetic distance calculated from representative OTU sequence alignments. By exploring the average distance of these groups within a site, we could evaluate the amount of phylogenetic diversity encompassed by these groups. If endemic taxa within a site are derived from a common source (for example evolved from a local source not historically present in marine iguanas or from an internal adaptive radiation of a single highly selected clade) these OTUs would be very phylogenetically similar to one another compared to the diversity within broadly distributed OTUs (i.e. endemic isolates would have decreased phylogenetic diversity). In contrast, if these OTUs evolved from more common extant or historical OTUs across the phylogeny or from acquisition from multiple local sources, then they should have higher phylogenetic variability as the branch tips of the phylogenetic tree expanded at multiple locations simultaneously. The latter pattern was predominant in four of five sites, suggesting that these rare OTUs that were detected at only a single site are derived from more than one lineage.

Increased phylogenetic diversity within endemic isolate pools might suggest that these isolates evolved from historical or extant members of the core marine iguana *Clostridium* community. However, this result alone cannot rule out the alternative that

these endemic taxa were instead derived from a variety of local sources, which might also lead to expanded phylogenetic diversity. In order to better discriminate between these alternatives, we then explored the relationships of each OTU to the rest by using genetic distance to identify the nearest genetic neighbor of each OTU. We then evaluated if the site distribution of these nearest neighbors of endemic OTUs differed from those that are more widespread.

The nearest neighbor distributions of both endemic and more widespread OTUs differed significantly from the site distribution of all OTUs but not from each other, suggesting that these associations were decidedly non-random but did not differ between endemic and all other OTUs. Nearest neighbors were more likely to be ubiquitous taxa and were less likely to be endemic taxa. These patterns suggest that rare endemic taxa are more likely derived from extant or historical diversity through adaptive radiation within the marine iguana community, rather than externally acquired from a broader environmental source.

In conflict with this finding, however, we also found a strong correlation between the average distance of endemic OTUs to their nearest neighbor and rank age of each island. This finding suggests that younger sites may have greater filling of empty niches by novel environmental sources which are more genetically disparate from the rest of the community than older sites, where adaptive radiation has had time to occur. Such age dependent cladogenesis was noted for niche filling in forest beetles on the Canary Islands, with the suggestion that anagenic versus cladogenic processes may be dependent on island age (Emerson and Oromí 2005). Thus, there is some evidence that both processes may be operating in this system. However, these analyses cannot definitively explore the sources of *Clostridium* community diversity without comparison to environmental *Clostridium* sequences from the Galápagos Islands against which to compare marine iguana-derived sequences. Additional study with increased sequence lengths and including environmental specimens would be required to solidify these assertions and to build reliable phylogenies from which to draw conclusions.

An additional possibility is that some of the rare, "endemic" OTUs may in fact be the result of sequencing error. Since the ubiquitous taxa also tended to be the most common, one might hypothesize that some rare OTU's may be in fact simply be incorrect reads of common taxa (leading to the close phylogenetic relationship between

the endemic and ubiquitous taxa). We attempted to minimize this possibility by using a conservative 95% similarity cut-off throughout these analyses. With an average read length of 370 nt, two reads must differ by around 18 nt to be considered distinct OTUs. With an average error rate for GS-FLX chemistry of ~0.5% (Niu et al. 2010), 1-2 nt error per read would be expected, making it relatively unlikely that sequence error alone would lead to a spurious OTU classification. For instance, assuming an error rate of 0.5% , a read length of 370 nt , and a Poisson distribution of errors, the probability of having 18 errors in a single read is only  $5.54 \times 10^{-12}$  (with a total of 245,685 reads, this predicts far less than one erroneous classification). At a less conservative similarity threshold, such as the commonly used “species” sequence similarity of 98%, the probability of a spurious OTU classification rises to 0.0034 (which would predict 844 erroneous classifications in this dataset).

#### *Conservatism in the face of OTU turnover*

In contrast to expectations of strong, systematic site divergence in *Clostridium* diversity, we found that the phylogenetic structure of these communities was conserved across sites despite evidence of on-going OTU gains and losses – processes which might be expected to lead to community divergences in allopatric host populations. These findings suggest that *Clostridium* diversity at this taxonomic level (i.e. 95% sequence similarity) may be strongly conserved, potentially due to the potential functional significance of these taxa to host digestion (Nelson et al.2010). Certainly enteric community composition has been shown to potentially strongly relate to digestive efficiency and host health (Turnbaugh and Gordon 2009, Goodwin and Gordon 2010). Whereas island communities of larger organisms often show reduced richness resulting in open niches which makes them susceptible to exotic species invasions (Emerson and Gillespie 2008), commensal communities may not have this capacity to allow open niche space due to functional constraints. Thus, enteric microbial communities may not behave as would be predicted based on previously described island taxa due to functional constraints not experienced by free-living organisms where open niche spaces may be tolerated.

While these findings were perhaps against initial expectations based on parallel to population genetic theory, a number of caveats of these analyses may shed light onto



this departure from anticipated results. Firstly, the findings in this study may be reliant on the taxonomic scale at which these analyses were performed. Given the diversity of this genus, we selected a relatively conservative sequence similarity level of 95% to explore this question within a limited number of OTUs. This level of similarity likely does not represent a true species-level, but rather a sub-genus level of organization that may represent broad functional differences within the genus which may be strongly conserved due to host-bacterial interactions, rather than specific species or strains that might be more susceptible to stochastic influences. In macrobial systems, consideration of the scale at which communities are defined has been shown to affect the degree of phylogenetic overdispersion or clustering detected in this type of community phylogenetic study (Cavender–Bares et al. 2006). Thus, these results cannot exclude the possibility of that *Clostridium* communities may show much stronger site patterns at a finer taxonomic scale or may have unique strains of each sub-genus type identified here that circulate on a very local level.

A second possibility that these results can neither rule in nor rule out, is that *Clostridium* taxa are broadly distributed throughout the island chain and disperse readily among locations sufficiently to homogenize local strain pools. There are no studies of the long-distance dispersal patterns of members of this genus, to our knowledge, making the initial assumption of extreme dispersal limitation on which our initial hypotheses were built potentially faulty. If this is the case, then an alternative possibility that could equally explain some of the patterns detected in this study is that *Clostridium* types are acquired from a common environmental species pool across the Galápagos island chain and the host-environment selects for a subset of these types as suites the digestive needs of the host. Fully exploring this possibility would require additional study of environmental sources such as water, soils and food sources, to understand environmental *Clostridium* exposures.

#### *Technological barriers to community biogeography of microbes*

Traditionally, studies of the phylogenetic structure of communities have required both well-resolved phylogenies and distributional data, which has limited these approaches to specific, well-studied groups such as birds, reptiles, plants and insects. The increasing application of next-generation sequencing of mixed microbial

communities raises the possibility of testing these concepts in much less malleable systems and in organisms where traditional approaches are not feasible (such as difficult to cultivate gastrointestinal bacteria). Thus, an additional goal of our study was to determine whether the short-sequence reads produced by 454 pyrosequencing would be sufficient to observe patterns and elucidate underlying ecological mechanisms which might be driving those patterns. We would posit that these analyses represent a relatively novel use of next-generation community sequence results, and demonstrate the potential of these methods for doing fine-scale community ecology on these minute organisms. However, we feel that these efforts were only partially successful, limited in phylogenetic resolution by the relatively short sequence reads obtained by pyrosequencing (<400 nt) and by the inability to draw an important link between phylogenetic and niche diversity, a important connection for fully exploring community assemblage and maintenance in a spatial context (Webb et al. 2002, Cavender-Bares et al. 2006, Caddotte et al. 2009, Leibold et al. 2010, Pavoine et al. 2011).

The latter issue may be especially important for linking ecological and evolutionary processes in microbial systems, as phylogenetic diversity within a given genus may be subservient to trait diversity for dictating local community dynamics due to the unique capacity of microbes to exchange genetic material across taxonomic boundaries (i.e. lateral or horizontal gene transfer; Ragan and Beiko 2009, Boto 2010). *Clostridium* species in this system have likely benefited from such lateral acquisition of genes contributing metabolic capacities for algal digestion (Nelson et al. 2010). Such lateral increases in niche breadth may confound phylogenetic relationships and expectations due to changes in fitness among strains unrelated to genetic lineage. As next-generation technologies continue to advance in read length and depth, obtaining this type of information on both a phylogenetic and functional level for microbial taxa which are difficult to study using traditional cultivation and population genetic approaches may allow for more advanced understanding of microbial community patterns and processes. Still, despite the limitations of current technology, we feel this study shows that these types of data have much to contribute for exploring the ecology and evolutionary mechanisms which shape microbial communities.

## **Conclusions**

In contrast to our initial expectation of strong site divergence reflecting the unique colonization history of the host, we found that the distribution and phylogenetic structure of *Clostridium* communities were conserved across sites despite evidence of both extinction and radiation of types – forces which might otherwise lead to vast divergence in enteric communities resident to allopatric host populations. At the sub-genus level explored in these analyses (i.e. 95% sequence similarity), *Clostridium* diversity appears to be strongly conserved, potentially due to the functional significance of these microbes to host digestion. However, these findings do not preclude the possibility of local acquisition of site specific strains within the conserved varieties that make up the core *Clostridium* community in marine iguanas. While current technologies are not quite advanced enough to explore such subtleties in bacterial groups which are challenging to cultivate, on-going advances in molecular methods may soon make such fine-scale studies at the juncture of ecological and evolutionary processes a realistic and intriguing prospect.

## **Chapter Acknowledgements**

We would like to express appreciation to the Galápagos National Park and Washington Tapia, and the Charles Darwin Research Station and Sonja Cisnera and Paulina Couenberg for facilitating sample collection. Many thanks to Augusto G. Haz Beltran, Lenin Cruz Beldon and the crew of the Pirata for help with sample collection. We also wish to thank the Department of Animal Sciences and Nancy Henry for administrative support. Many thanks to Peiyong Hong for her laboratory work and expertise in acquiring the 454 pyrosequencing data. Thanks to Yuejian Mao for providing bioinformatics support in handling these datasets. Thanks to Richard Lankau for statistical support, assistance with performing the simulation modeling, and constructive comments on early drafts of this manuscript. Thanks to Isaac Cann and Bryan White for access to laboratory resources.

Funding for this work was provided by the USDA National Research Initiative (Proposal no. AG-2008-35206-18784, R. I. M.) and by the US Environmental Protection Agency Science to Achieve Results Fellowship program (E.W.). This chapter was

developed with support from a STAR Research Assistance Agreement No. 91684301-1 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this chapter.

## References

- Ackerly DD. 2004. Adaptation, niche conservatism and convergence: comparative studies of leaf evolution in the California chaparral. *American Naturalist* 163: 654-671.
- Andersson AF, Lindberg M, Jakobsson H, Backhed F, Nyren P, Engstrand L. 2008. Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLoS One* 3: e2836. doi: 10.1371/journal.pone.0002836.
- Bala A, Murphy P, Giller KE. 2003. Distribution and diversity of rhizobia nodulating agroforestry legumes in soils from three continents in the tropics. *Molecular Ecology* 12: 917–929.
- Beheregaray LB, Gibbs JP, Havill N, Fritts TH, Powell JR, Caccone A. 2004. Giant tortoises are not so slow: rapid diversification and biogeographic consensus in the Galápagos. *Proceedings of the National Academy of Sciences* 101: 6514-6519.
- Benavides E, Baum R, McClellan D, Sites JW Jr. 2007. Molecular phylogenetics of the lizard genus *Microlophus* (Squamata: Tropiduridae): Aligning and retrieving indel signal from nuclear Introns. *Systems Biology* 56: 776-797.
- Boto L. 2010. Horizontal gene transfer in evolution: facts and challenges. *Proceedings of the Royal Society B* 277: 819-827.
- Caddotte ME, Cavender-Bares J, Tilman D, Oakley TH. 2009. Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. *PLoS One*. 4: e5695. doi:10.1371/journal.pone.0005695.
- Cavender-Bares J, Keen A, Miles B. 2006. Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology* 87: S109-S122.
- Cho JC, Tiedje JM. 2000. Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Applied and Environmental Microbiology*. 66: 5448-5456.

- Christie DM, Duncan RA, McBirney AR, Richards MA, White WM, Harpp KS, Fox CG. 1992. Drowned islands downstream from the Galápagos hotspot imply extended speciation times. *Nature* 355: 246-248.
- Ciofi C, Wilson GA, Beheregaray LB, Marquez C, Gibbs JP, Tapia W, Snell HL, Caccone A, Powell JR. 2006. Phylogeographic history and gene flow among giant Galápagos tortoises on southern Isabela Island. *Genetics* 172: 1727-1744.
- Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syex-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM. 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* 37: D141-145.
- Cox A. 1983. Ages of the Galápagos Islands. In *Patterns of Galápagos Organisms* (Bowman RI, Berson M, Leviton AE, eds.) pp 11-23. AAAS Pacific Division: San Francisco, CA.
- Emerson BC, Gillespie RG. 2008. Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution* 23: 619-630.
- Emerson BC, Oromí P. 2005. Diversification of the forest beetle genus *Tarphius* on the Canary Islands, and the evolutionary origins of island endemics. *Evolution* 59: 586-598.
- Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, Blaser MJ, Graham DY, Vacher S, Perez-Perez GI, Yamaoka Y, Mégraud F, Otto Kristina, Reichard U, Katzowitsch E, Wang X, Achtman M, and Suerbaum S. 2003. Traces of human migrations in *Helicobacter pylori* populations. *Science* 299: 1582-1585.
- Fenchel T, Finlay B. 2004. The ubiquity of small species: Patterns of local and global diversity. *BioScience* 54: 777-784.
- Franklin RB, Taylor DR, Mills AL. 2000. The distribution of microbial communities in anaerobic and aerobic zones of a shallow coastal plain aquifer. *Microbial Ecology* 38: 377-386.
- Gentile G, Fabiani A, Marquez C, Snell HL, Snell HM, Tapia W, Sbordoni V. 2009. An overlooked pink species of land iguana in the Galápagos. *Proceedings of the National Academy of Sciences* 106: 507-511.
- Goodwin AL, Gordon JI. 2010. Our unindicted coconspirators: Human metabolism from a microbial perspective. *Cell Metabolism* 12: 111-116.

- Warnes GR. 2010. Includes R source code and/or documentation contributed by (in alphabetical order): Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B. *gplots: Various R programming tools for plotting data*. R package version 2.8.0. <http://CRAN.R-project.org/package=gplots>.
- Hamady MJ, Walker J, Harris JK, Gold NJ, Knight R. 2008. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nature Methods* 5: 235-237.
- Hedlund BP, Staley JT. 2003. Microbial endemism and biogeography. In *Microbial Diversity and Bioprospecting* (Bull AT, ed.). American Society of Microbiology: Washington, DC.
- Holt JG. 1994. *Bergey's Manual of Determinative Bacteriology* (9<sup>th</sup> ed.). Lippincott Williams and Wilkins: Philadelphia, PA.
- Hong PY, Wheeler E, Cann IKO, Mackie RI. In press. Phylogenetic analysis of the intestinal microbial community in herbivorous land and marine iguanas of the Galapagos Islands using 16S rRNA-based pyrosequencing. *ISME Journal*.
- Hugenholtz P, Tyson GW. 2008. Microbiology: metagenomics. *Nature* 455: 481-483.
- Hughes C, Eastwood R. 2006. Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Science* 103: 10334-10339.
- Jackson MH. 2007. *Galápagos: A Natural History* (11<sup>th</sup> ed.). University of Calgary Press: Calgary, Alberta, Canada.
- Kizirian D, Trager A, Donnelly MA, Wright JW. 2004. Evolution of Galápagos Island Lava Lizards (Iguania: Tropiduridae: Microlophus). In *Molecular Phylogenetics and Evolution* 32: 761-769.
- Lack D. 1947. *Darwin's Finches*. Cambridge University Press: Cambridge, UK.
- Leibold MA, Economo EP, Peres-Neto P. 2010. Metacommunity phylogenetics: separating the roles of environmental filters and historical biogeography. *Ecology Letters* 13: 1290-1299.
- Losos JB, Jackman TR, Larson A, de Queiroz K, Rodríguez-Chettino L. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279: 2115-2118.

- Losos JB, Leal M, Glor RE, de Queiroz K, Hertz PE, Rodríguez-Chettino L, Chamizo-Lara A, Jackman TR, Larson A. 2003. Niche lability in the evolution of a Caribbean community. *Nature* 424: 542-545.
- Losos JB, Ricklefs RE. 2009. Adaptation and diversification on islands. *Nature* 457: 830-836.
- Mackie R, Nelson D, Wheeler E, Wikelski M. 2008. Fermentative digestion in herbivorous lizards. In *Proceedings of the 4<sup>th</sup> CPB Meeting in Africa: MARA*. (Morris S, Vosloo A, eds.) pp 193-202. Medimond Publishing Co.:Bologna, Italy.
- Mackie RI, Rycyk M, Ruemmler RL, Aminov RI, Wikelski M. 2004. Biochemical and microbiological evidence for fermentative digestion in free-living land iguanas and marine iguanas (*Amblyrhynchus cristatus*) on the Galápagos archipelago. *Physiological and Biochemical Zoology* 77: 127-138.
- McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, Liu Z, Lozupone CA, Hamady M, Knight R, Bushman FD. 2008. The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathogens* 4: e20. doi: 10.1371/journal.ppat.0040020.
- Nelson DM, Cann IKO, Altermann E, Mackie RI. 2010. Phylogenetic evidence for lateral gene transfer in the intestine of marine iguanas. *PLoS ONE* 5: e10785. doi:10.1371/journal.pone.0010785.
- Niu B, Fu L, Sun S, Li W. 2010. Artificial and natural duplicates in pyrosequencing reads of metagenomic data. *BMC Bioinformatics* 11: 187.
- Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2010. *vegan: Community Ecology Package*. R package version 1.17-4. <http://cran.r-project.org/>, <http://vegan.r-forge.r-project.org/>.
- Papke RT, Ramsing NB, Bateson MM, Ward DM. 2003. Geographical isolation in hot springs cyanobacteria. *Environmental Microbiology* 5: 650-659.
- Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M, Fire AZ. 2007. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Research* 35: e130.
- Parent CE, Caccone A, Petren K. 2008. Colonisation and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical Transactions of the Royal Society B* 363: 3347-3361.

- Pavoine S, Vela E, Gachet S, de Bélair G, Bonsall MB. 2011. Linking patterns in phylogeny, traits, abiotic variables and space: a novel approach to linking environmental filtering and plant community assembly. *Journal of Ecology* 99: 165-175.
- Petren K, Grant PR, Grant BR, Keller LF. 2005. Comparative landscape genetics and the adaptive radiation of Darwin's finches: the role of peripheral isolation. *Molecular Ecology* 14: 2943-2957.
- R Development Core Team. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Ragan MA, Beiko RG. 2009. Lateral genetic transfer: open issues. *Philosophical Transactions of the Royal Society B* 364: 2241-2251.
- Roesch LF, Fulthorpe RR, Riva A, Casella G, Hadwin AK, Kent AD, Daroub SH, Camargo FA, Farmerie WG, Triplett EW. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME Journal* 1: 283-290.
- Shepherd SA, Hawkes MW. 2005. Algal food preferences and seasonal foraging strategy of the marine iguana, *Amblyrhynchus cristatus*, on Santa Cruz, Galápagos. *Bulletin of Marine Science* 77: 51-72.
- Steinfartz S, Glaberman S, Lanterbecq D, Russello MA, Rosa S, Hanley TC, Marquez C, Snell HL, Snell HM, Gentile G, Dell'Olmo G, Powell AM, Caccone A. 2009. Progressive colonization and restricted gene flow shape island-dependent population structure in Galápagos marine iguana (*Amblyrhynchus cristatus*). *BMC Evolutionary Biology* 9: 297.
- Takacs-Vesbach C, Mitchell K, Jackson-Weaver O, Reysenbach A-L. 2008. Volcanic calderas delineate biogeographic provinces among Yellowstone thermophiles. *Environmental Microbiology* 10: 1681-1689.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*. <http://www.kumarlab.net/publications>.
- Trillmich KGK, Trillmich F. 1986. Foraging strategies of the marine iguana, *Amblyrhynchus cristatus*. *Behavioral Ecology and Sociobiology* 18: 259-266.



- Turnbaugh PJ, Gordon JI. 2009. The core gut microbiome, energy balance and obesity. *Journal of Physiology*. 587: 4153-4158.
- Tzika AC, Rosa SFP, Fabiani A, Snell HL, Snell HM, Marquez C, Tapia W, Rassmann K, Gentile G, Milinkovitch MC. 2008. Population genetics of Galápagos land iguana (genus *Conolophus*) remnant populations. *Molecular Ecology* 17: 4943-4952.
- Vos M, Velicer GJ. 2008. Isolation by distance in the spore-forming soil bacterium *Myxococcus xanthus*. *Current Biology* 18: 386-391.
- Wagner WL, Funk VA. 1995. Hawaiian biogeography: evolution on a hot spot archipelago. Smithsonian Institution Press: Washington, DC.
- Webb CO, Ackerly DD, McPeck MA, Donoghue MJ. 2002. Phylogenies and community ecology. *Annual Reviews in Ecology and Systematics* 33: 475-505.
- Wheeler E, Cann IKO, Mackie RI. 2011. Genomic fingerprinting and serotyping of *Salmonella* from Galápagos iguanas demonstrates island differences in strain carriage. *Environmental Microbiology Reports* 3: 166-173.
- Whitaker RJ, Grogan DW, Taylor JW. 2003. Geographic barriers isolate endemic populations of hyperthermophilic Archaea. *Science* 301: 976-978.
- White WM, McBirney AR, Duncan RA. 1993. Petrology and geochemistry of the Galápagos Islands – Portraits of a pathological mantle plume. *Journal of Geophysical Research* 98: 19533-19563.
- Wikelski M, Gall B, Trillmich F. 1993. Ontogenetic changes in food intake and digestion rate of the herbivorous marine iguana (*Amblyrhynchus cristatus*, Bell). *Oecologia* 94: 373-379.
- Zhang N, Blackwell M. 2002. Population structure of dogwood anthracnose fungus. *Phytopathology* 92: 1276–1283.

## Figures

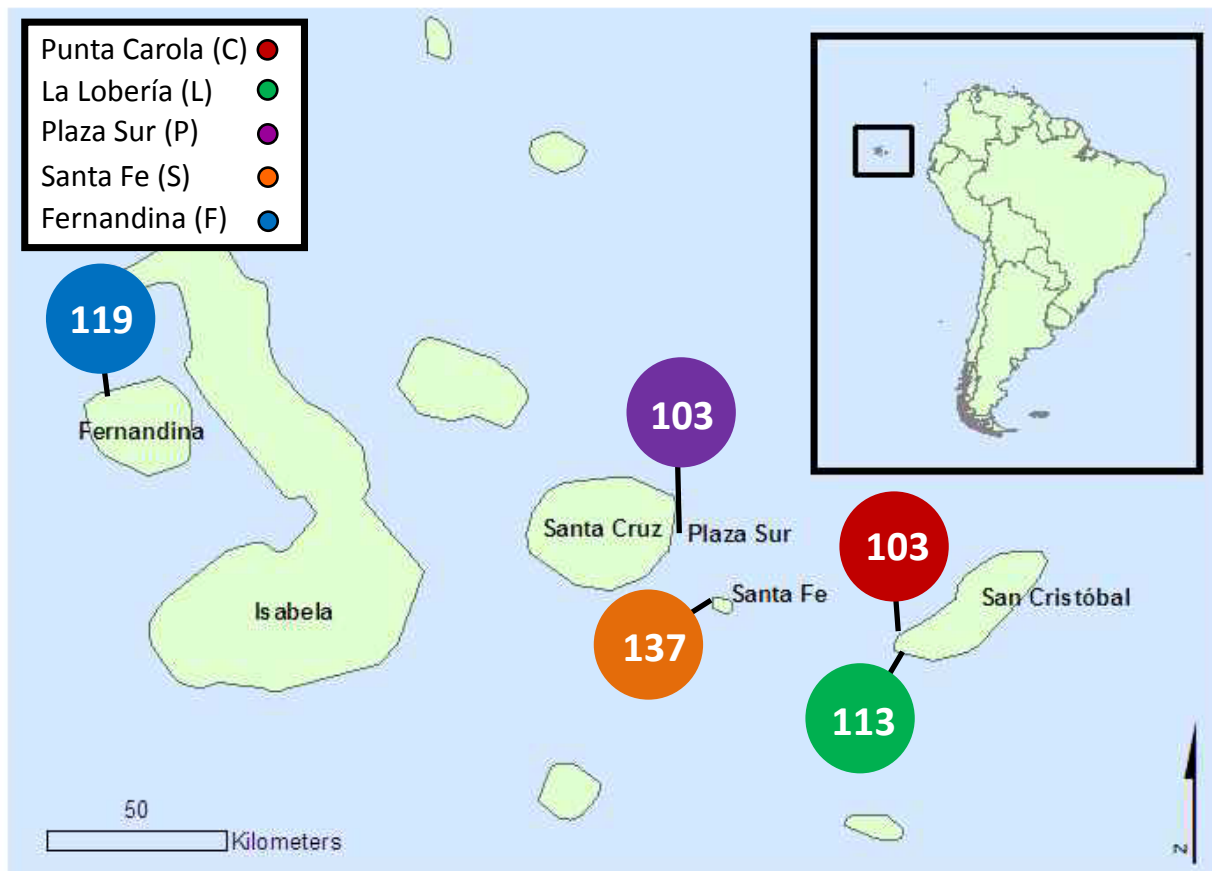


Figure 3.1. Map of study sites. A total of 212 *Clostridium* OTUs were detected at the 95% sequence similarity level. Site OTU richness is indicated by the number present in each site marker. There were no significant differences in OTU richness among sites (Fisher's Exact two-tailed test,  $p=0.365$ ).

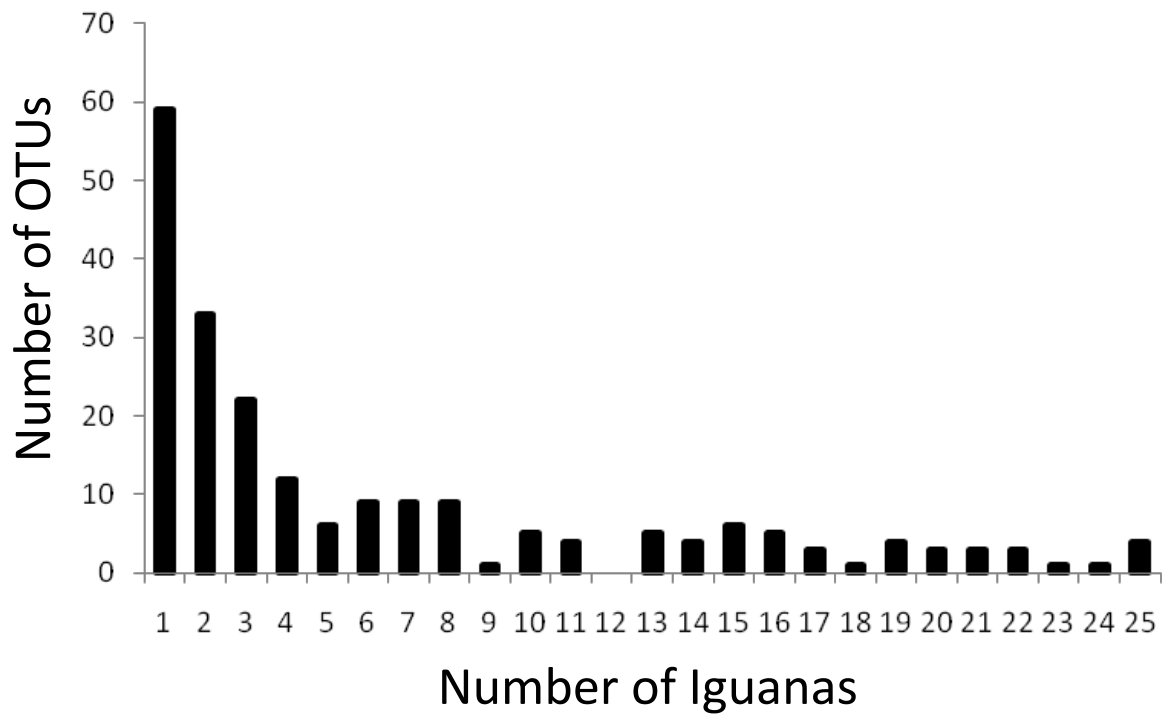


Figure 3.2. Frequency of *Clostridium* OTUs in individual marine iguana hosts. *Clostridium* OTUs demonstrated a skewed distribution with a relatively long right tail. This observed OTU distribution was used to parameterize a simulation model to predict island distributions of these OTUs if all islands shared a common community with this abundance structure.

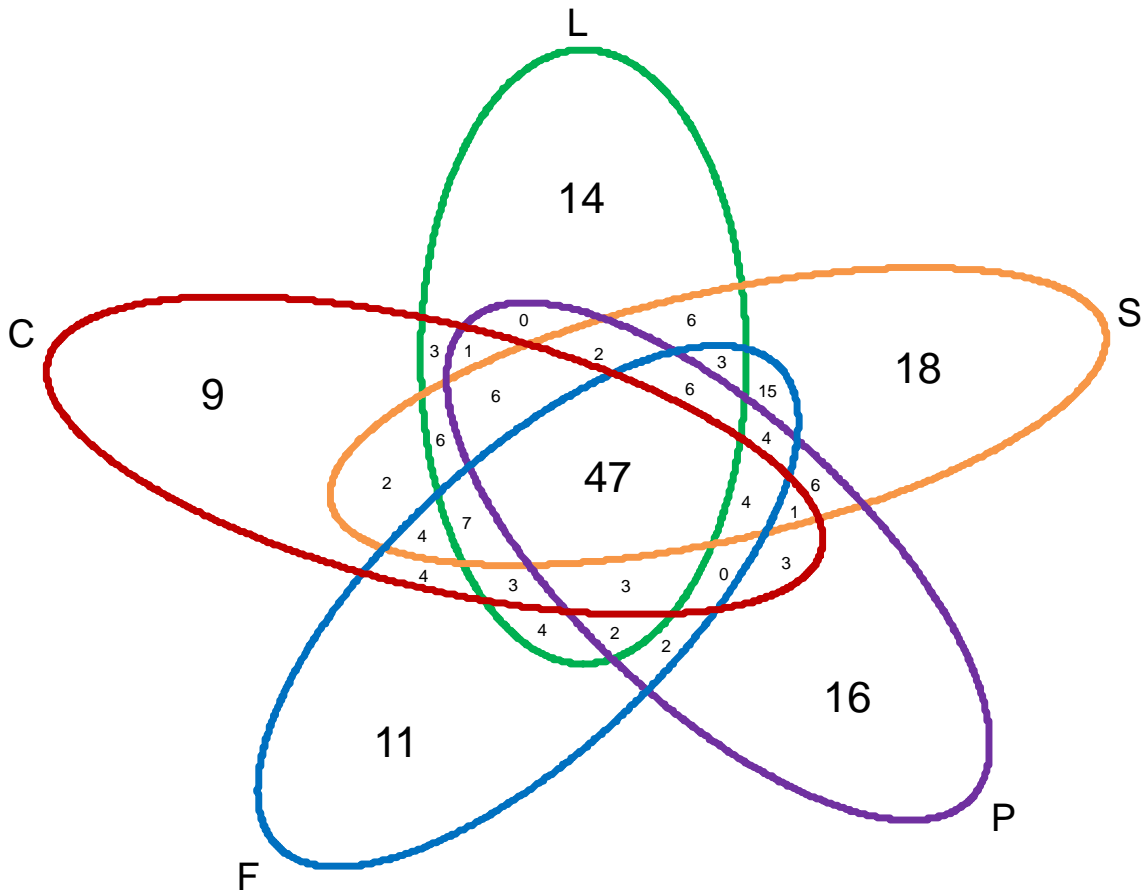


Figure 3.3. Venn diagram of shared OTUs across sites. Sampling sites are indicated by letter code (C=Punta Carola, L=La Lobería, S=Santa Fe, P=Plaza Sur, F=Fernandina). Numbers indicate count of OTUs overlapping in each section (sections not to scale due to complexity of overlap). Notably, a relatively large portion of OTUs were detected on all sites and a total of 68 OTUs were isolated to a single site. The remaining isolates were detected on 2, 3, or 4 sites.

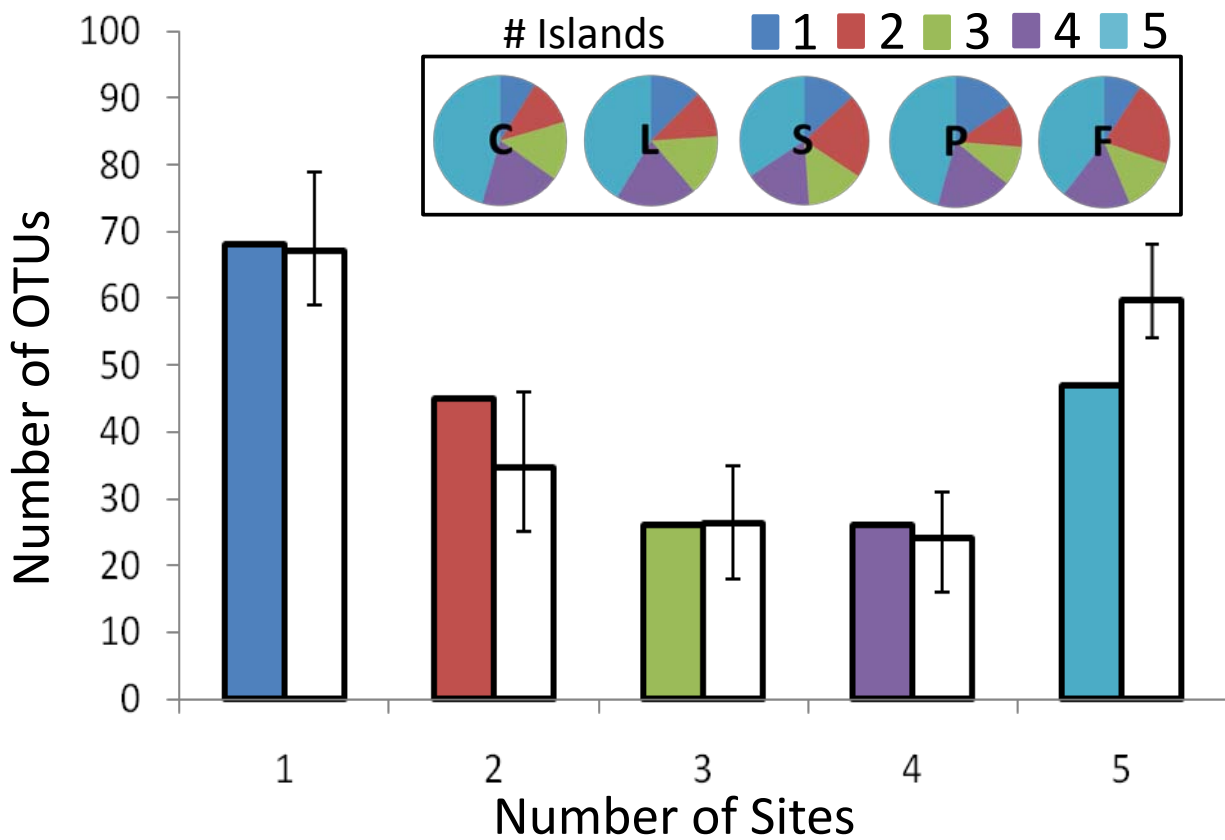


Figure 3.4. Comparison of observed and simulated *Clostridium* OTU frequencies among sites. Model results (open bars) shows the expected average number of OTUs +/- 1 SEM versus the number of islands on which each OTU is present as estimated by a simulation of 1000 sampling events of 212 OTUS distributed across 5 sites. The filled bars show the actual observed OTU frequencies at each site for comparison. Only comparison of model and simulations results for OTUs present at 5 sites showed significant differences. These frequency distributions did not differ across sites (inset, Fisher's Exact  $p > 0.05$ ).

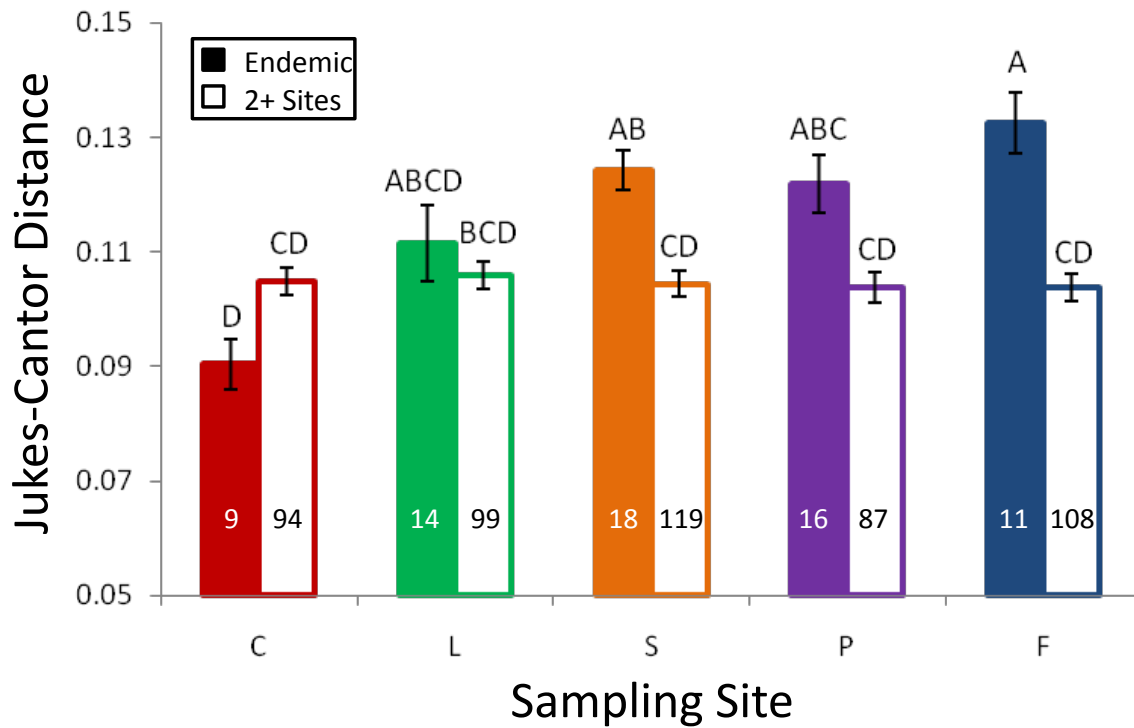


Figure 3.5. Comparison of phylogenetic variance within sites. Sampling sites are indicated by letter code (C=Punta Carola, L=La Lobería, S=Santa Fe, P=Plaza Sur, F=Fernandina). Endemic OTUs (filled bars) trended towards wider phylogenetic diversity than more widely distributed OTUs (found on two or more sites, open bars). The number of OTUs in each category are indicated within each bar.

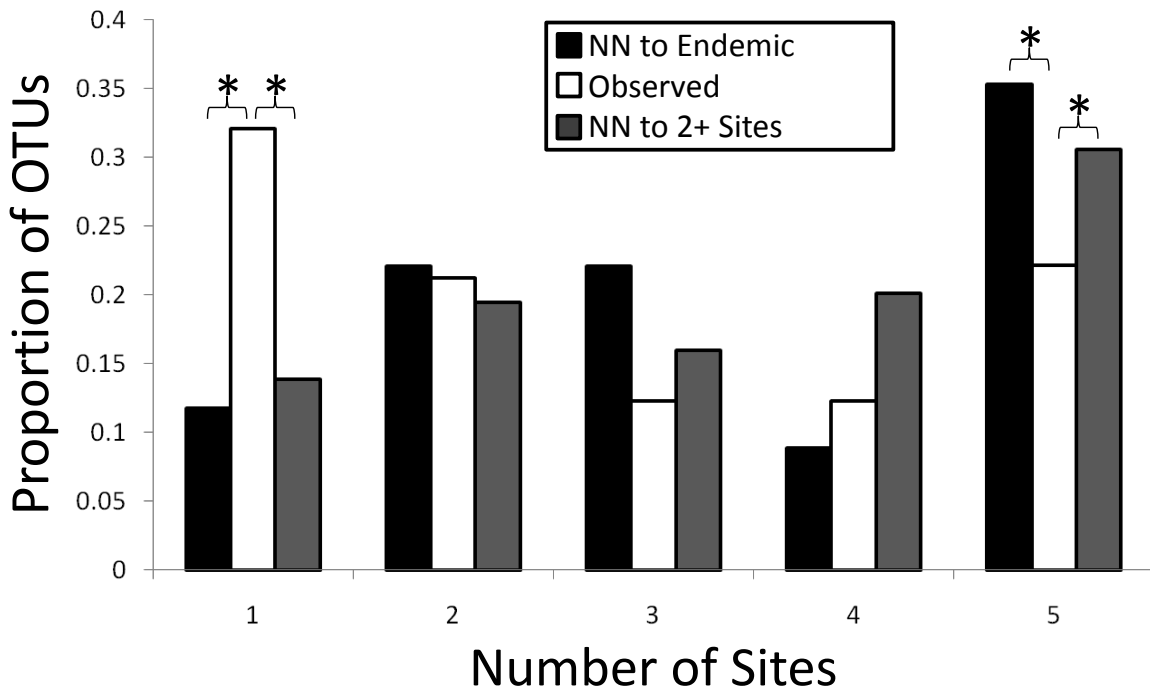


Figure 3.6. Site frequency of OTU nearest neighbors. The nearest neighbor (NN) of each OTU was identified as the OTU with the smallest phylogenetic distance. The site distributions of these NN OTUs are shown for endemic (black bars) and more widely distributed OTUs (grey bars). The observed distribution of OTUs (open bars) was used as a null model of the expected distribution if NN relationships are random. Both NN distributions (endemic and distributed OTUs) differed significantly from the random null model (Fisher's exact: Endemic  $p=0.003$ , 2+ sites  $p<0.001$ ). These distributions did not differ from each other (Fisher's exact:  $p=0.255$ ). Pairwise contrasts demonstrated greater than expected NN associations with widespread OTUs and a significant decrease in NN associations with endemic OTUs (z test:  $p<0.05$  marked with an \*).

## **CHAPTER 4: SEROTYPING AND GENOMIC FINGERPRINTING OF *SALMONELLA ENTERICA* FROM GALÁPAGOS IGUANAS REVEALS COMPLEX BIOGEOGRAPHICAL PROCESSES**

### **Chapter Summary**

Dispersal limitation is thought to primarily structure populations of host-associated bacteria because host distributions inherently limit transmission opportunities. However, enteric bacteria can quickly travel long distances on agricultural products during food-borne outbreaks. Thus, it is unclear if such long-distance dispersal events happen regularly in more natural systems, or if strong dispersal limitation is, in fact, the rule for host-associated microbes in nature. In this study, we explored population similarity and diversity patterns for *Salmonella enterica* isolated from the feces of free-living Galápagos land and marine iguanas from five sites on four islands. We evaluated *S. enterica* population patterns using phenotypic serotyping and genomic fingerprinting. Serotyping suggested a strong barrier to *S. enterica* transmission among sites, with each site hosting a unique and nearly mutually exclusive assemblage of strains. Genomic fingerprint analysis offered a more complex model of *S. enterica* biogeography. The primary pattern on genomic analysis agreed with serotyping results, suggesting site-specific strain assemblages. However, rather than demonstrating each site to be entirely independent of geographical forces as for serovar patterns, genomic fingerprinting also identified a clear geographical pattern. A southeast to northwest axis explained a significant portion of the variation in these banding patterns, indicating that some portion of the genetic variation driving site differences in these isolates follows a specific geographical pattern. These findings suggest that even relatively generalist enteric bacteria may be strongly limited in their ability to cross oceanic divides in a natural system, but yet a distinct and biologically meaningful geographical patterning is detectable within these strains, potentially indicating that the biogeography of the core and flexible portions of this species genome operate under different ecological processes.

### **Introduction**

Recent studies in bacterial biogeography have demonstrated that both dispersal-limitation and niche selection processes can contribute to bacterial distribution patterns (Cho and Tiedje 2000, Finlay 2002, Oda et al. 2003, Papke et al. 2003, Whitaker et al. 2003, Finlay and Fenchel 2004, Fontaneto et al. 2008, Takacs-Vesbach et al. 2008, Vos and Velicer 2008). The detection of characteristic population genetic patterns can indicate which of these processes is most dominant in a given system. When niche selection is the primary force shaping microbial distribution, a single strain or species is often widely or globally detected in patches of appropriate habitat (Finlay 2002, Finlay and Fenchel 2004, Fontaneto et al. 2008). In contrast, when stochastic



dispersal processes shape distributions, genetic patterning such as correlation between genetic and geographic distances among microbial populations can result (Cho and Tiedje 2000, Papke et al. 2003, Whitaker et al. 2003, Takacs-Vesbach et al. 2008, Vos and Velicer 2008). A distance-decay pattern (or isolation-by-distance relationship), where genetic similarity is highest in populations that are more geographically proximate, suggests that dispersal and genetic mixing occurs more often between sites that are closer together because of limitations in how far the organism is able to move (i.e. the converse of gene flow equals isolation-by-distance, Wright 1943).

Dispersal limitation has been proposed to dominate structuring of host-associated bacterial populations because of the inherent confines on microbial dispersal imposed by host species distributions (Fenchel and Finlay 2004, Falush et al. 2003, Hedlund et al. 2003, Zhang and Blackwell 2002, Bala et al. 2003, Papke and Ward 2004). While this may be especially true for obligatory or highly host-specific associations such as pathogens or co-evolved mutualists, this restriction may possibly be relaxed for host-associated microbes which have the ability to survive even temporarily in non-host environments or to associate with many different host species, resulting in increased transmission opportunities. For example, gastrointestinal bacteria responsible for food-borne outbreaks can persist for days to months in a variety of environmental sources (e.g. brackish water - Rhodes and Kator 1988, seawater - Rozen and Belkin 2001, swine manure - Guan and Holley 2003, produce- Harris et al. 2003, plants - Brandl 2006, fruit - Alegre et al. 2010, cattle manure - Semenov et al. 2010), resulting in wide and rapid dispersal by agricultural distribution systems (e.g. CDC 2006, 2007, 2009, 2010). Thus, for at least some host-associated bacteria, a single strain may quickly travel long distances, as during a food-borne outbreak, and thus would not necessarily be expected to demonstrate clear genetic-by-distance relationships. Thus, as for environmental microbes, both dispersal limitations and the distribution patterns such stochastic forces can create may also vary for host-associated microbes. However, it is unclear how often rapid long-distance dispersal events occur in more natural host-bacterial systems, or if such events are an artifact of anthropogenic acceleration via production systems.

*Salmonella enterica* is a gram negative Proteobacterium that primarily resides in the gastrointestinal tract of animals, but can also survive in environmental habitats,

foods and water (Foster and Spector 1995). While *Salmonella* are generally pathogenic to warm-blooded animals, reptiles are rarely reported to have *Salmonella*-associated illness (Scherer and Miller 2001, Geue and Löschner 2002). In reptile systems, the transmission ecology of *S. enterica* is not yet entirely understood. Studies in wild reptiles have suggested that commensal *S. enterica* populations in reptile hosts may be primarily structured by geographical barriers to migration with geographically proximate heterospecific reptile hosts sharing *S. enterica* strains (Briones et al. 2004, Wheeler et al. 2011). However, *S. enterica* is also a relatively ubiquitous gastrointestinal pathogen in warm-blooded animals, primarily associated with rapidly spreading food-borne outbreaks (e.g. CDC 2007, 2009). Further, *S. enterica*'s capacity for relatively lengthy extra-intestinal survival may also present increased opportunity to move among free-living hosts, even among host populations that are relatively isolated if ecological conditions facilitate such dispersal.

Island systems have long contributed to the study of how geographic isolation contributes to ecological and evolutionary processes, from Charles Darwin's early musings to contemporary microbial ecology theory (Darwin 1859, Mayr 1963, MacArthur and Wilson 1967, Diamond and May 1976, Krebs 1994, van der Gast 2008, Losos and Ricklefs 2009). The Galápagos Islands in Ecuador are particularly famous for these contributions and are host to 27 reptile species that are distributed throughout the island chain, including two ecologically distinct types of large herbivorous lizards, the marine and land iguanas. Marine iguanas (*Amblyrhynchus cristatus*) are locally abundant and widely distributed throughout the islands, living in dense colonies along the rocky coasts and feeding on marine algae (Trillmich and Trillmich 1986, Wikelski et al. 1993, Shepherd and Hawkes 2005). Land iguanas (*Conolophus subcristatus*, *C. pallidus* and the newly recognized *C. marthae*) are less social, often living alone or in small groups, (Kritcher 2006) and eat inland terrestrial vegetation (Christian et al. 1984). *C. subcristatus* is more widely distributed among the islands than *C. pallidus*, which is an ecologically similar sister-species endemic to the island of Santa Fe and the newly recognized pink land iguana species lives only on a Wolf Volcano, island of Isabella (Gentile et al. 2009).

Previous work in this system has demonstrated that these iguanas carry *Salmonella enterica* at high prevalence with unique, but not exclusive, strain subsets

present on the two islands sampled. In addition, site effects were a much stronger contributor to *S. enterica* population patterns than differences among host species, with land and marine iguanas from the same location sharing bacterial strains (Wheeler et al. 2011). These findings suggest that this system is a strong model for exploring the effect of strong host isolation on bacterial population structure, with reptile hosts across sites likely to provide selectively similar gut habitat across sites.

In this study, we ask the question: How isolated are *S. enterica* populations within hosts residing on different islands? In this study, we applied genomic fingerprinting and phenotypic strain typing (i.e. serotyping) to fecal-derived *S. enterica* isolates from Galápagos land and marine iguanas from five sampling sites on four islands. We hypothesized that if this strong host isolation results in a geographical barrier to enteric bacterial transmission, then each island should be host to a unique strain pool. However, if this isolation is not absolute, then we expected that dispersal would be more limited between more distant sampling sites, resulting in a genetic isolation-by-distance pattern. Alternatively, if dispersal among host populations is not the limiting factor in this system, then we would expect to find no significant divergence between *S. enterica* populations.

## **Methods**

### *Study design and sampling sites*

The Galapagos Islands are located approximately 1,000 km off the west coast of Ecuador and have a variety of unique endemic wildlife, including two unique types of large iguanas - land iguanas (*C. pallidus* and *C. subcristatus*) and marine iguanas (*Amblyrhynchus cristatus*). Five locations on four islands were sampled (Figure 4. 1, Table 4.1). Marine and land iguanas were observed from a distance for defecation and then fecal specimens were collected from the ground using a sterile wooden applicator. We intentionally selected fecal samples from different areas of each site and collected samples over a short period of time (24-48 hours) to avoid repeated sampling of the same individual. The slow gastrointestinal transit rates of herbivorous reptiles (e.g. estimated to be 5-10 days for marine iguanas, Wikelski et al. 1993), reduces the likelihood of observing a single individual defecating twice within a short sampling period. Fecal specimens were obtained from both iguana types on Plaza Sur ( $N_{\text{land}}=11$ ,

$N_{\text{marine}}=12$ ), Fernandina ( $N_{\text{land}}=6$ ,  $N_{\text{marine}}=12$ ), and Santa Fe ( $N_{\text{land}}=11$ ,  $N_{\text{marine}}=12$ ), where both species are present and only marine iguanas from the sampling sites on San Cristobal (Loberia  $N=12$ , Punta Carola  $N=12$ ), where land iguanas are not present (see sample summary in Table 4.1).

Samples were placed in sterile plastic tubes, stored at 4°C during travel and transported to Urbana, Illinois for storage at -20°C until cultivation could be completed.

#### *Isolation and identification of S. enterica*

*S. enterica* was isolated from frozen fecal samples using previously published protocols (Hoorfar and Baggesen 1998, Mitchell and Shane 2000, Corrente et al 2004, Wheeler et al.2011). Briefly, 0.5g of feces was pre-enriched in buffered peptone water (BPW) at 37°C for 24 hours followed by enrichment in a 1:9 ml dilution of turbid BPW:Rappaport-Vassiliadis broth (RVB) at 37°C for another 24 hours. Turbid RVB was then streak plated on selective-differential media (Xylose lysine deoxycholate, XLD). XLD plates demonstrating no *Salmonella*-characteristic growth at 24 hours were re-incubated for an additional 24 hours before being considered negative for growth. Samples negative for *Salmonella*-like isolates were retested from the original fecal samples before being considered negative for the bacterium.

Presumptive *S. enterica* isolates (clear to pink with or without a dark black center) were tested using standard biochemical metabolic assays (lysine iron agar and triple sugar agar) to confirm membership in the genus. Isolates testing positive on both of these assays were considered confirmed to the *Salmonella* genus, as previous work in this system has demonstrated a 100% concordance between confirmation of *Salmonella* using these two assays following isolation on XLD selective media when compared to identification using API 20E metabolic test strips (bioMerieux Inc., France; E. Wheeler, unpublished data). Up to three confirmed *Salmonella* isolates per individual fecal sample were retained frozen in 40% glycerol for further analysis

#### *Strain typing*

Serotyping is a phenotypic bacterial typing approach which identifies the combined O (cell membrane) and H (flagellar) antigenic variation among strains. Over 2500 of these antigenic variants (i.e. serovars) have been characterized based on cell

membrane (“O”) and flagellar (“H” phase 1 and 2) antigens (Popoff and Le Minor. 1997). Previous work has demonstrated strong concordance between sequence-based genotyping and serotype identity (Tankouo-Sandjong et al. 2007, Ben-Darif et al. 2010, Wheeler et al. 2011), suggesting that serotyping is a reasonable strong phenotypic marker for genetic strain relationships. One isolate per positive sample was submitted for serotyping at the National Veterinary Service Laboratory, Ames, IA.

While serotyping is often sufficient for drawing epidemiological links between sources and cases during an outbreak, higher resolution typing of strains is often desirable to better understand outbreak dynamics. Genotyping techniques which target genomic-level variation have proven most successful at differentiating closely related strains of *Salmonella*. While pulse field gel electrophoresis (PFGE), is the current “gold standard” used by the Centers for Disease Control for typing of *Salmonella* (Olsen et al. 1994), rep-PCR has been demonstrated to potentially have higher resolution power for closely related strains than PFGE (Beyer et al. 1998, Weigel et al. 2004) and presents an efficient and inexpensive approach for rapid genotyping of large numbers of isolates. Rep-PCR produces genomic fingerprints through amplification of intervening sequences between repetitive elements of the bacterial genome, resulting in numerous bands of varied size that can be separated by agarose gel electrophoresis.

In previous work, genomic fingerprinting and traditional serotyping were demonstrated to provide ecologically relevant resolution of *S. enterica* population patterns in this system (Wheeler et al. 2011), and thus these two methods were selected for use in this study. DNA was extracted from colonies grown on tryptic soy agar using a modification of the manufacturer’s protocol for Instagene chelex resin (Biorad, Hercules, CA). Briefly, 2-4 large, well-isolated colonies were placed in 100 µl of 6% chelex 100 resin and heated at 95°C for 10 minutes to lyse cells. Lysates were centrifuged for 5 minutes at 14000 g and 80µl of supernatant was transferred to a new storage tube.

Genomic fingerprinting was done by PCR of repetitive extragenic palindromic elements (rep-PCR, Versalovic et al. 1991). Rep-PCR was performed as in previously described protocols (Versalovic et al. 1991, Rademaker et al. 1998). Rep-PCR amplification was performed using primers ERIC1R (5’-ATG TAA GCT CCT GGG GAT TCA-3’) and ERIC2 (5’-AAG TAA GTG ACT GGG GTG AGC G-3’), targeting

enterobacterial repetitive intergenic consensus (ERIC) repetitive motifs (Versalovic et al. 1991) on a TGradient thermocycler (Biometra, Germany) with an initial denaturation at 95° C for 2 minutes followed by 30 cycles each of 94° C for 3 seconds, 92° C for 30 seconds, and 50° C for 1 minute and a final extension at 65° C for 8 minutes. Polymerase chain reaction amplification mixtures (25µl) included 1.75U of Takara Taq polymerase (Takara Bio Inc., Japan), 1X Takara PCR Buffer with 2.5mM (final concentration) of MgCl<sub>2</sub>, 2.0mM Takara dNTP mixture (0.5mM each), 1µM each of the forward and reverse primers, and approximately 50ng of template. PCR products were separated on a 2% Agarose gel in 0.75% TAE (Tris-Acetate EDTA) run at 80 V for 12 hours. Three lanes of a 1kb plus DNA ladder (Novartis, Carlsbad, CA) were included to allow for standardization of molecular weight assignments to DNA fragments, with one lane at each end of the gel and one located in the approximate middle to allow for correction of irregular gel runs. Gels were stained with ethidium bromide and digitally photographed using an AlphaImager™ 2200 (Alpha Innotec Co., San Leandro, CA).

### *Statistical analysis*

Rep-PCR band assignment and sizing from gel images was done using BioNumerics version 4.0 (Applied Maths, Belgium). Band assignments were exported as a binomial presence/absence matrix for statistical analysis. To control for non-independence (i.e. pseudoreplication, Hurlbert 1984) resulting from inclusion of multiple isolates per individual iguana fecal sample, fingerprint profiles of isolates from the same individual iguana were nested by averaging presence/absence within an individual host.

To evaluate the effects of explanatory variables (island, site, host ecotype, and geographical coordinates of site) on bacterial population structure, we used the *adonis* function in the *vegan* package in R (Oksanen et al. 2008) to perform a series of permutation or non-parametric multivariate analysis of variance analyses. These models partition the sums of squares of distance matrices among treatments and has relaxed assumptions relative to traditional MANOVA. Significance in the permutation tests was determined by comparing the observed effects against 5000 random permutations of the data for each model run independently. As island and site are nested subsets of one

another, they were first run in separate models including ecotype as a second factor and the models were compared to select the most significant of the two for additional geographic analyses. Interaction terms were not included as land iguanas were only present for sampling at 3 of 5 sites. Geographical location was included in subsequent permutation MANOVA models first as the approximate latitude and longitude coordinates of each site and then as a rotated axis (i.e. the principal component scores derived from the sites' latitudes and longitudes) to reflect the primarily northwest-southeast axis along which the Galápagos islands are distributed, via a principal components analysis of the original coordinates which produced two orthogonal axes oriented along a southeastern to northwestern axis (PC1) and along a northeastern to southwestern axis (PC2). PC1 and PC2 were also included as explanatory variables in permutation MANOVA. Genetic differences in many plant and animal species in the Galápagos follow a southeast-to-northwest directional gradient of genetic diversity, with the most divergent and isolated populations located on the older islands in the eastern and central islands and younger, more genetically recent populations in the northwestern sites (i.e. the "progression hypothesis," Parent et al. 2008, Steinfartz et al. 2009). Thus, this principal components rotation of the primary latitudinal and longitudinal axes has the potential to more fully capture the primary geographical pattern in this system by capturing the majority of the variation present in the geographical axis into a single variable (PC1).

We then performed multivariate ordinations on the genomic banding patterns to visualize the significant effects detected by permutation MANOVA, using both an unconstrained and a constrained correspondence analysis (CA and CCA, respectively, both performed with the *cca* function in the *vegan* package by Oksanen et al. 2010). CCA was performed using sampling site as the constraining variable, as this factor explained the largest amount of variation (i.e. had the highest  $R^2$ ) in the various permutation MANOVA analyses. We then performed an environmental fitting of the previously described PC1 and PC2 location axes on the CCA patterns using the *envfit* function in the *vegan* package of R (Oksanen et al. 2010). We used the constrained correspondence analysis to isolate those aspects of the genomic information which showed the strongest geographic structuring, and used the environmental fitting function to test whether this

geographically structured subset showed directional patterns of similarity as predicted by the nature of island development and ocean currents in the Galapagos islands.

A Jaccard dissimilarity matrix was calculated using the `vegdist` function in the `vegan` package in R (Oksanen et al. 2010). This matrix was then used to calculate the average genetic dissimilarity of isolates from within the same sampling sites for isolates derived from marine iguanas only. While previous work has demonstrated the geographical distance is a strong force for shaping these communities, thus allowing pooling of isolates from both host ecotypes for between site comparisons, that study also showed that on a more subtle level, strain carriage may also differ between host ecotypes, either due to ecological differences in bacterial exposures or due to host-bacterial interactions resulting in strain sorting (Wheeler et al. 2011). For this reason, to evaluate within-site patterns we performed this analysis only on strains isolated from marine iguanas in order to hold the selective environment constant, which resulted in the elimination of Santa Fe from the analysis due to low marine iguana sample size.

The average within-site dissimilarity was used to test for differences in strain diversity among sites using ANOVA with a Tukey-HSD multiple comparison to identify the sites which differed from one another in isolate similarities. For each isolate, we calculated the average dissimilarity of that isolate to all other isolates from the same site, resulting in one average dissimilarity per isolate. We then used ANOVA to test whether the sites differed in the degree of genetic dissimilarity among their component isolates. A significantly higher average dissimilarity among isolates from the same site suggests a larger variety of divergent fingerprint patterns, and thus, higher strain or genomic diversity at that site.

Basic statistical and graphing procedures were performed in either JMP 8 (SAS Institute, Cary, NC) or R statistical language (R Development Core Team 2008).

## **Results**

### *Prevalence*

Prevalence of *S. enterica* in Galápagos reptile populations ranged from 0% to 75%, with no significant difference in carriage rates between marine and land iguanas (Fisher exact test  $p=0.140$ , Table 1). *S. enterica* carriage varied significantly across sites



for marine iguanas (Fisher exact test  $p < 0.001$ ) but were not significantly different among sites for land iguanas (Fisher exact test  $p = 0.3427$ ).

### *Serotype patterns*

We identified a total of 23 unique serotypes (plus one untyped *Salmonella* presumptive isolate) distributed among the 5 sites sampled (Table 4.2). Serovar richness was similar for most sites when controlled for the number of positive individuals sampled, but was slightly higher at Santa Fe and Fernandina (serovars/positive samples: C=0.56, L=0.56, SF=0.67, P=0.47, and F=0.75). Fernandina also had a larger proportion of strains were not able to be fully serotyped (i.e. had a “rough” O antigen phenotype).

### *Genomic fingerprint patterns*

The dominant pattern detected in fingerprint assemblages reflects the strain or serogroup level differences seen in serotyping (correspondence analysis : Figure 4.1 a). In contrast, in this unconstrained analysis, grouping of isolates by site differences are readily not visible (Figure C.1). However, permutation MANOVA analysis of genomic fingerprint patterns does indicate significant differences among groups at both the site and island level of sampling (Table 4.3). Both island and site were significant factors for isolate similarity in these models, however site differences explained a higher proportion of the variation (Table 4.3,  $R^2 = 0.137$ ,  $p = 0.004$  for site versus  $R^2 = 0.089$ ,  $p = 0.053$  for island). Thus subsequent models included site and either latitude and longitude of the sampling locations or a rotated axis derived from the principal components of latitude and longitude representing approximately southeast-northwest (PC1) and southwest-northeast (PC2) directions. While latitude, longitude and PC2 were not significant, PC1 was a significant factor for explaining isolate similarities (Table 4.3). PC1 was also a significant environmental fit to the site constrained canonical correspondence analysis, indicating that site groupings align directionally in a southeast to northwest direction.

### *Within site variability*

Both ordination-based (Figure C.1) and dissimilarity-based (Figure C.2) analyses suggested visually that genomic fingerprint diversity differed among sites. Thus, this potential difference in variance within sites was evaluated by performing ANOVA by site on the average within-site Jaccard dissimilarities of isolates derived from marine iguana feces. Sites varied significantly in the average isolate dissimilarity within each site (ANOVA by site:  $R^2 = 0.6094$ ,  $p < 0.001$ ), with Punta Carola on San Cristóbal Island and the island of Plaza Sur having higher within-site variability compared to La Loberia on San Cristóbal and the Fernandina sampling site (Figure 4.3, Tukey's HSD).

## **Discussion**

### *S. enterica populations are geographically structured*

Reptile populations are often host to diverse *S. enterica* strain pools, and the serovar richness per positive individual detected in this study was comparable to that seen in previous studies (average of 7 studies 0.46 strains/positive sample, range of 0.19-0.75: Monzón-Moreno et al. 1995, Mitchell and Shane 2000, Briones et al. 2004, Corrente et al. 2004, Hidalgo-Vila et al. 2007, Pederson et al. 2009, Wheeler et al. 2011). However, this is, to the best of our knowledge, a relatively novel dataset in that we have extended such explorations of serovar diversity to understanding reptile host-*Salmonella* ecology on an explicitly spatial scale.

Both serovar and genomic fingerprints demonstrate strong site differences, suggesting that geographic isolation is the primary force structuring *S. enterica* populations in this system. The wide variation in serovar composition from site to site further suggests that stochastic processes, in particular, drive the assemblage of local *S. enterica* populations from the regional Galápagos strain pool by primarily stochastic migration events, rather than selection for particular strains. When stochastic dispersal limitation is a primary driver of population or community composition, there should be considerable site-to-site variation as bacterial strains arrive and establish by chance at each site independently (Condit et al. 2002, Chave and Leigh 2002). In contrast, when assembly of diversity is primarily driven by niche selection composition should be similar among similar selective environments (Chase 2003, Tuomisto et al. 20003, Dornelas et al. 2006). While we cannot with these data entirely rule out contributions of

either host-bacterial or strain-strain interactions to development and maintenance of strain assemblages, the observed pattern supports rare stochastic migration among sites as a driving force for determining *S. enterica* population structure.

While dispersal limitation does appear to be key for structuring *S. enterica* populations, strikingly, we did not see a strong correlation between genetic and geographic distances among sites that is often used as a marker for dispersal-limited systems (Figure C.1). In the case of isolation-by-distance, sites within the same island, such as La Loberia and Punta Carola, would be expected to be the most similar in strain identities, yet these sites at under 5 km from one another differed from each other as strongly as they both differed from the *S. enterica* population on Fernandina, over 200 km from the island of San Cristóbal.

The lack of simple distance-related similarity patterns in both serovar and non-constrained correspondence analysis of genomic fingerprints may be driven by two processes, which are not mutually exclusive in their application. Firstly, serovar differences among sites are very strong in this system, suggesting that the migration barrier among islands is quite strong, nearly absolute on ecological time scales detectable with these sample sizes. Thus, within-, rather than between-site dynamics may dominate, with each strain assembly changing over time due to within-site population drivers (stochastic or selective changes in strain dominance and retention). If these within-site differences occur more rapidly than rare migration events between sites, then any dispersal-based similarities between even relatively proximate bacterial populations would be lost to detection, swamped by the more rapid within-site effects and resulting in populations that appear equally different at the strain level, regardless of distance.

It is likely that genetic relationships among serovars represent a deeper evolutionary strain history, rather than the recent ecological interactions among these sites. We see suggestion of this in unconstrained ordination (correspondence analysis) of genomic fingerprint patterns, where isolates from the same serovar or O-antigen serogroup are generally located in proximate ordination space. A quantitative test of this hypothesis would be difficult with this diverse strain assemblage, but qualitatively we see support for this line of reasoning in our data that would warrant additional study

on the role of deep versus recent strain history in determining diversity patterns in this system.

### *Conflicting models of biogeographic distributions*

Given the strong site patterns driven by serovar composition seen on both serotyping and genomic fingerprint analysis, it is reasonable to conclude that the barrier to bacterial migration among sites is quite strong. Yet at least a small proportion of the genomic-level variation on fingerprint analysis demonstrates a striking directional geographic relationship among bacterial populations, with site similarities aligning along a southeast to northwest gradient. These findings together suggest that even if dispersal of specific strains of *S. enterica* among islands is rare, some aspect of *S. enterica* genetic diversity is free to move among islands and is under some specific influence of dispersal limitation. Given the often high degree of horizontal gene transfer often detected in bacterial communities, generally (as reviewed in Thomas and Nielsen 2005, Juhas et al. 2008, Wozniak and Waldor 2010), and within marine iguana intestinal communities (Nelson et al. 2010), it is possibly that these directional geographical relationships are driven by movement of extragenomic elements that may have higher mobility than their enteric bacterial hosts. Thus, within a given site, contiguous strains appear to share certain genomic elements or traits which are local to the strain pool, despite not sharing a deep genetic history on the serovar level.

This potential for disconnect between evolutionary and ecological history based on different typing methods was hinted at in a previous work in this system (Wheeler et al. 2011). Further, this finding is not inconsistent with current knowledge of *Salmonella* genetic diversity. *S. enterica* shares a large portion of its genome with related bacteria such as *Escherichia coli*, but has in addition the tendency to flexibly gain and lose large, mobile pathogenicity islands (Groisman and Ochman 1997). Whereas some *Salmonella* pathogenicity islands (SPI), such as SPI1 and SPI2, are common to all *Salmonella enterica* serovars, others are variably present only in some subspecies or strains (Morgan 2007). These variable genetic elements compose a major portion of genetic variation between strains. Full genome sequence comparisons of diverse *Salmonella* serovars reveals that genetic differences among strains is primarily based on insertions and deletions of such mobile chromosomal elements, rather than on coding region

sequence divergence of the types of genes typically targeted for sequence analysis (Edwards et al. 2002). Such discordance between genetic markers has also been observed for *Escherichia coli*, where multilocus enzyme electrophoresis of *E. coli* reference strains demonstrated six distinct phylogenetic groups while rep-PCR based analysis grouped these same isolates primarily by species of origin (Ishii and Sadowsky 2009).

If variation in phenotypic (i.e. serovar) and genomic (i.e. Rep PCR banding patterns) diversity are influenced by different ecological and evolutionary pressures or temporal scales, then this would explain the observed disjunction in biogeographical patterns seen in this study. Given *S. enterica*'s tendency to flexibly acquire genomic elements like pathogenicity islands, it would not be unreasonable to suspect that gains and losses of these elements may occur within local populations and may contribute to differences in genomic fingerprint patterns on a site-to-site basis. These mobile elements may represent *Salmonella*- or Enterobacteracea-specific genetic elements or more broadly transposable elements (e.g. transposons, viruses, etc.) that can be horizontally acquired from a wide diversity of bacterial interactions within the enteric or environmental habitat.

Even with this explanation, what is particularly striking about the detected site differences by genomic fingerprinting is the geographical directionality of site similarity patterns within the site constrained ordination. The directional pattern is biologically quite a meaningful finding for this system. While simple site differences would just suggest within-site horizontal gene transfer dynamics at play that parallel the detected serovar differences, such within-site dynamics in isolation would not necessarily result in a specific geographical association among the sites. Each site could operate independently on a genomic variation level, as appears to be true at the strain level, rather than demonstrate a geographically-based patterning. These population patterns in microbial species align with the historical development of the island chain and may be, in part, driven by ocean current movements in this region, or other directional processes. The Cromwell and Humbolt currents and associated trade winds sweep across the Galápagos Island chain in a primarily east-west or southeast-northwest direction for most of the year, and may potentially carry with them a variety of bacterial types or genetic elements in their travels. So it is possible that in this system dispersal is

stochastic in the sense of what traits or strains may move among sites, but non-random in the direction towards which dispersal most readily proceeds. Thus, detection of broad isolation by distance patterns may not be the most appropriate test for dispersal-limitation patterns if directionality is not taken into account in both this, and other similarly structured systems.

### *Local exposures drive within-site population dynamics*

As discussed previously, strong site differences in this study suggest that within, rather than between site process may be the primary drivers of *S. enterica* population diversity, a proposition further supported by the detection of site differences in genomic diversity. Holding selective environment as constant as possible for within site analysis by including only isolates obtained from marine iguanas allows us to focus primarily on the potential stochastic processes which might contribute to within-site diversity patterns.

In the classical ecological theory of island biogeography, the number of types present on a given island are determined by the equilibrium between gain of new species and the loss of existing ones (MacArthur and Wilson 1967). By analogy, the detected differences in genomic diversity are likely dependent on both the rate of strain or trait flow into a site and the chance or selective loss of these elements. In our system, two site-dependent factors, in particular, stand out as potential drivers of diversity gains and losses – proximity to sources of new genetic material, such as humans or other animal populations and host population size, which is likely a host-associated bacterium parallel of island size in terms of effect on extinction rate or loss of elements. The larger the host population, the more strains it should support as random loss of any one strain is less probable when distributed among more hosts.

Given this theoretical framework, we would predict that larger populations or those at closer proximity to sources of new variation should have more variable genomic traits than smaller populations or those farther from a source of new diversity (i.e. alternative host populations or environmental reservoirs of acquirable traits). While we cannot fully isolate the effect of these two elements in this system with samples from only four marine iguana populations, it is notable that the highest within-site diversity was detected at Punta Carola - a site with a very low marine iguana density and a high

likelihood of contact with human fecal waste from the sewage outlet of Puerto Baquerizo Moreno – and at Plaza Sur – a very small island with extremely high marine iguana and land iguana densities only 20 km northwest of a second major port town, Puerto Ayora. However, it is particularly striking that La Lobería – a site less than 2 km south of Puerto Baquerizo Moreno, but on the opposite side of the island from the main oceanic sewage pipe. However, additional sampling would be required to more fully document how these types of within site factors affect strain or trait diversity patterns in this system and to explain what processes at such proximate sites would explain this strong difference in population patterns.

## **Conclusions**

Our findings suggest that *S. enterica* genetic patterns in the Galápagos Islands derive from a complex overlap of both geographic and temporal processes, and fully exploring this complexity was dependent on evaluating strain diversity in more than one manner. Traditionally, biogeographical studies of plants and animals have relied on such combined approaches to fully explore how migration and selection combine to shape geographic patterns. In particular, it is often important to evaluate both neutral and trait-based markers as their geographic patterns will reflect different driving forces (stochastic versus deterministic, Hendry 2002, McKay and Latta 2002, Porcher et al. 2006, Steane et al. 2006, Gandour et al. 2008). Given the growing understanding of how horizontal gene transfer can contribute to patterns in bacterial genetics, it is a reasonable parallel to suggest that bacterial biogeography might also benefit from simultaneous measures of diversity at different loci or levels.

Additionally, this study demonstrates that dispersal-based biogeographical patterns can be more complex and subtle than simple distance decay, especially in systems with a strong directionality driven by abiotic conditions (e.g. streams, oceanic currents, human or animal transit routes). While our study is certainly limited both by sample size and the number of sites compared, our data suggest that in this system that stochastic dispersal limitations can produce patterning other than the typically-evaluated distance-decay relationships. Further, these data suggest that stochastic processes may act at different scales to produce complex diversity patterns both within and between sites. While the host relationship can provide additional complexity to

predicting and interpreting the contributions of neutral and selective forces on bacterial population patterns, the insights that can be gained from these well-studied systems certainly translates to other types of microbial systems. For example, expectations of increased strain diversity when a host population is either larger in number or more proximate to other host species populations, especially populations with highly mobile membership like human tourists, should run parallel to predicted effects of species-area and habitat patch proximity on diversity in a variety of environmental microbial systems.

Thus, these findings may offer two broader contributions to discussions of bacterial biogeographical patterns and the processes that drive these distributions (Bell et al. 2005, Fenchel and Finlay 2005, Martiny et al. 2006, Whitfield 2005). Firstly, as an occurrence common to the initial development of these classical community ecology theories in “macrobial” systems (Ricklefs 2004, Chase and Leibold 2003), this tension between the effect of dispersal processes and habitat or biotic selection on microbial movement and community assembly is certainly not new. However, in such classical ecological systems, recent synthesis increasingly supports a shift in emphasis from exclusive outcomes to exploring how specific ecosystem traits might result in difference in the relative contribution of neutral or selective forces on population and community processes (Chase 2007). Further, these processes may produce much more complex or subtle genetic patterns than simple isolation-by-distance if dispersal potential is asymmetrical (Cook and Crisp 2005); however, simple isolation-by-distance models certainly remains a reasonable initial model for consideration (see review by Jenkins et al. 2010). Only a few studies to date have begun to incorporate potential asymmetries in dispersal potentials into understanding microbial population or community biogeography (Boyer 2008, Fierer et al. 2007).

Our data demonstrate that gastrointestinal bacteria can and should be included in current discussions of microbial biogeography on the global scale. While it has been suggested that host-associated microbes are of no interest to such discussions due to the inherent dispersal limitation created by obligatory host associations (Fenchel and Finlay 2004), we would argue that well studied host-bacterial systems have much to contribute. While host-associated microbes may be more dispersal limited than some free living groups, these organisms, along with free-living microbes which live in highly



specialized and potentially spatially restricted habitats, represent one end of a gradient of dispersal limitation among microbial organisms. Even within host-associated microbial systems, there exists a wide variation in host-specificities and in ability to survive in the environmental matrix between hosts, a range which in itself may be useful a model for better understanding what microbial traits lead to more restricted versus catholic distributions due to dispersal restrictions. Comparison of different markers of bacterial divergence, especially in well-characterized systems in which understanding of the relationship of these markers to bacterial evolution processes is more advanced, can produce illuminating conflicts which may increase understanding not only of bacterial biogeographical patterns, but may also begin to reveal the complex processes which create them.

### **Chapter Acknowledgements:**

We would like to express appreciation to the Galápagos National Park and Washington Tapia, and the Charles Darwin Research Station and Sonja Cisneros and Paulina Couenberg for facilitating sample collection. Many thanks to Augusto G. Haz Beltran, Lenin Cruz Beldon and the crew of the Pirata for help with sample collection. We also wish to thank the Department of Animal Sciences and Nancy Henry for administrative support. Thanks to Richard Lankau for providing statistical support and constructive comments on early drafts of this chapter, and Katie Amato for constructive comments as the work developed. Thanks to Isaac Cann and Bryan White for access to laboratory resources.

Funding for this work was provided by the Conservation Medicine Center of Chicago Student Research Grant (E.W.) and by the United States Environmental Protection Agency Science to Achieve Results Fellowship program (E.W.). This chapter was developed with support from a STAR Research Assistance Agreement No. 91684301-1 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this chapter.

## References

- Alegre I, Abadias M, Anguera M, Usall J, Viñas I. 2010. Fate of *Escherichia coli* O157:H7, *Salmonella*, and *Listeria innocua* on minimally processed peaches under different storage conditions. *Food Microbiology* 27: 862-868.
- Bala A, Murphy P, Giller KE. 2003. Distribution and diversity of rhizobia nodulating agroforestry legumes in soils from three continents in the tropics. *Molecular Ecology* 12: 917–929.
- Bell T, Newman JA, Thompson IP, Lilley AK and van der Gast CJ. 2005. Bacteria and island biogeography. *Science* 309: 1998-1999.
- Ben-Darif E, De Pinna E, Threlfall EJ, Bolton FJ, Upton M, Fox AJ. 2010. Comparison of semi-automated rep-PCR system and multilocus sequence typing for differentiation of *Salmonella enterica* isolates. *Journal of Microbiological Methods* 81: 11-16.
- Beyer W, Mukendi FM, Kimmig P, Bohm R. 1998. Suitability of repetitive-DNA-sequence-based PCR fingerprinting for characterizing epidemic isolates of *Salmonella enterica* serovar Saintpaul. *Journal of Clinical Microbiology* 36: 1549-1554.
- Boyer DG. 2008. Fecal coliform dispersal by rain splash on slopes. *Agricultural and Forrester Meteorology* 148: 1395-1400.
- Brandl MT. 2006. Fitness of human enteric pathogens on plants and implications for food safety. *Annual Review of Phytopathology* 44: 367-392.
- Briones V, Téllez S, Goyache J, Ballesteros C, del Pilar Lanzarot M, Domínguez L, Fernández-Garayzábal JF. 2004. *Salmonella* diversity associated with wild reptiles and amphibians in Spain. *Environmental Microbiology* 6: 868-871.
- CDC. 2006. Ongoing multistate outbreak of *Escherichia coli* serotype O157:H7 infections associated with consumption of fresh spinach – United States, September 2006. *Morbidity and Mortality Weekly Reports* 55: 1-2.
- CDC. 2007. Multistate outbreak of *Salmonella* serotype Tennessee infections associated with peanut butter – United States, 2006-2007. *Morbidity and Mortality Weekly Reports* 57: 521-524.
- CDC. 2009. Outbreak of *Salmonella* Serotype Saintpaul Infections associated with eating alfalfa sprouts – United States, 2009. *Morbidity and Mortality Weekly Report* 58: 500-503.

- CDC. 2010. Two multistate outbreaks of shiga toxin–producing *Escherichia coli* infections linked to beef from a single slaughter facility – United States, 2008. *Morbidity and Mortality Weekly Report* 59: 557-560.
- Chase JM. 2007. Drought mediates the importance of stochastic community assembly. *Proceedings of the National Academy of Science* 104: 17430-17434.
- Chase JM, Leibold MA. 2003. *Ecological niches: Linking classical and contemporary approaches*. University of Chicago Press: Chicago, IL.
- Chase JM. 2003. Community assembly: When should history matter? *Oecologia* 236: 489-498.
- Chave J, Leigh EG, Jr. 2002. A spatially explicit neutral model of beta-diversity in tropical forests. *Theoretical Population Biology* 62: 153–168.
- Cho JC, Tiedje JM. 2000. Biogeography and degree of endemism of fluorescent *Pseudomonas* strains in soil. *Applied and Environmental Microbiology*. 66: 5448-5456.
- Christian KA, Tracy CR, Porter WP. 1984. Diet, digestion, and food preferences of Galápagos land iguanas. *Herpetologica* 40: 205-212.
- Condit R, Pitman N, Leigh EG, Jr, Chave J, Terborgh J, Foster RB, Núñez P, Aguilar S, Valencia R, Villa G, Muller-Landau HC, Losos E, Hubbell SP. 2002. Beta-diversity in tropical forest trees. *Science* 295: 666–669.
- Cook LG, Crisp MD. 2005. Directional asymmetry of long-distance dispersal and colonization could mislead reconstructions of biogeography. *Journal of Biogeography* 32: 741-754.
- Corrente M, Madio A, Friedrich KG, Greco G, Desario C, Tagliabue S, D’Incau M, Campolo M, and Bueonavoglia C. 2004. Isolation of *Salmonella* strains from reptile faeces and comparison of different culture media. *Journal of Applied Microbiology* 96: 709-715.
- Darwin C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray: London.
- Diamond JM, May RM. 1976. Island biogeography and the design of natural reserves. In *Theoretical Ecology: Principles and Applications* (May RM, ed.), pp. 163–186. Blackwell Scientific: Oxford, UK.

- Dornelas M, Connolly SR, Hughes TP. 2006. Coral reef diversity refutes the neutral theory of biodiversity. *Nature* 440: 80–82.
- Edwards RA, Olsen GJ, and Maloy SR. 2002. Comparative genomics of closely related salmonellae. *Trends in Microbiology* 10: 94-99.
- Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, Blaser MJ, Graham DY, Vacher S, Perez-Perez GI, Yamaoka Y, Mégraud F, Otto Kristina, Reichard U, Katzowitsch E, Wang X, Achtman M, and Suerbaum S. 2003. Traces of Human migrations in *Helicobacter pylori* populations. *Science* 299: 1582-1585.
- Fenchel T, Finlay B. 2004. The ubiquity of small species: Patterns of local and global diversity. *BioScience* 54: 777-784.
- Fenchel T, Finlay BJ. 2005. Bacteria and island biogeography. *Science* 309: 1997.
- Fierer N, Morse JL, Berthrong ST, et al. 2007. Environmental controls on the landscape-scale biogeography of stream bacterial communities. *Ecology*. 88: 2162-2173.
- Finlay BJ. 2002. Global dispersal of free-living microbial eukaryote species. *Science*. 296: 1061-1063.
- Finlay BJ, Fenchel T. 2004. Cosmopolitan metapopulations of free-living microbial eukaryotes. *Protist* 155: 237-244.
- Fontaneto D, Barraclough TG, Chen K, Ricci C, Herniou EA. 2008. Molecular evidence for broad-scale distributions in bdelloid rotifers: everything is not everywhere but most things are very widespread. *Molecular Ecology* 17: 3136-3146.
- Foster JW, Spector MP. 1995. How *Salmonella* survive against the odds. *Annual Review of Microbiology*. 49: 145-174.
- Gandour M, Hessini K, Abdelly C. 2008. Understanding the population genetic structure of coastal species (*Cakile maritima*): seed dispersal and the role of sea currents in determining population structure. *Genetical Research* 90: 167-178.
- Gentile G, Fabiani A, Marquez C, Snell HL, Snell HM, Tapia W, Sbordoni V. 2009. An overlooked pink species of land iguana in the Galápagos. *Proceedings of the National Academy of Science* 106: 507-511.
- Geue L, Löschner U. 2002. *Salmonella enterica* in reptiles of German and Austrian origin. *Veterinary Microbiology* 84: 79-91.

- Grimont PAD, Weill F-X. 2007. Antigenic Formulae of the *Salmonella* Serovars (9<sup>th</sup> ed.). WHO Collaborating Centre for Reference and Research on *Salmonella*. Institute Pasteur: Paris, France.
- Groisman EA, Ochman H. 1997. How *Salmonella* became a pathogen. Trends in Microbiology 5: 343-349.
- Guan TY, Holley RA. 2003. Pathogen survival in swine manure environments and transmission of human enteric illness – a review. Journal of Environmental Quality 32: 383-392.
- Harris LJ, Farber JN., Beuchat LR, Parish ME, Suslow TV, Garret EH, Busta FF. 2003. Outbreaks associated with fresh produce: incidence, growth, and survival of pathogens in fresh and fresh-cut produce. Comprehensive Reviews in Food Science and Food Safety 2: 78-141.
- Hedlund BP, Staley JT. 2003. Microbial endemism and biogeography. In Microbial Diversity and Bioprospecting (Bull AT, ed.). American Society of Microbiology: Washington, DC.
- Hendry AP. 2002. Q(ST) $\neq$ F-ST? Trends in Ecology & Evolution 17: 502.
- Hidalgo-Vila J, Díaz-Paniagua C, de Frutos-Escobar C, Jiménez-Martínez C, Pérez-Santigos N. 2007. *Salmonella* in free living terrestrial and aquatic turtles. Veterinary Microbiology 119: 311-315.
- Hoorfar J, Baggesen DL. 1998. Importance of pre-enrichment media for isolation of *Salmonella* spp. from swine and poultry. FEMS Microbiology Letters 169: 125-130.
- Hurlbert SH. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54: 187-211.
- Ishii S, Sadowsky MJ. 2009. Application of the rep-PCR DNA fingerprinting technique to study microbial diversity, ecology and evolution. Environmental Microbiology 11: 733-740.
- Jackson MH. 2007. Galapagos: A Natural History (11<sup>th</sup> ed.). University of Calgary Press: Calgary, Alberta, Canada.
- Jenkins DG, Carey M, Czemieska J, Fletcher J, Hether Y, Jones A, Knight S, Knox J, Long T, Mannino M, McGuire M, Riffle A, Segelsky S, Shappell L, Sterner A, Strickler T, Tursi R. 2010. A meta-analysis of isolation by distance: relic or reference standard for landscape genetics? Ecography 33: 315-320.

- Juhas M, van den Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. 2008. Genomic islands: Tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiology Reviews* 33: 376-393.
- Krebs CJ. 1994. *Ecology: The Experimental Analysis of Distribution and Abundance*. HarperCollins: New York, NY.
- Kritcher JC. 2006. Hideous Looking Creatures. *Galapagos: A Natural History*. pp. 81-97. Princeton University Press: Princeton, NJ.
- Losos JB, Ricklefs RE. 2009. Adaptation and diversification on islands. *Nature* 457: 830-836.
- MacArthur RH and Wilson EO. 1967. *The Theory of Island Biogeography*. Princeton University Press: Princeton, NJ.
- Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR, Morin PJ, Naeem S, Ovreas, Reysenbach AL, Smith VH, Staley JT. 2006. Microbial biogeography: Putting microorganisms on the map. *Nature Reviews in Microbiology*. 4: 102-112.
- Mayr E. 1963. *Animal Species and Evolution*. Harvard University Press: Cambridge, MA.
- McKay JK, Latta RG. 2002. Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution* 17: 285-291.
- Mitchell MA, Shane SM. 2000. Preliminary findings of *Salmonella* spp. in captive green iguanas (*Iguana iguana*) and their environment. *Preventive Veterinary Medicine* 45: 297-304.
- Monzón-Moreno C, Ojeda Varga MM, Echeita A, Usera MA. 1995. Occurrence of *Salmonella* in cold-blooded animals in Gran Canaria, Canary Islands, Spain. *Antonie van Leeuwenhoek*. 68: 191-194.
- Morgan E. 2007. *Salmonella* pathogenicity islands. In *Salmonella Molecular Biology and Pathogenesis* (Rhen M, Maskell D, Mastroeni P, Threlfall J, eds.). pp 67-88. Cromwell Press: Trowbridge, Great Britain.
- Nelson DM, Cann IKO, Altemann E, and Mackie RI. 2010. Phylogenetic evidence for lateral gene transfer in the intestine of marine iguanas. *PLoS One* 5: e10785. doi:10.1371/journal.pone.0010785.

- Oda Y, Star B, Huisman LA, Gottachal JC, Forney LJ. 2003. Biogeography of the purple nonsulfur bacterium *Rhodospseudomonas palustris*. *Applied and Environmental Microbiology*. 69: 5186-5191.
- Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2010. *vegan*: Community Ecology Package. R package version 1.17-4. <http://cran.r-project.org/>, <http://vegan.r-forge.r-project.org/>.
- Olsen JE, Skov MN, Threlfall EJ, Brown DJ. 1994. Clonal lines of *Salmonella enterica* serotype Enteritidis documented by IS200-, ribo-, pulsed-field gel-electrophoresis and RFLP typing. *Journal of Medical Microbiology* 40: 15-22.
- Papke RT, Ramsing NB, Bateson MM, Ward DM. 2003. Geographical isolation in hot springs cyanobacteria. *Environmental Microbiology* 5: 650-659.
- Papke RT, Ward DM. 2004. The importance of physical isolation to microbial diversification. *FEMS Microbiology Ecology* 48: 293–303.
- Parent CE, Caccone A, Petren K. 2008. Colonisation and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical Transactions of the Royal Society B-Biological Sciences* 363: 3347-3361.
- Pederson K, Lassen-Neilsen A-M, Nordentoft S and Hammer AS. 2009. Serovars of *Salmonella* from captive reptiles. *Zoonoses and Public Health* 56: 238-242.
- Popoff MY, Le Minor L. 1997. Antigenic formulas of the *Salmonella* serovars (7th revision). WHO Collaborating Center for Reference and Research on *Salmonella* Institute Pasteur: Paris, France.
- Porcher E, Giraud T, Lavigne C. 2006. Genetic differentiation of neutral markers and quantitative traits in predominantly selfing metapopulations: confronting theory and experiments with *Arabidopsis thaliana*. *Genetical Research* 87: 1-12.
- R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Rademaker JLW, Louws FJ, De Bruijn FJ. 1998. Characterization of the diversity of ecologically important microbes by rep-PCR genomic fingerprinting. In *Molecular Microbial Ecology Manual*. pp 1-27. Kluwer Academic Publishers: the Netherlands.
- Rhodes MW, Kator H. 1988. Survival of *Esherichia coli* and *Salmonella* spp. in estuarine environments. *Applied and Environmental Microbiology* 54: 2902-2907.

- Ricklefs RE. 2004. A comprehensive framework for global patterns in biodiversity. *Ecological Letters* 7: 1-15.
- Rozen Y, Belkin S. 2001 Survival of enteric bacteria in seawater. *FEMS Microbiology Reviews* 25: 513-529.
- Scherer CA and Miller SI. 2001. Molecular pathogenesis of *Salmonellae*. *Principles of Bacterial Pathogenesis*. pp 265-333. Academic Press: San Diego, CA.
- Semenov AM, Kuprianov AA, van Bruggen AHC. 2010. Transfer of enteric pathogens to successive habitats as part of microbial cycles. *Microbial ecology* 60: 239-249.
- Shepherd SA, Hawkes MW. 2005. Algal food preferences and seasonal foraging strategy of the marine iguana, *Amblyrhynchus cristatus*, on Santa Cruz, Galapagos. *Bulletin of Marine Science* 77: 51-72.
- Steane DA, Conod N, Jones RC, Vaillancourt RE, Potts BM. 2006. A comparative analysis of population structure of a forest tree, *Eucalyptus globulus* (Myrtaceae), using microsatellite markers and quantitative traits. *Tree Genetics & Genomes* 2: 30-38.
- Steinfartz S, Glaberman S, Lanterbecq D, Russello MA, Rosa S, Hanley TC, Marquez C, Snell HL, Snell HM, Gentile G, Dell'Olmo G, Powell AM, Caccone A. 2009. Progressive colonization and restricted gene flow shape island-dependent population structure in Galápagos marine iguana (*Amblyrhynchus cristatus*). *BMC Evolutionary Biology* 9: 297.
- Takacs-Vesbach C, Mitchell K, Jackson-Weaver O, and Reysenbach A-L. 2008. Volcanic calderas delineate biogeographic provinces among Yellowstone thermophiles. *Environmental Microbiology* 10: 1681-1689.
- Tankouo-Sandjong B, Sessitch A, Liebana E, Kornschöber C, Allerberger F, Hächler H, Bodrossy L. 2007. MLST-v, multilocus sequence typing based on virulence genes for molecular typing of *Salmonella enterica* subsp. *enterica* serovars. *Journal of Microbiological Methods* 69: 23-36.
- Thomas CM, Nielsen KM. 2005 Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews in Microbiology* 3: 711–721.
- Trillmich KGK, Trillmich F. 1986. Foraging strategies of the marine iguana, *Amblyrhynchus cristatus*. *Behavioral Ecology and Sociobiology* 18: 259-266.



- Tuomisto H, Ruokolainen R, Yli-Halla M. 2003. Dispersal, environment, and floristic variation of western Amazonian forests. *Science* 299: 241–244.
- van der Gast CJ. 2008. Chapter 6: Islands shaping thought in microbial ecology. *Advances in Applied Microbiology* 64: 167-182.
- Versalovic J, Koeuth T, Lupski JR. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Research* 19: 6823-6831.
- Vos M, Velicer GJ. 2008. Isolation by distance in the spore-forming soil bacterium *Myxococcus xanthus*. *Current Biology* 18: 386-391.
- Weigel RM, Qiao B, Teferedegne B, Suh DK, Barber DA, Isaacson RE, White BA. 2004. Comparison of pulsed field gel electrophoresis and repetitive sequence polymerase chain reactions as genotyping methods for detection of genetic diversity and inferring transmission of *Salmonella*. *Veterinary Microbiology* 100: 205-217.
- Wheeler E, IKO Cann, RI Mackie. 2011. Genomic fingerprinting and serotyping of *Salmonella* from Galápagos iguanas demonstrates island differences in strain carriage. *Environmental Microbiology Reports* 3: 166-173.
- Whitaker RJ, Grogan DW, Taylor JW. 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* 301: 976-978.
- Whitfield J. 2005. Biogeography: Is everything everywhere? *Science* 310: 960-961.
- Wikelski M, Gall B, and Trillmich F. 1993. Ontogenetic changes in food intake and digestion rate of the herbivorous marine iguana (*Amblyrhynchus cristatus*, Bell). *Oecologia* 94: 373-379.
- Wozniak RAF, Waldor MK. 2010. Integrative and conjugative elements: Mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nature* 463: 552-563.
- Wright, S. 1943. Isolation by distance. *Genetics* 28: 114-138.
- Zhang N, Blackwell M. 2002. Population structure of dogwood anthracnose fungus. *Phytopathology* 92: 1276–1283.

## Tables and Figures

Table 4.1. Sampling site characteristics and *S. enterica* detection in Galápagos iguanas

Island	Area km <sup>2</sup> *	Site	Species	N	% <i>S. enterica</i> positive (N)
San Cristóbal	558	Punta Carola	Marine Iguana	12	75.0% (9)
San Cristóbal	558	La Lobería	Marine Iguana	12	75.0% (9)
Plaza Sur	0.13	(whole island)	Marine Iguana	12	67% (8)
			Land Iguana**	11	54.5% (6)
Santa Fe	24	El Miedo	Marine Iguana	2	50.0% (1)
			Land Iguana <sup>§</sup>	11	45.5% (5)
Fernandina	642	Cape Douglass	Marine Iguana	12	50.0% (6)
			Land Iguana**	6	16.7% (1)

\* Land area estimates from Jackson 2007

\*\* *Conolophus pallidus*; <sup>§</sup>*Conolophus subcristatus*

Table 4.2. Serovar distributions among sites for *S. enterica* isolated from land and marine iguanas

Sub-species	Serotype	Serogroup*	Total (47)	C (9)	L (9)	P (15)	S (6)	F (8)
I	Muenchen	O:8 (C2-C3)	2	1		1		
I	Manhattan	O:8 (C2-C3)	2			2		
I	Sandiego	O:4 (B)	6			6		
I	Poona	O:13 (G)	3			3		
IV	53:z4,z23-	O:53	1			1		
IV	44:z36-	O:44	1			1		
I	Pomona	O:28 (M)	4			1	3	
I	Berta	O:9 (D1)	4		3		1	
I	Treforest	O:51	1				1	
I	Rough O:L,V:1,7	R	2				1	1
I	Montevideo	O:7 (C1)	2	2				
I	Bredeney	O:4 (B)	1	1				
I	Newport	O:8 (C2-C3)	2	2				
II	II 47:b:1,5	O:47 (X)	3	3				
I	Mjordan	O:30 (N)	2		2			
I	Saintpaul	O:4 (B)	1		1			
I	Reading	O:4 (B)	2		2			
I	Rubislaw	O:11 (F)	1		1			
I	Rough O:y:1,7	N/A	1					1
I	57:b:-	O:57	2					2
I	Rough O:c:enz15	R	1					1
I	Wedding	O:28 (M)	1					1
IV	Rough O:g,z,51:-	R	1					1
Untyped**	Untyped	Untyped	1					1

\*Serogroup as in Grimont and Weill 2007. \*\* untyped isolate confirmed as *Salmonella* spp. by metabolic assay but was contaminated on submission to NVSL and was not typed further.

C=Punta Carola, San Cristóbal, L=La Loberia, San Cristóbal, P=Plaza Sur, S=Santa Fe, F=Fernandina

Table 4.3. Permutation MANOVA of rep-PCR fingerprints evaluating the role of island, site and location on *Salmonella* similarities

Permutation MANOVA			
Source	F	R <sup>2</sup>	p
Island	1.408	0.089	0.053*
Ecotype	1.126	0.024	0.368
Site	1.687	0.137	0.004**
Ecotype	1.163	0.024	0.306
Site	1.666	0.137	0.004**
Latitude	1.181	0.024	0.323
Longitude	1.053	0.022	0.371
Site	1.698	0.070	0.018**
PC1 (SE-NW)	2.176	0.045	0.010**
PC2 (SW-NE)	1.072	0.022	0.381

\* Significant at  $\alpha=0.10$ ; \*\* Significant at  $\alpha=0.05$

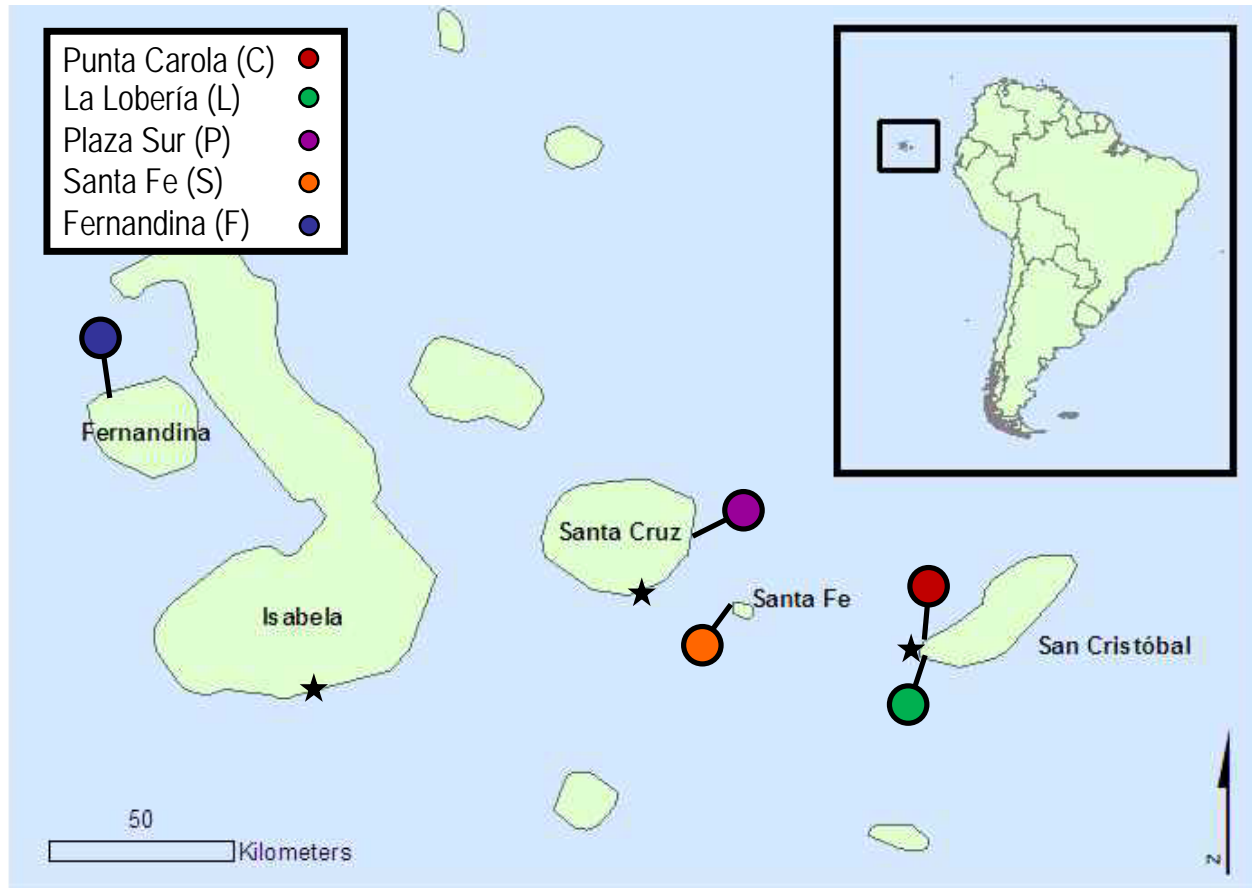


Figure 4.1. Sampling locations where fecal specimens were collected from marine iguanas (San Cristóbal sites Punta Carola, La Lobería, Santa Fe, Plaza Sur and Fernandina) land iguanas (Santa Fe, Plaza Sur and Fernandina). Black stars indicate the three major port towns of the Galápagos Islands, Puerto Baquerizo Moreno on Isla San Cristóbal, Puerta Ayora on Isla Santa Cruz, and Puerto Villamil on Isla Isabela

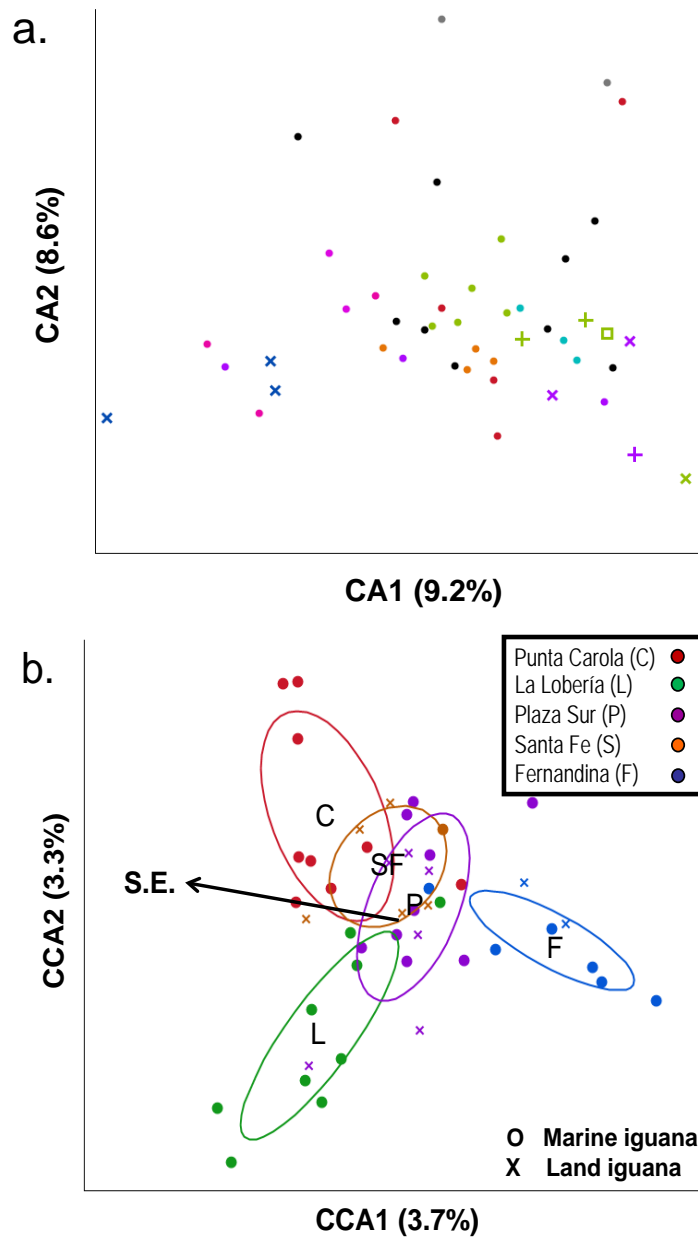


Figure 4.2. Unconstrained and constrained correspondence analysis of *Salmonella* genomic fingerprints. Panel a presents the unconstrained correspondence analysis (CA) with O antigen serogroups represented by color and different serovars within these groups with differing symbols. Panel b presents the same data constrained by sampling site (CCA; including 50% contour ellipses for sites with points and corresponding ellipses colored and labeled by site). In panel b, environmental fitting of a single principal component location axis (an axis rotation which captures latitudinal and longitudinal position of each sampling site along a northwestern to southeastern gradient) explains a significant portion of the variation in this ordination (PC1 environmental fitting  $R^2=0.678$ ,  $p<0.001$ ). The orthogonal principal component (PC2) was not a significant environmental fit to this ordination.

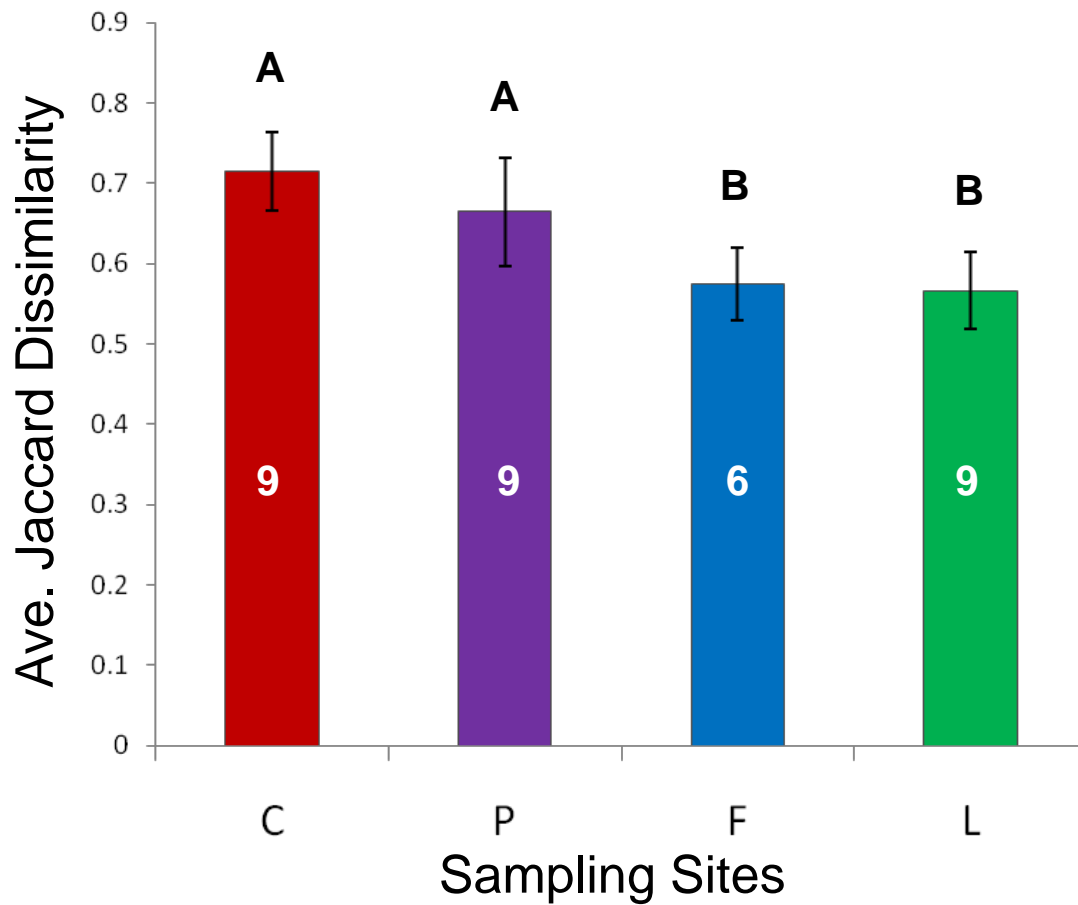


Figure 4.3. Site differences in the average ( $\pm 1$  SEM) within-site dissimilarities of *S. enterica* isolates from marine iguanas (ANOVA by site:  $R^2= 0.6094$ ,  $p<0.001$ ; N per site indicated in white text on each bar). Significant differences between locations are demonstrated by different letters above the error bars (Tukey's HSD test).

## **CHAPTER 5: CARRIAGE OF ANTIBIOTIC RESISTANT ENTERIC BACTERIA VARIES AMONG SITES IN GALÁPAGOS REPTILES**

### **Chapter Summary**

Increased overlap between human and wildlife populations has resulted in increased risk of novel disease emergence. Detecting contact points of increased risk for disease exchange requires identification of dependable measures of microbial exchange. In this study we evaluated antibiotic resistance as a potential marker for the intensity of human-wildlife microbial connectivity in the Galápagos Islands. We isolated *Escherichia coli* and *Salmonella enterica* from the feces of land iguanas (*Conolophus* sp.), marine iguanas (*Amblyrhynchus cristatus*), and giant tortoises (*Geochelone nigra*) and from sea water and tested these bacteria for resistance to ten antibiotics. Antibiotic resistance was found in reptile feces at two tourism sites (Isla Plaza Sur and La Galapaguerra on Isla San Cristóbal) and in sea water at a public-use beach near Puerto Baquerizo Moreno on Isla San Cristóbal, but no resistance was found at two protected beaches on more isolated islands (Isla Santa Fe and Isla Fernandina) and at a coastal tourism site (La Lobería on Isla San Cristóbal). A total of eighteen *E. coli* isolates from three locations, all sites relatively proximate to port town, showed resistance to combinations of ampicillin, doxycycline, tetracycline, and trimethoprim-sulfazoline. In contrast, only a few *S. enterica* isolates showed an intermediate resistance to doxycycline and tetracycline from these same sites, with no complete resistance detected in this bacterial species. These findings suggest that iguanas with increased human contacts potentially have higher exposure to bacteria of human origin but it is not clear how this potential exposure translates to on-going exchange of bacterial strains or genetic traits. Resistance patterns and bacterial exchange in this system warrant further investigation to more clearly understand how human associations influence disease risk in endemic Galápagos wildlife.

### **Introduction**

The emergence of antimicrobial resistant bacteria from medical and agricultural use of antibiotic compounds presents an ongoing global health concern (Bonomo and Rossolini 2008, Allen et al. 2010). The ecological processes underlying the dissemination of antimicrobial resistance are complex and are yet to be fully elucidated (Singer et al. 2006). Antibiotic resistant bacteria have been commonly detected in the environment and in wildlife populations, the latter of which could provide a biological mechanism for the spread of resistance from sites of origin across broad landscapes (reviewed in Allen et al. 2010). Proximity of wildlife populations to humans has been shown to correlate to the amount of antibiotic resistance detected in bacteria in wild hosts, with wildlife living closer to dense human settlements more likely to carry resistant bacteria (Souza et al. 1999, Gilliver et al. 1999, Oseterblad et al. 2001, Rolland



et al. 1985, Rwego et al. 2008, Goldberg et al. 2007, Cole et al. 2005, Dolejska et al. 2007, Skurnik et al. 2007). These findings suggest that antibiotic resistance might be a strong indicator of human influences on bacterial exposures in wild populations.

Increased overlap in habitat use among humans, domestic livestock and wildlife species, has been linked to increases in disease outbreaks and novel disease emergence (Daszak et al. 2000, Epstein and Price 2009). Recently, molecular epidemiology of model bacterial systems has been used to document the effects of such overlap on pathogen transmission risk using antibiotic resistance as a marker for contact with humans. For example, recent studies have employed documentation of strain similarity and antibiotic resistance patterns in *Escherichia coli* for investigating patterns of gastrointestinal bacterial exchange among humans and non-human primates in eastern Uganda's fragmented forests as a proxy for disease risk (Goldberg et al. 2007, Rwego et al. 2008).

Documentation of such changes in "microbial connectivity" between humans and wildlife, even in the absence of an immediate disease concern, may help with identifying interactions where the risk of new or emerging disease is increased and with recognizing contact points where increased surveillance may be warranted. In human-impacted systems, such understanding could be critical for the conservation and protection of imperiled wildlife species cohabitating with human residences. The Galápagos Islands, Ecuador, is one such system that is potentially highly susceptible to human impacts (Brockie et al. 1988, Watson et al. 2010), where increased understanding of human effects on microbial ecology might benefit conservation efforts and both human and animal health.

Aggressive management of the local human population and of introduced species has protected the Galápagos Islands, to some extent, from experiencing the vast extinctions of endemic species observed in other human-dominated island systems, such as Hawaii (de Groot 1983). However, despite historic successes in protecting these fragile ecosystems, resident populations continue to increase, despite governmental efforts to limit this expansion and tourism to the Galápagos Islands has become an increasing management concern as ecotourism has risen in popularity (Epler 2007). The Galápagos Islands currently support an estimated human population of about 25-30 thousand residents and temporary workers, living mainly on two of the larger islands,

San Cristóbal and Santa Cruz. Additionally, approximately 150,000 tourists from all over the world visit the Galápagos annually to see the unique endemic wildlife (Epler 2007, Watkins and Cruz 2007). These delicate island ecosystems are especially vulnerable to human impacts from garbage disposal and pollution, changes in land use, and introduction of non-native species (such as pigs, cattle, goats, dogs, cats and rats, Brockie et al. 1988).

Endemic reptile species, including the giant tortoise and two types of large herbivorous iguanas are a main attraction for tourists. These animals are often relatively unafraid of humans and live in close proximity to human residences and waste disposal sites on the inhabited islands. On islands zoned for tourism, many of these animals are tame and bask among the tourists as they explore the islands, offering opportunities for very close contact and potentially disease exchange between humans and wildlife (personal observation, E. Wheeler). While the potential for human influences on the behavior and physiology of endemic Galápagos wildlife has been studied (Wikelski et al. 2002, Berger et al. 2007, Tanner and Perry 2007), disease risks and bacterial exchange arising from human-wildlife interactions have only limited documentation to date (Thaller et al. 2010a, Thaller et al. 2010b, Chapter 4) . Specifically, recent work in this system has demonstrated significant site differences in cultivable bacteria carried by endemic wild iguanas that may be associated with variable human impacts on water quality and bacterial exposures among the different islands in this chain (Thaller et al. 2010a, Chapter 4). Further, low levels of antimicrobial resistant bacteria have been detected in Galápagos endemic land iguanas (*Conolophus pallidus*) from a research camp site on the relatively isolated island of Santa Fe (Thaller et al. 2010b). If human-wildlife interactions underlie some of these observed patterns, then we would expect that bacteria from more human-impacted sites should also have an increased carriage of phenotypic or genetic traits which might indicate recent origin in human populations, such as increased antimicrobial resistance or resistance genes. Thus, we expected to find increased severity of antibiotic resistance in enteric bacteria isolated from animals or sea water at sites with higher human contact compared to more protected areas (i.e. a dose-response relationship between the proximity to humans and amount of resistance detected).

In this study, we collected feces from free living marine iguanas (*Amblyrhynchus cristatus*), land iguanas (*Conolophus* spp.), and giant tortoises (*Geochelone nigra*), along with a sea water environmental control from six sites on four islands (Figure 5.1). These samples were cultivated for two species of enteric bacteria, *Escherichia coli* and *Salmonella enterica*, which were then tested for resistance to ten antibiotics, chosen to represent common antibiotics potentially carried by tourists or available to local residents and representing a variety of antimicrobial classes. The six sites selected for sampling provided a gradient of potential contact between humans and wild reptiles, based on proximity to port towns and site usage. Within this system, we explored the hypothesis that enteric bacteria from wild reptiles with increased human contacts would carry more antibiotic resistant bacteria than those living on more isolated sites.

## **Methods**

### *Study design and sampling sites*

Sampling sites differed in the potential for contact between wildlife and humans, ranging from close proximity to a port town sewage outlet to tourist visitation sites to protected research sites (site C; sites L, G, and P; sites S and F, respectively). Punta Carola on the island of San Cristóbal, was suspected to be the most highly contaminated site as it is both the location a popular public beach and is quite near the site of outflow of Puerto Baquerizo Moreno's main sewage pipe (Figure D.1). La Lobería and La Galapaguera, both on the island of San Cristóbal, and the small island of Plaza Sur are primarily tourist visitation sites. El Miedo on the island of Santa Fe and Cape Douglass on the island of Fernandina are protected research sites and are likely the least impacted by human use of these sampling locations.

Free-living reptiles (marine iguanas, land iguanas, and Galápagos tortoises) were observed from a distance for defecation and then fecal specimens were collected from the ground using a sterile wooden applicator. We intentionally selected fecal samples from different areas of each site and collected samples over a short period of time (24-48 hours) to avoid repeated sampling of the same individual. The slow gastrointestinal transit rates of herbivorous reptiles (e.g. estimated to be 5-10 days for marine iguanas, Wikelski et al. 1993), reduces the likelihood of observing a single individual defecating twice within a short sampling period. Fecal specimens were obtained from both iguana

types on Plaza Sur, Fernandina, and Santa Fe, where both species are present and only marine iguanas from the two sampling sites on the island of San Cristóbal, where land iguanas are not present. Galápagos tortoises (*Geochelone nigra*) were sampled at a single site, La Galapaguerra, the inland tourist and tortoise breeding facility located on San Cristóbal near El Junco.

Samples were placed in sterilized tubes, stored at 4°C to -20°C or on ice, as available during travel, and transported to Urbana, Illinois for laboratory analyses. As an environmental control, two 50ml sea water samples were obtained from Fernandina and both sites on San Cristóbal.

#### *Isolation of enteric bacteria*

*S. enterica* was isolated from frozen fecal samples using previously published protocols (Hoorfar and Baggesen 1998, Mitchell and Shane 2000, Corriente et al 2004, Wheeler et al. 2011). Briefly, 0.5 g of feces was pre-enriched in buffered peptone water (BPW) at 37°C for 24 hours followed by enrichment in a 1:9 ml dilution of turbid BPW:Rappaport-Vassiliadis broth (RVB) at 37°C for another 24 hours. Turbid RVB was then streak plated on selective-differential media (xylose lysine deoxycholate, XLD). XLD plates demonstrating no *Salmonella*-characteristic growth at 24 hours were re-incubated for an additional 24 hours before being considered negative for growth. Samples negative for *Salmonella*-like isolates were retested from the original fecal samples before being considered negative for the bacterium.

Presumptive *S. enterica* isolates (clear to pink with or without a dark black center) were tested using standard biochemical metabolic assays (lysine iron agar, LIA and triple sugar agar, TSA) to confirm membership in the genus. Isolates testing positive on both of these assays were presumptively confirmed to the *Salmonella* genus, as previous work in this system has demonstrated a 100% concordance between confirmation of *Salmonella* using these two assays following isolation on XLD selective media when compared to identification using API 20E (bioMerieux Inc., France) metabolic identification strips (E. Wheeler, unpublished data). Up to three confirmed *Salmonella* isolates per individual fecal sample were retained frozen in 40% glycerol for further analysis. One isolate per fecal sample was submitted to the National Veterinary Services Laboratory in Ames, Iowa for O and H antigen group serotyping.

Turbid buffered peptone water was also inoculated on eosin methylene blue (EMB) agar plates for isolation of *Escherichia coli*. Up to three purple colonies (lactose positive) with a metallic green sheen were selected from these plates as presumptive *E. coli*. Lactose positive isolates which tested positive for the production of indole and negative for metabolism of citrate were confirmed as *E. coli* (MacFaddin 1980) and were stored frozen in 40% glycerol for analysis.

Water samples were centrifuged at 14000 rpm for 10 minutes to collect bacterial cells and debris. The pellet was retained and reconstituted in 2 ml of sterile water. A 0.1 ml portion of each condensed water sample was inoculated into buffered peptone water and then cultivated for both *E. coli* and *S. enterica* as previously described. All six sea water samples tested negative for *Salmonella* but tested positive for *E. coli* at two sites, Isla Fernandina (one of two sea water samples) and Punta Carola on Isla San Cristóbal (both sea water samples).

#### *Antibiotic resistance testing by disk diffusion*

We tested confirmed *E. coli* and *S. enterica* isolates, one isolate of each bacterial species per sample when available, for resistance to ten antibiotics (ampicillin 10µg, amoxicillin with clavulanic acid 30µg, Cefazolin 30µg, Cefuroxime 30µg, Chloramphenicol 30µg, Ciprofloxacin 5µg, Doxycycline 30µg, Enrofloxacin 5µg, Gentamicine 10µg, Tetracycline 30µg, and Trimethoprim/Sulfazolin 25µg, all antibiotic resistance discs obtained from Remel, Lenexa, Kansas). A number of these antibiotics were selected because they are commonly prescribed to travelers for prophylactic or symptomatic use (e.g. tetracyclines, fluoroquinolones, and beta-lactams) or because they are inexpensive representatives of a variety of antibiotic classes. We performed antimicrobial susceptibility testing according to standard protocols for the disk diffusion method following National Committee for Clinical Laboratory Standards guidelines (National Committee for Clinical Laboratory Standards 2003). The disk diffusion method qualitatively evaluates antibiotic resistance by establishing a drug concentration gradient (via diffusion) around an antibiotic-permeated disc of filter paper. The wider the clear (no bacterial growth) zone around an antibiotic disc, the more susceptible the strain is to the antibiotic (i.e. growth is inhibited at lower concentrations of the drug if the clearing is wider). Published standards for *E. coli* and related Enterobacteraceae

were used to assign resistant, intermediate or susceptible phenotypes to each strain for each isolate tested. While minimum inhibitory concentration by broth dilution is considered the gold standard in antibiotic resistance testing, this method is potentially more time consuming or expensive compared to disc diffusion. As the goal of this testing is to document the range of phenotypes present in these isolates, rather than to establish appropriate clinical therapies for medical treatment of infection, utilization of a qualitative but more time and cost efficient method was deemed acceptable.

### *Identification of tetracycline resistance genes*

Due to the high frequency of tetracycline resistance in these isolates we further evaluated the genetic basis of this resistance phenotype by PCR identification of the resistance genes present in tetracycline resistant isolates using established protocols (Aminov et al. 2004). Briefly, DNA was extracted from all isolates demonstrating resistance (i.e. those included in Table 5.1) and from a site matched subset of non-resistant isolates using the Instagene chelex resin as previously described (Chapter 4, Biorad, Hercules, CA) Then PCR amplifications were performed using established protocols using primers specific for tetracycline efflux genes A, B, C, D, and E with positive and negative controls (Aminov et al. 2004). PCR products were visualized by agarose gel electrophoresis, with the presence of a strong band of the appropriate size indicating that an isolate was positive for the gene in question.

### *Statistical analysis*

The distribution of antibiotic resistance was modeled using generalized linear models with binomial errors. For this analysis, isolates were classified as either sensitive or resistant, if they were resistant to at least one antibiotic. Results were similar when modeling the number of different antibiotics to which a given isolate showed resistance (data not shown). Akaike's information criterion corrected for small sample sizes (AICc) was used to select the best fitting model from a set of candidate models (Burnham and Anderson 1998). Information criterion approaches allow for evaluating the fit of multiple models to select the model which best describes the data, rather than more traditional hypothesis testing of a single candidate model against the standard null hypothesis. Candidate models for this analysis included various combinations of four

explanatory variables: host species (marine iguana, land iguana, tortoise, or environmental sample), bacterial species (*S. enterica* or *E. coli*), the distance from the sampling site to the nearest permanent human settlement, and the geographic position of the sampling site. Statistical significance of the effects retained in the best fitting model was determined by likelihood ratio tests, and the relative importance of the different effects was determined by the weight of evidence (Burnham and Anderson 1998). Basic statistical and graphing procedures were performed in either JMP 8 (SAS Institute, Cary, NC) or R statistical language (R Development Core Team 2008).

## Results

### *Bacterial carriage patterns*

The prevalence of both *E. coli* and *S. enterica* differed among host types, with high *S. enterica* carriage in marine iguanas (73% of 48 positive), moderate in land iguanas (46% of 28 positive) and none detected in the giant tortoises. The carriage of *S. enterica* differed significantly between iguanas and tortoises (Fisher's exact with 1000 permutations:  $p < 0.001$ , OR=3.45) and between the two iguana types (Fisher's exact with 1000 permutations:  $p = 0.028$ , OR=3.05). In contrast, *E. coli* carriage was low in marine iguanas (25% of 48 positive) and relatively high in land iguanas and giant tortoises (64% of 28 positive and 100% of 10 positive, respectively). Iguanas and tortoises and marine and land iguanas both differed significantly in *E. coli* carriage (Fisher's exact with 1000 permutations:  $p < 0.001$ , OR < 0.001 and  $p = 0.001$ , OR = 0.19, respectively). All water samples tested negative for *S. enterica* but did test positive for *E. coli* at two sites, Fernandina and San Cristóbal-Punta Carola.

A total of 46 *S. enterica* and 59 *E. coli* (1 per individual fecal sample for each bacterial species when available and up to 6 isolates per water sample) were used for antibiotic resistance testing (By sampling site - *E. coli* from iguanas: C=4, L=2, P=17, S=6, F=2; *S. enterica* from iguanas C=9, L=9, P=15, S=6, F=7; *E. coli* from giant tortoises G=10; and *E. coli* from sea water: C=12, F=6). *E. coli* was not isolated from sea water samples obtained from La Lobería on Isla San Cristóbal. *S. enterica* was not isolated from the giant tortoise fecal samples or from sea water samples.

### *Antibiotic resistance patterns*

A total of eight unique resistance profiles were detected at three locations on two islands (San Cristóbal-Punta Carola, San Cristóbal-La Galapaguerra, and Plaza Sur, Table 5.1, Figure 5.2). Only *E. coli* isolates (18/59 total isolates) demonstrated resistance to these antibiotics, with only a few *S. enterica* isolates displaying an intermediate phenotype (i.e. a zone of inhibition less than the clinical susceptibility cut off established by the NCCLS (2003), but greater than the established diameter for clinical resistance). Three *S. enterica* isolates from Plaza Sur (3/15) and two *S. enterica* isolates from Punta Carola (2/9) demonstrated this intermediate phenotype (Table 5.1). The majority of resistant isolates were resistant to more than one of the antibiotics tested and the majority of these resistance profiles included resistance to tetracyclines. Due to the high frequency of tetracycline resistance in these isolates we further evaluated the genetic basis of this resistance phenotype by PCR identification of the resistance genes present in tetracycline resistant isolates using established protocols (Aminov et al. 2004). Tetracycline resistant *E. coli* isolates demonstrated site-dependent gene carriage, with isolates from Plaza Sur carrying a tetracycline efflux gene A (TetA), isolates from Punta Carola carrying both TetA and tetracycline efflux gene B (TetB) and isolates from La Galapaguera carrying TetB (Table 5.1). *S. enterica* isolates, in contrast, only demonstrated an intermediate resistance phenotype to doxycycline or to both tetracycline and doxycycline and did not carry any of the efflux resistance genes tested (TetA, TetB, TetC, TetD, or TetE).

To explore the contribution of explanatory factors for the observed distribution of resistant isolates we used a model selection approach. Initial binomial models (for the presence of resistance to one or more antimicrobial) were selected to explore the combined contributions of sampling site, site proximity to the nearest large town, host species and bacterial species to detection of resistance (Table 5.2). Comparison of these models indicated that a model including proximity to the nearest of the three major port towns, host species (land or marine iguana or sea water) and bacterial species provided the strongest explanation of resistance phenotype distribution (i.e. had the lowest AIC<sub>c</sub> score, Table 5.2). Significance values and total weights of evidence were then calculated for the factors included in this best model (Table 3.3). All three factors were significant explanatory variables for the presence of antimicrobial resistance, with the strongest



weight of evidence for bacterial species, followed by the proximity of the sampling site to port towns.

## **Discussion**

### *Bacterial carriage patterns vary by site and species*

We found strong host species and site differences in the carriage of two species of gram negative Proteobacteria, *E. coli* and *S. enterica*. This finding is consistent with previous work in this system that showed similar variability in bacterial carriage rates (*Vibrio alginolyticus*, *Salmonella enterica* and *Escherichia coli* in Thaller et al. 2010a, *Salmonella enterica* in Wheeler et al. 2011, Chapter 4) among reptile species and among sites in this system. Species-level differences in carriage of bacterial species may reflect underlying differences in diet, host-bacterial interactions, or exposure to different strains. It is not clear what drives site differences, but previous studies have suggested that human influences on the marine system, in terms of water quality, might drive observed site differences in marine iguana intestinal biota (Thaller et al. 2010a). It is also possible that rarer members of the gut biota may vary stochastically between locations within a species based on local site dynamics and natural exposure sources (e.g. overlap with other endemic wildlife or with humans, Chapter 2 and Chapter 3).

### *Antimicrobial resistance is associated with human proximity*

Simply detecting antibiotic resistance in a given animal population is neither sufficient to identify the likely source nor to draw conclusions concerning directionality of transmission among populations. This is especially true of wildlife populations, which may be highly mobile and have many local potential exposure sources. However, evaluation of wildlife populations across a spatial gradient, in which local populations likely experience differences in absolute exposure to human or livestock sources of resistant bacteria, can assist with demonstrating a causal “dose-response” effect of increasing severity of resistance for populations in closer proximity to the hypothesized source of exposure. Previous studies in a number of wild species support the idea that wild populations proximate to human or livestock sources of resistant bacteria often harbor the elevated resistance levels relative to populations more distant from this potential contamination. Both wild non-human primates and wild birds living in closer

proximity to human sources of resistant bacteria harbor more resistant isolates than respective populations more remote from human exposures (non-human primates: Rolland et al. 1985, Goldberg et al. 2007, Rwego et al. 2008; wild birds: Cole et al. 2005, Dolesjka et al. 2007).

The density and antibiotic use patterns of human populations with which wild species are in contact can also contribute to differences in the degree of antibiotic resistance exchange and to the severity of resistance carriage in wild populations (Allen et al. 2010). Thus, both proximity to and degree of exposure may produce a gradient effect in resistance levels, as has been seen at broad global scales, both within the same study and across similar studies in different areas of the globe. For example, wildlife species from remote areas (Antarctica and Gabon), a low human density area (Mountainous region of Pyrenees, France) or a high human density area (near Paris, France) showed a gradient of resistant isolates from none to nearly 20% presence across these sampling locations. Likewise, a comparative study of resistance in *E. coli* from wild animals in Mexico and Australia found higher frequency of antibiotic resistance in isolates from Mexico. This difference was attributed to differences in human population densities and possibly due to increased selection pressures from wide-spread use of antibiotics in Mexico (Souza et al. 1999).

More local-scale comparisons also support these patterns. For example, ninety-percent of gut bacteria of wild mice and voles living in relatively populous rural England carried resistance to  $\beta$ -lactam antibiotics, while wild mammals in the less densely populated area of Finland have almost no resistance to antimicrobials (Gilliver et al. 1999, Osterblad et al. 2001). Finally, while rural rodents in Germany in areas of high density livestock production had higher levels of resistance than rodents remote from agricultural production, but urban rodents had higher resistance levels still than their country cousins, regardless of proximity to livestock (Guenther et al. 2010a, Guenther et al. 2010b).

Similar to these established findings, in this study we found increased resistance in enteric bacterial isolates from sites located closer to the major port towns. While we did not detect antibiotic resistance at the more isolated sites at Fernandina and Santa Fe, it is important to note that sample sizes in this study limit conclusions on the presence or absence of resistance at these sites. The sampling design of this study can

only speak to differences in severity of resistance, which were significant, but does not allow for reliable estimation of the likelihood of antimicrobial resistance presence. It is likely that resistance is present at these more isolated sites, as antibiotic resistance at low levels has been detected at very isolated areas around the globe (Sjolund et al. 2008). It is important to note that for highly mobile species, these patterns may be weaker due to seasonal migration and transportation of resistant bacteria to more remote areas (Sjolund et al. 2008); the location where exposure occurs for mobile wildlife may, at times, be quite distant from the site of capture and sampling, confounding expected patterns.

Finally, it is important to note that for the sample sizes presented in this study, we cannot comment on whether or not remote sites have an absolute presence or absence of resistant bacteria. A recent study of resistance carriage in gram negative enteric bacteria of land iguanas on Santa Fe suggests that the frequency of resistant isolates at these more isolated sites is very low (approximately 1% of enteric bacteria tested, Thaller et al. 2010a), and therefore is below the detection limits of this study's sampling effort. However, these low levels of resistance at such an isolated site are consistent with both the data presented herein and the overall expectations of very low carriage in wildlife populations more remote from likely antibiotic resistance exposure sources in this system (e.g. port towns and popular tourism venues).

#### *Antimicrobial resistance levels differ for different bacteria*

While proximity to port towns was a significant factor for the presence of antimicrobial resistance, the factor with the strongest weight of evidence for predicting antibiotic resistance traits of bacterial isolates was bacterial species, with *E. coli* but not *S. enterica* demonstrating resistance traits. This pattern suggests that the biology of these two organisms in the Galápagos may be quite different, despite the close relationship of these organisms at the genetic level (Sharp 1991). The potential for differences between *E. coli* and *S. enterica* ecology outside the host environment has been noted previously, specifically drawing into question whether *E. coli* is a good model for source tracking of related pathogens like *S. enterica*, despite their genetic similarities (Winfield and Groisman 2003). Genetic and phenotypic patterns in our Galápagos

isolates at least partially supports the idea that these two species undergo different strain turnover processes in the Galápagos system (see Appendix D).

The lack of antimicrobial resistance in *S. enterica* isolates from sites in which *E. coli* carries multiple resistance profiles in either sea water or iguana fecal specimens suggests that the rates and mechanisms of turnover of these two bacterial species varies in this system. A number of circumstantial pieces of evidence would agree with this supposition. Firstly, prolific *E. coli* growth in the relatively small sea water samples obtained near the sewage outlet of Puerto Basquerizo Moreno suggests that the input of novel *E. coli* strains from the flux of human visitors to these port towns may be locally quite high. Secondly, the detection of *E. coli* in sea water at the more isolated island of Fernandina demonstrates that fecal bacteria from a variety of animals can be found at high numbers in sea water, providing the potential for movement between sites with currents or in the intestinal tracts of these mobile wildlife or with human visitors. Previous work has shown that antibiotic resistant *E. coli* isolates from phylotype B, presumably of human origin, can be found in wildlife, even on a relatively remote island (Thaller et al. 2010b).

While diverse strains of *E. coli* may be very mobile between hosts in the Galápagos Islands, *S. enterica* appears to be more dispersal limited, with strong genetic population structuring by site, even for sites that are relatively proximate (Wheeler et al. 2011, Chapter 4). Further, strain turnover within a sampling site appears to be quite slow, with a large portion of serotypes common between samples taken in 2005 and 2009 (Table D.1). *S. enterica* was not detected in sea water samples in this system and is likely a less common member of human enteric communities than *E. coli*.

Due to the variable detection of *E. coli* in different species and locations, *E. coli* isolates in this study are too unevenly distributed to perform a complete genetic analysis among sites to more formally test the idea of increased movement of *E. coli* versus *S. enterica* in this system. However, we were able to partially explore this question by comparing the average genetic diversity within versus between sites for both species (Figure D.2). Overall, *S. enterica* isolates were more similar to each other both between and within sites than *E. coli*, suggesting that *E. coli* isolates are overall more diverse in this system than *S. enterica*. However, these *E. coli* populations may also be quite structured by location as the within site and between site diversity varied significantly

for *E. coli* and not for *S. enterica*. Thus, further study of strain turnover is warranted to fully understand the complexities of enteric bacterial biogeography and ecology in this system.

## **Conclusions**

Tourism in the Galápagos is a growing industry, bringing visitors from around the globe at high rates, and these tourists bring with them both economic benefit and the potential for damage to the conservation of this delicate ecosystem (Taylor et al. 2009). In addition, the local human population continues to expand in the Galápagos, despite efforts to limit immigration from the main-land, representing an additional potential source of novel exposures and stressors for wildlife (Epler 2007, Watkins and Cruz 2007). A wide variety of methods have been applied to quantifying human-imposed stress on wild species (Tarlow and Blumstein 2007) and these measures can indicate a great deal in terms of the physiological effects of overlapping habitat use. One potentially hidden impact of human populations in this system, an effect that is yet only in nascent stages of exploration, is the health implications of human-associated microbiological flux (Thaller et al. 2010a). Given the increasing emergence of new diseases from changes in contacts among wildlife, livestock and human populations (Daszak et al. 2000), identifying simple measures that can indicate potential disease risks may prove useful for fully understanding human-wildlife interactions. However, the complexities reveal in this study suggest that identifying candidate model organisms for studying microbial exchange among humans and wild species may require a deeper understanding of the gastrointestinal communities of these unique animals. However, the results of this study do clearly indicate that even in the isolated Galápagos Islands, human influence can potentially alter bacterial traits and communities in wildlife with which they come in contact. Such studies have great potential to contribute to theoretical understanding of how and when disease agents move among populations and may reveal opportunities for early interventions to prevent or reduce cross-species disease transmission events.

## **Chapter Acknowledgements**

We would like to express appreciation to the Galápagos National Park and Washington Tapia, and the Charles Darwin Research Station and Sonja Cisneros and Paulina Couenberg for facilitating sample collection. Many thanks to Augusto G. Haz Beltran, Lenin Cruz Beldon and the crew of the Pirata for help with sample collection. We also wish to thank the Department of Animal Sciences and Nancy Henry for administrative support. Thanks to Richard Lankau for statistical support and constructive comments on early drafts of this chapter. Thanks to Isaac Cann and Bryan White for access to laboratory resources.

Funding for this work was provided by the Conservation Medicine Center of Chicago Student Research Grant (E.W.) and by the United States Environmental Protection Agency Science to Achieve Results Fellowship program (E.W.). This chapter was developed with support from a STAR Research Assistance Agreement No. 91684301-1 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this chapter.

## **References**

- Allen HK, Donato J, Wang HH, Cloud-Hansen KA, Davies J Handelsman J. 2010. Call of the wild: Antibiotic resistance genes in natural environments. *Nature Reviews Microbiol.* 8: 251-259.
- Aminov RI, Chee-Sandford JC, Garrigues N, Mehboob A, and Mackie RI. 2004. Detection of tetracycline resistance genes by PCR methods. *Methods in Molecular Biology*, vol. 268 *Public Health Microbiology: Methods and Protocols* (Spencer JFT, Tagout AL, eds.). pp 3-13. Humana Press Inc: Totowa, NJ.
- Berger S, Wikelski M, Romero LM, Kalko EKV, Rödl T. 2007. Behavioral and physiological adjustments to new predators in an endemic island species, the Galápagos marine iguana. *Hormones and Behavior* 52: 653-663.
- Bonomo R, Rossolini GM. 2008. Importance of antibiotic resistance and antibiotic and resistance mechanisms. *Expert Review of Anti-Infective Therapy* 6: 549–550.

- Brockie RE, Loope LL, Usher MB, Hamann O. 1988. Biological invasions of island nature reserves. *Biological Conservation* 44: 9-36.
- Burnham KP, Anderson DR. 1998. *Model Selection and Multimodal Inference: A practical Information-Theoretic Approach*. Springer-Verlag: New York, NY.
- Cole D, Drum DJV, Stallknecht DE, White DG, Lee MD, Ayers S, Sobsey M, Maurer JJ. 2005. Free-living Canada geese and antimicrobial resistance. *Emerging Infectious Disease* 11: 935-938.
- Daszak PA, Cunningham AA, and Hyatt AD. 2000. Emerging Infectious Diseases of Wildlife – Threats to Biodiversity and Human Health. *Science* 287: 443-449.
- de Groot RS. 1983. Tourism and conservation in the Galápagos Islands. *Biological Conservation* 26: 291-300.
- Dolejska M, Cizek A, Literak I. 2007. High prevalence of antimicrobial-resistant genes and integrons in *Escherichia coli* isolates from black-headed gulls in the Czech Republic. *Journal of Applied Microbiology* 103: 11-19.
- Epler B. 2007. *Tourism, the economy, population growth and conservation in Galapagos*. The Charles Darwin Foundation, Puerto Ayora, Santa Cruz Island: Galapagos Island, Ecuador. Last accessed online November 11, 2010: <http://www.galapagos.org/>.
- Epstein JH, Price JT. 2009. The significant but understudied impact of pathogen transmission from humans to animals. *Mount Sinai Journal of Medicine* 76: 448-455.
- Gilliver MA, Bennett M, Begon M, Hazel SM, Hart CA. 1999. Antibiotic resistance found in wild rodents. *Nature* 401: 233-234.
- Goldberg TL, Gillespie TR, Rwego IB, Wheeler E, Estoff EE, Chapman CA. 2007. Transmission of *Escherichia coli* between chimpanzees and humans engaged in research and tourism in western Uganda. *Biological Conservation*. 135: 511-517.
- Guenther S, Grobbel M, Beutilich J, Guerra B, Ulrich RG, Wieler LH, Ewers C. 2010a. Detection of pandemic B2-O25-ST131 *Escherichia coli* harbouring the CTX-M-9 extended-spectrum beta-lactamase type in a feral urban brown rat (*Rattus norvegicus*). *Journal of Antimicrobial Chemotherapy* 65: 582-584.
- Guenther S, Grobbel M, Heidemanns K, Schlegel M, Ulrich RG, Ewers C, Wieler LH. 2010b. First insights into antimicrobial resistance among faecal *Escherichia coli*

- isolates from small wild mammals in rural areas. *Science of the Total Environment*. 408: 3519-3522.
- MacFaddin JF. 1980. *Enterobacteraceae* and other intestinal bacteria. In *Biochemical Tests for Identification of Medical Bacteria* (MacFaddin JF, ed.). pp 439-464. Williams & Wilkins: Baltimore, MD.
- National Committee for Clinical Laboratory Standards. 2003. M2-A8 Performance standards for antimicrobial disk susceptibility tests (M2-A8) AND Disk Diffusion Supplemental Tables (M100-S13[M2]). NCCLS Approved Standards (8th ed.).
- Osterblad M, Norrdahl K, Korpimaki E, Huovinen P. 2001. Antibiotic resistance. How wild are wild mammals? *Nature* 409: 37-38.
- R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Rolland RM, Hausfater G, Marshall B, Levy SB. 1985. Antibiotic-resistant bacteria in wild primates: increased prevalence in baboons feeding on human refuse. *Applied and Environmental Microbiology* 49: 791-794.
- Rwego IB, Isabirye-Basuta G, Gillespie TR, Goldberg TL. 2008. Gastrointestinal bacteria transmission among humans, mountain gorillas and livestock in Bwindi Impenetrable National Park, Uganda. *Conservation Biology* 22: 1600-1607.
- Sharp PM. 1991. Determinants of DNA-sequence divergence between *Escherichia coli* and *Salmonella typhimurium* – Codon usage, map position, and concerted evolution. *Journal of Molecular Evolution* 33: 23-33.
- Singer RS, Ward MP, Maldonado G. 2006. Can landscape ecology untangle the complexity of antibiotic resistance? *Nature Reviews in Microbiology* 4: 943–952.
- Sjolund M, Bonnedahl J, Hernandez J, Bengtsson S, Cederbrant G, Pinhassi J, Kahlmeter G, Olsen B. 2008. Dissemination of multidrug-resistant bacteria into the Arctic. *Emerging Infectious Disease* 14: 70–72.
- Skurnik D, Ruimy R, Andremont A, Amorin C, Rouquet P, Picard B, Denamur E. 2006. Effect of human vicinity on antimicrobial resistance and integrons in animal faecal *Escherichia coli*. *Journal of Antimicrobial Chemotherapy* 57: 1215-1219.



- Souza V, Rocha M, Valera A, Eguiarte LE. 1999. Genetic structure of natural populations of *Escherichia coli* in wild hosts on different continents. *Applied and Environmental Microbiology* 65: 3373-3385.
- Tanner D and Perry J. 2007. Road effects on abundance and fitness of Galápagos lava lizards (*Microlophus albemarlensis*). *Journal of Environmental Management* 85: 270-278.
- Tarlow EM, Blumstein DT. 2007. Evaluating methods to quantify anthropogenic stressors on wild animals. *Applied Animal Behavior Science* 102: 429-451.
- Taylor JE, Hardner J, Stewart M. 2009. Ecotourism and economic growth in the Galápagos: an island economy-wide analysis. *Environment and Development Economics* 14: 139-162.
- Thaller MC, Ciambotta M, Sapochetti M, Migliore L, Tapia W, Cedeño V, Gentile G. 2010a. Uneven frequency of *Vibrio alginolyticus*-group isolates among different populations of Galápagos marine iguana (*Amblyrhynchus cristatus*). *Environmental Microbiology Reports* 2: 179-184.
- Thaller MC, Migliore L, Marquez C, Tapia W, Cedeño V, Rossolini GM, Gentile G. 2010b. Tracking acquired antibiotic resistance in commensal bacteria of the Galápagos land iguanas: No man, no resistance. *PLoS One* 5: e8989, doi: 10.1371/journal.pone.0008989.
- Watkins G and Cruz F. 2007. Galapagos at risk: A socioeconomic analysis. The Charles Darwin Foundation, Puerto Ayora, Santa Cruz Island: Galapagos Islands, Ecuador. Last accessed online November 11, 2010. <http://www.galapagos.org/>.
- Watson J, Trueman M, Tufet M, Henderson S, Atkinson R. 2010. Mapping terrestrial anthropogenic degradation on the inhabited islands of the Galápagos archipelago. *Oryx*. 44: 79-92.
- Wheeler E, Cann IKO, Mackie RI. 2011. Genomic fingerprinting and serotyping of *Salmonella* from Galápagos iguanas demonstrates island differences in strain carriage. *Environmental Microbiology Reports* 3: 166-173.
- Wikelski M, Gall B, Trillmich F. 1993. Ontogenetic changes in food intake and digestion rate of the herbivorous marine iguana (*Amblyrhynchus cristatus*, Bell). *Oecologia* 94: 373-379.

Wikelski M, Wong V, Chevalier B, Rattenborg N, and Snell HL. 2002. Galapagos Islands: Marine iguanas die from trace oil pollution. *Nature* 417: 607-608.

Winfield MD, Groisman EA. 2003. Role of nonhost environments in the lifestyle of *Salmonella* and *Escherichia coli*. *Applied and Environmental Microbiology* 69: 3687-3694.

## Tables and Figures

Table 5.1. Antibiotic resistance profiles of *E. coli* and *S. enterica*, grouped by site and sample

Island	Host species*	Bacterial species	# Isolates	AM**	CIP	CZ	D	GM	TE	SXT	Tet Gene\$
Plaza Sur	LI	<i>E. coli</i>	1	R	S	S	I	S	R	S	A
	LI	<i>E. coli</i>	1	R	S	I	S	S	S	S	Neg
	MI	<i>E. coli</i>	1	R	S	S	I	S	R	R	A
	MI	<i>S. enterica</i> ss. I ser. Sandiego	3	S	S	S	I	S	S	S	Neg
San Cristóbal Punta Carola	Env	<i>E. coli</i>	1	S	S	S	I	S	R	S	A
	Env	<i>E. coli</i>	2	S	S	S	R	S	R	S	B
	Env	<i>E. coli</i>	1	S	S	S	S	S	I	S	Neg
	Env	<i>E. coli</i>	4	R	S	S	R	S	R	S	A
	Env	<i>E. coli</i>	2	R	I	S	R	S	R	S	A
	MI	<i>S. enterica</i> ss. II ser. 47:b:1,5	2	S	S	S	I	S	I	S	Neg
San Cristóbal La Galapaguerra	T	<i>E. coli</i>	5	S	S	S	R	S	R	S	B

R=Resistant, I =Intermediate, S=Susceptible as determined by disc diameter thresholds (NCCLS 2003)

\*LI=land iguana, MI=marine iguana, Env=sea water, T=Galápagos tortoise \*\*AM=Ampicillin, CIP=Ciprofloxacin, CZ=Cefazolin, D=Doxycycline, GM=Gentamicin, TE=Tetracycline, SXT=Trimethoprim/Sulfazoline; Other antibiotics tested AmC=Amoxicillin with Clavulinic Acid, C=Chloramphenacol, Cxm=Cefuroxime

\$ Neg indicated that the isolates tested negative for the presence of TetA, TetB, TetC, TetD, and TetE efflux protein genes as detected by gene-specific PCR amplication (Aminov et al. 2004), while A or B indicates presence of tetracycline resistance genes TetA or TetB, respectively.

Table 5.2. Model selection for factors contributing to the presence of isolates resistant to one or more antibiotics

Model*	AIC	# Parameters	AIC <sub>c</sub>	Weights	Deltas
Town distance+Host species+Bacterial species	61.44	6	62.30	0.3940	0.00
Town distance +Bacterial species	63.29	3	63.53	0.2130	1.23
SE-NW PC1+Host species + Bacterial species	63.83	6	64.69	0.1193	2.39
Town distance + Host species	64.28	5	64.89	0.1080	2.59
SE-NW PC1 + Bacterial species	64.54	3	64.78	0.1140	2.48
SE-NW PC1 + Host species	66.76	5	67.37	0.0312	5.07
Host species + Bacterial species	68.26	5	68.87	0.0148	6.57
Host species	71.06	4	71.46	0.0040	9.16
Bacterial species	72.96	2	73.08	0.0018	10.78
SE-NW PC1	82.55	2	82.67	0.0000	20.37
Town distance	83.65	2	83.77	0.0000	21.47

\* Factor descriptions: “Town distance” is an estimate of proximity to large human populations measured as the geographical distance between each sampling site and the nearest of the three major port towns indicated in Figure 5.1. “Host species” is the sample sources from which resistant bacteria were isolated and included marine iguanas, land iguanas, Galápagos tortoises and sea water samples; “Bacterial species” were either *E. coli* or *S. enterica*. “SE-NW PC” is the first principal components score produced from the latitude and longitude of each sampling site, which condenses geographical location of the sampling sites onto a single axis oriented along a south eastern to northwestern axis.

Table 5.3. Significance tests for factors contributing to antibiotic resistance patterns

Factor*	Degrees of freedom	Log likelihood ratio (D)	p value	Weight of evidence**
Distance from City**	1	9.98	0.002	0.7149
Bacterial Species	1	20.67	<0.001	0.8567
Source of Bacteria	3	26.57	<0.001	0.6712

\* Weight of evidence = the sum of weights for all models in Table 5.2 which contain the factor of interest. This value indicates the relative importance of each factor for explaining the observed antibiotic resistance patterns. This can be interpreted that, while all three factors in the most well supported model have a significant influence on antibiotic resistance patterns, the factor with the largest weight of evidence score (Bacterial species) has the strongest association with the likelihood of antibiotic resistance being present.

\*\* These factors are described in Table 5.2

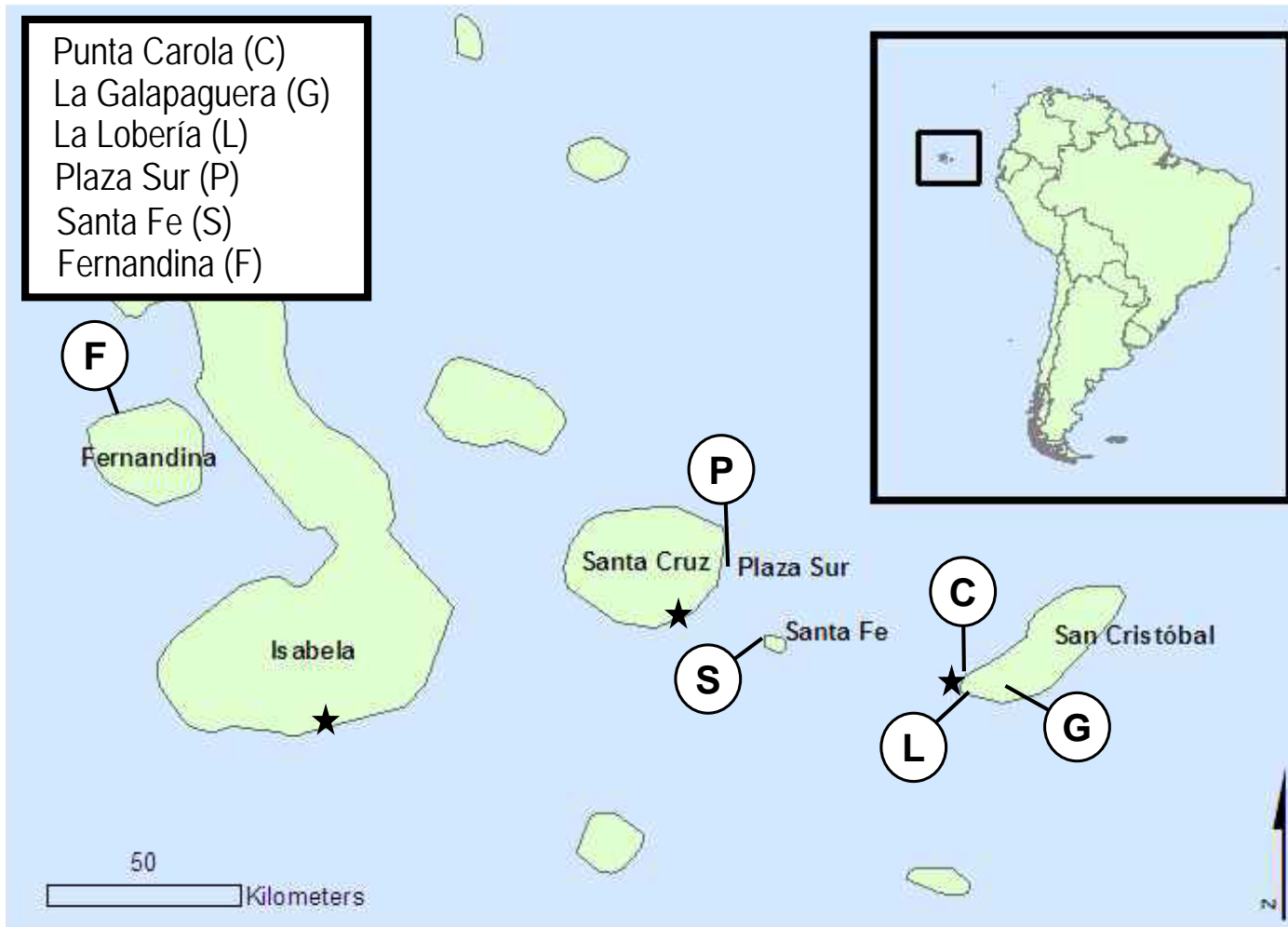


Figure 5.1. Sampling sites in the Galápagos Islands. Marine iguanas were sampled on sites C (N=12), L (N=12), S (N=2), P (N=12) and F (N=12). Land iguanas were sampled at sites S (N=11), P (N=11) and F (N=6). Galápagos tortoises were sampled only site G (N=10).

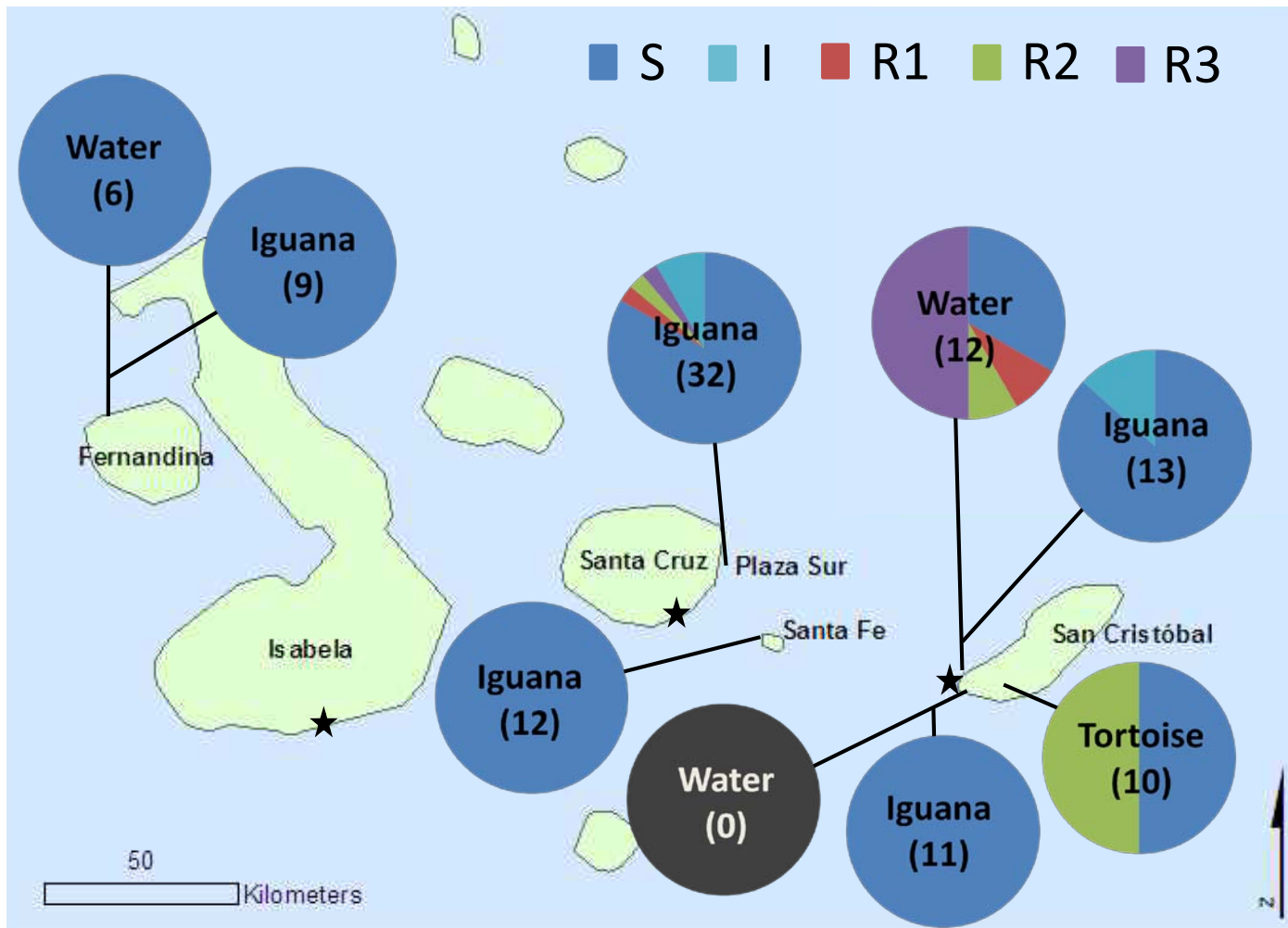


Figure 5.2. Antibiotic resistance patterns for *E. coli* and *S. enterica* isolated from land and marine iguanas, Galápagos tortoises and sea water. Black stars show approximate location of the three major port towns in the Galápagos, Puerto Baquerizo Moreno on the island of San Cristóbal, Puerto Ayora on the island of Santa Cruz and Puerto Villamil on the island of Isabela. Pie-charts show the proportion of isolates demonstrating each resistance phenotype with S=Susceptible to all antibiotics tested, I=Intermediate to one or more antibiotics, R1=Resistant to one antibiotic, R2=Resistant to two antibiotics, R3=Resistant to three antibiotics. The number of isolates evaluated for each is shown in parentheses.

## CHAPTER 6: CONCLUSIONS

### Summary of Findings

The study of microbial biogeography has made vast strides in recent years due, in part, to advances in the technological capacity to document microscopic biodiversity at the community, population and genomic levels. In this dissertation, we applied a combination of molecular and traditional cultivation-based approaches to the study of enteric microbial diversity in a unique system of sister iguanid species – the Galápagos land and marine iguanas (*Conolophus* spp. and *Amblyrhynchus cristatus*, respectively). We explored the spatial diversity of enteric bacteria in these two species across a variety of scales, from the community level to the level of mobile genetic traits, using the unique geography of this island chain to formulate expectations for biogeographical patterns and processes.

First, we explored species and site patterns of gastrointestinal bacterial communities in marine and land iguana populations using 454 pyrosequencing to document bacterial taxonomic diversity and abundance among populations. This study suggested that while species differences in carriage of bacterial genera primarily reflected host or dietary selection, stochastic dispersal and demographic processes and local variation in exposures may also contribute to community differences among sites. Enteric community diversity at the genus level primarily reflected host differences, but within each host ecotype, variation among sites were noted, in particular at a single site, Isla Plaza Sur, where land and marine iguanas experience increased ecological overlap due to island size and dense populations of both species. A number of genera were identified that demonstrated distribution patterns reflective of unique, possibly stochastic site exposures among sampling locations (Chapter 2).

Evidence of such geographical and local site effects at the genus level might lead to the prediction that species within a given genus might also be highly divergent among sites and should mirror the host genetic patterns often seen in such island systems. For example, in island systems, host population genetic patterns often demonstrate distinct patterns such as decreased diversity across a temporal colonization gradient due to founder events. In Chapter 3, we explored the question of how host colonization and population divergence affected the phylogenetic diversity and site similarity within the



genus *Clostridium* in marine iguana populations using 454 pyrosequencing data. *Clostridium* spp. are known to be highly abundant and diverse in this host species, but whether this diversity varied widely among locations had not previously been investigated. In contrast to the expectation of strong site divergence and possible nesting of communities by site age, we found that the phylogenetic structure of *Clostridium* communities was conserved across sites. This phylogenetic conservation was apparent despite concurrent evidence suggesting extinction and speciation events – forces which might be expected to lead to divergences in allopatric host populations. These findings suggest that *Clostridium* diversity may be quite functionally significant, and thus strongly selected for in this system. However, these findings do not necessarily preclude the possibility of local acquisition of site-specific strains within these conserved core taxa.

To further explore the effect of geography on strain diversity patterns, we utilized a cultivation-based approach to explore the population diversity and similarity of a single intestinal bacterial species, *Salmonella enterica*, among host ecotypes and locations. This study demonstrated that while strain movements among sampling sites is likely very rare, movement of genomic elements is much more common and produces a distinctly directional biogeographical relationship among *S. enterica* populations. Further, iguana populations living in areas with heavier human use were also host to more diverse *S. enterica* populations than iguanas living on more isolated islands, which may suggest that local site interactions may also play a key role in genetic trait diversity patterns (Chapter 4).

Finally, we explored this possibility that these site differences in microbial traits could be driven by local exposure dynamics by evaluating the impact of proximity to the Galápagos' larger port towns on antibiotic resistance traits of enteric bacteria in wild iguanas. We found that wild reptiles more proximate to humans had higher potential exposure to antibiotic resistant bacteria and were more likely to carry resistant isolates. These findings together suggest that wildlife with increased human contacts may be exposed to many novel bacteria or bacterial traits from the fluxing tourist and resident human populations on these islands (Chapter 5).

## **Broader Significance**

Through this series of studies we demonstrated that biogeographical patterns and mechanisms can vary widely across levels of organization even within the same study system, with both ecological and evolutionary forces acting in concert to shape enteric biodiversity in a variety of ways. Increased understanding of the interplay among these forces for shaping microbial community form and function across taxonomic scales has potential to improve not only theoretical understanding of microbial ecology but also to progress management of pressing issues such as novel disease emergence or antibiotic resistance dissemination.

### *Theoretical contributions*

The growing consensus is that microbes likely follow the same ecological principles that have been demonstrated in larger organisms, but may not manifest this adherence in the ways we initially expect. Issues of scale, both spatial and temporal, may pose novel challenge to this work due to the effects of small size, large populations, rapid evolution, asexual reproduction and the potential for widespread horizontal gene transfer among species. Yet even without seeing our subjects, we can make great strides and continue to do so as new methods help with documenting the structure of microbial populations and communities more clearly.

This dissertation makes a novel contribution to the broader study of microbial biogeography by exploring patterns within the same sample set across scales of organization. Within a single experiment, we found evidence of strong selective effects and local propagule pressures (whole community level), phylogenetic conservatism (genus level), strong site isolation (strain level) and distance-decay (strain and trait level). Such seemingly inconsistent patterns are not easily resolved and are not even fully resolved through the completion of this work, but such inconsistencies do illuminate a yet untapped source of complexity in microbial biogeography – the interacting effects of ecological and evolutionary processes. Studies of community phylogenetics across scales in plants, for example have revealed that expected spatial patterns may vary across such organizational scales due to the opposing effects of ecological and evolutionary forces (Cavender-Bares et al. 2006).

Microbial biogeography might benefit from increased focus on understanding not whether ecological rules apply, but on incorporating the specific biological traits of systems to understand patterns and mechanisms within the complex context of contemporary and historical processes entangled.

### *Application to health and disease*

Disease agents have at times been treated as a “special case” in microbial systems, studied by epidemiologists and medical researchers without a strong tie to the broader ecological theory of microbial movements. However, interest in the ecology of diseases, in particular, those that have emerged from increasing contacts among wildlife, human and livestock populations, has made it clear that disease organisms exist and persist within broader communities of micro-organisms (Belden and Harris 2007). Improved understanding of both the ecology of known pathogens and of the microbial communities of which these taxa are a part can work in conjunction with more traditional epidemiological approaches to inform management of health and disease. Application of ecological principles to understanding patterns of host-associated microbial population and community assembly has larger consequences than mere theoretical understanding.

Molecular epidemiology approaches traditionally applied to outbreak investigations have been increasingly applied to model bacterial systems to document the effects of environmental changes on pathogen transmission risk (Goldberg et al. 2007, Rwego et al. 2008). Changes in the way landscapes are used for agriculture, urban development, and recreation have resulted in increased overlap among humans, livestock, and wildlife populations, leading to a surge in cross-species disease exchange across the globe (Daszak et al. 2008). However, exploring these effects by detecting new disease incidence relies on passive surveillance of relatively rare events, requiring huge samplings efforts to produce a substantial dataset of disease positive samples. However, commensal organisms may share biological traits such as transmission routes or environmental tolerances which allow for more focused sampling efforts to document “microbial connectivity” among populations. Groups with higher overall “microbial connectivity” may also be at risk for increased exchange of disease agents, and thus represent a key starting place for proactive disease management efforts. By determining

how and when bacteria move among human and animal populations, such studies can reveal new opportunities to prevent illness, even before an outbreak occurs.

### *Future directions*

Documenting microbial community structure at fine taxonomic and genetic levels is becoming increasingly approachable with advances in the sequence lengths and depths that can be obtained from a variety of mixed microbial sources. However, fully incorporating microbial diversity patterns with understanding of the metabolic and ecological contributions of these communities remains a challenge for applying microbial ecology knowledge to broader issues such as host health or ecosystem function. Such advances are quite dependent not only on technology but on careful study design to allow for causal reasoning in such challenging systems where true experimental manipulation may be difficult or impossible (Plowright et al. 2008).

The Galápagos Islands are a challenging place to work as both political, logistical and conservation concerns can place limitations on scientific research. But these challenges are companion to the benefits of working in a system with such a unique geological and biological history. However, the greatest strengths of any model system cannot be fully realized without clear and intentioned study design, and certainly the Galápagos Islands are no exception. To further progress work in this system will require more direct correlation of microbial community and population data to variables which likely play a strong role in shaping these communities – including host genetics, host population density, spatial overlap of host populations and sources of environmental exposures. While improved understanding of microbial systems has potential application for benefiting the health and conservation of endemic Galápagos wildlife, a great deal of work lies ahead for translating the findings presented in this dissertation to practical management decisions for conserving the health and well-being of creatures great and small in this historical biodiversity hotspot.

## **Conclusions**

Together, these studies suggest that enteric microbial community diversity and assemblage involves a complex interplay of evolutionary and ecological process which vary in strength at different taxonomic scales. However, elucidating patterns and mechanisms of diversity is only a first step in using ecological theory to make applied advances towards bettering human and animal health (Robinson et al. 2010). Continuing to document enteric microbial diversity patterns and to link such patterns to community functions such as nutrition, immunity and pathogen resistance has great promise for managing disease risk and impacts in a changing world.

## **References**

- Belden LK, Harris RN. 2007. Infectious diseases in wildlife: the community ecology context. *Frontiers in Ecology and the Environment* 5: 533-539.
- Cavender-Bares J, Keen A, Miles B. 2006. Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology* 87: S109-S122.
- Daszak PA, Cunningham AA, Hyatt AD. 2000. Emerging Infectious Diseases of Wildlife – Threats to Biodiversity and Human Health. *Science* 287: 443-449.
- Goldberg TL, Gillespie TR, Rwego IB, Wheeler ER, Estoff EE, and Chapman CA. 2007. Patterns of gastrointestinal bacterial exchange between chimpanzees and humans involved in research and tourism in western Uganda. *Biological Conservation* 135: 527-533.
- Plowright RK, Sokolow SH, Gorman ME, Daszak P, Foley JE. 2008. Causal inference in disease ecology: investigating ecological drivers of disease emergence. *Frontiers in Ecology and the Environment* 6: 420-429.
- Robinson CJ, Bohannan BJM, Young VB. 2010. From structure to function: the ecology of host-associated microbial communities. *Microbiology and Molecular Biology Reviews* 74: 452-476.
- Rwego IB, Isabirye-Basuta G, Gillespie T, Goldberg TL. 2008. Gastrointestinal bacterial transmission among humans, mountain gorillas, and livestock in Bwindi Impenetrable National Park, Uganda. *Conservation Biology* 22: 1600-1607.

**APPENDIX A: SUPPORTING MATERIALS FOR CHAPTER 2**

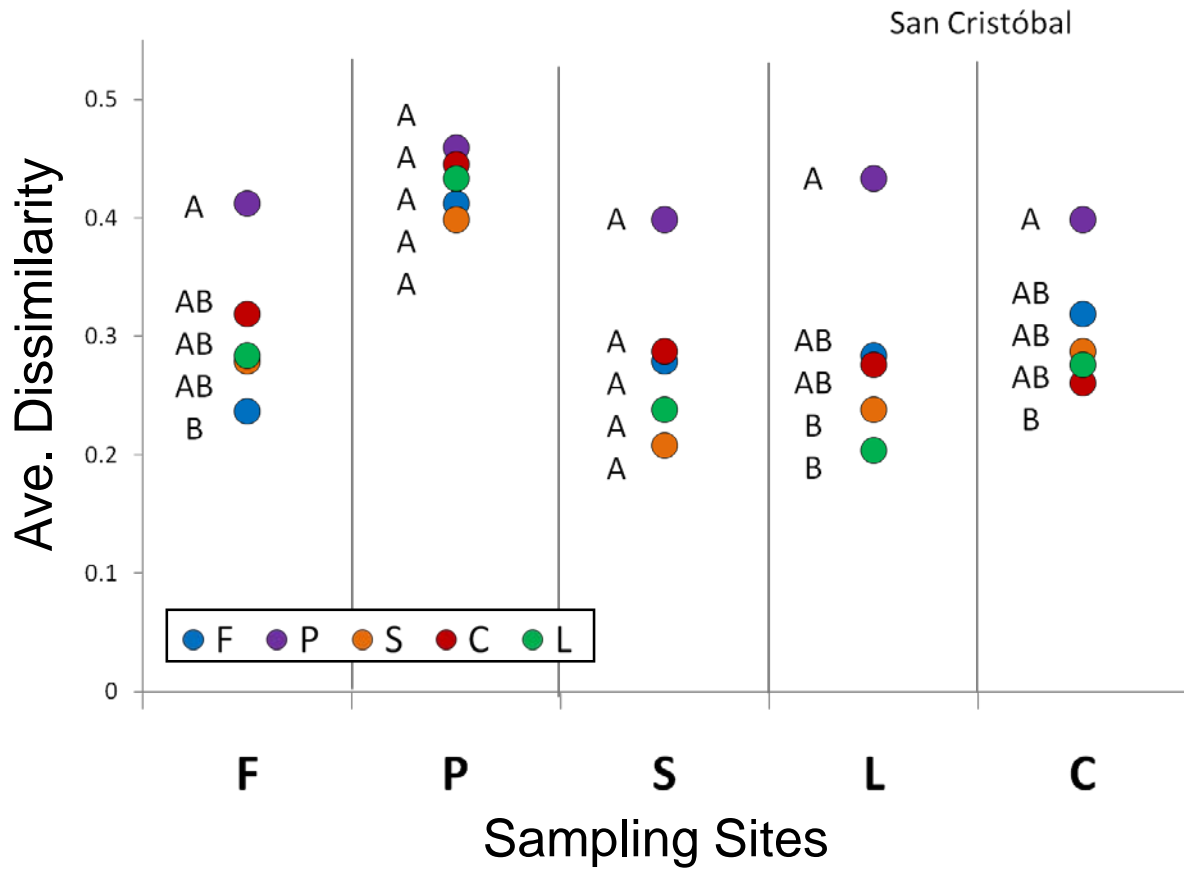


Figure A.1. Patterns in marine iguana gastrointestinal community dissimilarity among sampling sites. The average microbial community dissimilarities between each site compared pair-wise to all other sites including self (compared site identified by the color of the dots within each column as indicated in the legend; C=Punta Carola, L=La Lobería, S=Santa Fe, P=Plaza Sur and F=Fernandina). Smaller letters next to each point represents significant differences among groups within each column (Tukey's HSD adjustment for multiple comparisons).

**APPENDIX B: SUPPORTING MATERIALS FOR CHAPTER 3**

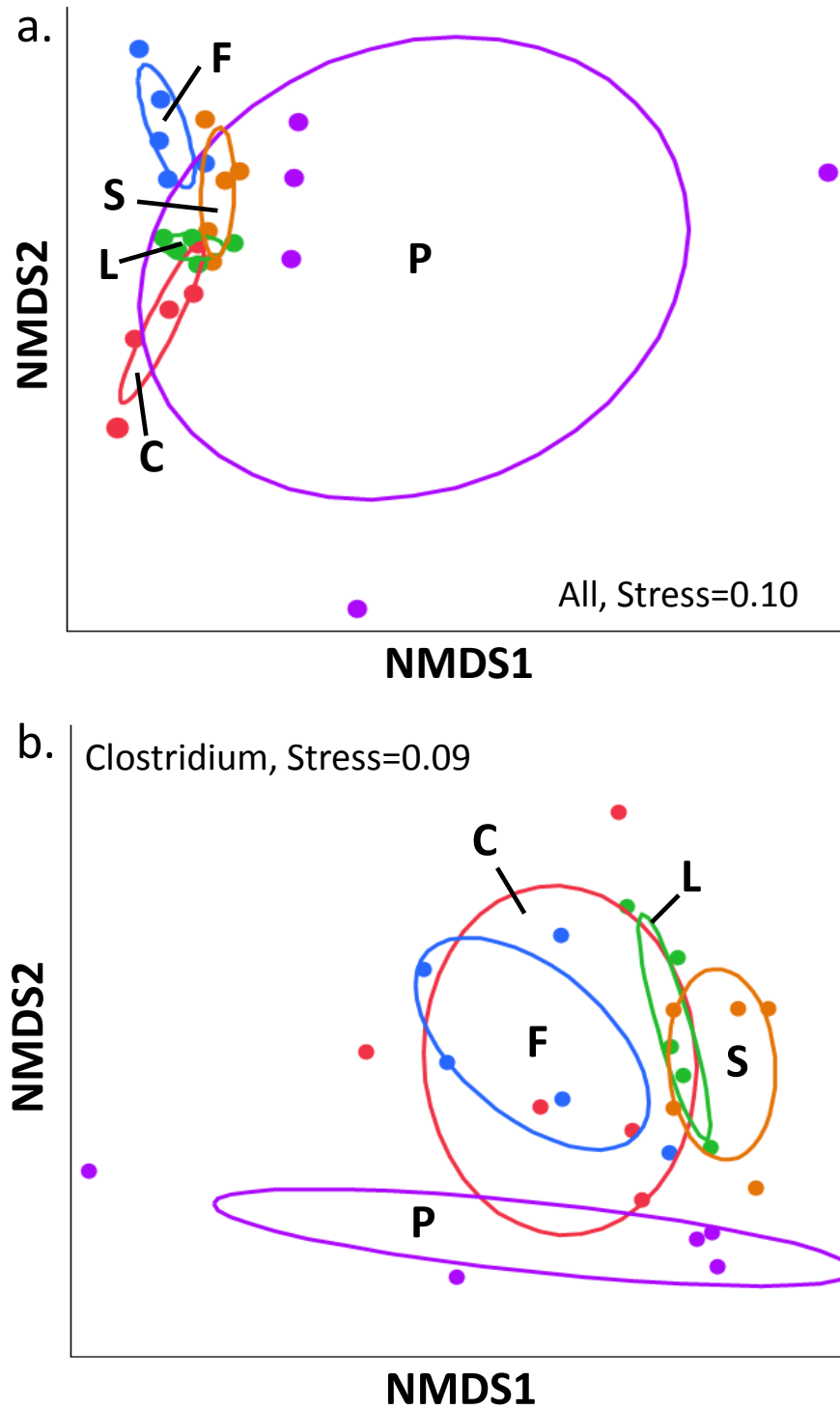


Figure B.1. Non-metric multidimensional scaling plots of 95% sequence similarity OTUs. Permutation MANOVA analysis by site of all taxa and of the *Clostridium* subset were both significant (All taxa:  $R^2=0.411, p<0.001$ ; *Clostridium*  $R^2=0.327, p=0.009$ ).

## APPENDIX C: SUPPORTING MATERIALS FOR CHAPTER 4

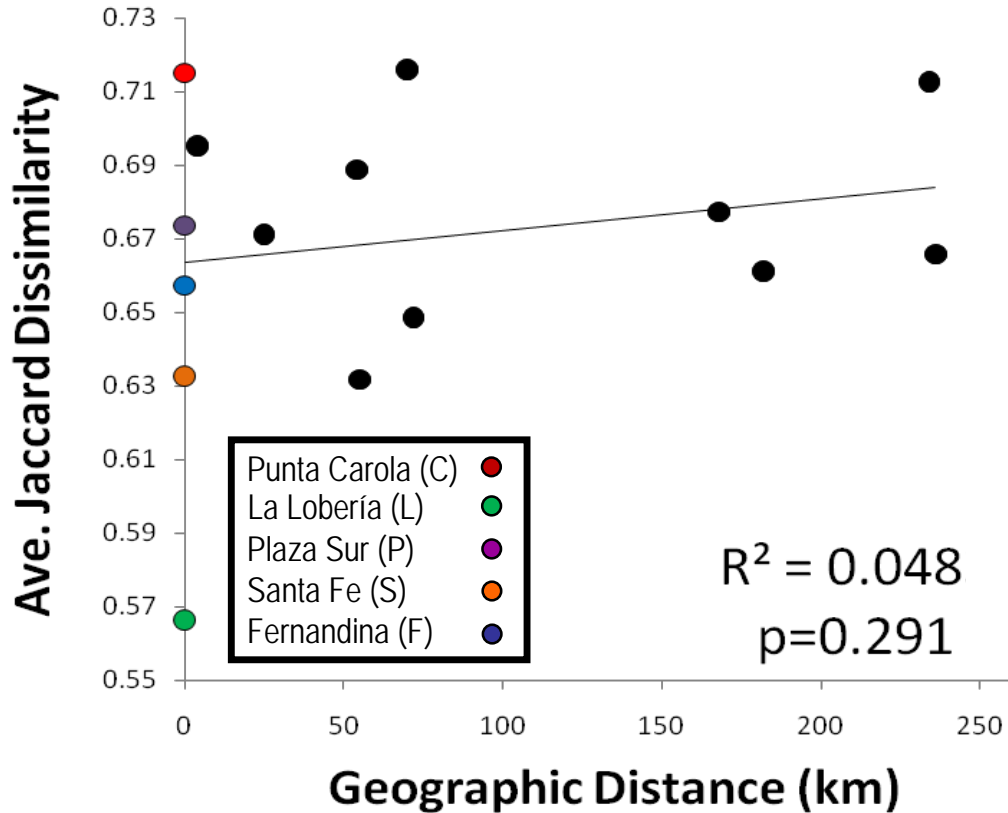


Figure C.1 Correlation of geographical distance and the average dissimilarity of sites from genomic fingerprints. The relationship between geographical and genetic distance was not significant. Removal of the highly variable within-site comparisons (colored dots as indicated by the legend) results in further reduction in the strength of association between geographic and genetic distances. Black dots represent pair-wise comparisons between locations among the five sites.



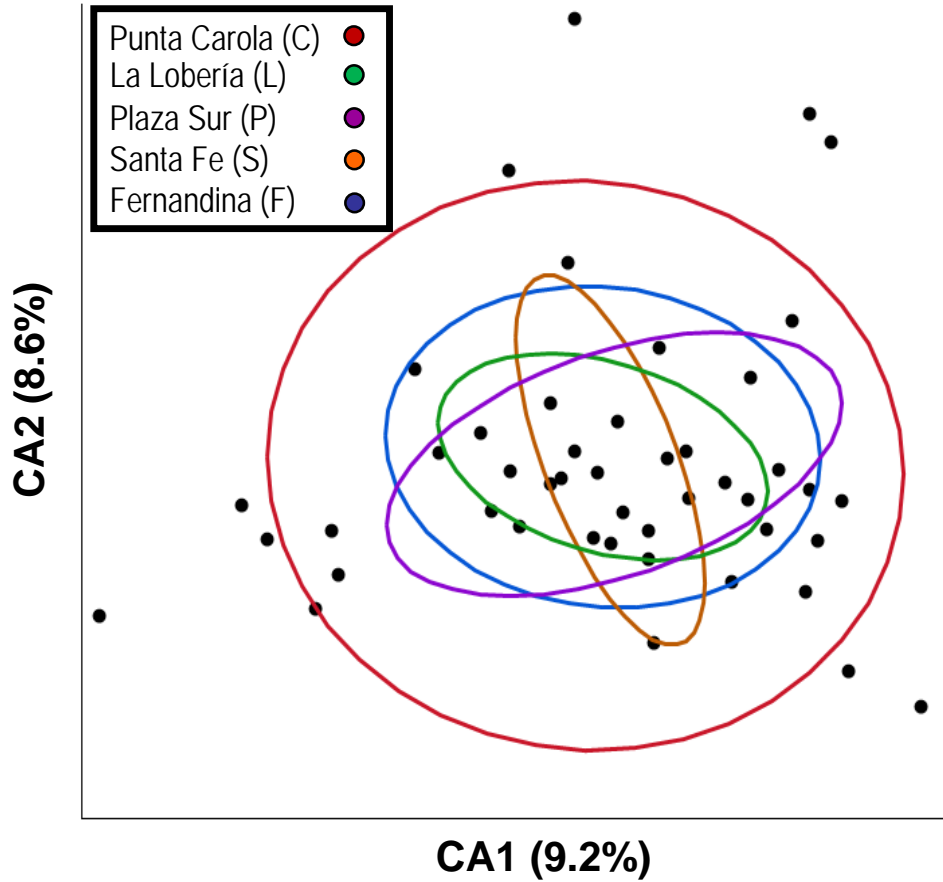


Figure C.2. Unconstrained correspondence analysis of *S. enterica* genomic fingerprints (as in Figure 4.2a). Grouping of isolates in this figure is shown by sampling site (as 50% contour ellipses colored by sampling site). Sampling sites do not demonstrate unique clustering but do exhibit site differences in strain diversity on the genomic level, ranging from high strain variability at Punta Carola (red) to relatively low strain diversity at La Lobería (green).

## **APPENDIX D: SUPPORTING MATERIALS FOR CHAPTER 5**

### **Supplemental Methods**

Genomic fingerprinting of *E. coli* and *S. enterica* isolates was done by PCR of repetitive extragenic palindromic elements (rep-PCR, Versalovic et al. 1991). Rep-PCR was performed as in previously described protocols (Versalovic et al. 1991, Rademaker et al. 1998). DNA was extracted from colonies grown on tryptic soy agar using a modification of the manufacturer's protocol for Instagene chelex resin (Biorad, Hercules, CA). Briefly, 2-4 large, well-isolated colonies were placed in 100 µl of 6% chelex 100 resin and heated at 95°C for 10 minutes to lyse cells. Lysates were centrifuged for 5 minutes at 14000 g and 80µl of supernatant was transferred to a new storage tube.

Rep-PCR amplification was performed using primers ERIC1R (5'-ATG TAA GCT CCT GGG GAT TCA-3') and ERIC2 (5'-AAG TAA GTG ACT GGG GTG AGC G-3'), targeting enterobacterial repetitive intergenic consensus (ERIC) repetitive motifs (Versalovic et al. 1991) on a TGradient thermocycler (Biometra, Germany) with an initial denaturation at 95° C for 2 minutes followed by 30 cycles each of 94° C for 3 seconds, 92° C for 30 seconds, and 50° C for 1 minute and a final extension at 65° C for 8 minutes. Polymerase chain reaction amplification mixtures (25µl) included 1.75U of Takara Taq polymerase (Takara Bio Inc., Japan), 1X Takara PCR Buffer with 2.5mM (final concentration) of MgCl<sub>2</sub>, 2.0mM Takara dNTP mixture (0.5mM each), 1µM each of the forward and reverse primers, and approximately 50ng of template. PCR products were separated on a 2% Agarose gel in 0.75% TAE (Tris-Acetate EDTA) run at 80 V for 12 hours. Three lanes of a 1kb plus DNA ladder (Novartis, Carlsbad, CA) were included to allow for standardization of molecular weight assignments to DNA fragments, with one lane at each end of the gel and one located in the approximate middle to allow for correction of irregular gel runs. Gels were stained with ethidium bromide and digitally photographed using an AlphaImager™ 2200 (Alpha Innotec Co., San Leandro, CA).

One isolate per iguana from a subset of randomly selected iguanas was submitted for serotyping at the National Veterinary Service Laboratory, Ames, IA. Serotype identities were compared between this study and a previous study encompassing two of the same sampling sites from 2005 (Table D.1).

## Supplemental Data Analysis

Rep-PCR band assignment and sizing from gel images was done using BioNumerics version 4.0 (Applied Maths, Belgium). Band assignments were exported as a binomial presence/absence matrix for ordination by both constrained and unconstrained correspondence analysis (CA and CCA, respectively) to visualize patterns. To control for non-independence (i.e. pseudoreplication, Hurlbert 1984) resulting from inclusion of multiple isolates per individual iguana fecal sample, fingerprint profiles of isolates from the same individual iguana were nested by averaging presence/absence within an individual host. As previous work in this system demonstrated that sampling location is a stronger driver of *S. enterica* population patterns than host species (Wheeler et al. 2011), isolates from land and marine iguanas from a given site were pooled to obtain increased power for exploring site-associated bacterial population patterns. Previous analyses of the *S. enterica* isolates included in this study demonstrated that the host species from which isolates were obtained was not a significant factor (Chapter 4), further supporting an analysis focused on site differences with pooling by host source.

A complete Jaccard dissimilarity matrix was calculated from Rep-PCR fingerprints for *E. coli* isolates and *S. enterica* isolates independently using the `vegdist` function in the `vegan` package in R (Oksanen et al. 2008). This matrix was then used to calculate average dissimilarities to compare patterns of strain diversity within and between sampling sites. For each isolate, we calculated the average dissimilarity of that isolate to all other isolates from the same species within and between islands, resulting in one average dissimilarity per isolate. We then used ANOVA to test whether the sites differed in the degree of genetic dissimilarity among their component isolates. A significantly higher average dissimilarity among isolates from the same site suggests a larger variety of divergent fingerprint patterns, and thus, higher strain or genomic diversity within that comparison (Figure D.2).

## Supplemental References

Hurlbert SH. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, 54: 187-211.

- Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2008. vegan: Community Ecology Package. R package version 1.15-0. <http://cran.r-project.org/>, <http://vegan.r-forge.r-project.org/>.
- Rademaker JLW, Louws FJ, De Bruijn FJ. 1998. Characterization of the diversity of ecologically important microbes by rep-PCR genomic fingerprinting. Molecular Microbial Ecology Manual. Kluwer Academic Publishers, the Netherlands: 1-27.
- Versalovic J, Koeuth T, Lupski JR. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Research 19: 6823-6831.
- Wheeler E, Cann IKO, Mackie RI. 2010. Genomic fingerprinting and serotyping of *Salmonella* from Galápagos iguanas demonstrates island differences in strain carriage. Environmental Microbiology Reports 3: 166-173.

## Supplemental Tables and Figures

Table D.1 Comparison of *S. enterica* serotypes from Plaza Sur and Santa Fe in 2005 and 2009

Serotype	Serogroup	Total (55)	Plaza Sur (# isolates)		Santa Fe (# isolates)	
			2005 (21)	2009 (15)	2005 (13)	2009 (6)
Treforest	O:51	1				1
Berta	O:9 (D1)	1				1
Rough O:z10:enx	Rough	1			1	
Montevideo	O:7 (C1)	1			1	
Rough O:z10:enz15	Rough	3			3	
<b>Rough O:L,V:1,7</b>	<b>Rough</b>	<b>4</b>			<b>3</b>	<b>1</b>
SSI 28:v:-	O:28 (M)	1			1	
Muenchen	O:8 (C2-C3)	1		1		
IV 53:z4,z23:-	O:53	1		1		
IV 44:z36:-	O:44	1		1		
<b>Sandiego</b>	<b>O:4 (B)</b>	<b>16</b>	<b>7</b>	<b>6</b>	<b>3</b>	
<b>Pomona</b>	<b>O:28 (M)</b>	<b>7</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>3</b>
<b>Manhattan</b>	<b>O:8 (C2-C3)</b>	<b>4</b>	<b>2</b>	<b>2</b>		
<b>Poona</b>	<b>O:11 (G)</b>	<b>5</b>	<b>2</b>	<b>3</b>		
Panama	O:9 (D1)	6	6			
Oranienburg	O:7 (C1)	2	2			



Figure D.1. Suspected sewage leak at Punta Carola on the island of San Cristóbal. A large diameter pipe from the city of Puerto Baquerizo Moreno exits into the ocean (arrow) along the rocky coast at Punta Carola where marine iguanas, sea lions and shore birds congregate. The brown discoloration of the water proximate to the shore at this site is suspected to be sewage leaking from the pipeline which empties into the ocean a few kilometers distance from the coast. Two small (50 ml) samples of sea water from this location grew numerous diverse colonies on gram-negative specific nutrient media, including isolates of *Escherichia coli*, which suggests that this area is a site of heavy fecal contamination. In contrast, sea water samples from Isla Fernandina and from La Lobería on the southern side of San Cristóbal had low or no growth (respectively) of gram negative bacteria from two 50ml water samples taken at each site.

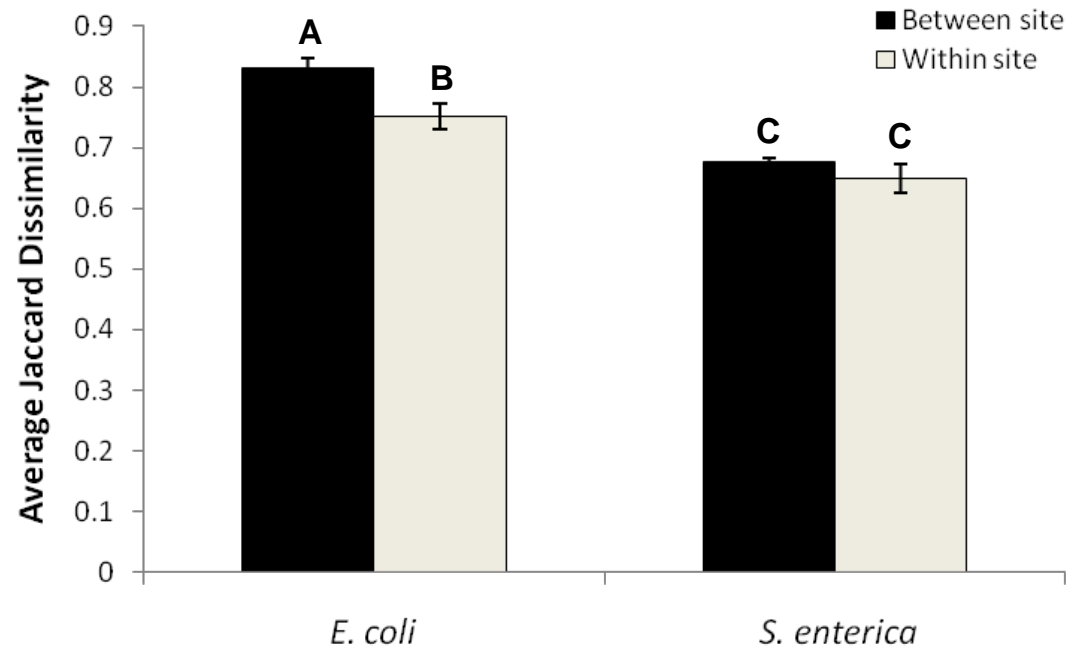


Figure D.2. Within and between site diversity patterns for *E. coli* and *S. enterica* isolated from land and marine iguanas (shown as average +/- 1 SEM) A fully factorial ANOVA model of bacterial species and type of comparison demonstrated significant differences between bacterial species and within versus between site comparisons (Full model  $R^2=0.740$ ,  $p<0.001$ ; Bacterial species  $p<0.001$ , Type of comparison  $p=0.036$  and their interaction  $p=0.1464$ ). While *E. coli* show significantly higher strain diversity when comparing isolates from different sites versus isolates from host species living at the same site (Tukey's HSD: significant differences indicated by different letters over the error bars), *S. enterica* shows an overall lower diversity compared to *E. coli* and no difference for between versus within site comparisons. These findings suggest that *E. coli* and *S. enterica*, while relatively genetically similar organisms, may have functionally very different population ecology in this system.