Effects of Environment on the Evolution of Variation

Presented in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy

with a Major in

Bioinformatics and Computational Biology

in the

College of Graduate Studies

University of Idaho

by

Genevieve Ann Metzger

Major Professors
Eva M. Top, Ph.D.
Jack M. Sullivan, Ph.D.

Committee
Benjamin J. Ridenhour, Ph.D.
Holly A. Wichman, Ph.D.

Department Administrator
David C. Tank, Ph.D.

December 2016

Authorization to Submit Dissertation

This dissertation of Genevieve Ann Metzger, submitted for the degree of Doctor of Philosophy with a Major in Bioinformatics and Computational Biology and titled "Effects of Environment on the Evolution of Variation," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professors:	Eva M. Top, Ph.D.	Date
	Jack M. Sullivan, Ph.D.	Date
Committee Members:	Benjamin J. Ridenhour, Ph.D.	Date
	Holly A. Wichman, Ph.D.	Date
Department Administrator:	David C. Tank, Ph.D.	Date

Abstract

The role of spatial structure on the patterns and maintenance of diversity in populations is a long-standing area of research in evolutionary biology. The effects of spatial structure have been well documented in large eukaryotes but questions still remain about the how specific environmental factors and historic patterns of spatial structure influence modern distributions of diversity. At the level of microorganisms, research into the influence of spatial structure on diversity has recently begun to develop at a rapid pace. Previous studies have shown that spatial structure prevents selective sweeps in bacterial populations, increasing diversity by limiting competition between genotypes to a local, rather than global, scale. In this dissertation I seek to address questions of the influence of the environment, especially spatial structure, on the maintenance and pattern of diversity in two organisms: *Ascaphus montanus*, the Rocky Mountain tailed frog, and *Acinetobacter baumannii*, a biofilm-forming Gram-negative bacterium.

In *A. montanus* I addressed the influence of environmental variables, incorporated through the use of Species Distribution Models, on the distribution of diversity at multiple spatial scales, from the entire species range, to within local clusters. Further, I used modeling based on estimates of past environmental conditions to investigate the role of historic separation of the species range into distinct glacial refugia affects current patterns of genetic diversity. I found that the influence of current vs. historic conditions varied based on spatial scale, with historic factors being most important at the largest spatial scale and modern environmental conditions being increasingly important at smaller spatial scales.

In *A. baumannii* I utilized a large, replicated, experimental evolution design to address the role of spatial structure resulting from biofilm growth and the presence or absence of an environmental variable, tetracycline, on evolution of both phenotype and genotypes of *A. baumannii* and the pB10 plasmid it carried. The presence of tetracycline did increase improvement of plasmid persistence in biofilms but did not alter genetic diversity of the plasmid or host. When compared to growth in planktonic populations, growth in the spatially structured biofilm environment increased phenotypic diversity in the form of plasmid persistence, though it also limited the average strength of improvement in persistence. Biofilm growth also resulted in markedly different patterns of genetic diversity in the plasmid, with most clones that were isolated from the biofilm

populations containing transferrable pB10. In contrast, only two plasmids isolated from the planktonic populations contained transferrable pB10. In the remaining plasmids large portions of the plasmid genome had been lost, resulting in loss of the genes involved in conjugation and making plasmid transfer impossible. This result suggests that spatial structure may dramatically modify the availability of plasmid genes in a population of bacteria compared to expectations based on studies performed with planktonic populations. Finally, I found that there were potential small differences in genetic diversity of *A. baumannii* itself, with a tendency for more unique mutations to be found when comparing bacteria isolated from biofilms to those isolated from planktonic populations.

As a whole, these results confirm the importance of spatial structure and environmental variables on the evolution of diversity across multiple spatial and temporal scales and within widely differing organisms.

Acknowledgments

I am extremely grateful for my advisors, Eva Top and Jack Sullivan, who have both taken chances on me at times during my degree. Thank you, Jack, for taking me in as a transfer student and encouraging me to work on plasmids when it became clear I did not share your passion for chipmunks. Thank you, Eva, for letting me jump headlong into a large experimental evolution project, for guiding me through (many) mistakes, and for always taking my thoughts about the direction of the project seriously.

I am also glad to have had the support and advice of Ben Ridenhour and Holly Wichman, my other committee members, and many other faculty members at the University of Idaho with whom I interacted during my degree. I have been grateful to always feel like your doors were open to me, no matter how large or minor my questions.

Special thanks to all the people involved in these projects, including multiple undergraduate students I had the pleasure to mentor during my time in Idaho: Rachana, Taylor, Mitch, Sean, Lindsey, and Andrew. Chris Drummond collected most of the data for the *Ascaphus* project and performed many analyses that, while not directly used in our final publication, were nonetheless incredibly informative as we decided which direction to take with our analyses and manuscript. Anahi Espindola ran MAXENT for the /textitAscaphus project and we spent countless hours in ArcGIS creating the maps that feature in the final manuscript. Additionally, Michael France and Thibault Stalder were an amazing help setting up and running the biofilms, Karol Gliniewicz ran most DNA extractions and qPCR, and Jack Millstein helped me manage the undergraduate students and did a tremendous amount of bench work for the *Acinetobacter baumannii* project. They stuck with me through the steep learning curve while I transitioned into microbiology and project management. Thank you as well to the other members of the Forney lab who provided ideas and skilled hands in the lab at various points.

Although my dissertation touches only slightly on anything phylogenetic, I spent many a Friday morning at PuRGe and always left feeling more excited to be doing science. After completing a Master's Degree in a phylogenetics lab as part of a department that considered phylogenetics to be practically a dirty word I cannot describe how amazing it was to spend 5 years surrounded by some of the brightest minds in phylogenetics.

I am extremely lucky to have been born into a family where science and medicine are a way of life. I suspect my parents knew I was going to be a scientist even before I did. I have been grateful for their love and support throughout my education.

One of the best things about being a part of the BCB program at University of Idaho was that it meant spending time with the best group of students I have ever had the fortune to meet. They have been there for countless poster and presentation prep sessions, have always been up for a trip to the Alehouse after a hard day, and have made my time at UI one of the best of my life.

Lisha Abendroth deserves a special mention for her tireless attention to detail that has prevented me from missing many a deadline. She has always been willing to help me out with paperwork or problems that were not even her responsibility and she provides the fastest reimbursements in the West.

And finally, and most importantly, I have been fortunate to get to spend my time in grad school beside Tyler Hether, an amazing friend, partner, husband, and father. Although we are (finally) concluding the "school" chapter of our lives I look forward to a lifetime of scientific discussion with him.

Dedication

For my children, Hannah and Torin, and my husband Tyler. Without their love and support this dissertation would not have been possible.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

General Introduction

Genetic diversity is generated through mutation. Once generated, the processes by which certain genotypes are preserved in or lost from the population, selection, drift, and migration, are heavily influenced by environmental factors. Environmental variables like precipitation define areas of suitable and unsuitable habitat and divide or limit interactions between populations of a species leading to spatial structure (Thomas et al., 2001). The structure limits interactions between genotypes, preventing selective sweeps that would otherwise reduce variation to the most favored genotype (Koch, 1974; Levin, 1981). Smaller subpopulations created by spatial structure increase the effects of genetic drift and allow local adaptation. Migration, mediated by environmental factors that enhance or limit movement of individuals, introduces alleles evolved in one subpopulation to other subpopulations (Ally et al., 2014). In my dissertation I address the influence of the environment on diversity at multiple spatial and temporal scales using two very different organisms: *Ascaphus montanus*, the Rocky Mountain tailed frog and *Acinetobacter baumannii*, a biofilm-forming, Gram-negative bacterium.

1.0.1 *Landscapes, Genes, and the Structure of Diversity*

In eukaryotic organisms habitat variables such as land use, precipitation, forest cover, and elevation define regions of suitable and unsuitable habitat and influence both the routes and frequency of migration between subpopulations (Nevo, 1978; Fore et al., 1992; Schmidt et al., 2009; Mosca et al., 2012). With the increased availability of molecular markers to measure diversity the field of landscape genetics has developed. This field attempts to connect the patterns of habitat variables, both past and present, with patterns of genetic diversity in organisms. These studies allow the identification of variables that are highly influential in the patterns of diversity for particular organisms, information that can then be used to guide decisions about conservation and land use. For example, the presence of impermeable surfaces such as roads was shown to be a significant predictor at small spatial scales for genetic variation in *Bufo boreas* in Yellowstone National Park (Murphy et al., 2010). This can be taken into consideration when plans are made for additional

roads to be constructed in areas that represent important habitat for *B. boreas*. Similar techniques that incorporate information about past climatic fluctuations and projected future climatic shifts can elucidate historical distributions of organisms, helping to explain long term effects of the environment on diversity and providing insight into how future changes in climate may affect species distributions and gene flow (Rossetto et al., 2012). In this dissertation I use current and historic data on environmental variables in combination with a set of microsatellite markers to assess the role of the environment on patterns of diversity in *A. montanus* at multiple spatial scales.

1.0.2 *Structure and Diversity at the Microscale*

Although frequently studied in planktonic populations with minimal or no spatial structure, genetic diversity in microorganisms such as bacteria is also strongly influenced by spatial structure. Previous work has identified increases in both phenotypic and genotypic diversity in bacteria in the presence of spatial structure (Boles et al., 2004; Rainey and Travisano, 1998; Korona et al., 1994). What is less clear is the influence that spatial structure may have on the evolution of plasmids carried within bacteria growing in spatially structured populations. Conjugative plasmids carrying genes for antibiotic resistance can be passed from one bacterium to another and may result in terrifyingly rapid spread of antibiotic resistance (Mazel and Davies, 1999; Mathers et al., 2011). Understanding how their diversity and persistence is affected by spatial structure is an important part of the fight against antibiotic resistant bacteria. Within this dissertation I investigate the impact of biofilm growth on (1) the evolution of plasmid persistence for the plasmid pB10 in *A. baumannii*, (2) genetic diversity and patterns of diversity of pB10, and (3) genetic diversity of *A. baumannii* itself.

1.1 LAYOUT OF CHAPTERS IN THIS DISSERTATION

Chapter 2 discusses the influences of environmental variables, both current and historical, on the spatial structure and extent of genetic diversity in populations of the Rocky Mountain Tailed Frog, *A. montanus*, in the Inland Northwest (INW) of the North America. The species is specialized to fast-flowing streams in montane areas of the INW. Phylogenetic data suggest that *A.*

montanus has inhabited the INW since the Pliocene and likely inhabited 2-3 glacial refugia during the climatic fluctuations of the Pliestocene. We used microsatellite markers and species distribution models (SDMs) that incorporate a number of environmental variables to address three hypotheses about the interaction between environment and genetic diversity in *A. montanus*. First, that the historic influence of refugial structure would be detectable on the range-wide spatial scale. Second, that environmental variables influence migration, resulting in a correlation between the representation of those variables in the SDMs and the spatial distribution of genetic diversity. Third, that the influence of historical factors would be most obvious at the largest spatial scales while the influence of current environmental variables would be most obvious at the smallest spatial scale.

Chapter 3 summarizes the results of an experimental study on the evolution of persistence of a multidrug (MDR) resistance plasmid in a novel bacterial host, *A. baumannii*. Genes coding for antibiotic resistance may be transferred between different bacteria via conjugative plasmids. Multidrug resistance conjugative plasmids are particularly important from a medical standpoint because they allow rapid dissemination of resistance genes to new hosts, dramatically speeding up evolution of resistance (Mazel and Davies, 1999; Mathers et al., 2011). Once in a novel host, the rate of loss of the MDR plasmid may be rapid until adaptation reduces the cost of the plasmid (Heuer et al., 2007; Subbiah et al., 2011; Sota et al., 2010). The plasmid loss rate has implications for the availability of resistance genes for transfer to new hosts. Multiple studies have addressed the issue of plasmid-host adaptation resulting in improved plasmid persistence but none of these studies have addressed the influence of spatial structure as a result of growth of the host in biofilms (Heuer et al., 2007; Subbiah et al., 2011; Sota et al., 2010). Because biofilms are widespread in environmental and clinical populations of bacteria this represents an important knowledge gap. We used a replicated experimental evolution design to compare the evolution of plasmid persistence between structured populations of *A. baumannii* growing in biofilms and unstructured populations growing in well-mixed batch cultures. We hypothesized that clones isolated from biofilms would show less dramatic improvements in plasmid persistence due to relaxed selection in the spatially structured environment but also that the diversity in plasmid persistence phenotypes would be higher in biofilms as a result of the same forces.

Chapter 4 utilizes an expanded version of the experimental evolution setup used in Chapter 3 to address how growth in spatially structured biofilms with and without the presence of a plasmid-selective antibiotic influenced evolution of plasmid genetic diversity. Based on the results from Chapter 3 we expected that patterns of plasmid genetic diversity would differ between biofilms and planktonic populations. We hypothesized that the presence of plasmid-selective antibiotics during evolution would result in evolved clones with higher levels of plasmid persistence regardless of whether the source population was biofilm or planktonic, and that biofilm growth would lead to greater genotypic diversity among evolved plasmids than planktonic growth.

Chapter 5 summarizes the amount of genetic diversity found in clones of *A. baumannii* sequenced during the evolution experiment described in Chapters 3 & 4. We hypothesized that genetic diversity of *A. baumannii* would be higher in biofilm populations than in liquid due spatial structure in biofilms preventing selective sweeps. Although there were fewer unique mutations in the planktonic populations than the biofilms the differences did not rise to the level of statistical significance.

CHAPTER 2

Genetic Structure Across Broad Spatial and Temporal Scales: Rocky Mountain Tailed Frogs (*Ascaphus montanus*; Anura: Ascaphidae) in the Inland Temperate Rainforest[1]

2.1 SUMMARY

Contemporary and historical processes interact to structure genetic variation, however discerning between these can be difficult. Here, we analyze range-wide variation at 13 microsatellite loci in 2098 Rocky Mountain tailed frogs, *Ascaphus montanus*, collected from 117 streams across the species distribution in the Inland Northwest (INW) and interpret that variation in light of historical phylogeography, contemporary landscape genetics, and the reconstructed paleodistribution of the species. Further, we project species distribution models (SDMs) to predict future changes in the range as a function of changing climate. Genetic structure has a strong spatial signature that is congruent with a deep (1.8 MY) phylogeographic split in mtDNA when we partition populations into 2 clusters (K=2), and is congruent with refugia areas in paleo-range reconstructions. There is a hierarchical pattern of geographic structure as we permit additional clusters, with populations clustering following mountain ranges. Nevertheless, the species exhibits classical population structure, where genetic diversity is highest in populations at the center of the range and is attenuated in populations closer to the range edges. Similarly, geographic distance is the single best predictor of pairwise genetic differentiation, but connectivity also is an important predictor. At intermediate and local geographic scales, deviations from Isolation-by-Distance are more apparent. These results indicate that both historical and landscape factors are contributing to the genetic structure and diversity of tailed frogs in the Inland Northwest, and that all these variables should be taken into account to define appropriate conservation measures.

[1]This is a pre-copyedited, author-produced PDF of an article accepted for publication in the Journal of Heredity following peer review. The version of record Genevieve Metzger, Anahi Espindola, Lisette P. Waits, and Jack Sullivan. Genetic Structure across Broad Spatial and Temporal Scales: Rocky Mountain Tailed Frogs (*Ascaphus montanus*; Anura: Ascaphidae) in the Inland Temperate Rainforest. J Hered (2015). 106 (6): 700-710. doi:10.1093/jhered/esv061 is available online at: http://jhered.oxfordjournals.org/content/106/6/700.short.

2.2 INTRODUCTION

Understanding the interactions between historical, evolutionary processes and contemporary, landscape processes in structuring patterns of genetic diversity is one of the biggest challenges in ecological and landscape genetics (Balkenhol et al., 2009; Sork and Waits, 2010). That is, processes acting across diverse temporal and geographic scales influence genetic diversity in ways that can be difficult to disentangle (Anderson et al., 2010). This is critical, however, because different questions in evolution and conservation may focus on different temporal or regional scales. For example, conservation biologists may be interested in current processes such as population connectivity and gene flow, whereas phylogeographers may be interested in historical processes, such as vicariance and dispersal from refugia, which often have acted across range-wide scales, and this has impeded the integration of the inference across scales.

However, historical and contemporary processes can each be important at both local and range-wide scales (Table 2.1). Thus, investigations into the genetic diversity and structure of a taxon should be interpreted at multiple time scales, regardless of the geographic scale in which the researcher is primarily interested. This is especially true for species that have a long evolutionary history in a particular ecosystem. Here, we provide a broad, multiscale investigation into ecological and population genetics of *Ascaphus montanus* Mittleman and Myers 1949, the Rocky Mountain tailed frog, a species endemic to inland rainforests of the northwestern US (Figure 2.1), and elucidate the interaction of these factors in governing the genetic structure of the species across its range.

The cedar-hemlock forests of the Inland Northwest of North America represent the largest inland temperate rainforest in the world and, along with the coastal rainforest, forms a large disjunction that occurs in >156 species or species complexes (e.g., Nielson et al., 2001; Gavin, 2009). Inland and coastal rainforest ecosystems are separated by >160 km of the xeric Columbia Basin, and comparative distributional and phylogeographic studies (e.g., Carstens et al., 2005, 2013; Gavin, 2009; Björk, 2010) have indicated that the inland rainforest ecosystem is composed of both old endemics (the result of a pre-Pleistocene vicariance) and recent arrivals (i.e., post-Pleistocene dispersers). Furthermore, phylogeographic studies of old endemics such as Constance's bittercress (*Cardamine constancei*; Brunsfeld and Sullivan, 2005), Coeur d'Alene salamanders (*Plethodon idahoensis*; Carstens et al., 2004) and Rocky Mountain tailed frogs (see

Table 2.1: Temporal and spatial scales of phenomena that structure patterns of genetic diversity.

	Local-Scale	**Range-wide**
Contemporary	SDM/Connectivity	Isolation by distance
	Gene Flow	Patterns of diversity
Historical	Introgression	Refugium/refugia
	Suture zones	Climatic fluctuations

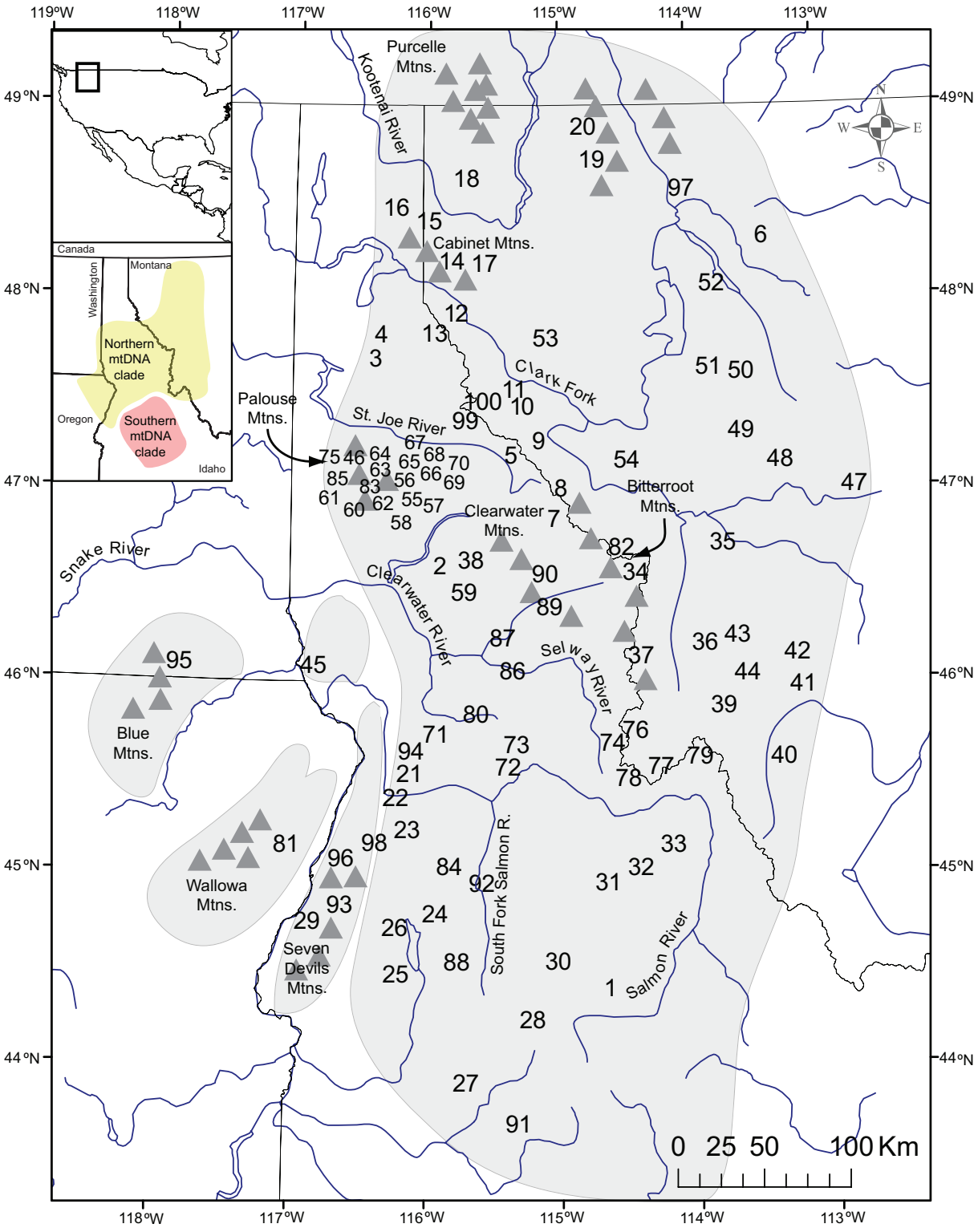

FIGURE 2.1: Range of *Ascaphus montanus* (shaded) and collection localities. The inset shows the same area but illustrates the distribitons of northern and southern mtDNA clades reported in Nielson et al. (2006). The entire region is mountainous and subranges discussed in the text are indicated by triangles.

below; Nielson et al., 2006) have indicated that there was more than a single inland Pleistocene refugium for rainforest taxa.

The genus *Ascaphus* Stejneger represents an ancient lineage that is either sister group to all other frogs (e.g., Ford and Cannatella, 1993) or is sister to the New Zealand genus *Leiopelma*, which then is sister to the remaining frogs (e.g., Pyron and Wiens, 2011; Zhang et al., 2013). Regardless of which hypothesis represents the actual order of divergences, *Ascaphus* occupies a critical phylogenetic position relative to anuran amphibians. Tailed frogs occur in and around high-gradient mountain streams and suffer from anthropogenic habitat alteration (e.g., Spear and Storfer, 2008, 2010). In spite of retention of several primitive features as adults (e.g., aspects of the palatoquadrate), they exhibit a number of derived features associated with their unusual habitat (e.g., oral suction discs in larvae and internal fertilization via an intromission organ).

Within *Ascaphus*, coastal and inland entities have been recognized as distinct species (*A. truei* and *A. montanus*, respectively) based on mitochondrial DNA (mtDNA; Nielson et al., 2001), morphology and allozymes (Nielson et al., 2006). Using coalescent analyses, Carstens et al. (2005) suggested that *A. montanus* has persisted in the inland region since the Pliocene and throughout Pleistocene climatic fluctuations. Further, Carstens and Richards (2007) suggested, based on projections of ecological niche models (ENMs) onto past climatic reconstructions, that *A. montanus* persisted in two to three refugia (a northern refugium in the Clearwater Drainage and possible southern refugia in the Salmon River Drainage and the Blue Mountains). The hypothesis of two refugia is supported by reciprocally monophyletic mtDNA clades that differ by as much as 2-3% uncorrected sequence divergence, and the divergence between these clades originated approximately 1 MYA (with a credibility interval of 0.092 - 5.4 MY; Nielson et al., 2006). The contact zone between northern and southern mtDNA clades is in the East Fork of the South Fork of the Salmon River (Figure 2.1, inset; Nielson et al., 2006). Thus, given this ancient phylogeograhic structure, long term occupation of the Inland Northwest, and the habitat specialization described above, current patterns of genetic variation and diversity are likely the result of complex interactions between both ancient and contemporary processes and between landscape and range-wide phenomena.

Here, we address the extent to which these different processes have affected the genetic structure and diversity of this ecologically specialized frog species. To do so, we address three hypothe-

ses regarding geographic patterns of genetic diversity in *A. montanus*. First, we hypothesize that refugial structure will be detectable at the range-wide spatial scale (e.g., geographic congruence between microsatellite and mtDNA data). Second, habitat variables should influence dispersal in such a way that a correlation between gene flow and environmental variation will be detectable across the range. Third, landscape variables may have a stronger effect on gene flow at local geographic scales, whereas historical factors (i.e., refugial structure) will be less important locally. We address these hypotheses with fine-scale, range-wide sampling using nuclear DNA microsatellite loci and high-resolution species distribution models (SDMs).

2.3 METHODS

2.3.1 *Sampling*

Two hundred twenty-eight streams were sampled for the presence of *Ascaphus* (Figure 2.1); this entailed kick-sampling for an hour at each site. If no tadpoles were found, frogs were considered to be absent at that site. Sampling typically occurred over a 40-meter stretch of a stream and we collected multiple age-classes at each site. Site selection was designed to cover the entire distribution of *A. montanus*, but the very small portion of the distribution in Canada (i.e., the Yaak River drainage) was not sampled. In addition, several paired headwater sites in different drainages were sampled, and one region in northern Idaho was sampled very intensively to provide for a future analysis of local-scale landscape genetics. All sampling was conducted under University of Idaho ACUC protocol 2007-14 and collecting permits from Idaho, Montana, Oregon, and Washington.

We extended the SDM for *Ascaphus* conducted by Carstens and Richards (2007) by refining the spatial resolution of 19 bioclimatic variables to 1 km2 (WORLDCLIM; Hijmans et al., 2005) and by adding one land-cover (i.e., coniferous cover) variable (USGS). We restricted our SDM to *A. montanus*, used presences previously known (11) in addition to those obtained from our kick-sampling of 228 sites (176 presences total), and 10000 background samples. Background samples included 126 locations where we searched but did not observe the species (i.e., true absences), and a set of 9874 background points that we been selected in the study area using geospatial manipulations. We used an approach similar to that used by Warren et al. (2014) to do this; we first created a continuous density surface that represented the sampling effort, with higher values

in better-sampled areas. We then sampled background points using the values of this surface to affect the probability of a point being drawn, such that a well-sampled region had a higher probability of contributing to the background set (Searcy and Shaffer, 2014).

The SDM was generated using the maximum entropy machine-learning algorithm implemented in Maxent (Phillips et al., 2006) randomly selecting 30% of the samples for cross-validation. In addition, in order to address the correlation among climatic variables, we conducted the SDM on a reduced set of uncorrelated variables (Peterson et al., 2011; Warren et al., 2014). The SDM calculated from the reduced data set were nearly identical to the full data set (see below).

As in Carstens and Richards (2007), we also projected a paleodistribution for *A. montanus* using 19 bioclimatic variables estimated for the last glacial maximum (LGM; 21 Kya) under the Community Climate System Model (CCM3; Collins et al., 2006). Further, to estimate future variation in the range of the species, we used climatic projections of the same 19 bioclimatic variables (WORLDCLIM; Hijmans et al., 2005). These projections correspond to predictions for the year 2070, using a Global Circulation Model (MRI-CGCM3), a worst-case CO_2-emission scenario (8.5 RCPs), and a 30-seconds resolution. Because strong range changes affect the spatial genetic structure of species, and can influence the survival of particular genetic groups in the absence of adaptation (Espíndola et al., 2012), we used the outputs from the current and forecasted projections to quantify the spatial and temporal variation of niche suitabilities for the total range as well as for each genetic cluster at K = 2 (see below). Using ArcGIS tools, we extracted all grid values within either the study area or in a polygon containing all sampled locations included in each genetic cluster for K=2. To improve the biological reality of this last analysis, and to implement some measure of dispersal potential, we added a 20km ("low") and a 100km ("high") buffer area to the polygons, and we also extracted all suitability values contained under these new polygons. We then calculated the proportion of grids harboring suitability values higher than 0.5 (>0.5 dataset) or 0.7 (>0.7 dataset) in the future, respective to current values. This was done in R, using custom scripts.

2.3.2 *Genetic Data Collection*

Genomic DNA was extracted from approximately 25 mg of tail tissue from 2098 individuals (primarily tadpoles) collected at 117 sites using standard protocols for the DNEasy Tissue kit (Qiagen).

Following the procedure of Spear and Storfer (2008) for *A. truei*, we performed three multiplex PCR panels on each sample using the Qiagen Multiplex PCR kit and included a negative control in each run, for a maximum of 13 microsatellite loci amplified for each sample. Products from each PCR panel were run on an ABI 3730 automated sequencer (Applied Biosystems, Inc.) at the University of Idaho College of Natural Resources and genotyped using GeneMapper 3.6 Software (Applied Biosystems, Inc.). To improve the accuracy of allele calls, alleles were binned using Flexibin (Amos et al., 2007) and, genotyping was repeated for 250-600 individuals for each locus. We used FreeNA (Chapuis and Estoup, 2007) to calculate the proportion of null alleles present for each microsatellite locus and each locus-population combination. We genotyped a range of 1-40 individuals from each site with an average of 18.4. Seventeen sites with < 10 individuals genotyped were excluded from the STRUCTURE and regression analyses described below.

2.3.3 *Genetic Diversity and Structure*

Standard descriptive statistics (expected and observed heterozygosity, allelic richness, F_{ST}) were calculated with GenePop 4.0.10 (Raymond and Rousset, 1995; Rousset, 2008). We also calculated F_{ST}, corrected for the presence of null alleles, using FreeNA (Chapuis and Estoup, 2007). We then evaluated population structure with no *a priori* constraints using the program STRUCTURE (Pritchard et al., 2000) under its admixture model. Because several populations exhibited extremely low allelic diversity (Figure 2.2A), we conducted separate analyses assuming co-dominance and recessive expression, but based most of our inferences from the analyses (correctly) assuming co-dominance. We varied K (the number of clusters of multilocus genotypes) from 1-15, and conducted ten replicate runs at each value of K. Each of the 150 analyses was run for one million generations prior to sampling the posterior distributions for another one million generations. We assessed the relative importance of each value of K using two methods. First, we used the ΔK criterion, which was calculated by dividing the mean difference in likelihood for successive values of K by the standard deviation of $\mathcal{L}(K)$ (Evanno et al., 2005). Second, we assessed qualitatively $P(D|K)$ and then compared that to the geographic distribution of the clusters for each value of K. This occurred very close to the value of K that maximized the average $P(D|K)$. This approach is particularly appropriate when there is hierarchical structure in the data

and we also used FreeNA (Chapuis and Estoup, 2007) to visualize this structure with a NJ-tree calculated from chord distances (Cavalli-Sforza and Edwards, 1967).

2.3.4 *Geography and Genetics*

We performed linear and multivariate regression analyses using R to examine correlations between geographic variables and genetic distance (linearized F_{ST}). We first performed linear regressions to evaluate the correlations between different distance metrics (see below) and genetic divergence. Inter-population geographic distances were obtained following different approaches. Results obtained from our SDMs were used to calculate least cost path (LCP) distances between populations. We then used circuit theory and the program CircuitScape (CS; McRae et al., 2008) to calculate environmental resistance between populations. To do so, we used the SDM as a univariate measure of resistance and calculated costs using four values: 1x (values the same as the SDM), 10X, 100X, and 1000X. Third, we used ArcMap 9.3 (ESRI, Redlands, CA) to obtain Euclidean geographic distance between sites. Fourth, we used ArcMap 9.3, the Pacific North West River Reach database (PNWRF3, StreamNet) and the ArcMap toolbox FLoWS v1 (Theobald et al., 2006) to calculate aquatic distance between sites, as the shortest distance between a pair of sites that involved only travel along water sources (rivers, streams, lakes). Where there was no water-based connection between population pairs, those populations were considered unconnected and N/A values were used for analyses (following Spear and Storfer, 2008). We then performed multiple regression analyses in R between allelic richness, edge distance, elevation, and latitude to accommodate associations among these variables. We also performed multiple regression analyses of the relationships between linearized F_{ST} and the distance metrics (geographic vs. LCP and geographic vs. CS distances).

In order to assess the effect of scale on landscape genetic inference, we ran each analysis on three data sets. To assess range-wide patterns, we used all the data (called Range-wide). To assess regional, intermediate-scale patterns, we analyzed northern and southern populations separately, with the criterion of congruent genetic breaks observed in mtDNA (Nielson et al., 2006) and our data at $K = 2$ (see below; called Regional). Finally, to assess local small-scale phenomena, we analyzed just the densely sampled populations from the North Fork of the Clearwater, St. Joe, St. Maries vicinities (called Local).

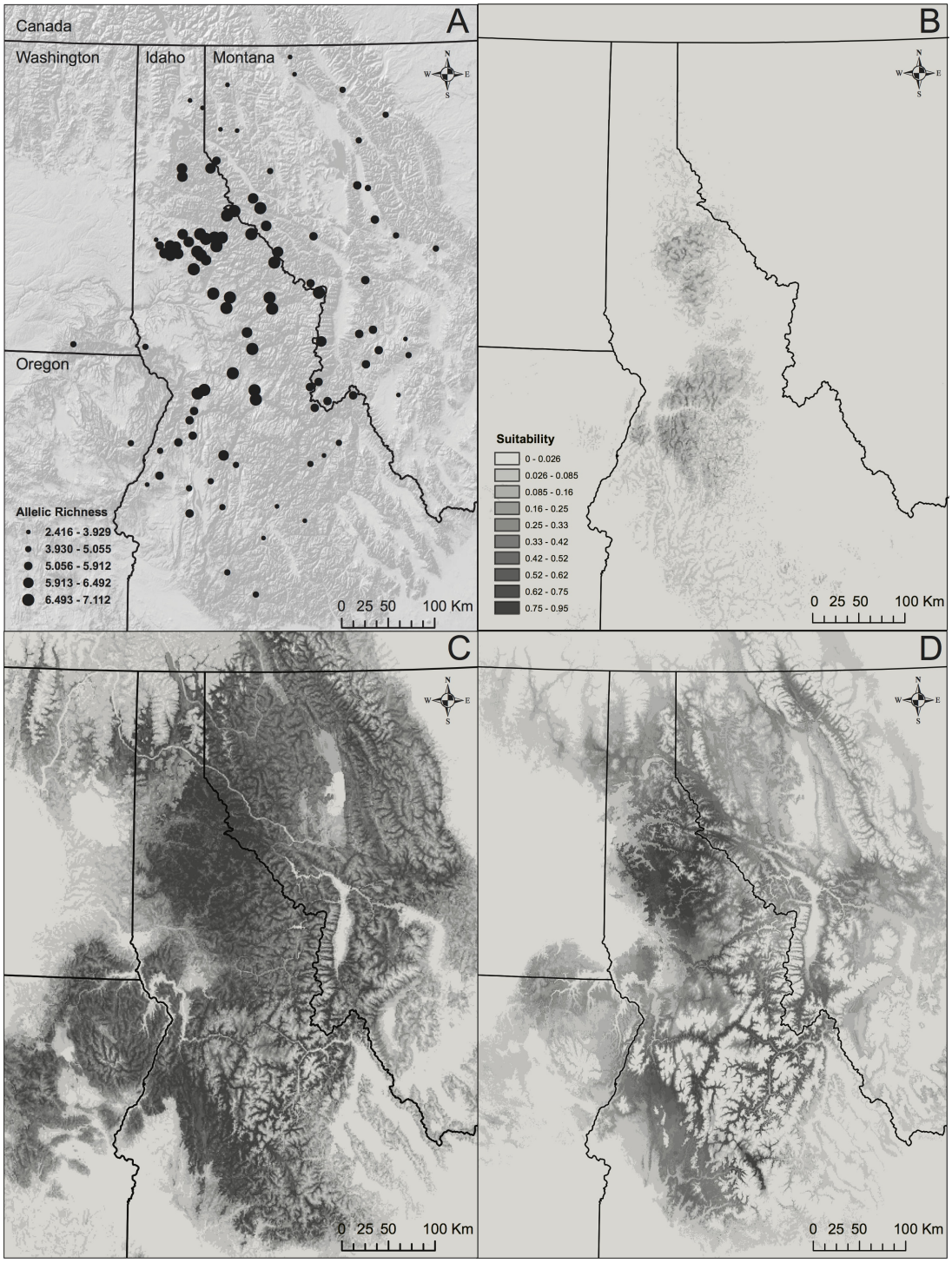

FIGURE 2.2: Distribution of suitable habitat for *A. montanus* derived from our SDM. (a) Suitabilities inferred by projecting our SDM onto climate reconstructions for the Last Glacial Maximum. (b) Distribution of suitabilities for current habitat. (c) Suitabilities inferred by projecting or SDM onto climate projections for 2070 under a worst-case scenario. (d) The pattern of allelic richness observed at our 99 sampling locations.

2.4 RESULTS

2.4.1 *Species Distribution Model*

The SDM approach provided a high AUC value (0.9668), indicating a very good model fit. The variables most strongly contributing to the model were coniferous cover (a preference for higher covers; 30.41%), flow accumulation (a preference for headwaters; 17.28%) and maximum temperature during the warmest month (12.98%). The SDMs estimated from the reduced data (seven uncorrelated climatic variables) are nearly identical (Figure A.1; correlation coefficients of suitabilities scores per pixel calculated from full vs. reduced data is 0.92, 0.96, and 0.80 for current, future and past SDMs); therefore we focus on the SDM from the full data. Our SDM for current conditions (Figure 2.2C) is similar to the *A. montanus* portion of the projections produced by Carstens and Richards (2007), but has a higher resolution. It broadly captures the species range of *A. montanus*, however it predicts areas of suitable habitat in some locations where *A. montanus* is known to be absent, notably in central Oregon. On a finer scale, poor habitat is often associated with deep canyons, such as the Hells and Salmon River Canyons, and the high resolution of our SDM renders those canyons visible in our projections (Fig 2.2C).

Our paleodistribution reconstruction based on climatic variables (Figure 2.2B) suggests that two areas of high climatic suitability (i.e., primary refugia) existed for *A. montanus* during the LGM, both contained within the current range of the species. However, the location of one of the refugial areas differed from that inferred by Carstens and Richards (2007). We infer the southern refugium to have been higher in the Salmon River drainage than was inferred by Carstens and Richards (2007), which was estimated to be outside the current distribution.

Under current CO_2 emission rates, our model predicts that there will be a severe reduction in suitable habitat for *A. montanus* (Figure 2.2D). This is particularly true for the southern portion of the current distribution; tailed frog populations in this region will likely be extremely fragmented, and this inference is robust to different dispersal rates and suitability thresholds (Figure A.2). Our predictions are not simply that suitable habitat will be shifted northward; indeed there is very little predicted latitudinal shift. Instead, suitable habitat is predicted to become restricted within the current species range, with a (perhaps) counterintuitive concentration to lower elevations (Figure 2.2D; Figure A.2).

2.4.2 *Genetic Diversity and Structure*

Multilocus genotypes for 2098 individuals are available on Dryad. The number of alleles per locus ranged from 1-15 across all samples. An edge effect was apparent in both allelic richness (Figure 2.2A) and, to a lesser extent, heterozygosity (not shown); populations from the center of the distribution had higher diversity than populations from the edges of the range, whereas many edge populations exhibited very low allelic richness (r^2 = 0.75; Table 2.2, Figure A.2). Overall levels of allelic richness and heterozygosity were moderate-to-high but lower than those observed in studies of *A. truei* (Spear and Storfer, 2008, 2010). Pairwise estimates of F_{ST} suggest low to moderate genetic differentiation between sites (deposited Dryad).

In the STRUCTURE analyses, likelihood scores, $P(D|K)$, increased with increasing K until K=10, with further increases leading to diminishing improvements (Figure 2.3A), and a plateau at higher values. The most dramatic increase in $P(D|K)$ was associated with increasing from a single cluster to K=2 (Figure 2.3A), while $P(D|K)$ continued to increase until K=10. At K=2, individuals were clustered into northern versus southern populations (red vs. yellow in Figure 2.3B,C) in a manner very similar to the north/south phylogeographic split in mtDNA haplotypes (Figure 2.1, inset; Nielson et al., 2006). This is particularly true for analyses using a recessive inheritance model (Figure 2.3C). Analyses run assuming co-dominant inheritance and K = 2 (Figure 2.3B) tended to cluster five northern samples with the southern cluster; we view this as spurious and the north-south clustering (Figure 2.3C) is more plausible (see Discussion).

At K=3 (Figure A.4A), the individuals from the northern cluster were split into northeastern (purple) and northwestern (yellow), at K=4 (Figure A.4B), individuals from the purple northeastern cluster were further split into a north-northeastern cluster (light blue) and an east-central cluster (purple). At K=5 (Figure A.4C), the red southern cluster split into southwestern (green) and southeastern clusters (red), and at K=6 (Figure 2.3D), the light blue north-northeastern cluster was split into clusters found in the Cabinet and Purcell Mountains (orange) versus those collected east of the Kootenay River (light blue). At K=7 (Figure A.4D), individuals from the yellow north-northwestern group were split into clusters from the Palouse Range (dark blue) versus Clearwater/Bitterroot Ranges (yellow). Increasing K further simply resulted in splitting individuals into peripheral populations (not shown).

Table 2.2: Correlations (above diagonal) and partial correlations (below diagonal) between geographic variables.

	Allelic Richness	**Latitude**	**Elevation**	**Edge Distance**
Allelic Richness	0	0.0946	-0.4302	0.7509
Latitude	-0.0936	0	-0.5804	0.0779
Elevation	-0.2254	-0.6044	0	-0.4014
Edge Distance	0.6855	-0.0866	-0.1567	0

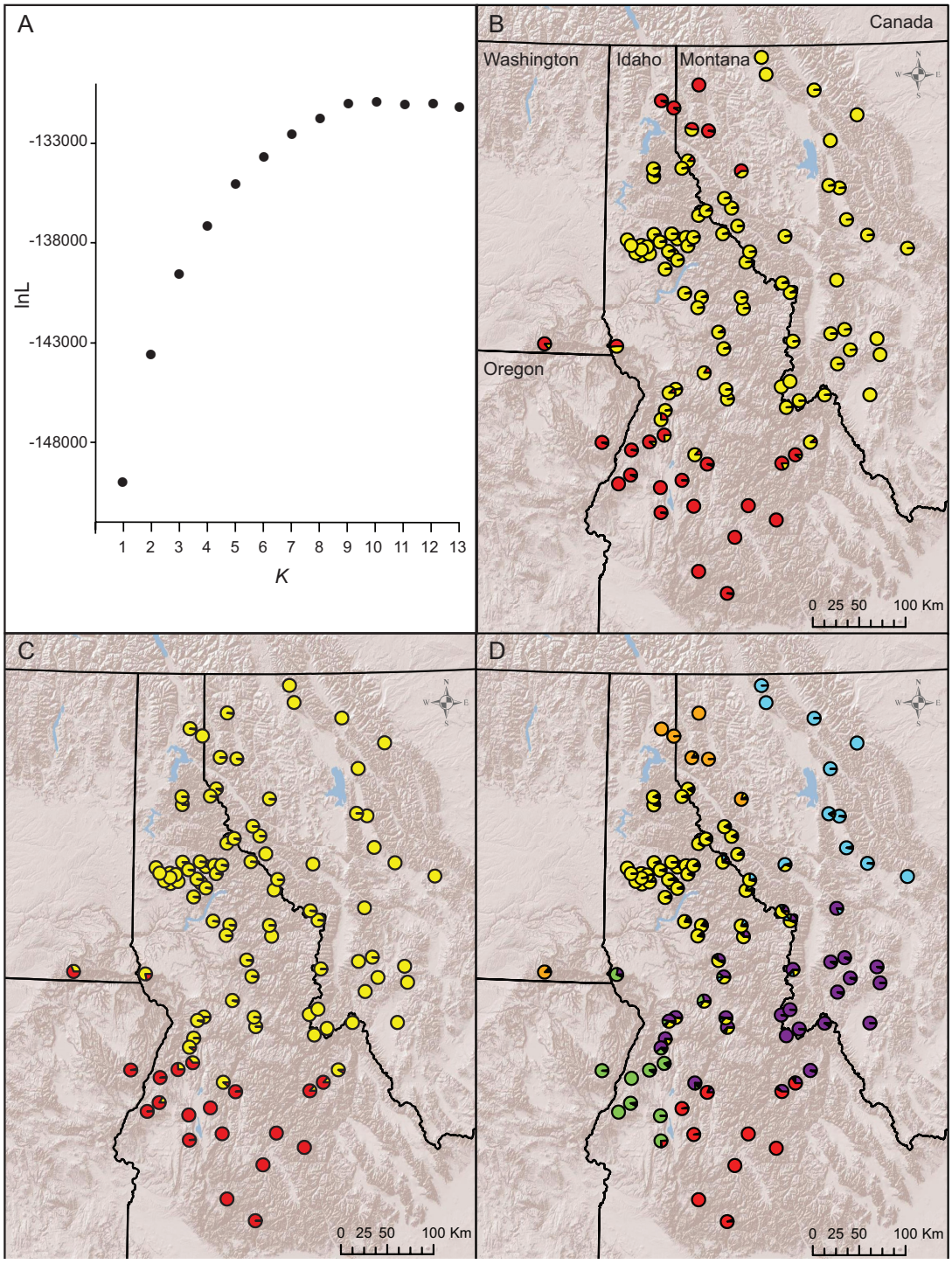

FIGURE 2.3: (a) Average log likelihood for each value of K used for STRUCTURE. (b) STRUCTURE results for K=2 based on microsatellite data under a co-dominant mode of inheritance, or (c) assuming recessive inheritance, both of which show Northern (yellow) and Southern (red) clusters. (d) STRUCTURE results for K=6 based on microsatellite data assuming codominant inheritance.

2.4.3 *Geography and Genetics*

At the range-wide scale, Euclidean distance between populations (i.e., Isolation-by-Distance; IBD) explained much of the variance in pairwise differentiation between populations (as measured by linearized F_{ST} ; Figure 2.4A). Addition of the SDM information to the model through LCP and CS distances (visualized in Figure A.5) increased its explanatory power slightly relative to Euclidean distance (Figure 2.4B,C). Additionally, for both LCP and CS distances, the distance measures obtained using the lower costs (1X or 10X) explained more of the variation than those obtained using higher (100X or 1000X) costs (data not shown). Aquatic distance explained very little of the variation in genetic distances between populations (Figure 2.4D). Furthermore, our multivariate analysis of environmental variables showed a positive correlation between distance from the nearest range edge and allelic richness, as well as a negative correlation between elevation and latitude. Multivariate analysis of the range-wide distance metrics illustrated that that CS and LCP distances explain an increasing portion of variation in linearized F_{ST} values after the effect of geographic distance are removed (r^2 = 0.42; Table 2.3).

In analyses at the intermediate scale, the SDM-based distances explained more of the variation in linearized F_{ST} than did simple geographic distance for the northern populations (r^2 = 0.16 for the IBD, and r^2 = 0.28 for CS distance; Figure 2.5A). This was not the case for the analyses of the southern populations, where simple geographic distance and CS distance have similar r^2 values (0.24 for geographic distance vs. 0.26 for the CS distance). Thus, at least for the northern populations, patterns of genetic variation are more strongly associated with climatic and environmental variables than at the range-wide scale, and this is particularly true once the effect of geographic distance in removed; partial correlation coefficients for CS-distance are 0.422 and 0.231 for northern and southern populations respectively.

This effect is seen more strongly at the local scale for the densely sampled populations in the Clearwater Drainage (the most densely-clustered yellow populations in Figure 2.3D). At this local scale, geographic distances explain little variation (r^2 = 0.115) whereas the CS distance explains substantially more (r^2 = 0.32; Figure 2.5C). In this analysis, a single population appears to be an outlier with relatively higher F_{ST} values than comparisons involving the other populations in the local sample (probably due to its extremely low allelic diversity). The correlation between

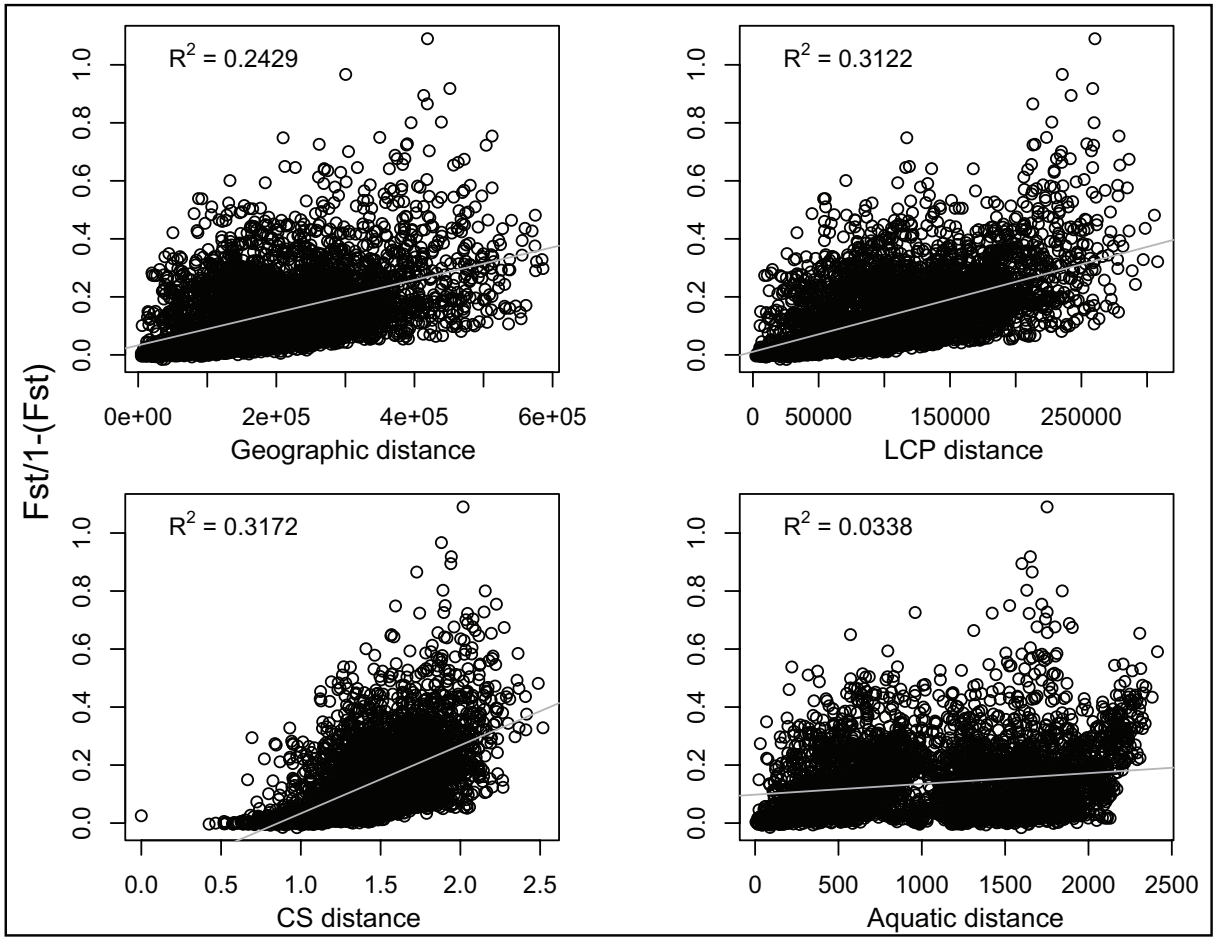

FIGURE 2.4: Range-wide linear regression of linearized F_{ST} and (a) Euclidean geographic distance calculated in ArcGIS, (b) Least Cost Path (LCP) distance calculated from the SDM at cost of 1X, (c) Lowest resistance (CS) distance calculated from the SDM using CircuitScape at cost of 1X, and (d) Aquatic distance calculated in ArcGIS. All correlations are significant ($p < 0.001$).

TABLE 2.3: Multiple regression of geographic with CS (top) and LCP (bottom) distances.

All pops	**F_{ST}**	**Euclidean Distance**	**CS1X Distance**
F_{ST}	0	0.2429	0.3178
Euclidean	0.0016	0	0.7530
CS1X	0.3479	0.8130	0
North	**F_{ST}**	**Euclidean Distance**	**CS1X Distance**
F_{ST}	0	0.1604	0.2789
Euclidean	-0.1226	0	0.7395
CS1X	0.4218	0.8256	0
South	**F_{ST}**	**Euclidean Distance**	**CS1X Distance**
F_{ST}	0	0.2370	0.2641
Euclidean	0.1131	0	0.7149
CS1X	0.2312	0.7920	0

All pops	**F_{ST}**	**Euclidean Distance**	**CS1X Distance**
F_{ST}	0	0.2429	0.3122
Euclidean	-0.0960	0	0.8714
CS1X	0.3057	0.8881	0
North	**F_{ST}**	**Euclidean Distance**	**CS1X Distance**
F_{ST}	0	0.1604	0.2660
Euclidean	-0.2270	0	0.8577
CS1X	0.4276	0.9124	0
South	**F_{ST}**	**Euclidean Distance**	**CS1X Distance**
F_{ST}	0	0.2370	0.2738
Euclidean	0.0400	0	0.8191
CS1X	0.2264	0.8723	0

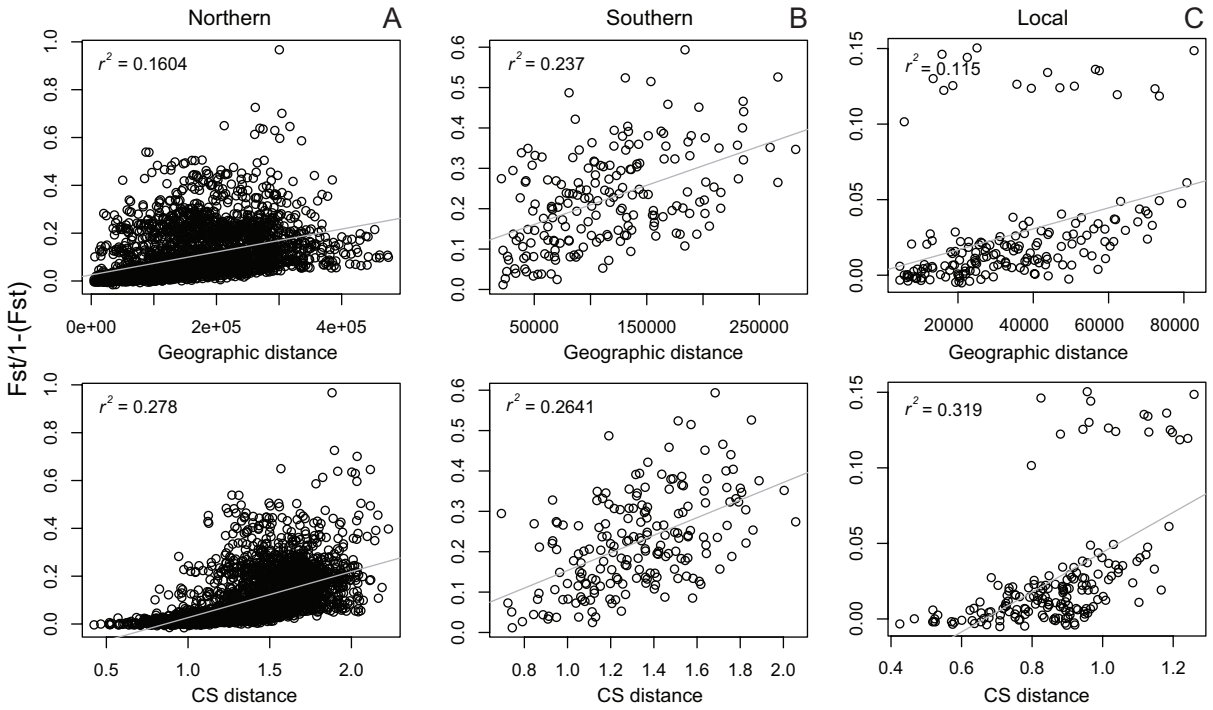

FIGURE 2.5: Linear regressions of linearized F_{ST} and geographic or CS distance for (a) northern, (b) southern, and (c) local samples.

geographic distance and differentiation is much higher at the local level once this differentiated population is removed (r^2 = 0.487). Thus, analyses at this local scale reveal patterns that are not detectable in analyses at larger spatial scales, especially identification of populations that deviate from Isolation-by-distance.

2.5 DISCUSSION

Genetic structure and patterns of gene flow are often characterized as being either governed by an IBD model or by being mediated by habitat features (i.e., Isolation-by-Environment, or IBE; e.g., Wang and Summers, 2010; Sexton et al., 2014). However, the geographic structure of genetic variations across a species' range is a multi-scale phenomenon, both temporally and spatially. Thus, understanding the geographic structure of genetic variation requires the inference of interactions between historical and contemporary processes and analyses at multiple spatial scales.

2.5.1 *Genetic Diversity and Structure: Broadest Scale*

At the broadest scale, allelic richness and heterozygosity are highest in populations at the center of the species' range and decline in populations sampled closer to the range edge, a pattern expected under equilibrium conditions. We hypothesized that our combination of molecular and SDM approaches would allow us to successfully identify the location of ancient refugia for the species. Accordingly, *A. montanus* exhibits a hierarchical pattern of geographic structure (Figure 2.3). As indicated above, at the lowest level (K=2), the genetic clusters correspond almost precisely (Figure 2.3C) with the deep phylogeographic split in mtDNA identified by Nielson et al. (2006), and this partitioning of populations is also the deepest split in our tree-based analysis of our microsatellite data (Figure A.4). With K=6-9, clusters of multilocus genotypes roughly correlate with mountain ranges. In addition, as we increased the number of clusters in the STRUCTURE analyses from 1 to 9, new clusters of populations were split from clusters identified by a previous run (i.e., a run with one fewer cluster).

The northern-most populations sampled occur at localities that were glaciated during the LGM; these populations therefore are necessarily the result of expansion from the Clearwater refugium. Counter-intuitively, however, the route of expansion appears to not have been directly

northward from the adjacent Clearwater Drainage, but based on the NJ tree (Figure A.6), appears to have occurred in a counter-clockwise fashion around the Missoula Basin (which is defined by the extent of pro-glacial Lake Missoula, and from which there are no current records of *A. montanus*). Thus, these populations in the Cabinet and Purcell ranges have become established by relatively long-range dispersal after glacial retreat. This explains the dramatic reduction in allelic diversity exhibited by these populations (Figure 2.2A), which in turn may be causing them to cluster spuriously with the southern populations in STRUCTURE analyses assuming co-dominant inheritance (Figure 2.3B). Populations with low allelic richness due (as we hypothesize) to the serial range expansions we hypothesize will have a stochastically reduced pool of genotypes. Our intuition suggests that applying a recessive alleles model (admittedly incorrectly) may reduce the sampling error in estimating coancestry, and the average coancestries of the Purcell and Cabinet mountains populations estimated this way (Figure 2.2C) make far more geographic sense than those estimated with the (appropriate) co-dominance model.

As a result of our extensive sampling within the species range, we were able to detect associations of small effect between genetic distance and geographic variables. We had, however, hypothesized that habitat variables play a role in tail frog dispersal, and that we would be able to identify correlations between habitat variables and genetic diversity. We chose therefore to focus on the magnitude of the association between variables. Our regressions and multivariate analysis indicate that inclusion of SDM-based distance measures (CS and LCP distances) into models explains more variation in the genetic distances between populations than models with geographic distance alone at the broadest spatial scale. The additional variation explained by inclusion of IBE is modest, but persists when the effect of geographic distance is removed in our partial regressions. Additionally, for both CS and LCP distances, the maximum proportion of variation in genetic distance between populations was explained when both CS and LCP costs were lowest for the environmental variables (1X). Together, these findings all support the conclusion that environmental variables included in the SDM affect gene flow in *Ascaphus*, but that geographic distance between sites is perhaps the most influential determinant of differentiation between populations. One important caveat to our analyses is that we accounted for multiple environmental variables by using a single metric, the SDM, to calculate resistance for our analyses. This prevents us from being able to determine which environmental variables may

be most important in mediating gene flow and may disguise the effect of environment if variables have opposing influences.

As was seen by Spear and Storfer (2008) on a more local scale, riparian distances seem to be wholly unrelated to population differentiation (Figure 2.4D); although the regression is significant (due to our very large sample size), the effect size is negligible ($r^2 < 0.034$). This was true across at all three scales examined, and indicates that dispersal and gene flow are concentrated to post-larval life stages; dispersal by tadpoles would be expected to lead to an association between riparian and genetic distances. More specifically, given that mark-recapture data suggest strong philopatry of breeding adults (Daugherty and Sheldon, 1982), dispersal in tailed frogs appears to be restricted to post-metamorphic sub-adults and to be occurring over land rather than along water-based corridors (i.e., saddle hopping). Although there is no direct evidence for this, it is more plausible to us that dispersal of small sub-adults has gone undetected than that dispersal of large adult frogs has.

2.5.2 *Accounting for Deep History: Regional Scale*

Our third hypothesis predicted that landscape variables may strongly influence gene flow at local scales, and are less important at the total range level. Thus, while geographic distance explains a large portion of the variation in genetic distance between populations at broad spatial scales (i.e., in the range-wide analyses), environmental factors may be more important at smaller spatial scales. To assess this for *Ascaphus*, we conducted separate analyses on northern and southern groups, as defined by the deep phylogeographic divergences in the mitochondrial data (Nielson et al., 2001, 2006) and the congruent results of the STRUCTURE analyses of the current microsatellite data with K=2 (Figure 2.3C). The results of our analyses at this intermediate spatial scale, with the deepest historical divergence factored out, are largely consistent with those for the species range as a whole; geographic distance still explains much of the variation in genetic distance between populations. However the SDM-based distance metrics explain more variation in genetic divergence than they do in the range-wide analysis (Figure 2.5A,B). This is particularly true in the northern cluster, and less so in the southern cluster, perhaps due to the more patchy habitat in the southern portion of the range (Figure 2.2B). This suggests that that environmental

variables are likely to have an increasing effect on gene flow as spatial scale decreases in this species.

2.5.3 *Landscape Genetics at Local Scale*

Further confirming our third hypothesis, even more variation in F_{ST} is explained in the analyses at the local scale (Figure 2.5C). First, there is a marked increase in the r^2 value when the SDM-based distance is included, compared to simple geographic distance, suggesting that deviation from IBE is more easily detected at the smallest scale. Second, the impact of a single aberrant population is detectable only at this local scale; all the highest F_{ST} values at this scale involve pairwise comparisons of a single population collected at locality 85 (Latah Creek) with the rest in the cluster. There are several unique features of this population; it is the only locality we examined from a drainage that flows directly into the Spokane River, it is separated from the rest of the nearby populations in the cluster by State Highway 6, and it exhibits extremely low genetic diversity (Figure 2.2D). Application of our SDM-based (CS) distance shifts the comparisons involving this population on the x-axis to the right compared to strictly geographic distance (Fig 2.5D), suggesting that the low allelic diversity in this population (and consequently high F_{ST}s) may be mediated by ecological differences.

2.5.4 *Modern and Paleo-Distribution Models of A. montanus*

Our SDM for *A. montanus* (Figures 2.2C, A.1B) is broadly consistent with that constructed by Carstens and Richards (2007), although we find a wider range of suitable habitat, especially within the Clearwater River drainage area, and portions of northwestern Montana. This difference is likely the result of our expanded sampling of *A. montanus*, and thus more complete presence/absence data used to construct the SDM.

The hindcast of our SDM using paleoclimatic reconstructions (Figures 2.2B, A.1A) provides strong support that *A. montanus* has had a very long history in the Inland Northwest; further, these projections support the persistence of two distinct refugia (one in the Clearwater Drainage and one in the Salmon River Drainage) during Pleistocene glacial maxima. These refugia are consistent both with previous SDM-based reconstruction (Carstens and Richards, 2007) and

the deep divergence the mtDNA sequence data identified by (Nielson et al., 2006, Figure 2.1, inset). Furthermore, the signature of these dual refugia is seen in our microsatellite data, both by the results of our STRUCTURE analysis at the level of K=2 (Figure 2.3C) and by our NJ tree (Figure A.6). In contrast to the analysis by Carstens and Richards (2007), we do not find evidence for refugial habitat present in the Blue Mountains of Oregon, perhaps due to a combination of differences in the spatial scale of the environmental variables used in construction of our SDM (which allowed for a finer scale analysis than was possible for Carstens and Richards 2007), and the updated presence/absence data we used to construct our SDM that arose from our extended sampling efforts. Additionally, whereas Carstens and Richards (2007) found evidence for this third refugium in their paleoclimate model, DNA evidence does not support the hypothesis of three refugia; thus the two-refugia model is consistently supported by available evidence.

Finally, projections into predictions of future climatic conditions (Figure 2.2D) indicate that the range of the species is expected to reduce drastically and, more importantly, become highly fragmented. Since environmental variables have been shown to affect species connectivity at local scales, this is expected to affect the genetic diversity of the species through reduced gene flow. Although the region with the highest allelic richness (Fig 2.2A) is expected to harbor high ecological suitability (Fig 2.2C), the southern range of the species will become extremely fragmented and reduced. This will likely drastically affect the survival and persistence of the southern genetic clusters (Figure 2.3B-D).

2.6 CONCLUDING REMARKS

The structure of genetic diversity within *A. montanus* is the result of processes acting across multiple geographic and temporal scales (Table 2.1). Here, we show a strong geographic structure across the species range, with the deep phylogeographic split between northern and southern range portions detected in mtDNA (Nielson et al., 2006) representing the deepest divergence in microsatellite loci as well. This is likely the result of the species restriction to two Pleistocene refugia, the expansion from which has produced a contact zone in the South Fork of the Salmon River detectable in both mtDNA and our data. Analyses at decreasing spatial scales indicate that the smaller the scale, the more additional variance in F_{ST} is explained by our SDM-based

CircuitScape and Least Cost Path distances, and only at the smallest scale are outlier populations identified. This suggests that meta-analyses that attempt to identify whether Isolation by Distance or Isolation by Environment is more common (e.g., Sexton et al., 2014) should include multiple geographic spatial scales to account for scale-specific phenomena.

CHAPTER 3

Persistence of Antibiotic Resistance Plasmids in Bacterial Biofilms

3.1 SUMMARY

The emergence and spread of antibiotic resistance is a crisis in health care today. Antibiotic resistance is often transferred to susceptible bacteria by means of horizontal gene transfer mediated by plasmids. In clinical settings bacterial pathogens typically grow as biofilms, and the evolution of plasmids in biofilms is poorly understood. We experimentally compared the evolution of plasmid persistence in the clinical pathogen *Acinetobacter baumannii* when grown in biofilms versus well-mixed liquid cultures. Plasmids persisted in *A. baumannii* in both environments even when antibiotics were not present. By the end of the experiment, plasmid persistence increased more in liquid cultures, but variation in the degree of persistence was greater among biofilm derived clones. The results of this study show for the first time that the persistence of MDR plasmids improves in biofilms, and furthermore that the evolutionary potential of plasmids is undiminished in biofilms.

3.2 INTRODUCTION

The emergence and spread of antibiotic resistant bacteria is a crisis faced by healthcare today, and the factors influencing these processes are poorly understood (Centers for Disease Control and Prevention, 2013). Resistance to antibiotics can be obtained either through mutations in key genes or by the acquisition of resistance genes via horizontal gene transfer (Benveniste and Davies, 1973), which is often mediated by transmissible plasmids (Mazel and Davies, 1999; Mathers et al., 2011). Broad-host-range plasmids that can be transferred to a wide range of bacterial species (Frost et al., 2005; De Gelder et al., 2007) are the most clinically threatening plasmids because they potentially carry genes that confer resistance to multiple drugs (MDR plasmids). Consequently, a single plasmid transfer event can turn a drug sensitive bacterium into a multiply drug resistant strain. A current alarming example is the rapid worldwide plasmid-mediated

spread of antibiotic resistance to colistin, an antibiotic of last resort (Liu et al., 2016; Poirel et al., 2016).

Research suggests MDR plasmids are often not stably maintained in novel hosts (De Gelder et al., 2007). Instability is overcome by exposure to antibiotics, which imposes strong natural selection for mutations in the plasmid, host or both, which will rapidly improve plasmid persistence (Bouma and Lenski, 1988; Heuer et al., 2007; De Gelder et al., 2008; Sota et al., 2010; San Millan et al., 2014; Harrison et al., 2015; Loftie-Eaton et al., 2016). We use the term 'plasmid persistence' but this trait of plasmid-host pairs is also referred to as plasmid maintenance, plasmid stability, or plasmid retention. In order to prevent (or at least hinder) the spread the of MDR plasmids, we need an understanding of how well plasmids persist in pathogenic bacteria in the absence of antibiotic-derived selection.

The evolution of MDR plasmid persistence has been studied in well-mixed liquid batch cultures (ex. Bouma and Lenski, 1988; Heuer et al., 2007; De Gelder et al., 2008; Sota et al., 2010; San Millan et al., 2014; Harrison et al., 2015; Loftie-Eaton et al., 2016), but such culture conditions are not typical within a clinical setting where bacteria more naturally occur in biofilms. Until this study, there has been a critical gap in our knowledge of plasmid evolution in biofilms. It is logical to assume to plasmid evolution in biofilms will be different because, in general, bacterial evolution exhibits notably different patterns biofilms (Rainey and Travisano, 1998; Lewis and Lewis, 2001; Donlan, 2002; Boles et al., 2004; Ponciano et al., 2009). Biofilms are fundamentally different from well-mixed liquid batch cultures because cells are fixed in one spatial location. The environment a cell experiences (e.g. nutrient availability or antibiotic exposure) is therefore locally variable, which leads to the possibility of spatially heterogeneous natural selection and ecology. Thus, one cell might experience particularly strong selection for antibiotic resistance – while another might experience relatively weak selection for antibiotic resistance – simply owing to its location within the biofilm.

To better understand how biofilm growth affects the evolution of plasmid persistence we performed an experimental evolution study using the Gram-negative pathogen *Acinetobacter baumannii*. Infections caused by *A. baumannii* are an emerging healthcare threat because the organism readily becomes resistant to multiple antibiotics and even pan-drug resistant strains have been reported (Villers et al., 1998; Hsueh et al., 2002; Perez et al., 2007). As a result this

species has emerged as an important cause of nosocomial infections. Several studies document longer hospital stays and more severe outcomes for patients with A. baumannii infections than with many other bacterial pathogens (Jerassy et al., 2006; Sunenshine et al., 2007). Specifically, we compare the evolution of plasmid persistence in *A. baumannii* when grown in biofilms and well-mixed liquid batch cultures by comparing the persistence of plasmids in evolved clones derived using these two culture conditions. We expected that clones that stably maintain an MDR plasmid would emerge under both conditions. More specifically, by quantitatively analyzing plasmid persistence in individual clones from each population we were able to test two hypotheses. First, weakened selective pressures experienced by bacteria in biofilms will lead to lower average levels of plasmid persistence. Second, evolutionary processes will generate and maintain a broader diversity of plasmid persistence phenotypes in biofilms than in liquid cultures, due to the inherent spatial structure of biofilms (Donlan, 2002; Boles et al., 2004). Our findings support both of these hypotheses.

3.3 METHODS

3.3.1 *Bacteria and Plasmid*

The experimental evolution of plasmid persistence in well-mixed liquid cultures and biofilms was done using *Acinetobacter baumannii* ATCC 17978, which was obtained from the American Type Culture Collection (Rockville, MD). ATCC 17978 is sensitive to tetracycline so plasmid bearing cells could readily be obtained by plating cells on tetracycline containing media. From here on we refer to this specific strain simply as *A. baumannii*.

For this study we used the well characterized IncP-1 plasmid pB10 (Schlüter et al., 2003) because it has a broad host-range and encodes resistance to four antibiotics (tetracycline, streptomycin, amoxicillin and sulfonamide). The ancestral strain used in all the experiments described here was constructed by electroporation of pB10 into *A. baumannii* (see APPENDIX B).

3.3.2 *Culture Media and Conditions*

For the evolution experiments *A. baumannii* (pB10) was grown in mineral basal medium (MBM) of M9 salts (Sambrook and Russell, 2001) and water supplemented with 18.5 mM succinate, 2 g/L

casamino acids, and 10 μg/mL tetracycline (tet), and trace element and mineral mixtures (Wolin et al., 1963) which is hereafter referred to as MBMS-tet. Plasmid persistence assays were done in the same MBMS medium without antibiotics. Dilution plating to obtain individual clones was done using Luria Bertani (LB) agar supplemented with 10 μg/mL tet (LB-tet). All experiments were done at 37°C.

3.3.3 *Experimental Evolution Protocol*

A. baumannii (pB10) was evolved in parallel in both biofilms and liquid batch cultures. Figure 3.1 is a timeline of the evolution experiments. To initiate the experiment, aliquots of an overnight culture grown in MBMS-tet were used to inoculate nine biofilm flow cells that contained 13 mL of the same medium. These were incubated for 24 h as batch cultures before the flow of fresh medium (5.2 mL/h) was initiated. After four days of growth, three flow cells were harvested; we designate these samples as our initial population state (t_0). For harvesting, the flow cells were moved to a biosafety cabinet where the seal between the lid and body of the flow cells were broken using sterile scalpel blades and the lids were put aside. Supernatant media were removed from the exposed flow cells using a pipette. Afterwards, the biofilms were re-suspended by adding 1 mL of 0.85% saline solution and subsequently transferred to 2 mL microcentrifuge tubes in order to disperse cells using a vortex mixer. Microscopic examination verified that the cells had dispersed and only free floating cells and small clumps of cells were present. The cell suspensions were serially diluted and plated on LB-tet agar. After overnight incubation, six clones were randomly chosen from each plate and archived at -70°C. This entire biofilm harvesting procedure was repeated on days 14 (t_{14}) and 28 (t_{28}). Because t_{14} revealed little additional information, we only report on time points t_0 and t_{28} for the sake of brevity.

Liquid batch cultures of *A. baumannii* (pB10) were started by inoculating three replicate test tubes containing five ml of MBMS-tet with 4.9 μL of the cell suspensions from the three t_0 biofilms. Liquid cultures were continuously mixed on rotary shakers at 200 RPM, and 4.9 μL of each culture was transferred daily to 5 mL of fresh media, a procedure which yields approximately 10 generations of growth per day. As with the biofilm culture, aliquots of these cultures were taken on days 14 and 28, serially diluted, and plated on LB-tet. Likewise, six clones from each plate were selected after overnight growth and archived at -70°C.

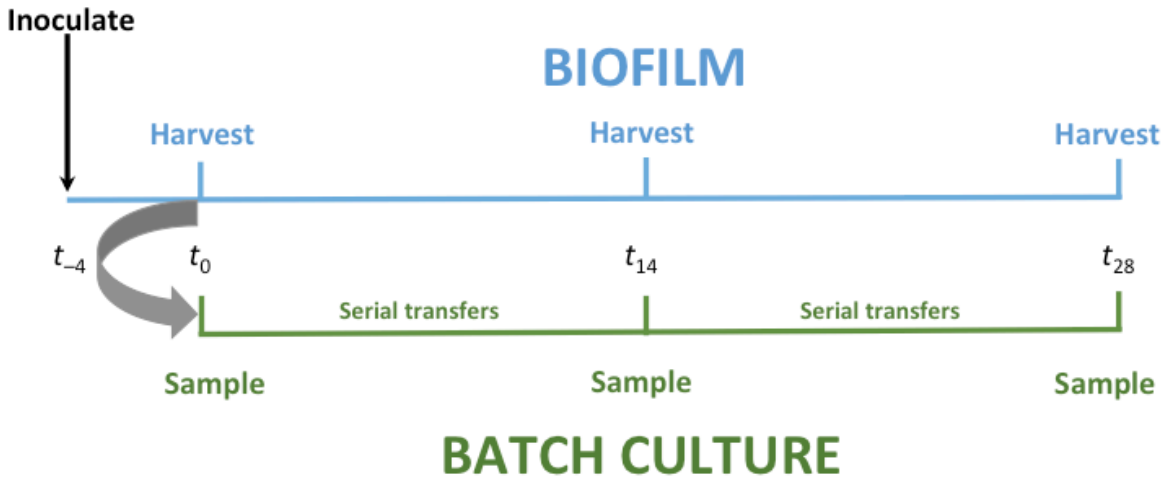

FIGURE 3.1: Schematic of experimental evolution experiments in liquid batch cultures (green) and biofilm cultures (blue). Biofilm flow cells were inoculated using an overnight grown culture of the ancestor (t_{-4}). Four days later (t_0), three randomly selected flow cells were harvested. The resulting cell suspensions were used to inoculate triplicate liquid batch cultures. They were also diluted and plated on LBA-tet to isolate plasmid-bearing clones. The remainder of the cell suspensions were archived at -70°C. The liquid batch cultures were transferred to new media daily for 28 days. On days 14 (t_{14}) and 28 (t_{28}), triplicate biofilm cultures were harvested, and samples were taken from the liquid batch cultures.

3.3.4 *Plasmid Persistence Assays*

The persistence of plasmids in clonal populations was quantified and compared as follows. Triplicate tubes containing 5 mL of MBMS (without antibiotic) were inoculated with 4.9 μL of each archived clone (representing day 0 of the plasmid persistence assays), and grown overnight. For the next 8 days these cultures were serially transferred to fresh media (4.9 μL into 5 mL). On days 0, 5, and 8, a 200 μL sample of each culture was removed, centrifuged, and stored at -20°C. Total DNA was isolated from each culture using a QIAsymphony DSP DNA Mini Kit on a QIAsymphony SP platform (QIAGEN, Inc.). Extraction yields were measured fluorometrically using a PicoGreen dsDNA kit.

The fraction of plasmid-bearing cells in each culture was estimated via quantitative PCR (qPCR; Loftie-Eaton et al., 2014) of the plasmid encoded trfA gene (encoding the replication initiation protein) and the chromosomally encoded 16S rRNA genes. Plasmid pB10 has a low copy number (2 per cell, data not shown), and the number of 16S rRNA gene copies in *A. baumannii* is 5 (Center for Microbial Systems, 2016; Maslunka et al., 2006). The ratio of these two genes was therefore used as a proxy for the fraction of plasmid-bearing cells in the populations, and was expressed as a plasmid:chromosome ratio. These qPCR assays were done in triplicate using a StepOnePlus real-time PCR system and a Power SYBR™ Green PCR master mix (Applied Biosystems Inc.) and by following the manufacturer's instructions. Details of the protocol are described in Appendix B.

Analysis of the qPCR data occurred in two stages. The first stage encompassed analysis of the raw qPCR data (i.e. raw fluorescence values for a given PCR cycle in a reaction) for the three qPCR replicates run for each plasmid persistence assay. Given that three replicate plasmid persistence assays were done per clone, 9 qPCR reactions were done for days 5 and 8 of these assays. Because all samples on day 0 came from the same archived glycerol stock and should be homogeneous, there were only three qPCR replicates for this day. Statistical analysis of the raw data is described in detail in the Technical Appendix. In a second stage, the qPCR-based estimates of the plasmid:chromosome ratio were used as a measure of the fraction of plasmid-bearing cells. Tracking this ratio over time allowed us to examine differences in plasmid persistence in our 2

different culturing conditions. The log-linear model we used to estimate the rate of plasmid loss over time is described in Appendix b.

Due to variance in our estimates of plasmid:chromosome ratios from the first stage of the analysis, bootstrapping techniques were used to provide more robust confidence intervals in subsequent analyses. In brief, values for ratios were drawn from normal distributions based on the ratios' estimated means and variances. These randomized values then served as the dependent variable for our log-linear model of plasmid persistence. Additionally, we utilized a Brown-Forsythe-Levene procedure on the residuals of the log-linear model in order to examine the level phenotypic diversity in biofilm versus liquid cultures (Brown and Forsythe, 1974). Such tests are based on group medians to test for linear trends in variances (function 'ltrend.test' of the R package 'lawstat'). The strength and significance of the linear trend between groups is measured using a correlation statistic (negative correlations indicate downward trends and positive correlations indicate upward trends. For both the log-linear models and the Brown-Forsythe-Levene tests, we used a minimum of 1000 bootstrap replicates to determine the significance of our results.

3.4 RESULTS

A. baumannii ATCC 17978 containing the MDR plasmid pB10 was grown in biofilms and liquid serial batch cultures for four weeks in the presence of antibiotics selecting for the plasmid (Figure 3.1). Plasmid pB10 was shown to be highly unstable in the ancestral strain of *A. baumannii*; the fraction of plasmid bearing cells in the ancestor at the end of persistence assays was $9.0x10^{-5}$ (data not shown). Clones isolated from t_0 biofilms showed levels of instability that were similar to those of the ancestor; our results compare t_0 to host-plasmid pairs found in liquid and biofilm cultures at t_{28}.

On average, after 28 days host-plasmid pairs showed significant improvement in plasmid persistence (Figure 3.2). However, the persistence was on average higher in liquid cultures than in biofilms, supporting our first hypothesis: natural selection is weaker in biofilms. This was shown in two ways. First, the fraction of plasmid-bearing cells observed in the plasmid persistence assays declined more slowly (Figure 3.2, Table 3.1). Second, the average plasmid bearing fraction

at the end of the 8-day plasmid persistence assays was approximately 782 times higher at t_{28} in liquid cultures, while it was only 25 times higher in biofilms (Table 3.1). To our knowledge, these results show for the first time that plasmid persistence can improve over time in the clinically relevant environment of biofilms, but not to the extent observed in liquid batch cultures.

Although plasmid persistence was, on average, higher in liquid cultures, the variability of plasmid persistence was significantly higher in biofilms (Table 3.1). These findings provide evidence for our second hypothesis: biofilms maintain broader diversity. A visual inspection of the distribution of residuals in the three environments shows liquid cultures had a much stronger central tendency and less diversity (Figure 3.3). A Brown-Forsythe-Levene test indicate that there no significant difference in the diversity of plasmid persistence when comparing t_0 biofilms and t_{28} biofilms. In contrast, the diversity in plasmid persistence in t_0 liquid cultures was significantly lower than the t_0 cultures used to seed them. Furthermore, we found a significant overall downward trend in diversity from t_0 (most diverse), to t_{28} biofilm cultures, to the t_{28} liquid cultures (least diverse) ($H_A : \sigma_{t_0} > \sigma_{t_{28,biofilm}} > \sigma_{t_{28,liquid}}$; $\rho = -0.2032$; $p = 0.0120$). As hypothesized, these results demonstrate that growth of *A. baumannii* in biofilm cultures maintains higher levels of diversity than liquid cultures.

3.5 DISCUSSION

The emergence and spread of antibiotic resistant bacteria is one of the greatest crises facing healthcare today, yet the factors influencing these processes are still poorly understood. Since self-transmissible MDR plasmids play a key role in the spread of antibiotic resistance (Frost et al., 2005), we need to better understand how well they persist in clinically relevant pathogens in the absence of natural selection, and, perhaps more importantly, how this persistence evolves over time in clinically relevant conditions such as biofilms. So far as we know, all former studies addressing plasmid evolution were done using well-mixed liquid cultures (ex. Bouma and Lenski, 1988; Heuer et al., 2007; De Gelder et al., 2008; Sota et al., 2010; San Millan et al., 2014; Harrison et al., 2015; Loftie-Eaton et al., 2016). Our findings show that the evolution of clinical pathogens and their plasmid-derived antibiotic resistance is markedly different in biofilms compared to the typically utilized well-mixed liquid culture environment. Specifically, the results of this study

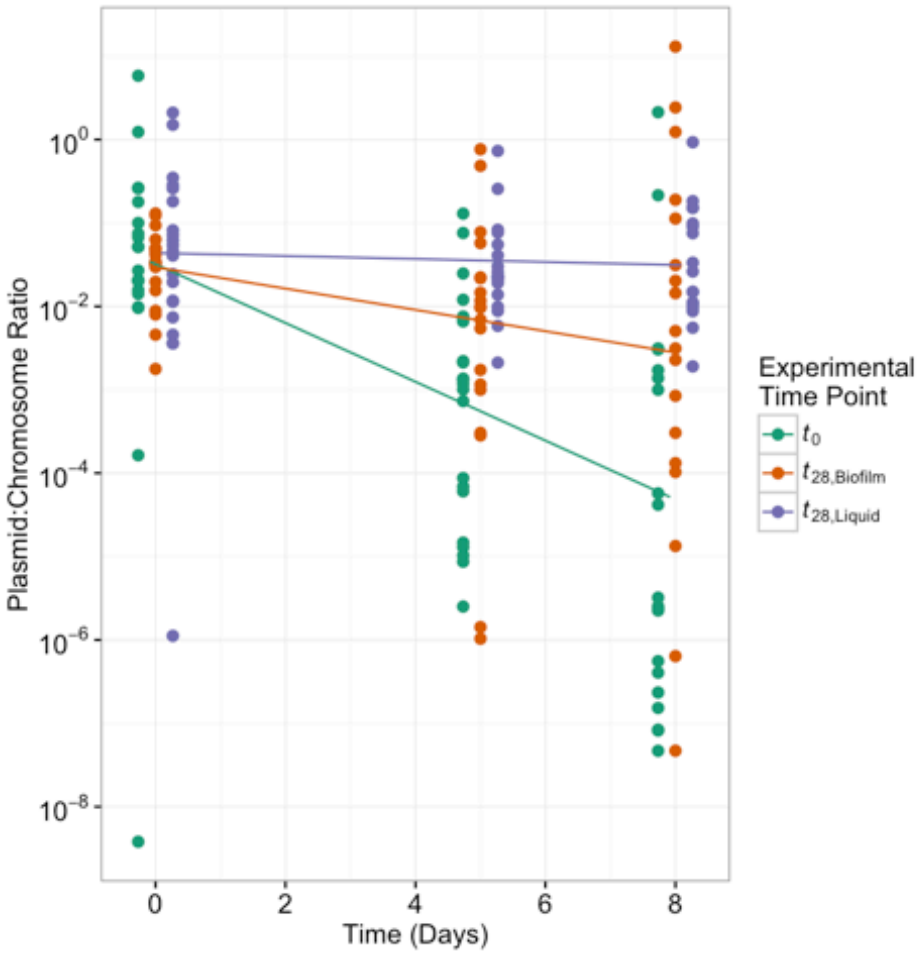

FIGURE 3.2: The persistence of plasmid pB10 in clones from liquid batch cultures and biofilm cultures after 28 days of experimental evolution. The fraction of plasmid-bearing cells was determined by quantitative PCR of the plasmid encoded trfA gene and chromosomally encoded 16S rRNA genes and expressed as the plasmid:chromosome ratio. The lines show the loss of plasmids over time in populations of clones isolated from biofilms at t_0 (green) and t_{28} (red), and liquid batch cultures at t_{28} (purple). For each group of samples, the spread of points around their respective lines reflects the diversity of plasmid persistence among the clones from a particular environment.

TABLE 3.1: Comparison of plasmid loss rates in liquid and biofilm cultures. Parameter β is the difference in mean rates of decline of the plasmid:chromosome ratio with its associated p-value. The fold loss is the relative difference in the fraction of plasmid bearing cells at day 8 of the plasmid persistence assays. The last three columns show differences in diversity between cultures. The first of these (difference in σ_{error}) shows the magnitude of the difference in diversity, ρ is the test statistic where $H_A : \sigma_1 > \sigma_2$, and the last column is the p-value for ρ.

Comparison	β	p	Fold loss	Diff. in σ_{error}	ρ	p
t_0 vs. t_{28}, liquid	-0.8327	0.0008	781.8	1.6957	-0.2504	0.0097
t_0 vs. t_{28}, biofilm	-0.4003	0.0497	24.6	0.3612	-0.0565	0.2984
t_{28}, biofilm vs. t_{28}, liquid	-0.4324	0.0320	31.8	1.3344	-0.2170	0.0367

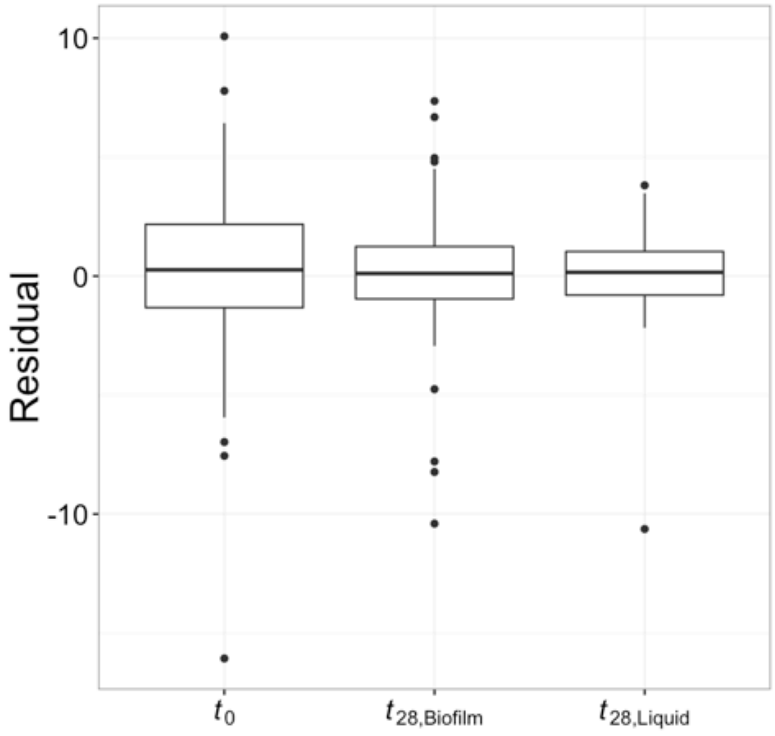

FIGURE 3.3: Empirical distributions of residuals within environments. Residuals are the difference between the modeled mean day 8 plasmid:chromosome ratio for a given clone and the modeled mean for a particular environment (e.g. t_0). The variance in residuals (i.e. diversity) between environments was significantly different when comparing t_0 to t_{28} liquid cultures. An ordered heterogeneity test supported the hypothesis of a significant overall downward trend in diversity from t_0 (most diverse), to t_{28} biofilm cultures, to the t_{28} liquid cultures (least diverse).

show for the first time that: (1) the persistence of MDR plasmids improves in biofilms though much less than in liquid cultures (Figure 3.2), and (2) phenotypic diversity is maintained biofilms while it is lost in liquid cultures (Figure 3.3). However, it should be noted that plasmid persistence assays were all performed in liquid batch cultures; therefore, we cannot exclude the possibility that biofilms might show higher average persistence and/or altered diversity if methods were available to test plasmid persistence directly in a biofilm environment.

Part of the difference between liquid cultures and biofilm cultures is that bacterial growth rates vary widely in biofilms. This is due to the inherent heterogeneity of spatially structured biofilm environments which leads to gradients of nutrients, electron acceptors and metabolic waste products that govern bacterial metabolism (Boles et al., 2004). Cells at the biofilm surface grow more readily than those buried deep in the biofilm matrix where nutrient limitation hinders cell growth. Reduced cell growth could be important because plasmid loss only occurs during cell division when plasmids are not properly apportioned between the daughter cells. Thus, one potential explanation is that MDR plasmids may persist in biofilms simply by remaining in non-dividing cells until natural selection for the plasmid is imposed by introducing antibiotics into the environment.

Another important difference is that natural selection is expected to operate differently within clinical biofilm environments in comparison to liquid cultures. Cells are fixed in space in a biofilm, which limits competitive interactions to a cell's immediate neighbors; natural selection therefore occurs on a local scale. This localized scale of natural selection in biofilms has been shown to increase the accumulation of chromosomal genetic diversity (Boles et al., 2004; Ponciano et al., 2009). Our large, well-replicated study design allowed us to assess the diversity generated by antibiotic application, rather than simply estimating means. Similar what has been observed for chromosomal genetic diversity, our evolution experiment showed that biofilm populations harbor a greater diversity of plasmid persistence compared well-mixed liquid batch cultures (Figure 3.3). Under the conditions used in this study, the variation in plasmid persistence was roughly the same in four day old biofilms (t_0) and those grown for another four weeks in the presence of tetracycline (t_{28}). In contrast, this diversity at t_0 was drastically reduced when the populations were used to found liquid batch cultures that were subsequently grown for four weeks (Table 3.1, Figure 3.3). Our findings confirm that growth in biofilms either protracts or

prevents the selective sweeps commonly observed in liquid cultures (Martens and Hallatschek, 2011). Thus, biofilm maintain MDR plasmid diversity that would not have been observed in the traditional experimental evolution studies which utilize liquid batch cultures.

Maintaining phenotypic diversity is important because it indicates that the evolutionary potential – in other words the ability to respond to natural selection – of biofilms remains high. Therefore, the maintenance of diversity may reduce the effectiveness of future antibiotic treatments because pre-adapted clones may be present in biofilms, whereas such clones may have been lost in selective sweeps in liquid cultures. This phenomenon can be thought of as a 'seed bank' from which new, possibly more dangerous, mutants can arise in the presence of novel natural selection pressures resulting antibiotic application (Boles et al., 2004). This highlights the critical need to understand how growth within a biofilm affects the evolution and persistence of MDR plasmids. Filling this knowledge gap will provide fundamental insights to the prevention of antibiotic recalcitrant bacterial infections in clinical settings. Because heritable genetic diversity must underlie phenotypic diversity to respond to natural selection, we are currently pursuing the next logical step and documenting the genetic basis of improved plasmid persistence using high-throughput sequencing technologies.

In conclusion, our study shows that the persistence of MDR plasmids can readily improve in biofilms. The results presented support a hypothesis that is of medical concern: biofilms generate a broad diversity of antibiotic resistant bacteria that better retain an MDR plasmid than the ancestor, but with variable success. To combat the emergence of antibiotic resistance we should not only find alternative therapies that limit MDR plasmid persistence in biofilms of pathogens, but investigate whether the genetic diversity generated in biofilms facilitates outbreaks of MDR pathogens in the future.

CHAPTER 4

Biofilms Limit Loss of Transmissibility of a Multidrug Resistance Plasmid

4.1 SUMMARY

Conjugative multidrug resistance (MDR) plasmids are a major human health concern because they allow rapid spread of antibiotic resistance genes among bacterial species. Currently, little is known about how MDR plasmids evolve when the host bacterium grows in biofilms rather than well-mixed planktonic populations. Closing this knowledge gap is important, due to the prevalence of biofilm formation among bacteria in general and pathogens in particular. We sought to understand the influence of biofilm growth, with and without antibiotics, on the evolution of the MDR plasmid pB10 in *Acinetobacter baumannii*. Biofilm growth was expected to result in greater genotypic diversity of the plasmid but less improvement in plasmid persistence than growth in planktonic populations. Although we found greater genotypic diversity in plasmids from biofilms, the more striking difference was in the pattern of genotypes underlying the diversity. Loss of plasmid transmissibility as a result of large deletions in the plasmid occurred rapidly and repeatedly in planktonic populations. Biofilm growth sharply limited this loss and most plasmids isolated from biofilms were transmissible, regardless of the presence of antibiotics. Additionally, the effect of the large deletions on plasmid persistence was highly host-specific, as only one of the three other bacterial species we tested retained the truncated plasmid better than ancestral pB10. Plasmid persistence improved more in planktonic populations than biofilms, regardless of antibiotic presence, and improved in planktonic populations even in the absence of antibiotics. These results highlight the need for further research into how plasmid evolution is influenced by biofilm growth.

4.2 INTRODUCTION

Antibiotic resistance is widely recognized as one of the most serious problems facing healthcare today, even meriting mention in a recent State of the Union Address (2014) and the attention of

the United Nations General Assembly. Resistance became a problem shortly following the introduction of the first antibiotics, and antibiotic resistant bacteria are now found worldwide, both in hospitals and the community (e.g., Ammerlaan et al., 2013; Braykov et al., 2013). Some types of resistance can develop due to relatively simple mutations in chromosomal genes. Increasingly however, resistance is the result of acquisition of specific resistance genes that code for a product that can, for example, deactivate or export the antibiotic (Davies, 1994). When present as part of mobile genetic elements such as a plasmids, these resistance genes may be rapidly spread between different species or strains of bacteria (Mazel and Davies, 1999; Mathers et al., 2011). This makes multidrug resistance (MDR) plasmids, particularly those with a broad host range, a target of particular concern in the fight to reduce the spread of antibiotic resistance.

In the presence of antibiotics MDR plasmids are maintained in cells via positive selection (Dionisio et al., 2005). When antibiotics are removed, however, plasmids are expected to be costly to the host. In most cases they should be removed from the host via purifying selection unless the cost of plasmid carriage has been reduced or eliminated by evolution (Heuer et al., 2007; Subbiah et al., 2011; Sota et al., 2010). The degree to which the plasmid is retained during growth without antibiotics is referred to as plasmid persistence and is related to several factors, including the cost of plasmid carriage. Previous studies that utilized liquid serial batch culture conditions have repeatedly shown that plasmid persistence or cost can be substantially improved in these planktonic populations after periods of growth in the presence of antibiotics (Bouma and Lenski, 1988; Loftie-Eaton et al., 2016; De Gelder et al., 2008; Sota et al., 2010).

While multiple studies have addressed the evolution of plasmid persistence in planktonic populations, much less is known about how plasmid persistence evolves during growth in spatially structured populations. Understanding this is important because in the natural and clinical settings the majority of bacteria live in biofilms or other spatially structured communities (Vlamakis et al., 2013). In particular, many medically relevant bacteria are known to form biofilms both in the environment and in the course of infection (Costerton et al., 1999; Parsek and Singh, 2003) and to acquire resistance genes via plasmids (Galimand et al., 1997; Gay et al., 2006). Due to the spatial structure, bacteria growing in a biofilm do not compete directly with all other bacteria. Instead, they compete only with their neighbors, which prevents selective sweeps and results in retention of more variability (Boles et al., 2004; Ponciano et al., 2009; Tyerman et al.,

2013). Biofilms are also expected to have undergo fewer generations on average and experience more variation in the number of generations across the populations than planktonic populations (Poulsen et al., 1993; Wang et al., 2015). This is because certain regions of the biofilm have very little opportunity for growth and reproduction due to space and nutrient limitation. In contrast, cells near the edges of the biofilm will not experience the same restrictions and may divide rapidly. As a result of these factors, biofilms are expected to harbor more diversity than planktonic populations (Boles et al., 2004). This increased diversity allows access to more peaks within a rugged adaptive landscape (Wright, 1932; Kauffman and Levin, 1987). This includes both the landscape in which the diversity evolved and future landscapes that may be encountered, for example due to movement of the bacteria into a new location or infiltration of the population by other species of bacteria. These effects of biofilm growth are expected to influence diversity of any plasmids contained within the bacterial hosts. However, no studies to date have specifically addressed how the evolution of plasmid persistence is affected by biofilm growth.

The biofilm forming bacterium *Acinetobacter baumannii* is an emerging threat in the United States and worldwide. This Gram-negative bacterium causes wound infections, ventilator associated pneumonia, and sepsis, and rapidly acquires resistance to antibiotics (Richet and Fournier, 2006). This combination of high resistance and biofilm formation makes infections with *A. baumannii* particularly difficult to treat. Infection with *A. baumannii* is associated with a higher rate of death, longer hospitalization, and other poor outcomes (Jerassy et al., 2006; Sunenshine et al., 2007), and the species is of special interest to the CDC (Pendleton et al., 2013). Although *A. baumannii* is known to acquire new forms of antibiotic resistance rapidly, and pan-drug resistant strains have been documented (Hsueh et al., 2002), relatively little is known about the mechanisms by which it acquires these resistances. Horizontal gene transfer is known to play a role, as both plasmids containing resistance genes and genes of plasmid origin located within the chromosome of multiple strains of *A. baumannii* have been found (Poirel and Nordmann, 2006; Adams et al., 2008). However, this information is based on studies of the genomes of isolated strains and there is very little information available on how *A. baumannii* interacts with a recently acquired MDR plasmid.

Here we utilize a large, replicated, experimental evolution design to investigate the roles of biofilm growth and the presence of antibiotics on the evolution of a broad host range, MDR

plasmid, pB10 in *A. baumannii*. Previously, we showed that biofilm growth increased the diversity of plasmid persistence profiles among clones evolved in biofilms compared to liquid batch cultures during growth with tetracycline, but we did not characterize the genetic basis of these changes (CHAPTER 3). Here we expand that work by seeking to understand the underlying genetic changes and how the presence/absence of the plasmid-selective antibiotic tetracycline can influence the evolution of plasmid persistence and genotypic diversity of the plasmid. We hypothesize that (1) the presence of plasmid-selective antibiotics during evolution will result in evolved clones with higher levels of plasmid persistence regardless of whether the source population was biofilm or planktonic, and (2) biofilm growth will lead to greater genotypic diversity among evolved plasmids than planktonic growth. We also investigate whether plasmid modifications that enhance plasmid persistence in *A. baumannii* result in similar improvements in persistence in other bacterial hosts. Our results mostly support our hypotheses, show that large deletions in the plasmid improve its persistence, and suggest that plasmid adaptation to one bacterial pathogen can either hamper or promote its persistence in other bacteria.

4.3 RESULTS

4.3.1 *Establishment and Harvest of Biofilm and Planktonic Populations*

In order to investigate the influence of antibiotic presence and growth environment (biofilm vs. planktonic) on the evolution of plasmid persistence in *A. baumannii* we set up 4 evolution experiments divided into two arms (Figure 4.1). The first arm included the biofilm and planktonic populations grown in the presence of tetracycline (tet) for the entire duration of the evolution experiment (Tet+). The second consisted of biofilm and planktonic populations that were evolved in the absence of tet, except for during the initial four days, intended to allow biofilms to establish with plasmid-containing cells. After this four-day establishment phase, at t_0, three replicate flow cells were harvested from each arm and used to inoculate the three replicate planktonic populations for each arm, respectively. Three more flow cells were harvested from each arm after 14 additional days of growth (t_{14}) and 2 mL of culture was harvested from each planktonic population at the same time point. For the Tet+ arm we continued the experiment for an additional 14 days and performed a final harvest of three flow cells and the three replicate

planktonic populations at t_{28}. We obtained six plasmid-containing clones from each harvested population. We also estimated the proportion of plasmid-containing cells remaining in each replicate for the Tet- arm and determined that the plasmid was lost much more rapidly from planktonic populations than from biofilms (Table 4.1). For those planktonic populations in which plasmid-containing clones were no longer detected at t_{14} we obtained plasmid-containing clones from the last day they could be detected.

4.3.2 *Evolution of Plasmid Persistence*

To assess if plasmid persistence had increased over time among individual clones that had evolved in the biofilm and planktonic populations, we performed persistence assays on six clones isolated at various time points from each replicate population. Some of the results for clones grown in biofilm and planktonic populations containing tet have been previously described, but are also used here for comparison purposes (Chapter 3). Here we first tested plasmid persistence for the ancestral host and for all clones isolated at t_0 (Figure 4.2A). There was no significant change in average plasmid persistence during biofilm establishment (for t_0, Tet+ = 0.65, for t_0, Tet- = 0.55), thus we used plasmid persistence data from the t_0 clones as the comparison for clones from the later time points.

We then compared the change in plasmid persistence relative to t_0 for clones isolated at t_{14} or earlier from the biofilm and planktonic populations grown both in the absence and presence of tet (Figure 4.2C). After 6-14 days in the absence of tet, plasmid pB10 showed improved persistence in *A. baumannii* clones from planktonic populations (Fig. 4.2D; p = 0.0002), but not in clones isolated from biofilms (Figure 4.2E; p = 0.336). When compared to plasmid persistence in the presence of tet (Figure 4.2B; Chapter 3) our results indicate that plasmid persistence improved in planktonic populations regardless of the presence of tet, but only improved in biofilms when they were exposed to tet. This may suggest that selection for plasmid persistence was stronger or able to work more efficiently in planktonic populations.

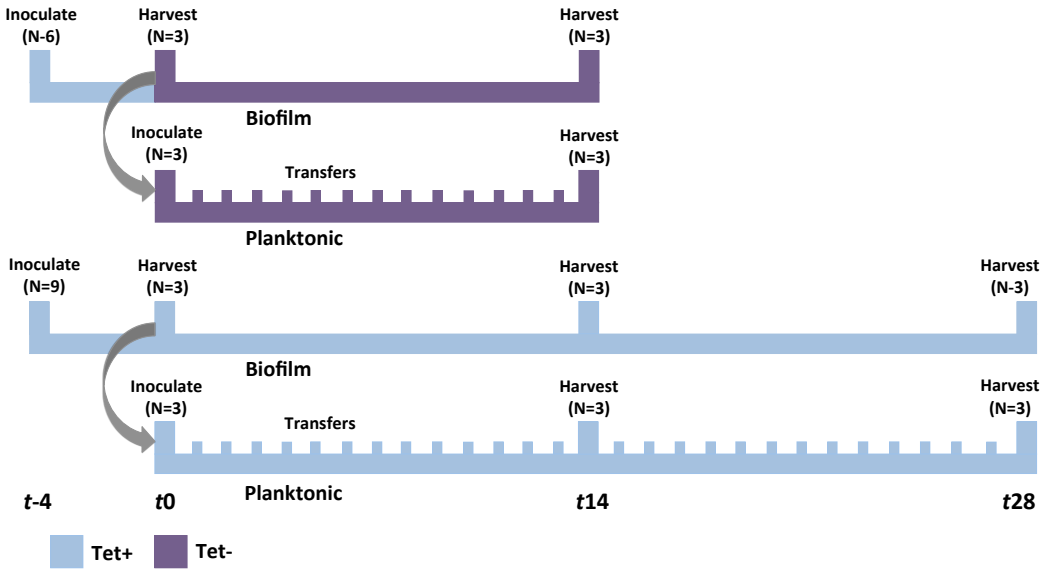

FIGURE 4.1: Timelines of the evolution experiments. Blue represents time periods when media contained tetracycline. Large ticks represent inoculation and harvest events. At t_{-4} the biofilm flow cells were inoculated with the ancestral culture. Four days later, at t_0, the first set of randomly selected flow cells (n=3) were harvested. They were plated on LBA with and without tet to determine the proportion of plasmid containing cells, number of viable cells, and to isolate plasmid-containing clones that were later assayed for plasmid persistence and sequenced (n=6 per replicate). A subsample of the cell suspension harvested from each t_0 replicate was used to inoculate each of three planktonic populations. Small tick marks indicate the daily transfers of planktonic populations. We used the same procedures as at t_0 for the harvests at t_{14} and t_{28}, each involving three randomly chosen biofilms and samples taken from the planktonic populations. We terminated the experiments at t_{14} for the Tet- arm and t_{28} for the Tet+ arm.

TABLE 4.1: Total cell count (microscopy), viable cell count, plasmid-containing (P+) cell count, and fraction of P+ cells for each replicate biofilm or planktonic populations grown without tetracycline. Note that total cell counts estimated via microscopic examination were only performed for biofilms. n/a = count not performed.

t_0	Total cells/mL	Viable cells (cfu/mL)	P+ cells/mL	Fraction P+ cells
Rep 1	$9.6x10^9$	$6.1x10^6$	$5.9x10^6$	0.98
Rep 2	$3.9x10^9$	$4.7x10^6$	$3.8x10^6$	0.82
Rep 3	$3.5x10^9$	$3.4x10^6$	$3.4x10^6$	1.00
t_0 Avg	$5.6x10^9$	$4.7x10^6$	$4.4x10^6$	0.93
t_{14} Biofilm				
Rep 1	$1.3x10^{11}$	$2.9x10^6$	$5.5x10^5$	0.19
Rep 2	$1.3x10^{10}$	$3.1x10^6$	$3.6x10^5$	0.11
Rep 3	$1.1x10^{10}$	$1.8x10^6$	$1.3x10^5$	0.07
Biofilm Avg	$5.1x10^{10}$	$2.6x10^6$	$3.5x10^5$	0.12
t_{14} Planktonic				
Rep 1	n/a	$2.0x10^8$	$3.2x10^4$	0.00016
Rep 2	n/a	$1.6x10^8$	<1	0
Rep 3	n/a	$2.0x10^8$	<1	0
Planktonic Avg	n/a	$1.9x10^8$	$1.1x10^4$	0.00005

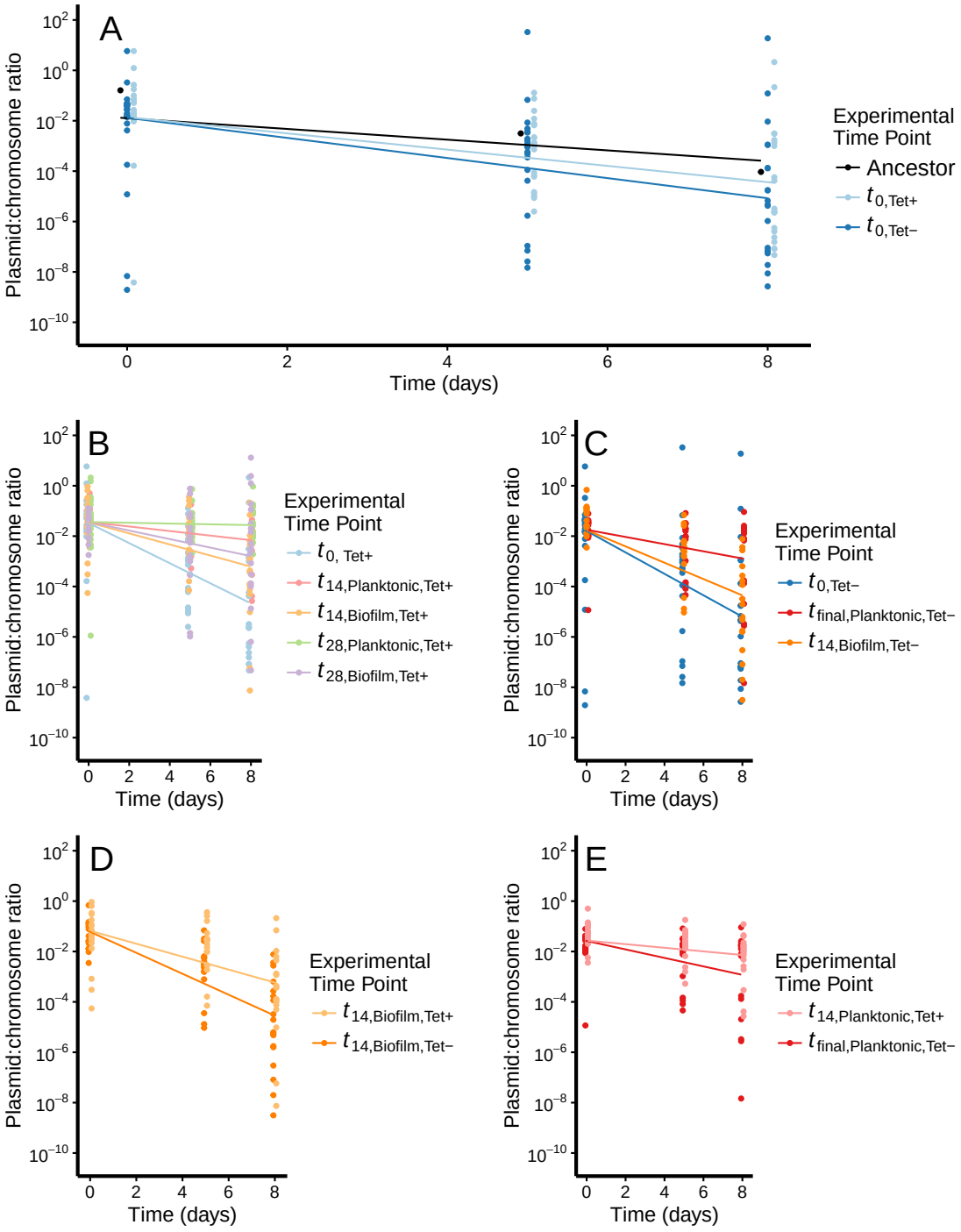

FIGURE 4.2: Plasmid persistence shown as the estimated ratio of *trfA*/*16S* rRNA genes for A) the ancestor used to inoculate the flow cells and clones isolated from the t_0 biofilms from both the Tet+ and Tet- arms of the experiment, B) clones isolated from each time point of the Tet+ evolution experiment, C) clones isolated from each time point of the Tet- evolution experiment, D) a comparison between clones isolated on day 14 from biofilms evolved with and without tet, and E) a comparison between clones isolated on day 14 or earlier from the planktonic populations evolved with and without tet. Lines are the results of the log-linear model for each group.

4.3.3 *Plasmid Genotypes*

To determine the genotypic changes in the plasmid pB10 that occurred during the evolution experiments we resequenced the pB10 variants from six clones isolated from each replicate biofilm and planktonic population in the presence and absence of tet at times t_0, t_{14} and also t_{28} (Tet+ arm only). Thus we sequenced a total of 18 clones per population type under each treatment at each time point, for a total of 144 plasmids. We identified three types of mutations in pB10. A visual summary of these mutations is depicted in Figure 4.3A-C, and a list of them with the genes affected is given in Tables C.1 & C.2. By far the most common were large deletions focused in the area of the genes involved in conjugative transfer (*tra*), mate pair formation (*trb*), and an intervening integron containing the sulfonamide and amoxicillin resistance genes. These deletions also often included a few genes thought to be involved in plasmid persistence and central control (Tables C.1 & C.2). We also identified nine smaller deletions all located in maintenance/control genes *kfrA*, *klcB*, or one of the *trb* genes. None of these smaller deletions involved any part of the integron. Finally, we identified two single nucleotide polymorphisms (SNPs), one in an intergenic region near the *tetR* gene and one in a *trb* gene. Thus large deletions removing virtually all the genes required for horizontal transfer as well as the two intervening resistance genes, were by far the most frequent genetic changes observed, suggesting they were under strong selection.

Among the large deletions, each genotype was found in at least 1 and up to 12 clones. The specific bounds of the deletions varied by thousands of base pairs on either end, and the number of genes deleted ranged from 24 to 35. There were 20 genes (*trbF – traD*) that were at least partially deleted in every large deletion variant and 39 genes (*trbE – incC2*) that were at least partially deleted in one or more of the large deletion variants (Figure 4.3, Tables C.1 & C.2). The size of the large deletions ranged from 23,369 bp to 34,594 bp. Thus, some of these large deletions represent more than half the total plasmid genome.

4.3.4 *Phenotypic Effects of a Large Plasmid Deletion*

To determine the effect of the large deletions in pB10, we compared plasmid persistence for clones carrying either truncated or full-length plasmids in two ways. First we compared persistence of truncated and full-length plasmids in their co-evolved hosts. On average clones with trun-

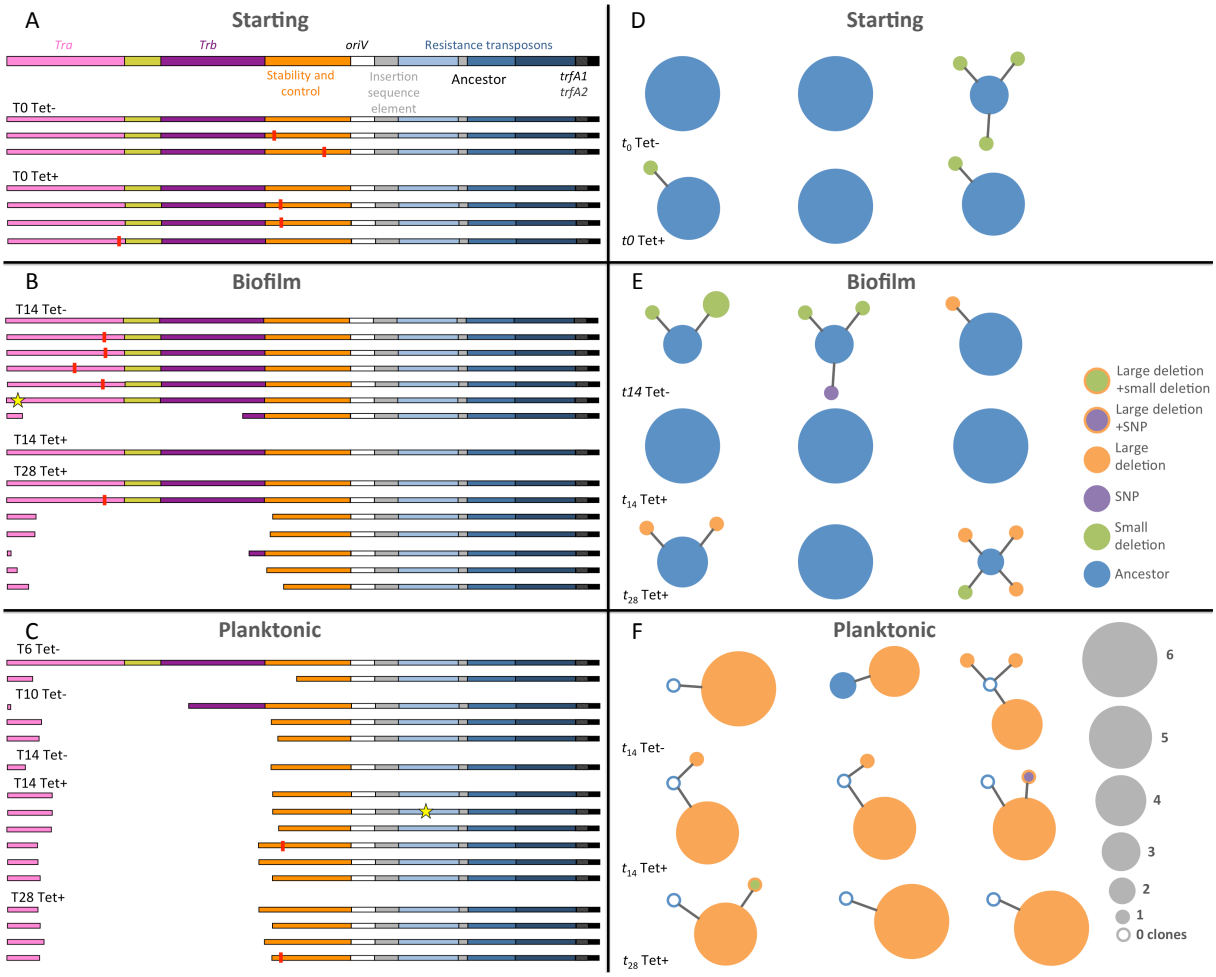

FIGURE 4.3: Map of the gene regions of pB10 plasmid, summary of the mutations found in the plasmid in the starting (A) and biofilm (B) and planktonic (C) replicates, and genotype networks for the starting (D), biofilm (E), and planktonic (F) replicates showing the relationships between different genotypes present in each replicate. Empty areas in the plasmid maps represent large (>2 kbp) deletions, red lines represent smaller deletions, stars represent SNPs. Each set of linked circles in the genotype networks represents the frequency of the plasmid genotypes isolated from a replicate and their relationships to each other. Each line linking genotypes within the networks represents a single mutation occurring between the linked genotypes. The size of the circle for a particular genotype indicates how many clones that genotype was found in within that replicate.

cated plasmids had significantly improved plasmid persistence ($p = 0.007$), while those with full-length plasmids had plasmid persistence similar to the ancestor (Figure 4.4A; $p = 0.8$). However, some clones with full-length plasmids demonstrated high plasmid persistence, suggesting that chromosomal changes in the host also contributed to plasmid persistence. To ensure that the observed improvement in persistence for truncated plasmids was not due only to chromosomal changes in their associated hosts we also transformed the ancestral host with one full-length pB10, and several truncated plasmids isolated from their co-evolved hosts at t_{28} of the Tet+ arm. Plasmid persistence was higher for the truncated plasmids than for the full-length ($p < 0.0001$) or ancestral pB10 (Figure 4.4B; $p = 0.003$). Thus we determined that the large deletions we observed in pB10 were a major driver of improved plasmid persistence in *A. baumannii*.

4.3.5 *Plasmid Genotypic Diversity in Biofilm and Planktonic Populations*

To determine the effect of biofilm growth on plasmid genotypic diversity we defined plasmid genotypes using the *breseq* data and compared patterns of diversity between biofilm and planktonic populations. We found higher average genotype diversity (calculated as haplotype diversity, H; Nei and Tajima, 1981) of pB10 in biofilm populations ($H = 0.38$ for biofilm populations and 0.29 for planktonic populations). There were also striking differences between the patterns of genotypes found in the different population types (Figure 4.3D-F). Only in one instance was a particular evolved pB10 genotype found in more than one clone sampled from biofilms (out of a total of 12 clones that contained mutated pB10). This is in contrast to the planktonic populations, where 7 of the 12 evolved genotypes were not unique, but shared between at least 2 and up to 12 clones. In three of nine planktonic populations (three each from t_{14} and t_{28} of the Tet+ arm, and one each from t_6, t_{10} and t_{14} of the Tet- arm) all six clones sampled contained the same large pB10 deletion, and in the remaining six replicates a single genotype was shared between at least four of the six clones (Figure 4.3E). Moreover, among all the clones from planktonic populations we found only two copies of ancestral pB10, both in the population from the Tet- arm that was sampled at t_6, the last day we could detect the plasmid in that replicate. In four of the nine biofilm populations (three each from t_{14} and t_{28} of the Tet+ arm and three from t_{14} of the Tet- arm) all six of the sampled clones still contained the ancestral genotype of pB10, and in the five remaining biofilms at least three of the six clones still contained ancestral pB10 (Figure 4.3F). Not only did

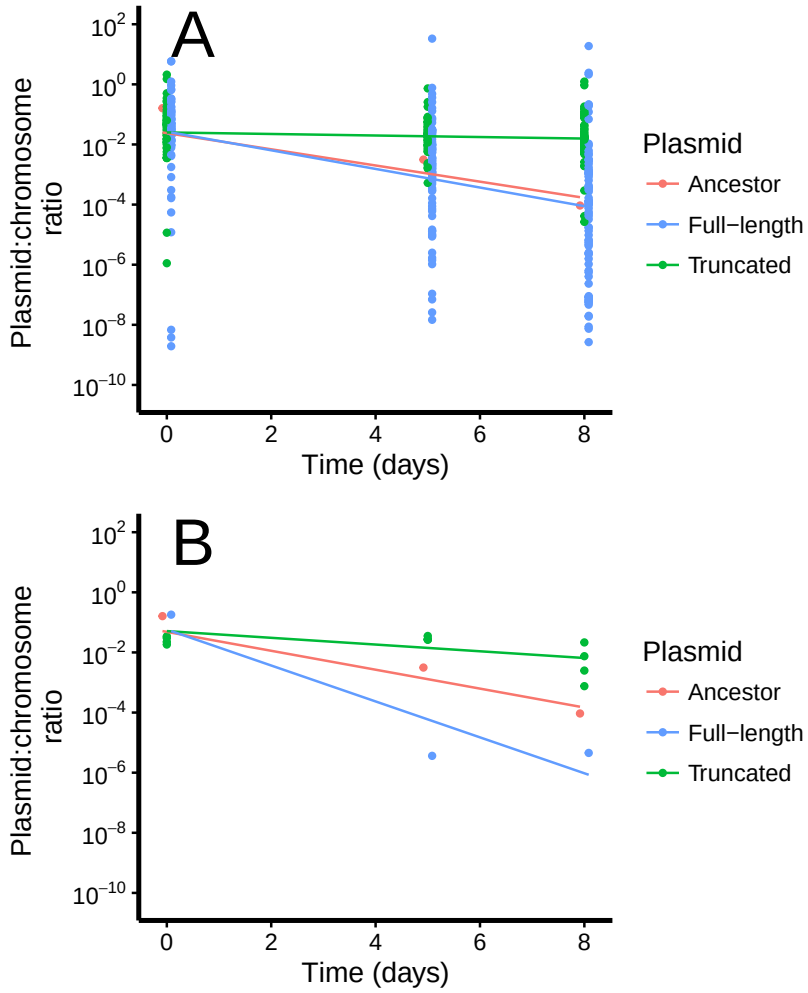

FIGURE 4.4: Plasmid persistence shown as the estimated ratio of *trfA*/*16S* rRNA genes for A) ancestral, truncated, and full-length plasmids in their associated evolved hosts and B) ancestral, truncated, and full-length plasmids in the ancestral host. Lines are the results of the log-linear model for each group.

biofilms contain more plasmid diversity than planktonic populations but importantly they still contained a large proportion of transmissible plasmids, whereas these mobile MDR elements were no longer detected in eight of the nine planktonic populations.

4.3.6 *Persistence of Evolved pB10 in Other Naïve Hosts*

In order to understand the broader effect of large deletions in pB10 on plasmid persistence, we tested one evolved truncated version of pB10 ($pB10_{ev}$) obtained from a t_{28} biofilm clone (Tables C.1 & C.2) in three different species. All three strains have been shown previously to poorly maintain plasmid pB10 (De Gelder et al., 2007). Surprisingly, in the two *Pseudomonas* hosts, $pB10_{ev}$ was even less persistent than ancestral pB10, while in *Stenotrophomonas maltophilia* P21 it was much more persistent. The size of the negative effect was stronger in *P. nov* H2 than in *P. moraviensis* R28. Thus a large deletion in pB10 that removed the transfer genes affects plasmid persistence in a highly species-specific manner but can clearly contribute to the persistence of MDR plasmids in more than one bacterial pathogen.

4.4 DISCUSSION

We found support for our hypothesis that biofilms exposed to an antibiotic that selects for the plasmid for two weeks would result in greater improvements in plasmid persistence than biofilms grown in the absence of the drug. However, this was not the case for planktonic populations. After 14 days of growth in biofilms in the presence of the antibiotic tetracycline (tet) there was a significant improvement in average plasmid persistence among the clones when compared to plasmid persistence at t_0, whereas no improvement was observed for clones evolved in the absence of the antibiotic. In contrast, in the planktonic populations plasmid persistence was improved after two weeks or less of evolution both in the presence and absence of tet. This may be due to a variety of factors. Fewer generations and the ability of cells to remain viable but not dividing for long periods of time in biofilms would be expected to limit opportunities for selection to improve plasmid persistence in biofilms compared to daily cycles of approximately 10 doublings in planktonic populations. Also, the spatial structure of the biofilm would be expected to limit competition between different plasmid genotypes, slowing down the rate of evolution to-

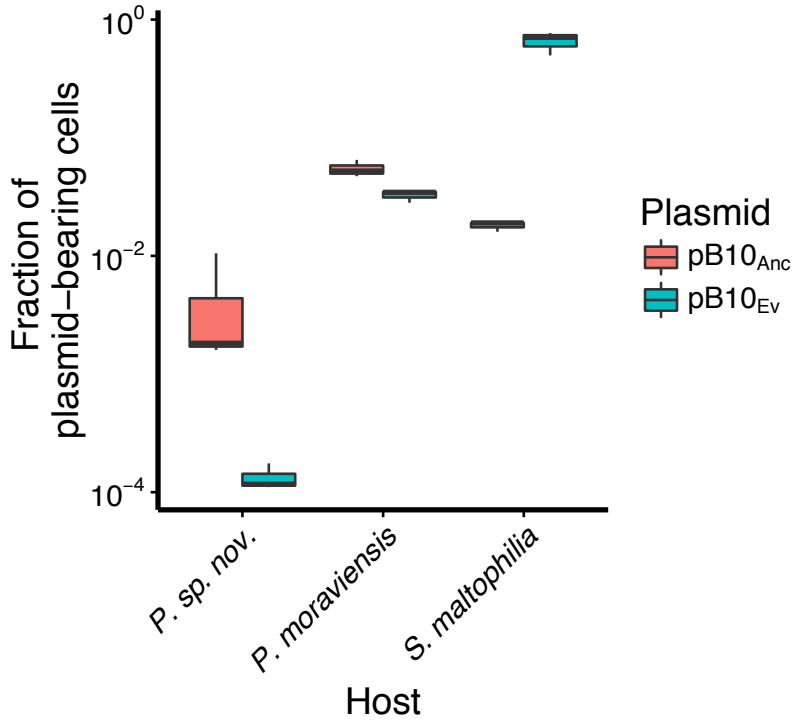

FIGURE 4.5: Plasmid persistence of pB10 and truncated plasmid $pB10_{ev}$ in three additional hosts, as determined by plate counting.

wards improved persistence. Meanwhile in planktonic populations, selection for mutations that improve plasmid persistence may have been stronger in the Tet- arm because plasmid-containing cells had to compete against plasmid-free cells that would have been eliminated from the Tet+ arm. Our results suggest that reduction of antibiotic use may be help reduce the number of copies of resistance genes in populations but does not eliminate selection for persistence of those genes when they are carried by plasmids.

Genetic diversity among the plasmids was moderately higher in biofilms than in planktonic populations, but perhaps more importantly the pattern of plasmid genotypes present in biofilms was markedly different than in planktonic populations. In each of the biofilms the ancestral plasmid genotype was retained in a large portion of the population, and even after 28 days of biofilm growth evolved plasmid genotypes made up only a third of the plasmids that were analyzed. There was only a single biofilm population where evolved genotypes outnumbered ancestral genotypes. In contrast, evolved genotypes, mostly with large deletions, represented a large majority of plasmids for all of the planktonic populations. This is likely at least partially due to differences in the number of generations, with the higher number of doublings in planktonic populations allowing for much more evolution to take place in the same time frame as the biofilm populations. It has to be noted that while the total number of generations in the planktonic populations at each time point is known, it is impossible to estimate this number for the biofilm populations. Additionally the spatial structure of biofilms preventing selective sweeps likely contributed (Martens and Hallatschek, 2011). However, we cannot rule out that there were also differences in selective pressures in biofilms that may have actively favored retention of the ancestral plasmid genotype. For example previous studies have found evidence that conjugative plasmids can stabilize biofilms during the early stages of biofilm formation due to expression of the sex pilus, the genes for which were often lost in our experiments (Ghigo, 2001; Molin and Tolker-Nielsen, 2003). If similar conditions favoring conjugation existed in our biofilms it could have favored retention of ancestral plasmids. The stark difference between planktonic and biofilm populations in our experiment suggests that it is critical to study plasmids in the context of biofilms in order to understand their evolution and the availability of plasmid-carried genes within natural, spatially structured populations.

Regardless of presence of tetracycline or whether growth occurred in biofilm or planktonic populations, the most common modification of pB10 in during all evolution experiments was the loss of a large portion of the plasmid that encompassed the conjugation and mating pair formation genes, and the integron containing the genes for sulfonamide and amoxicillin resistance. The loss of these genes resulted in a loss of transmissibility of the plasmid by conjugation (data not shown). Previous research has observed loss of plasmid transmissibility and deletion of transfer regions during bacteria-plasmid evolution, and suggested that a tradeoff exists for plasmids between horizontal and vertical transmission (Turner et al., 1998; Dahlberg and Chao, 2003; Porse et al., 2016). Although we did not directly measure the plasmid cost, our results appear to be only partially consistent with this hypothesis. Opportunities for horizontal transfer are thought to be more frequent in biofilms than in planktonic populations (Hausner and Wuertz, 1999; Molin and Tolker-Nielsen, 2003), although this is not confirmed by all studies (Stalder and Top, 2016). Because horizontal transfer is expected to be more beneficial when opportunities for transfer are more common the higher rate of transfer loss in planktonic populations would be expected under this hypothesis. On the other hand, we still saw loss of transfer in all but two clones obtained from the planktonic populations that were not exposed to tetracycline, despite the presence of extremely high numbers of plasmid-free hosts for conjugation. These results do not fit with the expectation from the tradeoff hypothesis that horizontal transfer should be more strongly favored in conditions in which plasmid-free hosts are common. This suggests that in planktonic populations the cost to *A. baumannii* of maintaining the transfer regions and integron is higher than the benefit of retaining the plasmid in the population through horizontal transfer, regardless of the availability of uninfected hosts.

The specific bounds of the large deletions in pB10 varied substantially between the 20 large plasmid deletion variants that we observed. This is in contrast to Porse et al. (2016) who identified a similar pattern of large deletions of the plasmid's conjugative transfer regions in an IncN plasmid from *Klebsiella pneumoniae* that they evolved in two strains of *Escherichia coli* under antibiotic selection. In that case the deletions were very similar to one another, including the specific boundaries of the deletion, which corresponded to the presence of an IS26 insertion element. Our findings are similar in that they are highly repeated across multiple independent evolved populations of *A. baumannii* grown under different environmental conditions. However,

in pB10 the deleted regions were not flanked by IS elements. Moreover, the high variability in the boundaries of the deletions and the fact that individual deletion variants are not shared across replicate populations suggest that the underlying mechanism that is causing the deletions is different and currently unknown.

Importantly, while Porse et al. (2016) found that the deletion created in one strain was also beneficial in the other hosts in which they tested it, we found mixed results in the alternative hosts we tested. The deletion we tested caused a large improvement in plasmid persistence in two pathogens, *A. baumannii* and *S. maltophilia*, but not in the two environmental *Pseudomonas* isolates we tested. It is clear from our results, then, that although deletions that eliminate conjugative transfer of the plasmid may be beneficial in one species we should be careful not to expect that benefit to extend to all other species.

In a previous experiment in which pB10 was evolved in *S. maltophilia* there was no evidence of deletions of the conjugative transfer regions in the evolved plasmids, despite the large improvement in persistence that we demonstrated here for one pB10 deletion variant in this host (De Gelder et al., 2008). This is similar to the finding of Porse et al. (2016) that the deletion that occurred in the plasmid that was evolved in their *E. coli* strains was beneficial in a *Klebsiella* host, even though such deletions never arose when *Klebsiella* was evolved with this plasmid. Thus our findings are consistent and provide additional support for the hypothesis that bacterial species may differ in the evolutionary trajectories of their plasmids, independent of whether the corresponding mutations are beneficial in that host.

Altogether, our findings suggest that biofilms are likely to play an important role in the maintenance and spread of MDR plasmids in natural populations of bacteria. They increase the time period over which MDR plasmids can persist, even in hosts in which they would normally be lost rapidly. They also allow for relatively rapid evolution of improved plasmid persistence within individual biofilm clones (though slower than in planktonic populations). Together these qualities mean that MDR plasmids can likely persist in biofilms for extended periods of time without antibiotics present and could subsequently evolve improved persistence once antibiotics are present. Finally, despite rapid loss of transmissibility in planktonic populations, almost 90% of clones isolated from biofilms retained the genes for conjugative transfer. This suggests biofilms could play a critical role in the horizontal transfer of MDR plasmids that would have been an-

ticipated from results of experiments limited to planktonic populations. Our findings show that while experimental evolution of bacteria in planktonic populations have addressed important basic evolutionary questions, we should remember that the growth conditions are far removed from those of natural bacterial populations, including pathogens. For these reasons it is vital that future research on the ecology and evolution of MDR plasmids include environments in which biofilm growth occurs.

4.5 METHODS

4.5.1 *Bacteria and Plasmid*

Acinetobacter baumannii strain ATCC 17978 was used as the wild-type strain (Accession #CP000521). In this study we simply refer to this strain as *A. baumannii*. Other strains used were *Pseudomonas sp.* nov. H2 (Loftie-Eaton et al., 2015), *Pseudomonas moraviensis* R28 (Hunter et al., 2014), and *Stenotrophomonas maltophilia* P21 (De Gelder et al., 2007).

Plasmid pB10 is an approximately 65 kbp broad-host-range IncP-1 plasmid, chosen for this study because it was poorly maintained in naïve *A. baumannii* and encodes resistance to tetracycline, an antibiotic to which the ATCC 17978 strain is sensitive (Schlüter et al., 2003).

Information regarding construction of the ancestral host and media and culture conditions used in this study can be found in Chapter 3 and Appendix B.

4.5.2 *Experimental Evolution Protocol*

A timeline of our experimental plan is shown in Figure 4.1. This study contained two arms: one in which the presence of antibiotics was maintained during the entire length of the experiment (Tet+) and one in which antibiotics were removed following the biofilm establishment period (Tet-). The conditions of our Tet+ arm have been previously described Chapter 3 so here we describe the conditions only briefly except where they differ between arms. To begin each arm, we inoculated 8 (Tet-) or 12 (Tet+) flow cells using our archived ancestral stock.

We had three predetermined sampling points for the Tet+ (t_0, t_{14}, and t_{28}) and two predetermined sampling points for the TetÂň- arm (t_0 and t_{14}). For the Tet+ we sampled at t_0, four days after flow was started through the cells following inoculation and represented our baseline

biofilm; t_{14}, which represented the halfway point of our experiment; and t_{28}, our end point. The Tet- arm was shorter, ending at t_{14}, due to difficulties in maintaining biofilms for periods longer than approximately 20 days without contamination and limitations on the time available to run the experiment. At each of these sampling points we harvested three flow cells, for a total of six (Tet-) or nine (Tet+) harvested flow cells through the course of the experiment. The harvest procedures are detailed in Chapter 3.

We initiated liquid batch cultures (planktonic populations) of *A. buamannii* (pB10) by inoculating three replicate test tubes with cell suspension from the three t_0 biofilms from that arm. Each of these three planktonic populations was then grown alongside the remaining biofilms. For the Tet- arm we began running media without tet through the flow cells at t_0, while in the Tet+ arm we continued to run media with tet. We used t_0 biofilms to initiate our planktonic populations in order to control for the starting level of diversity found in our biofilms and any evolution that had taken place during the establishment phase of biofilm growth. Following inoculation, the planktonic populations were grown 24 hours (± 1 hour) between transfers of a small portion of the culture to a new test tube containing fresh media. For the Tet+ arm the planktonic populations were archived at t_{14} and t_{28} and plated from serial dilutions on LBA and LBA$_{tet}$ as occurred during the biofilm harvest procedure. For the Tet- arm the planktonic populations were archived daily, due to our inability to predict how long plasmid-containing (tet resistant) cells would be maintained in those populations. They were initially plated at t_5, t_{10}, and t_{14} on LBA and LBA$_{tet}$ in the same manner as the biofilm populations. If we could not detect plasmid-containing cells we then plated the archived samples from that planktonic population until we located the latest time point in which plasmid-containing cells could be detected.

Plates from both the biofilm and liquid harvests were incubated overnight and we selected six resistant clones per replicate biofilm or planktonic population to use for phenotyping and sequencing. Descriptions of the procedures used for archiving clones and plasmid preparations from the clones can be found in Chapter 3. We also inserted the plasmids obtained from the plasmid preparation of five of the resistant clones obtained from t_{28} of the Tet+ arm into the ancestral *A. baumannii* host to isolate the effect of large plasmid deletions on persistence.

4.5.3 *Plasmid Persistence Assays*

To determine plasmid persistence we performed plasmid persistence assays as described in spaceds-mallcapsChapter 3. Briefly, each archived clone was used to inoculate 3 test tubes, then grown for 24 hours (± 1 hour) before the next transfer into new media. This was repeated for a total of 8 days in a manner similar to the growth of the planktonic populations during the evolution experiments. We refer to time points from the persistence assays using the letter *d*, in contrast to time points in the evolution experiments (*t*). Each replicate was archived at days 4 and 7. On days 5 and 8 a cell pellet was frozen at -20°C for later DNA extraction and qPCR. In order to minimize differences between rounds of DNA extraction and qPCR, DNA extractions were performed using the QIAsymphony SP (QIAgen Inc.) and the QIAsymphony DSP DNA Mini Kit (192).

Following extraction, the samples were quantified using a PicoGreen dsDNA kit. qPCR was performed to determine the number of copies of pB10 present in each clone relative to the number of *16S* rRNA copies over the course of our phenotyping assay. PCR was performed on a StepOnePlus real-time PCR system utilizing a Power SYBR™ Green PCR master mix (Applied Biosystems Inc.) and following the manufacturer's instructions.

4.5.4 *Analysis of Plasmid Persistence Data*

Analysis of the qPCR data was divided into two stages that have been previously described (see Chapter 3). Briefly, the first stage involved analysis of the raw qPCR data from which we obtained the fluorescence data for the *16S* gene of the *A. baumannii* host and the trfA gene of the pB10 plasmid and determine the ratio of these values for each time point of the persistence assay. In the second stage, we use the ratio of these ratios as a measure of the proportion of plasmid bearing cells present at each time point. We tracked this ratio over the course of each assay to compare the persistence of the plasmid in clones derived from our biofilm and planktonic populations grown with and without the presence of antibiotics. In order to account for the variance present in the raw qPCR data we used bootstrapping techniques that draw from normal distributions of the ratios' estimated mean and variance which were then used as the dependent variable in the linear model described in Chapter 3.

4.5.5 *Sequencing and Sequence Analysis*

For purposes of DNA sequencing, a sample from each clone that had been phenotyped was extracted using the GenElute Bacterial Genomic DNA kit (Sigma-Aldrich) and samples were submitted to the University of Idaho Genomic Resources Core facility for library preparation and sequencing using an Illumina MiSeq and associated chemistry. Following sequencing the sequence data were screened to remove low-quality reads. High quality reads were then analyzed against the plasmid reference using breseq (Deatherage and Barrick, 2014) to identify predicted mutations in the plasmids of each clone. We used output from breseq to define plasmid genotypes, which were used to construct so-called haplotype networks and calculate haplotype diversity for each population (Nei and Tajima, 1981). Here haplotypes correspond to plasmid genotypes to assess plasmid genotypic diversity.

4.5.6 *Plasmid Persistence in Other Hosts*

We extracted DNA of ancestral pB10 and one evolved truncated pB10 variants from a clone obtained from a t_{28} biofilm (Tables C.1 & C.2), and used electroporation to insert each into three additional bacterial strains. Plasmid persistence assays for these clones were performed in a manner similar to that described above but for a total of seven days. They were sampled at days 0, 4, 7, and instead of using qPCR we used serial dilutions and plate counting on LBA and LBA$_{tet}$ to determine plasmid presence over the course of the assay.

CHAPTER 5

The Influence of Biofilm Growth on Genetic Diversity of *Acinetobacter baumannii*

5.1 SUMMARY

The astonishing amount of genetic diversity that exists in natural populations despite selection that is expected to often favor one genotype, is one of the longest standing questions in evolutionary biology. Previous research suggests that spatial structure may play a critical role. Here we use clones of *Acinetobacter baumannii* obtained during a replicated, long-term evolution experiment to attempt to determine the effect of growth in biofilms, a spatially structured environment, on genetic diversity at the individual and population level. Individual diversity was minimally affected by growth condition, while population level diversity was higher in biofilms than in planktonic populations after approximately 14 days of growth. These results are consistent with experimental evolution studies in bacteria and support the hypothesis that spatial structure is important in maintenance of genetic diversity in *A. baumannii*.

5.2 INTRODUCTION

The generation and maintenance of diversity in natural populations is a central problem of evolutionary biology. Selection acts to remove less fit organisms from populations and is expected to reduce diversity to the single genotype that conveys the highest fitness. Despite this, incredible levels of diversity persist in natural populations. Research into this apparent paradox has indicated that spatial structure is an important component in the maintenance of genetic diversity. For example, spatial structure has been found to limit the influence of selection by limiting competition between different genotypes and preventing selective sweeps (Ally et al., 2014), to increase the effect of genetic drift by isolating smaller populations (Felsenstein, 1976), and to increase the number of available niches in an environment (Rainey and Travisano, 1998; Korona et al., 1994). These effects of spatial structure can have varied results on the genetic diversity of populations. Some, such as increased genetic drift, will tend to eliminate diversity. Others,

like a reduction in selective sweeps, help to maintain it. In general, a moderate level of spatial structure that allows for some gene flow but maintains the opportunity for local adaptation should maximize the maintenance of diversity (Slatkin, 1987).

The pattern of extensive diversity is evident in both eukaryotic and prokaryotic organisms. The advent of genomics has expanded our knowledge of bacterial diversity beyond the few culturable species previous known and revealed startling amounts of diversity. For example, recent studies of the soil and the microbiomes living within humans and other animals have documented thousands of species per cubic cm of soil or living within a single host (Roesch et al., 2007; Gill et al., 2006). In these natural populations of bacteria spatial structure predominates. However, most laboratory studies of bacterial evolution are performed on populations of bacteria grown in well-mixed liquid cultures (e.g., Lenski et al., 1991; Vogwill et al., 2016). Under such culture conditions there is no spatial structure and maintenance of genetic diversity is impeded by selective sweeps that take place regularly. These studies thus present an incomplete picture of how diversity evolves and is maintained in natural bacterial populations.

Increasingly, researchers are seeking to study bacterial populations in the laboratory under more naturalistic conditions. Because spatial structure is such a widespread part of natural bacterial communities this has resulted in a focus on experimental conditions that include the opportunity for the development of spatial structure (e.g., Rainey and Travisano, 1998; Ponciano et al., 2009; Jezequel et al., 2013). These include methods as simple as liquid cultures grown without shaking and as complex as elaborate flow cell setups that allow formation of bacterial biofilms that can then be supported through days or weeks of growth. These studies are beginning to shed light on the influence of spatial structure on establishment and maintenance of genetic diversity in bacteria. For example, previous studies on *Pseudomonas* have found that spatial structure repeatedly allowed the evolution of both phenotypic and genotypic diversity, while populations grown in unstructured populations did not diversify (Rainey and Travisano, 1998). Furthermore, placing diverse populations evolved under conditions of spatial structure into unstructured culture conditions resulted in rapid loss of diversity. This underscores just how critical spatial structure is to the development and maintenance of diversity. Biofilms are an especially interesting area of research because many bacteria of clinical relevance form biofilms during both environmental growth and infection (Parsek and Singh, 2003). Furthermore, they

add a considerable burden to the healthcare system, with an estimated half a million deaths in the USA each year attributable to biofilms (Wolcott et al., 2010). Understanding the role of biofilm growth on genetic diversity of bacteria may thus be useful for understanding some of the dynamics of infection.

Acinetobacter baumannii is a biofilm forming bacterium frequently found in hospitals. It causes opportunistic infections such as ventilator associated pneumonia, wound infections, and infections of implanted medical devices. To date there have been very few studies of *A. baumannii* in the laboratory setting and nothing is known about how biofilm growth influences genetic diversity in the species (Hammerstrom et al., 2015).

We performed a large, replicated, experimental evolution study that allowed us to compare genetic diversity in *A. baumannii* grown as planktonic populations in well-mixed liquid batch culture and in biofilms growing inside of flow cells. Previously we investigated the influence of biofilm growth of *A. baumannii* on diversity of a phenotypic trait, plasmid persistence. Phenotypic diversity was enhanced in biofilms relative to planktonic populations and that diversity was lost in planktonic populations compared to early biofilms (Chapter 3) and two clones isolated from planktonic populations were more likely to contain the same plasmid genotype than clones isolated from a biofilm population (Chapter 4). We hypothesized that patterns of genetic diversity in *A. baumannii* would be similar, with higher genetic diversity in biofilms. Notably, the number of mutations per clone relative to the ancestor was similar for clones isolated from biofilm and planktonic populations. However, in line with our hypothesis, population level diversity was higher in biofilms than in planktonic populations due to more shared alleles in clones isolated from planktonic populations.

5.3 Methods

5.3.1 *Strain*

The evolution experiments were conducted with the bacterial strain *Acinetobacter baumannii* ATCC 17879, in which we introduced plasmid pB10 as described in Chapter 4. This strain was obtained from the American Type Culture Collection (Rockville, MD), and a reference sequence is available (Accession #CP000521).

5.3.2 *Evolution Experiment*

A timeline of the experiment can be found in Figure 3.1. The experimental evolution setup, media used, sampling design, and sequencing methods used have all been previous described in Chapter 3, Appendix B, and Chapter 4. Briefly, we utilized a two armed study design where in one arm populations were grown in the presence of the antibiotic tetracycline (tet) for the duration of the experiment (Tet+ arm) and in the other arm tet was removed after a brief biofilm establishment phase of four days (Tet- arm). The flow cells used to grow biofilms were inoculated with the ancestral strain at t_{-4}. After four days of growth with MBMS$_{tet}$ the biofilm establishment phase was ended and three flow cells in each arm were harvested (t_0). These were then used to inoculate the three planktonic populations for each arm, which were grown in test tubes in a shaking incubator to prevent the formation of spatial structure. Harvests of three flow cells took place on day 14 (t_{14}) for both arms of the study and 28 (t_{28}) for the Tet+ arm only. We also sampled the planktonic populations at those times. The Tet- arm was terminated at t_{14}.

5.3.3 *DNA Extraction, Sequencing, and Sequence Analysis*

For each treatment, six clones were selected from each replicate population at each time point. As we had three replicates of each type of population this meant that there were a total of 18 clones sampled for each condition. The experiment had multiple goals, some specifically requiring that the selected clones contain the plasmid pB10 (see Chapters 3 &4). Therefore, all clones that were sequenced contained the plasmid. Rapid loss of plasmid-containing cells in the Tet- planktonic populations led us to isolate clones from the last day in which these plasmid-containing clones could still be identified for two replicates. For these two replicate populations the clones were isolated at t_6 and t_{10}. For all analyses these clones are included with the other planktonic clones isolated at t_{14}. The clones were grown overnight with tet and genomic DNA was isolated using the GenElute Bacterial Genomic DNA kit (Sigma-Aldrich). Samples were submitted to the University of Idaho Genomic Resources Core facility for library preparation, sequencing using an Illumina MiSeq, and subsequent processing to remove low quality reads. To identify mutations that occurred during our experiment we constructed a reference sequence using high quality reads from our ancestral strain, sequenced using a portion of the same culture used to inoculate

the flow cells. These reads were assembled using the GenBank reference sequence and the BAsys utility (Van Domselaar et al., 2005) was used to annotate the resulting reference. We analyzed high quality reads with the program breseq (Deatherage and Barrick, 2014), which compared reads from each clone against our reference sequence. After mutations had been identified for each clone and mutations present in less than 90% of reads had been removed the total number of mutations per clone and total number of unique mutations per replicate were calculated. We conducted a series of t-tests to compare the number of mutations per individual and number of unique mutations per replicate for each condition.

5.4 RESULTS

To determine how much genetic diversity was present in each population we first determined how many mutations were present in each individual clone when compared to the ancestor, then determined how many unique mutations were found in each population. There were a total of 175 sites where at least one sampled clone contained a derived allele. Figure 5.1 provides an overview of the allele status for each clone at each site. Table 5.1 provides a summary of the number of chromosomal mutations found per clone in each replicate and Table 5.2 provides a summary of the number of unique mutations per replicate for each condition. These mutations were of three types: SNPs, small indels (2-7 bp), and new junctions representing rearrangements of parts of the chromosome. A detailed table that includes the location of each mutation and presence or absence for each clone is available upon request. The presence of tetracycline did not affect the number of mutations found at either the individual or population level (data not shown). Each clone contained an average of 10.8 mutations per clone. At the level of the individual clone there was no difference in the number of mutations between biofilms and planktonic populations (Table 5.3).

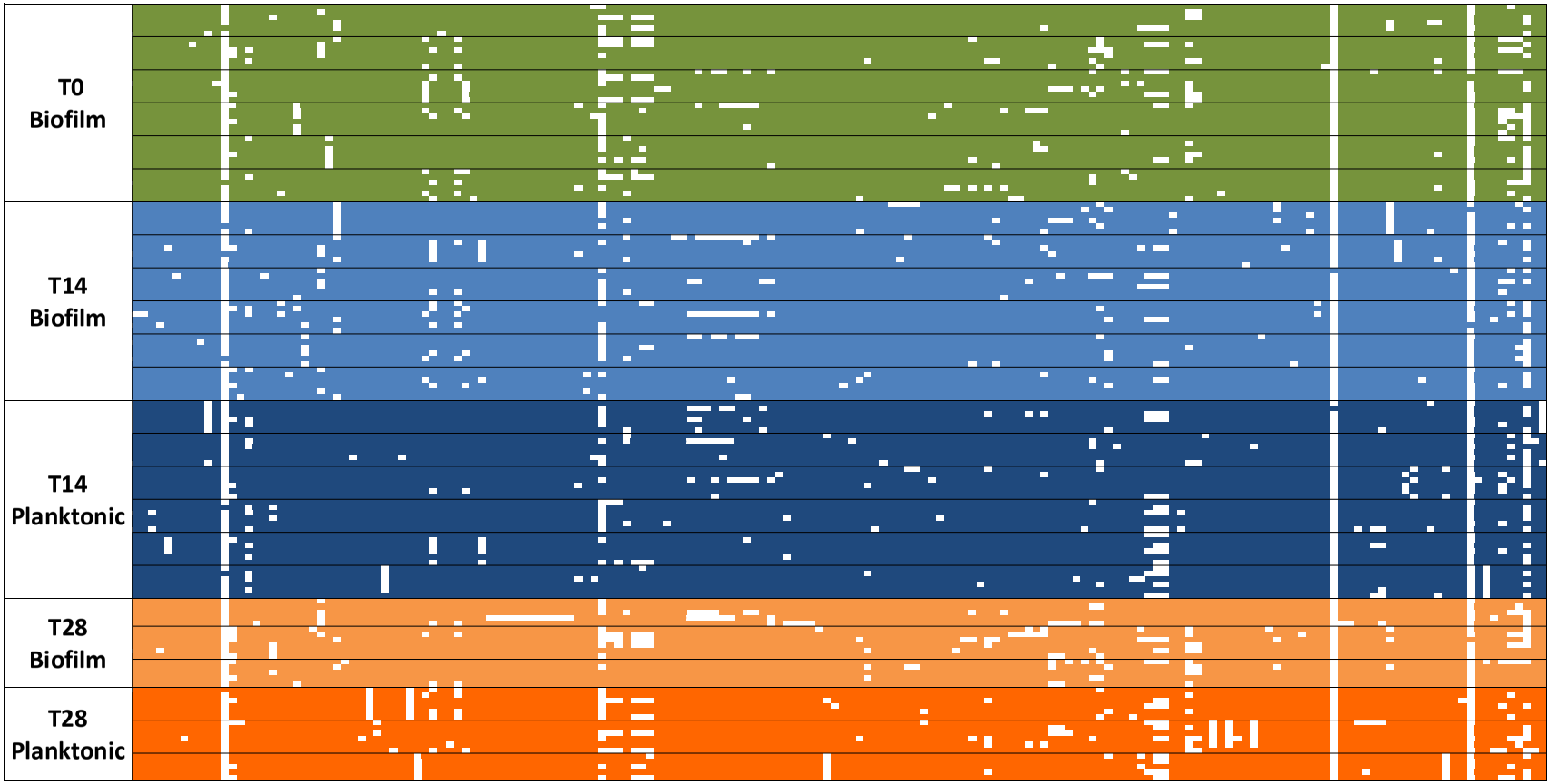

FIGURE 5.1: Overview of 175 sites with derived alleles in at least one clone of *A. baumannii* ordered as they appear in the reference sequence. Filled areas represent ancestral alleles. Empty areas represent derived alleles. Areas that did not contain derived alleles in any clone are not depicted.

Table 5.1: Number of mutations per clone for each replicate population of *A. baumannii*.

Clone ID	Time point	Growth condition	Antibiotics	Replicate	Mutations
DOD_321	t_0	Biofilm	Tet+	1	6
DOD_322	t_0	Biofilm	Tet+	1	6
DOD_323	t_0	Biofilm	Tet+	1	13
DOD_324	t_0	Biofilm	Tet+	1	8
DOD_325	t_0	Biofilm	Tet+	1	12
DOD_326	t_0	Biofilm	Tet+	1	7
				Replicate average	8.7
DOD_327	t_0	Biofilm	Tet+	2	17
DOD_328	t_0	Biofilm	Tet+	2	16
DOD_329	t_0	Biofilm	Tet+	2	14
DOD_330	t_0	Biofilm	Tet+	2	8
DOD_331	t_0	Biofilm	Tet+	2	11
DOD_332	t_0	Biofilm	Tet+	2	8
				Replicate average	12.3
DOD_333	t_0	Biofilm	Tet+	3	18
DOD_334	t_0	Biofilm	Tet+	3	10
DOD_335	t_0	Biofilm	Tet+	3	11
DOD_336	t_0	Biofilm	Tet+	3	14
DOD_337	t_0	Biofilm	Tet+	3	11
DOD_338	t_0	Biofilm	Tet+	3	13
				Replicate average	12.8
DOD_1508	t_0	Biofilm	Tet-	1	18
DOD_1509	t_0	Biofilm	Tet-	1	14
DOD_1510	t_0	Biofilm	Tet-	1	13
DOD_1511	t_0	Biofilm	Tet-	1	7

Table 5.1 Continued: Number of mutations per clone for each replicate population of *A. baumannii*.

Clone ID	Time point	Growth condition	Antibiotics	Replicate	Mutations
DOD_1512	t_0	Biofilm	Tet-	1	8
DOD_1513	t_0	Biofilm	Tet-	1	8
				Replicate average	11.3
DOD_1518	t_0	Biofilm	Tet-	2	10
DOD_1519	t_0	Biofilm	Tet-	2	6
DOD_1520	t_0	Biofilm	Tet-	2	9
DOD_1521	t_0	Biofilm	Tet-	2	9
DOD_1522	t_0	Biofilm	Tet-	2	13
DOD_1523	t_0	Biofilm	Tet-	2	7
				Replicate average	9.0
DOD_1528	t_0	Biofilm	Tet-	3	10
DOD_1529	t_0	Biofilm	Tet-	3	13
DOD_1530	t_0	Biofilm	Tet-	3	8
DOD_1531	t_0	Biofilm	Tet-	3	12
DOD_1532	t_0	Biofilm	Tet-	3	9
DOD_1533	t_0	Biofilm	Tet-	3	10
				Replicate average	10.3
				t_0 biofilm average	10.8
DOD_339	t_{14}	Biofilm	Tet+	1	13
DOD_340	t_{14}	Biofilm	Tet+	1	9
DOD_341	t_{14}	Biofilm	Tet+	1	8
DOD_342	t_{14}	Biofilm	Tet+	1	12
DOD_343	t_{14}	Biofilm	Tet+	1	8
DOD_344	t_{14}	Biofilm	Tet+	1	11

Table 5.1 Continued: Number of mutations per clone for each replicate population of *A. baumannii*.

Clone ID	Time point	Growth condition	Antibiotics	Replicate	Mutations
				Replicate average	10.2
DOD_345	t_{14}	Biofilm	Tet+	2	19
DOD_346	t_{14}	Biofilm	Tet+	2	12
DOD_347	t_{14}	Biofilm	Tet+	2	14
DOD_348	t_{14}	Biofilm	Tet+	2	8
DOD_349	t_{14}	Biofilm	Tet+	2	11
DOD_350	t_{14}	Biofilm	Tet+	2	4
				Replicate average	11.3
DOD_351	t_{14}	Biofilm	Tet+	3	7
DOD_352	t_{14}	Biofilm	Tet+	3	11
DOD_353	t_{14}	Biofilm	Tet+	3	13
DOD_354	t_{14}	Biofilm	Tet+	3	8
DOD_355	t_{14}	Biofilm	Tet+	3	8
DOD_356	t_{14}	Biofilm	Tet+	3	6
				Replicate average	8.8
DOD_1667	t_{14}	Biofilm	Tet-	1	12
DOD_1668	t_{14}	Biofilm	Tet-	1	10
DOD_1669	t_{14}	Biofilm	Tet-	1	22
DOD_1670	t_{14}	Biofilm	Tet-	1	8
DOD_1671	t_{14}	Biofilm	Tet-	1	8
DOD_1672	t_{14}	Biofilm	Tet-	1	6
				Replicate average	11.0
DOD_1673	t_{14}	Biofilm	Tet-	2	15
DOD_1674	t_{14}	Biofilm	Tet-	2	6

Table 5.1 Continued: Number of mutations per clone for each replicate population of *A. baumannii*.

Clone ID	Time point	Growth condition	Antibiotics	Replicate	Mutations
DOD_1675	t_{14}	Biofilm	Tet-	2	10
DOD_1676	t_{14}	Biofilm	Tet-	2	9
DOD_1677	t_{14}	Biofilm	Tet-	2	9
DOD_1678	t_{14}	Biofilm	Tet-	2	9
				Replicate average	9.7
DOD_1679	t_{14}	Biofilm	Tet-	3	6
DOD_1680	t_{14}	Biofilm	Tet-	3	9
DOD_1681	t_{14}	Biofilm	Tet-	3	12
DOD_1682	t_{14}	Biofilm	Tet-	3	10
DOD_1683	t_{14}	Biofilm	Tet-	3	5
DOD_1684	t_{14}	Biofilm	Tet-	3	8
				Replicate average	8.3
				t_{14} biofilm average	9.9
DOD_375	t_{14}	Planktonic	Tet+	1	6
DOD_376	t_{14}	Planktonic	Tet+	1	13
DOD_377	t_{14}	Planktonic	Tet+	1	12
DOD_378	t_{14}	Planktonic	Tet+	1	13
DOD_379	t_{14}	Planktonic	Tet+	1	8
DOD_380	t_{14}	Planktonic	Tet+	1	11
				Replicate average	10.5
DOD_381	t_{14}	Planktonic	Tet+	2	8
DOD_382	t_{14}	Planktonic	Tet+	2	17
DOD_383	t_{14}	Planktonic	Tet+	2	8
DOD_384	t_{14}	Planktonic	Tet+	2	4

Table 5.1 Continued: Number of mutations per clone for each replicate population of *A. baumannii*.

Clone ID	Time point	Growth condition	Antibiotics	Replicate	Mutations
DOD_385	t_{14}	Planktonic	Tet+	2	10
DOD_386	t_{14}	Planktonic	Tet+	2	11
				Replicate average	9.7
DOD_387	t_{14}	Planktonic	Tet+	3	10
DOD_388	t_{14}	Planktonic	Tet+	3	8
DOD_389	t_{14}	Planktonic	Tet+	3	17
DOD_390	t_{14}	Planktonic	Tet+	3	9
DOD_391	t_{14}	Planktonic	Tet+	3	7
DOD_392	t_{14}	Planktonic	Tet+	3	10
				Replicate average	10.2
DOD_1722	t_6	Planktonic	Tet-	2	7
DOD_1723	t_6	Planktonic	Tet-	2	7
DOD_1724	t_6	Planktonic	Tet-	2	10
DOD_1727	t_6	Planktonic	Tet-	2	8
DOD_2115	t_6	Planktonic	Tet-	2	7
DOD_2116	t_6	Planktonic	Tet-	2	16
				Replicate average	9.2
DOD_1728	t_{10}	Planktonic	Tet-	3	5
DOD_1729	t_{10}	Planktonic	Tet-	3	9
DOD_1730	t_{10}	Planktonic	Tet-	3	11
DOD_1731	t_{10}	Planktonic	Tet-	3	9
DOD_1732	t_{10}	Planktonic	Tet-	3	6
DOD_1733	t_{10}	Planktonic	Tet-	3	12
				Replicate average	8.7

Table 5.1 Continued: Number of mutations per clone for each replicate population of *A. baumannii*.

Clone ID	Time point	Growth condition	Antibiotics	Replicate	Mutations
DOD_1698	t_{14}	Planktonic	Tet-	1	7
DOD_1699	t_{14}	Planktonic	Tet-	1	9
DOD_1700	t_{14}	Planktonic	Tet-	1	11
DOD_1701	t_{14}	Planktonic	Tet-	1	9
DOD_1702	t_{14}	Planktonic	Tet-	1	7
DOD_1703	t_{14}	Planktonic	Tet-	1	11
				Replicate average	9.0
				t_{14} planktonic average	9.5
DOD_357	t_{28}	Biofilm	Tet+	1	5
DOD_358	t_{28}	Biofilm	Tet+	1	7
DOD_359	t_{28}	Biofilm	Tet+	1	19
DOD_360	t_{28}	Biofilm	Tet+	1	26
DOD_361	t_{28}	Biofilm	Tet+	1	22
				Replicate average	15.8
DOD_363	t_{28}	Biofilm	Tet+	2	14
DOD_364	t_{28}	Biofilm	Tet+	2	23
DOD_365	t_{28}	Biofilm	Tet+	2	22
DOD_366	t_{28}	Biofilm	Tet+	2	15
DOD_367	t_{28}	Biofilm	Tet+	2	11
DOD_368	t_{28}	Biofilm	Tet+	2	9
				Replicate average	15.7
DOD_369	t_{28}	Biofilm	Tet+	3	17
DOD_371	t_{28}	Biofilm	Tet+	3	14

Table 5.1 Continued: Number of mutations per clone for each replicate population of *A. baumannii*.

Clone ID	Time point	Growth condition	Antibiotics	Replicate	Mutations
DOD_372	t_{28}	Biofilm	Tet+	3	6
DOD_373	t_{28}	Biofilm	Tet+	3	7
DOD_374	t_{28}	Biofilm	Tet+	3	9
				Replicate average	10.6
				t_{28} biofilm average	14.1
DOD_393	t_{28}	Planktonic	Tet+	1	8
DOD_394	t_{28}	Planktonic	Tet+	1	10
DOD_395	t_{28}	Planktonic	Tet+	1	16
DOD_396	t_{28}	Planktonic	Tet+	1	11
DOD_397	t_{28}	Planktonic	Tet+	1	11
DOD_398	t_{28}	Planktonic	Tet+	1	14
				Replicate average	11.7
DOD_399	t_{28}	Planktonic	Tet+	2	16
DOD_400	t_{28}	Planktonic	Tet+	2	11
DOD_401	t_{28}	Planktonic	Tet+	2	17
DOD_402	t_{28}	Planktonic	Tet+	2	20
DOD_403	t_{28}	Planktonic	Tet+	2	12
DOD_404	t_{28}	Planktonic	Tet+	2	18
				Replicate average	15.7
DOD_405	t_{28}	Planktonic	Tet+	3	11
DOD_406	t_{28}	Planktonic	Tet+	3	7
DOD_407	t_{28}	Planktonic	Tet+	3	12
DOD_408	t_{28}	Planktonic	Tet+	3	11

Table 5.1 Continued: Number of mutations per clone for each replicate population of *A. baumannii*.

Clone ID	Time point	Growth condition	Antibiotics	Replicate	Mutations
DOD_409	t_{28}	Planktonic	Tet+	3	15
				Replicate average	11.2
				t_{28} planktonic average	12.9

In contrast, at the population level there was a trend toward fewer unique mutations present in planktonic populations at t_{14} than in biofilms at t_{14} or t_0. This was a result of greater sharing of mutations between clones in the planktonic populations (Table 5.4). This suggests that there was a loss of diversity over time in planktonic populations but not in biofilm populations. The difference between biofilm and planktonic populations was in the expected direction but the value of p was larger at t_{28}. This may be due to a reduction in the influence of spatial structure or simply to the reduced sampling at t_{28} relative to t_{14} limiting statistical inference.

5.5 DISCUSSION

Although spatial structure has been previously documented to be an important driver of diversity for many different organisms (Tyerman et al., 2013; Rainey and Travisano, 1998; Boles et al., 2004; Korona et al., 1994) we found limited support for the hypothesis that growth in spatially structured biofilms would increase genetic diversity relative to planktonic populations of *Acinetobacter baumannii*. At the level of individual clones there were only small differences between conditions and the direction of the effect was inconsistent. This is not unexpected since the effect of biofilm growth on mutations in individuals is unclear and is potentially explained by the fact that biofilms are expected to undergo far fewer generations than planktonic populations and thus might be expected to have a reduced opportunity for mutation accumulation in individual clones. This would be predicted to reduce the number of mutations per clone in biofilm populations but might be offset by mutations accumulated during times in which the cell is not dividing, something that has previously been documented in bacteria (Bull et al., 2001). Thus, in general spatial structure did not increase the accumulation of genetic diversity within clones of *A. baumannii*.

In contrast, we found the expected pattern of genetic diversity when comparing genetic diversity at the population level, although our results did not rise to the level of statistical significance. There appear to be more unique mutations present in t_0 and t_{14} biofilm populations than t_{14} planktonic populations. This occurs, despite a similar number of mutations at the level of individual clones, because of greater sharing of mutations between clones in the planktonic populations. Sharing of mutations is expected to be higher during planktonic growth if selective sweeps are occurring more frequently in those populations than in the spatially structured population.

Table 5.2: Unique mutations per replicate population of *A. baumannii*.

Time point	Growth condition	Antibiotics	Replicate	Mutations
t_0	Biofilm	Tet+	1	26
t_0	Biofilm	Tet+	2	35
t_0	Biofilm	Tet+	3	40
t_0	Biofilm	Tet-	1	32
t_0	Biofilm	Tet-	2	23
t_0	Biofilm	Tet-	3	35
			Average	31.8
t_{14}	Biofilm	Tet+	1	29
t_{14}	Biofilm	Tet+	2	38
t_{14}	Biofilm	Tet+	3	28
t_{14}	Biofilm	Tet-	1	37
t_{14}	Biofilm	Tet-	2	30
t_{14}	Biofilm	Tet-	3	28
			Average	31.7
t_{14}	Planktonic	Tet+	1	26
t_{14}	Planktonic	Tet+	2	30
t_{14}	Planktonic	Tet+	3	32
t_6	Planktonic	Tet-	2	25
t_{10}	Planktonic	Tet-	3	21
t_{14}	Planktonic	Tet-	1	21
			Average	25.8
t_{28}	Biofilm	Tet+	1	48
t_{28}	Biofllm	Tet+	2	46
t_{28}	Biofllm	Tet+	3	32
			Average	42.0
t_{28}	Planktonic	Tet+	1	29
t_{28}	Planktonic	Tet+	2	46
t_{28}	Planktonic	Tet+	3	20
			Average	31.7

TABLE 5.3: Average number of mutations per clone for biofilms and planktonic populations of *A. baumannii* at t_0, t_{14}, and t_{28} and p values resulting from t-tests to compare them. Values in boldface are significant at the 0.05 level.

Group	**Average # mutations**	**p value**
t_0 biofilm	10.8	
t_{14} biofilm	9.9	0.304
t_0 biofilm	10.8	
t_{14} planktonic	9.5	0.112
t_{14} biofilm	9.9	
t_{14} planktonic	9.5	0.652
t_0 biofilm	10.8	
t_{28} biofilm	14.1	0.075
t_0 biofilm	10.8	
t_{28} planktonic	12.9	**0.043**
t_{28} biofilm	14.1	
t_{28} planktonic	12.9	0.541

TABLE 5.4: Average number of unique mutations for biofilm and planktonic populations of *A. baumannii* at t_0, t_{14}, and t_{14} and p values resulting from t-tests to compare them.

Group	**Average # mutations**	**p value**
t_0 biofilm	31.8	
t_{14} biofilm	31.7	0.9594
t_0 biofilm	31.8	
t_{14} planktonic	25.8	0.09074
t_{14} biofilm	31.7	
t_{14} planktonic	25.8	0.05117
t_0 biofilm	31.8	
t_{28} biofilm	42.0	0.167
t_0 biofilm	31.8	
t_{28} planktonic	31.7	0.985
t_{28} biofilm	42.0	
t_{28} planktonic	31.7	0.3301

Although the results presented here are not significant, all differences in population level diversity between biofilm and planktonic populations are in the expected direction (lower diversity for planktonic populations). Additionally, the results here are consistent with our previous analyses that showed higher levels of phenotypic diversity in plasmid persistence in biofilm populations (Chapter 3) and in genotypic diversity of plasmid pB10 carried by *A. baumannii* (Chapter 4). These parallel lines of evidence suggest that additional sampling would provide stronger support for increased genetic diversity of *A. baumannii* in biofilms relative to planktonic populations.

This experiment demonstrates the feasibility of growing long-term biofilms of *A. baumannii* to address evolutionary questions. The sampling strategy used here was not ideal to address issues of population level diversity. A more appropriate course would be to utilize population level sequencing of each replicate population, which would allow for dramatically better estimation of the frequency of mutations in the population, especially for rare alleles. Future work with population level sequencing and a higher number of replicates would allow for enhanced understanding of the role of spatial structure on the generation and maintenance of diversity in *A. baumannii*.

CHAPTER 6

Concluding Remarks and Future Directions

This dissertation provides additional support for the role of both modern and historic environmental variables on patterns of diversity and gene flow in a eukaryotic species, *Ascaphus montanus*. It also fills in critical knowledge gaps in the understanding of how spatial structure in microbial communities influences the evolution of both bacterial hosts and the plasmids they carry. It supports the role of spatial structure in maintaining diversity in both systems, highlighting how similar evolutionary processes play out at widely divergent spatial and temporal scales.

6.1 FUTURE DIRECTIONS

Although this research helps to fill a critical knowledge gap about the effects of spatial structure on plasmid evolution there are still many questions left unanswered. During these experiments we were not able to sample from within the spatial structure of the biofilm due to problems with cell viability after biofilm stabilization to allow spatially explicit sampling. As a result we were unable to assess how plasmid evolution varied with position in the biofilm. One goal of future experiments in this area should then be to sample the biofilms in a spatially explicit way so that more can be learned about the influence of position within the biofilm on plasmid and host evolution.

Additionally, our sampling was not well designed to capture and assess differences in diversity of *A. baumannii* between the biofilm and planktonic populations. With additional funding it would be useful to conduct population level sequencing on the preserved samples from the populations in order to provide a more complete picture of diversity than it was possible to achieve with only six clones sequenced per population.

Bibliography

Adams, M. D., K. Goglin, N. Molyneaux, K. M. Hujer, H. Lavender, J. J. Jamison, I. J. MacDonald, K. M. Martin, T. Russo, A. A. Campagnari, A. M. Hujer, R. A. Bonomo, and S. R. Gill. 2008. Comparative genome sequence analysis of multidrug-resistant Acinetobacter baumannii. Journal of Bacteriology 190:8053–8064.

Ally, D., V. R. Wiss, G. E. Deckert, D. Green, P. Roychoudhury, H. A. Wichman, C. J. Brown, and S. M. Krone. 2014. The impact of spatial structure on viral genomic diversity generated during adaptation to thermal stress. PLoS ONE 9:e88702.

Ammerlaan, H. S. M., S. Harbarth, A. G. M. Buiting, D. W. Crook, F. Fitzpatrick, H. Hanberger, L. A. Herwaldt, P. H. J. Van Keulen, J. A. J. W. Kluytmans, A. Kola, R. S. Kuchenbecker, E. Lingaas, N. Meessen, M. M. Morris-Downes, J. M. Pottinger, P. Rohner, R. P. Dos Santos, H. Seifert, H. Wisplinghoff, S. Ziesing, A. S. Walker, and M. J. M. Bonten. 2013. Secular trends in nosocomial bloodstream infections: Antibiotic-resistant bacteria increase the total burden of infection. Clinical Infectious Diseases 56:798–805.

Amos, W., J. I. Hoffman, A. Frodsham, L. Zhang, S. Best, and A. V. S. Hill. 2007. Automated binning of microsatellite alleles: Problems and solutions. Molecular Ecology Notes 7:10–14.

Anderson, C. D., B. K. Epperson, M. J. Fortin, R. Holderegger, P. M. A. James, M. S. Rosenberg, K. T. Scribner, and S. Spear. 2010. Considering spatial and temporal scale in landscape genetic studies of gene flow. Molecular Ecology 19:3565–3575.

Balkenhol, N., F. Gugerli, S. Cushman, L. Waits, A. Coulon, J. Arntzen, R. Holderegger, and H. Wagner. 2009. Identifying future research needs in landscape genetics: where to from here? Landscape Ecology 24:455–463.

Benveniste, R. and J. Davies. 1973. Mechanisms of antibiotic resistance in bacteria. Annual Review of Biochemistry 42:471–506.

Björk, C. R. 2010. Distribution patterns of disjunct and endemic vascular plants in the interior wetbelt of northwest North America. Botany 88:409–428.

Boles, B. R., M. Thoendel, and P. K. Singh. 2004. Self-generated diversity produces "insurance effects" in biofilm communities. Proceedings of the National Academy of Sciences of the United States of America 101:16630–16635.

Bouma, J. E. and R. E. Lenski. 1988. Evolution of a bacteria/plasmid association. Nature 335:351–352.

Braykov, N. P., M. R. Eber, E. Y. Klein, D. J. Morgan, and R. Laxminarayan. 2013. Trends in resistance to carbapenems and third-generation cephalosporins among clinical isolates of Klebsiella pneumoniae in the United States, 1999-2010. Infection Control & Hospital Epidemiology 34:259–268.

Brown, M. B. and A. B. Forsythe. 1974. Robust tests for the equality of variances. Journal of the American Statistical Association 69:364–367.

Brunsfeld, S. J. and J. Sullivan. 2005. A multi-compartmented glacial refugium in the northern Rocky Mountains: Evidence from the phylogeography of Cardamine constancei (Brassicaceae). Conservation Genetics 6:895–904.

Bull, H. J., M. J. Lombardo, and S. M. Rosenberg. 2001. Stationary-phase mutation in the bacterial chromosome: recombination protein and DNA polymerase IV dependence. Proceedings of the National Academy of Sciences of the United States of America 98:8334–8341.

Carstens, B. C., R. S. Brennan, V. Chua, C. V. Duffie, M. G. Harvey, R. A. Koch, C. D. McMahan, B. J. Nelson, C. E. Newman, J. D. Satler, G. Seeholzer, K. Posbic, D. C. Tank, and J. Sullivan. 2013. Model selection as a tool for phylogeographic inference: An example from the willow Salix melanopsis. Molecular Ecology 22:4014–4028.

Carstens, B. C., S. J. Brunsfeld, J. R. Demboski, J. M. Good, and J. Sullivan. 2005. Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: hypothesis testing within a comparative phylogeographic framework. Evolution 59:1639–1652.

Carstens, B. C. and C. L. Richards. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. Evolution 61:1439–1454.

Carstens, B. C., A. L. Stevenson, J. D. Degenhardt, and J. Sullivan. 2004. Testing nested phylogenetic and phylogeographic hypotheses in the Plethodon vandykei species group. Systematic Biology 53:781–792.

Cavalli-Sforza, L. L. and A. W. F. Edwards. 1967. Phylogenetic analysis. Models and estimation procedures. The American Journal of Human Genetics 19:233–257.

Center for Microbial Systems. 2016. The Ribosomal RNA Database. Retrieved from https://rrndb.umms.med.umich.edu/genomes/.

Centers for Disease Control and Prevention. 2013. Antibiotic resistance threats in the United States. Retrieved from http://www.cdc.gov/drugresistance/threat-report-2013/.

Chapuis, M. P. and A. Estoup. 2007. Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution 24:621–631.

Collins, W. D., C. M. Bitz, M. L. Blackmon, G. B. Bonan, C. S. Bretherton, J. a. Carton, P. Chang, S. C. Doney, J. J. Hack, T. B. Henderson, J. T. Kiehl, W. G. Large, D. S. McKenna, B. D. Santer, and R. D. Smith. 2006. The Community Climate System Model version 3 (CCSM3). Journal of Climate 19:2122–2143.

Costerton, J. W., P. S. Stewart, and E. P. Greenberg. 1999. Bacterial biofilms: a common cause of persistent infections. Science 284:1318–1322.

Dahlberg, C. and L. Chao. 2003. Amelioration of the Cost of Conjugative Plasmid Carriage in Eschericha coli K12. Genetics 165:1641–1649.

Daugherty, C. H. and A. L. Sheldon. 1982. Age-specific movement patterns of the frog Ascaphus truei. Herpetologica 38:468–474.

Davies, J. 1994. Inactivation of antibiotics and the dissemination of resistance genes. Science 264:375–382.

De Gelder, L., J. M. Ponciano, P. Joyce, and E. M. Top. 2007. Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. Microbiology 153:452–463.

De Gelder, L., J. J. Williams, J. M. Ponciano, M. Sota, and E. M. Top. 2008. Adaptive plasmid evolution results in host-range expansion of a broad-host-range plasmid. Genetics 178:2179–2190.

De Gregoris, T. B., N. Aldred, A. S. Clare, and J. G. Burgess. 2011. Improvement of phylum- and class-specific primers for real-time PCR quantification of bacterial taxa. Journal of Microbiological Methods 86:351–356.

Deatherage, D. E. and J. E. Barrick. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods in Molecular Biology 1151:165–188.

Dionisio, F., I. C. Conceição, A. C. R. Marques, L. Fernandes, and I. Gordo. 2005. The evolution of a conjugative plasmid and its ability to increase bacterial fitness. Biology Letters 1:250–252.

Donlan, R. M. 2002. Biofilms: microbial life on surfaces. Emerging Infectious Diseases 8:881–890.

Espíndola, A., L. Pellissier, L. Maiorano, W. Hordijk, A. Guisan, and N. Alvarez. 2012. Predicting present and future intra-specific genetic structure through niche hindcasting across 24 millennia. Ecology Letters 15:649–657.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology 14:2611–2620.

Felsenstein, J. 1976. The theoretical population genetics of variable selection and migration. Annual Review of Genetics 10:253–280.

Ford, L. S. and D. C. Cannatella. 1993. The major clades of frogs. Herpetological Monographs 7:94–117.

Fore, S. A., R. J. Hickey, J. L. Vankat, S. I. Guttman, and R. L. Schaefer. 1992. Genetic structure after forest fragmentation - a landscape ecology perspective on Acer saccharum. Canadian Journal of Botany 70:1659–1668.

Frost, L. S., R. Leplae, A. O. Summers, and A. Toussaint. 2005. Mobile genetic elements: the agents of open source evolution. Nature Reviews Microbiology 3:722–732.

Galimand, M., A. Guiyoule, G. Gerbaud, B. Rasoamanana, S. Chanteau, E. Carniel, and P. Courvalin. 1997. Multidrug resistance in Yersinia pestis mediated by a transferable plasmid. New England Journal of Medicine 337:677–681.

Gavin, D. G. 2009. The coastalâĂŘdisjunct mesic flora in the inland Pacific Northwest of USA and Canada: refugia, dispersal and disequilibrium. Diversity and Distributions 15:972–982.

Gay, K., A. Robicsek, J. Strahilevitz, C. H. Park, G. Jacoby, T. J. Barrett, F. Medalla, T. M. Chiller, and D. C. Hooper. 2006. Plasmid-mediated quinolone resistance in non-Typhi serotypes of Salmonella enterica. Clinical Infectious Diseasess 43:297–304.

Ghigo, J. M. 2001. Natural conjugative plasmids induce bacterial biofilm development. Nature 412:442–445.

Gill, S. R., M. Pop, R. T. Deboy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. a. Relman, C. M. Fraser-Liggett, and K. E. Nelson. 2006. Metagenomic analysis of the human distal gut microbiome. Science 312:1355–1359.

Hammerstrom, T. G., K. Beabout, T. P. Clements, G. Saxer, and Y. Shamoo. 2015. Acinetobacter baumannii repeatedly evolves a hypermutator phenotype in response to tigecycline that effectively surveys evolutionary trajectories to resistance. PLoS ONE 10:e0140489.

Harrison, E., D. Guymer, A. J. Spiers, S. Paterson, and M. A. Brockhurst. 2015. Parallel compensatory evolution stabilizes plasmids across the parasitism-mutualism continuum. Current Biology 25:2034–2039.

Hausner, M. and S. Wuertz. 1999. High rates of conjugation in bacterial biofilms as determined by quantitative in situ analysis. Applied and Environmental Microbiology 65:3710–3713.

Heuer, H., R. E. Fox, and E. M. Top. 2007. Frequent conjugative transfer accelerates adaptation of a broad-host-range plasmid to an unfavorable Pseudomonas putida host. FEMS Microbiology Ecology 59:738–748.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965–1978.

Hoffman, M. D. and A. Gelman. 2014. The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research 15:1593–1623.

Hsueh, P.-R., L.-J. Teng, C.-Y. Chen, W.-H. Chen, C.-J. Yu, S.-W. Ho, and K.-T. Luh. 2002. Pandrug-resistant Acinetobacter baumannii causing nosocomial infections in a university hospital, Taiwan. Emerging Infectious Diseases 8:827–832.

Hunter, S. S., H. Yano, W. Loftie-Eaton, J. Hughes, L. De Gelder, P. Stragier, P. De Vos, M. L. Settles, and E. M. Top. 2014. Draft Genome Sequence of Pseudomonas moraviensis R28-S. Genome Announcements 2:153–154.

Jerassy, Z., A. M. Yinnon, S. Mazouz-Cohen, S. Benenson, Y. Schlesinger, B. Rudensky, and D. Raveh. 2006. Prospective hospital-wide studies of 505 patients with nosocomial bacteraemia in 1997 and 2002. Journal of Hospital Infection 62:230–236.

Jezequel, N., M. C. Lagomarsino, F. Heslot, and P. Thomen. 2013. Long-term diversity and genome adaptation of Acinetobacter baylyi in a minimal-medium chemostat. Genome Biology and Evolution 5:87–97.

Kauffman, S. and S. Levin. 1987. Towards a general theory of adaptive walks on rugged landscapes. Journal of Theoretical Biology 128:11–45.

Koch, A. L. 1974. The pertinence of the periodic selection phenomenon to prokaryote evolution. Genetics 77:127–142.

Korona, R., C. H. Nakatsu, L. J. Forney, and R. E. Lenski. 1994. Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. Proceedings of the National Academy of Sciences of the United States of America 91:9037–41.

Lenski, R. E., M. R. Rose, S. C. Simpson, and S. C. Tadler. 1991. Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations. The American Naturalist 138:1315–1341.

Levin, B. R. 1981. Periodic selection, infectious gene exchange and the genetic structure of E. coli populations. Genetics 99:1–23.

Lewis, K. and K. I. M. Lewis. 2001. Riddle of Biofilm Resistance MINIREVIEW. Antimicrobial Agents and Chemotherapy 45:999–1007.

Liu, Y.-Y., Y. Wang, T. R. Walsh, L.-X. Yi, R. Zhang, J. Spencer, Y. Doi, G. Tian, B. Dong, and X. Huang. 2016. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. The Lancet Infectious Diseases 16:161–168.

Loftie-Eaton, W., H. Suzuki, K. Bashford, H. Heuer, P. Stragier, P. De Vos, M. L. Settles, and E. M. Top. 2015. Draft Genome Sequence of Pseudomonas sp. nov. H2. Genome Announcements 3:e00241–15.

Loftie-Eaton, W., A. Tucker, A. Norton, and E. M. Top. 2014. Flow cytometry and real-time quantitative PCR as tools for assessing plasmid persistence. Applied and Environmental Microbiology 80:5439–5446.

Loftie-Eaton, W., H. Yano, S. Burleigh, R. S. Simmons, J. M. Hughes, L. M. Rogers, S. S. Hunter, M. L. Settles, L. J. Forney, and J. M. Ponciano. 2016. Evolutionary paths that expand plasmid host-range: implications for spread of antibiotic resistance. Molecular Biology and Evolution 33:885–897.

Martens, E. A. and O. Hallatschek. 2011. Interfering waves of adaptation promote spatial mixing. Genetics 189:1045–1060.

Maslunka, C., E. Carr, V. Gürtler, P. Kämpfer, and R. Seviour. 2006. Estimation of ribosomal RNA operon (rrn) copy number in Acinetobacter isolates and potential of patterns of rrn operon-containing fragments for typing strains of members of this genus. Systematic and Applied Microbiology 29:216–228.

Mathers, A. J., H. L. Cox, B. Kitchel, H. Bonatti, A. K. C. Brassinga, J. Carroll, W. M. Scheld, K. C. Hazen, and C. D. Sifri. 2011. Molecular dissection of an outbreak of carbapenem-resistant Enterobacteriaceae reveals intergenus KPC carbapenemase transmission through a promiscuous plasmid. MBio 2:e00204–11.

Mazel, D. and J. Davies. 1999. Antibiotic resistance in microbes. Cellular and Molecular Life Sciences 56:742–754.

McRae, B. H., B. G. Dickson, T. H. Keitt, and V. B. Shah. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89:2712–2724.

Molin, S. and T. Tolker-Nielsen. 2003. Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. Current Opinion in Biotechnology 14:255–261.

Mosca, E., A. J. Eckert, E. A. Di Pierro, D. Rocchini, N. La Porta, P. Belletti, and D. B. Neale. 2012. The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. Molecular Ecology 21:5530–5545.

Murphy, M. A., J. S. Evans, and A. Storfer. 2010. Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. Ecology 91:252–261.

Nei, M. and F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97:145–163.

Nevo, E. 1978. Genetic variation in natural populations: Patterns and theory. Theoretical Population Biology 13:121–177.

Nielson, M., K. Lohman, C. H. Daugherty, F. W. Allendorf, K. L. Knudsen, J. Sullivan, D. J. Ellis, B. T. Firth, and I. Belan. 2006. Allozyme and Mitochondrial Dna Variation in the Tailed Frog (Anura: Ascaphus): the Influence of Geography and Gene Flow. Herpetologica 62:235–258.

Nielson, M., K. Lohman, and J. Sullivan. 2001. Phylogeography of the tailed frog (Ascaphus truei): implications for the biogeography of the Pacific Northwest. Evolution 55:147–160.

Obama, B. 2014. State of the Union Address. Retrieved from https://www.whitehouse.gov/the-press-office/2014/01/28/president-barack-obamas-state-union-address.

Parsek, M. R. and P. K. Singh. 2003. Bacterial biofilms: an emerging link to disease pathogenesis. Annual Review of Microbiology 57:677–701.

Pendleton, J. N., S. P. Gorman, and B. F. Gilmore. 2013. Clinical relevance of the ESKAPE pathogens. Expert Review of Anti-Infective Therapy 11:297–308.

Perez, F., A. M. Hujer, K. M. Hujer, B. K. Decker, P. N. Rather, and R. A. Bonomo. 2007. Global challenge of multidrug-resistant Acinetobacter baumannii. Antimicrobial Agents and Chemotherapy 51:3471–3484.

Peterson, A. T., J. Soberón, R. G. Pearson, R. P. Anderson, E. Martínez-Meyer, M. Nakamura, and M. Bastos Araujo. 2011. Ecological niches and geographic distributions, vol. 49. Princeton University Press, Princeton.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231–259.

Poirel, L., N. Kieffer, N. Liassine, D. Thanh, and P. Nordmann. 2016. Plasmid-mediated carbapenem and colistin resistance in a clinical isolate of Escherichia coli. The Lancet Infectious Diseases 16:281.

Poirel, L. and P. Nordmann. 2006. Carbapenem resistance in Acinetobacter baumannii: Mechanisms and epidemiology. Clinical Microbiology and Infection 12:826–836.

Ponciano, J. M., H.-J. La, P. Joyce, and L. J. Forney. 2009. Evolution of diversity in spatially structured Escherichia coli populations. Applied and Environmental Microbiology 75:6047–6054.

Porse, A., K. Schønning, C. Munck, and M. O. A. Sommer. 2016. Survival and evolution of a large multidrug resistance plasmid in new clinical bacterial hosts. Molecular Biology and Evolution 33:2860–2873.

Poulsen, L. K., G. Ballard, and D. A. Stahl. 1993. Use of rRNA fluorescence in situ hybridization for measuring the activity of single cells in young and established biofilms. Applied and Environmental Microbiology 59:1354–1360.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–59.

Pyron, R. A. and J. J. Wiens. 2011. A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. Molecular Phylogenetics and Evolution 61:543–583.

Rainey, P. B. and M. Travisano. 1998. Adaptive radiation in a heterogeneous environment. Nature 394:69–72.

Raymond, M. and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248–249.

Richet, H. and P. E. Fournier. 2006. Nosocomial Infections Caused by Acinetobacter baumannii A Major Threat Worldwide. Infection Control 27:645–646.

Roesch, L., R. Fulthorpe, A. Riva, G. Casella, A. Hadwin, A. Kent, S. Daroub, F. Camargo, W. Farmerie, and E. Triplett. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. The ISME Journal 1:283–290.

Rossetto, M., C. Allen, K. Thurlby, P. Weston, and M. Milner. 2012. Genetic structure and bioclimatic modeling support allopatric over parapatric speciation along a latitudinal gradient. BMC Evolutionary Biology 12:149.

Rousset, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8:103–106.

Sambrook, J. and D. W. Russell. 2001. Molecular cloning: a laboratory manual 3rd edition. Coldspring-Harbour Laboratory Press, Coldspring Harbour.

San Millan, A., R. Peña-Miller, M. Toll-Riera, Z. V. Halbert, A. R. McLean, B. S. Cooper, and R. C. MacLean. 2014. Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids. Nature Communications 5:5208.

Schlüter, A., H. Heuer, R. Szczepanowski, L. J. Forney, C. M. Thomas, A. Pühler, and E. M. Top. 2003. The 64 508 bp IncP-1β antibiotic multiresistance plasmid pB10 isolated from a wastewater treatment plant provides evidence for recombination between members of different branches of the IncP-1β group. Microbiology 149:3139–3153.

Schmidt, T., P. Arens, M. J. Smulders, R. Billeter, J. Liira, I. Augenstein, and W. Durka. 2009. Effects of landscape structure on genetic diversity of Geum urbanum L. populations in agricultural landscapes. Flora - Morphology, Distribution, Functional Ecology of Plants 204:549–559.

Searcy, C. A. and H. B. Shaffer. 2014. Field validation supports novel niche modeling strategies in a cryptic endangered amphibian. Ecography 37:983–992.

Selim, S. A. and N. I. Hagag. 2013. Analysis of Plasmids and Restriction Fragment Length Polymorphisms of Acinetobacter baumannii Isolated from Hospitals-AL Jouf Region-KSA. *In* Proceedings of World Academy of Science, Engineering and Technology, 78, p. 153. World Academy of Science, Engineering and Technology (WASET).

Sexton, J. P., S. B. Hangartner, and A. A. Hoffmann. 2014. Genetic isolation by environment or distance: Which pattern of gene flow is most common? Evolution 68:1–15.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. Science 236:787–792.

Sork, V. L. and L. Waits. 2010. Contributions of landscape genetics - Approaches, insights, and future potential. Molecular Ecology 19:3489–3495.

Sota, M., H. Yano, J. M. Hughes, G. W. Daughdrill, Z. Abdo, L. J. Forney, and E. M. Top. 2010. Shifts in the host range of a promiscuous plasmid through parallel evolution of its replication initiation protein. The ISME Journal 4:1568–1580.

Spear, S. F. and A. Storfer. 2008. Landscape genetic structure of coastal tailed frogs (Ascaphus truei) in protected vs. managed forests. Molecular Ecology 17:4642–4656.

———. 2010. Anthropogenic and natural disturbance lead to differing patterns of gene flow in the Rocky Mountain tailed frog, Ascaphus montanus. Biological Conservation 143:778–786.

Stalder, T. and E. Top. 2016. Plasmid transfer in biofilms: a perspective on limitations and opportunities. npj Biofilms and Microbiomes 2:16022.

Subbiah, M., E. M. Top, D. H. Shah, and D. R. Call. 2011. Selection pressure required for long-term persistence of blaCMY-2-positive IncA/C plasmids. Applied and Environmental Microbiology 77:4486–4493.

Sunenshine, R. H., M.-O. Wright, L. L. Maragakis, A. D. Harris, X. Song, J. Hebden, S. E. Cosgrove, A. Anderson, J. Carnell, and D. B. Jernigan. 2007. Multidrug-resistant Acinetobacter infection mortality rate and length of hospitalization. Emerging Infectious Diseases 13:97–103.

Theobald, D. M., J. B. Norman, E. Peterson, S. Ferraz, A. Wade, and M. R. Sherburne. 2006. Functional Linkage of Water basins and Streams (FLoWS) v1 User's Guide: ArcGIS tools for Network-based analysis of freshwater ecosystems. Natural Resource Ecology Lab, Colorado State University, Fort Collins, CO.

Thomas, J. A., N. A. D. Bourn, R. T. Clarke, K. E. Stewart, D. J. Simcox, G. S. Pearman, R. Curtis, and B. Goodger. 2001. The quality and isolation of habitat patches both determine where butterflies persist in fragmented landscapes. Proceedings of the Royal Society of London. Series B: Biological Sciences 268:1791–1796.

Turner, P., V. Cooper, and R. Lenski. 1998. Tradeoff between horizontal and vertical modes of transmission in bacterial plasmids. Evolution 52:315–329.

Tyerman, J. G., J. M. Ponciano, P. Joyce, L. J. Forney, and L. J. Harmon. 2013. The evolution of antibiotic susceptibility and resistance during the formation of Escherichia coli biofilms in the absence of antibiotics. BMC Evolutionary Biology 13:1–7.

Van Domselaar, G. H., P. Stothard, S. Shrivastava, J. A. Cruz, A. C. Guo, X. Dong, P. Lu, D. Szafron, R. Greiner, and D. S. Wishart. 2005. BASys: A web server for automated bacterial genome annotation. Nucleic Acids Research 33.

Villers, D., E. Espaze, M. Coste-Burel, F. Giauffret, E. Ninin, F. Nicolas, and H. Richet. 1998. Nosocomial Acinetobacter baumannii infections: microbiological and clinical epidemiology. Annals of Internal Medicine 129:182–189.

Vlamakis, H., Y. Chai, P. Beauregard, R. Losick, and R. Kolter. 2013. Sticking together: building a biofilm the Bacillus subtilis way. Nature Reviews Microbiology 11:157–168.

Vogwill, T., R. L. Phillips, D. R. Gifford, R. C. MacLean, G. Conte, M. Arnegard, C. Peichel, D. Schluter, D. Stern, T. Wood, J. Burke, L. Rieseberg, T. Lieberman, R. Marvig, D. Dolce, L. Sommer, B. Petersen, O. Ciofu, S. Campana, S. Molin, G. Taccetti, H. Johansen, R. Marvig, L. Sommer, S. Molin, H. Johansen, R. Dahan, R. Duncan, A. Wilson, L. Dávalos, J. McCutcheon, B. McDonald, N. Moran, J. Wernegreen, J. Blair, M. Webber, A. Baylay, D. Ogbolu, L. Piddock, T. Vogwill, M. Kojadinovic, V. Furió, R. MacLean, J. Losos, T. Ord, T. Summers, E. Breidenstein, B. Khaira, I. Wiegand, J. Overhage, R. Hancock, A. Dötsch, T. Becker, C. Pommerenke, Z. Magnowska, L. Jänsch, S. Häussler, A. Fajardo, A. Liu, L. Tran, E. Becket, K. Lee, L. Chinn, E. Park, K. Tran, J. Miller, J. de Visser, J. Krug, J. D. Visser, D. Rozen, I. Szendro, J. Franke, J. de Visser, J. Krug, N. Colegrave, W. Hill, A. Robertson, P. Gerrish, R. Lenski, G. Bell, R. Gomulkiewicz, R. Holt, L. Wahl, P. Gerrish, L. Wahl, P. Gerrish, I. Saika-Voivod, J. Lachapelle, J. Reid, N. Colegrave, D. Rozen, M. Habets, A. Handel, J. de Visser, S. Schoustra, T. Bataillon, D. Gifford, R. Kassen, C. Miller, P. Joyce, H. Wichman, S. Bailey, N. Rodrigue, R. Kassen, J. Dettman, N. Rodrigue, A. Melnyk, A. Wong, S. Bailey, R. Kassen, A. S. Millan, R. Peña-Miller, M. Toll-Riera, Z. Halbert, A. McLean, B. Cooper, R. MacLean, M. DePristo, H. Li, K. Chen, K. Ye, M. Schulz, Q. Long, R. Apweiler, Z. Ning, V. Boeva, T. Popova, K. Bleakley, P. Chiche, J. Cappo, G. Schleiermacher, I. Janoueix-Lerosey, O. Delattre, E. Barillot, P. Cingolani, A. Platts, L. Wang, M. Coon, T. Nguyen, L. Wang, S. Land, X. Lu, D. Ruden, G. Winsor, D. Lam, L. Fleming, R. Lo, M. Whiteside, N. Yu, R. Hancock, F. Brinkman, P. Jaccard, M. Anderson, M. Anderson, B. Goldstein, C. Dorel, P. Lejeune, A. Rodrigue, T. Mahoney, T. Silhavy, S. Suzuki, T. Horinouchi, C. Furusawa, N. Weatherspoon-Griffin, D. Yang, W. Kong, Z. Hua, Y. Shi, E. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb, S. Darst, R. MacLean, A. Buckling, T. Bataillon, T. Zhang, R. Kassen, R. Kassen, T. Bataillon, M. Webber, R. Whitehead, M. Mount, N. Loman, M. Pallen, L. Piddock, S. Kryazhimskiy, J. Plotkin, and M. Farhat. 2016. Divergent evolution peaks under intermediate population bottlenecks during bacterial experimental evolution. Proceedings of the Royal Society B: Biological Sciences 283:751–764.

Wang, I. J. and K. Summers. 2010. Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. Molecular Ecology 19:447–458.

Wang, L., D. Fan, W. Chen, and E. M. Terentjev. 2015. Bacterial growth, detachment and cell size control on polyethylene terephthalate surfaces. Scientific Reports 5:1–11.

Warren, D. L., A. N. Wright, S. N. Seifert, and H. B. Shaffer. 2014. Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. Diversity and Distributions 20:334–343.

Wolcott, R. D., D. D. Rhoads, M. E. Bennett, B. M. Wolcott, L. Gogokhia, J. W. Costerton, and S. E. Dowd. 2010. Chronic wounds and the medical biofilm paradigm. Journal of Wound Care 19:45–46.

Wolin, E. A., M. Wolin, and R. S. Wolfe. 1963. Formation of methane by bacterial extracts. Journal of Biological Chemistry 238:2882–2886.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *In* Proceedings of The Sixth Congress on Genetics, pp. 356–366.

Zhang, P., D. Liang, R.-L. Mao, D. M. Hillis, D. B. Wake, and D. C. Cannatella. 2013. Efficient sequencing of anuran mtDNAs and a mitogenomic exploration of the phylogeny and evolution of frogs. Molecular Biology and Evolution 30:1899–1915.

APPENDIX A

Supplementary Information to Chapter 2

The following pages provide supplemental data for Chapter 2.

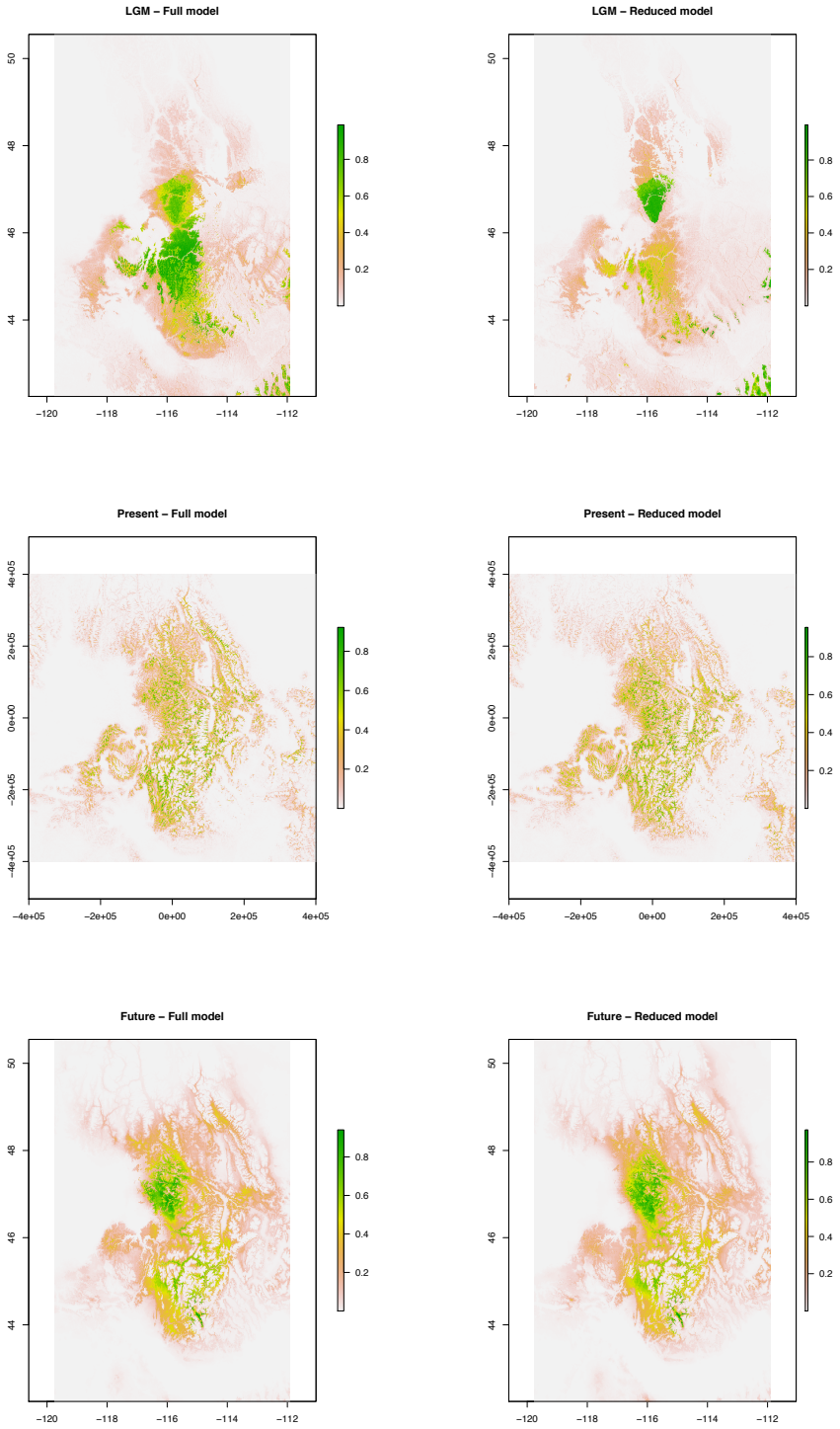

FIGURE A.1: Comparison of SDMs calculated from the full set vs. reduced set of climatic variables. Top. Full vs. Reduced data sets for LGS SDM; Middle. Full vs. Reduced data sets for current SDM; Bottom. Full vs. Reduced data sets for 2070 SDM.

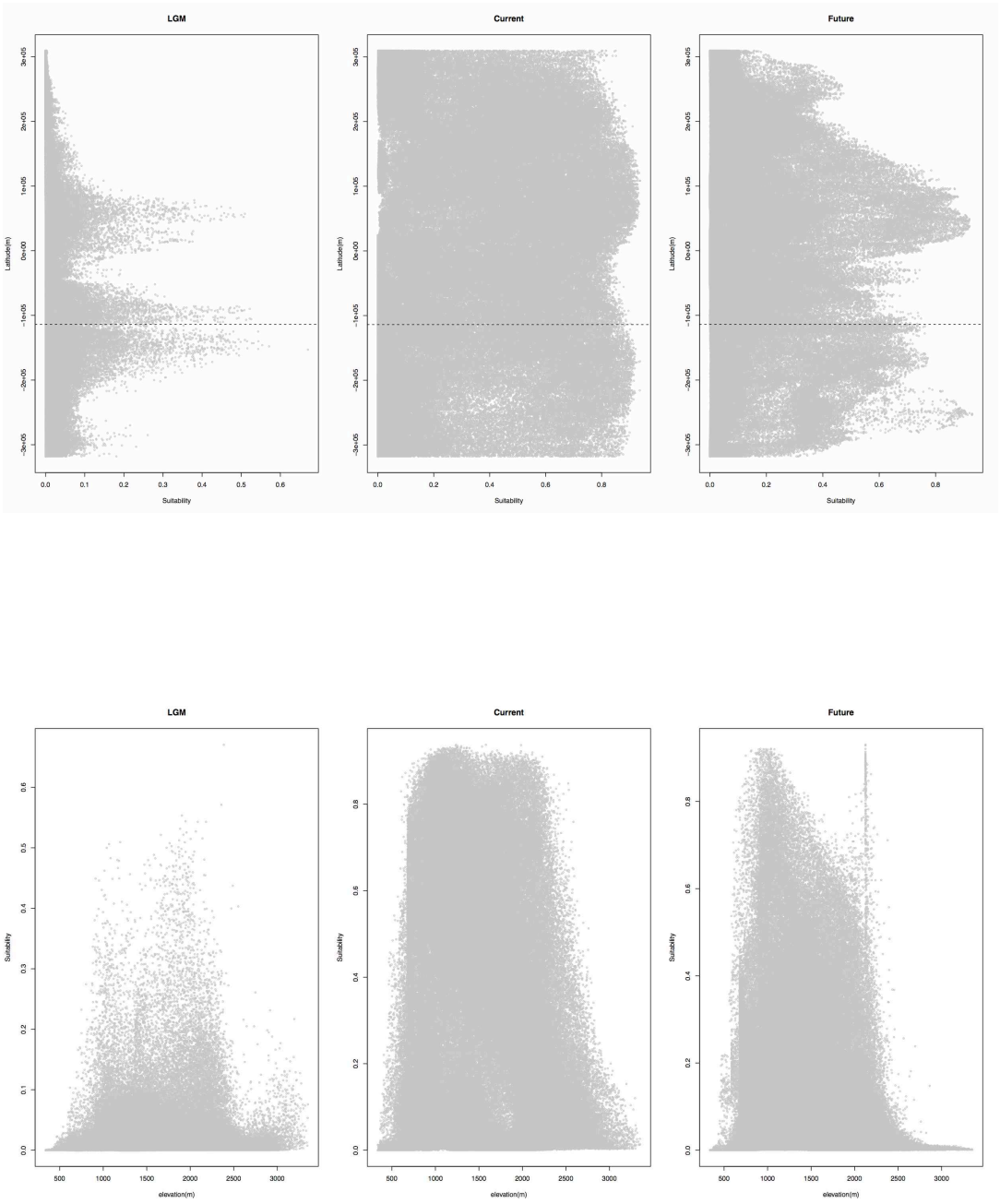

FIGURE A.2: Top. Distribution of suitabilities by latitude for (left) the Latest Glacial Maximum, (middle) the current conditions, and (right) the 2070 ENM forecast. The horizontal line is at the contact between northern and southern genetic clusters at K=2, as well as mtDNA clades.

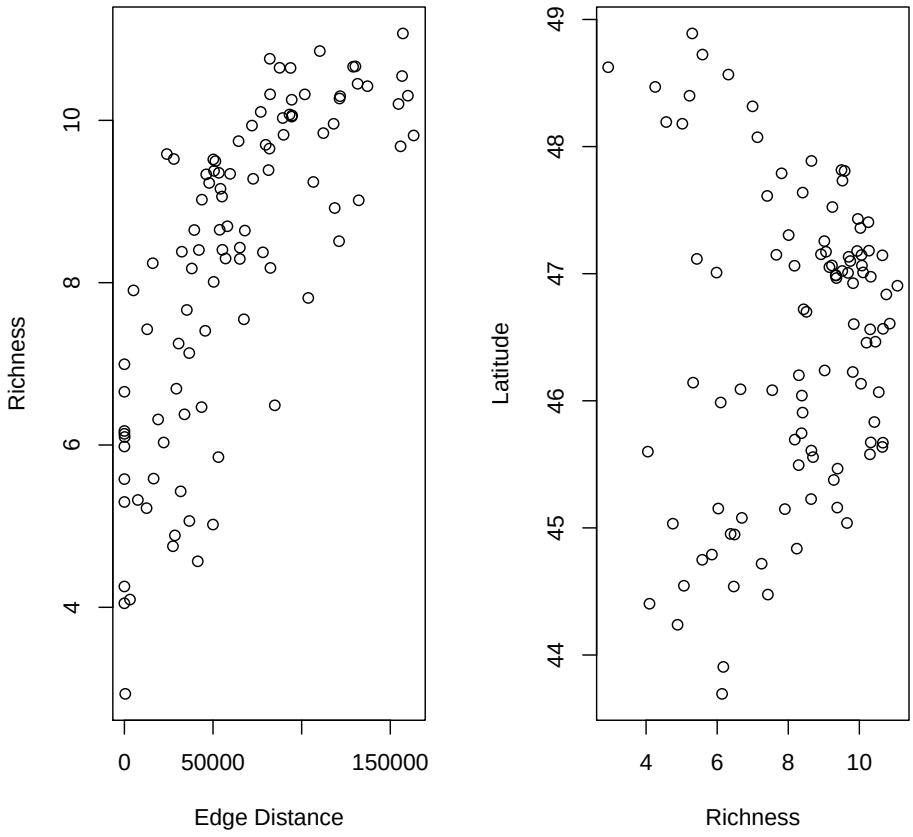

FIGURE A.3: Allelic richness versus A. Distance from edge of range, and B. Latitude.

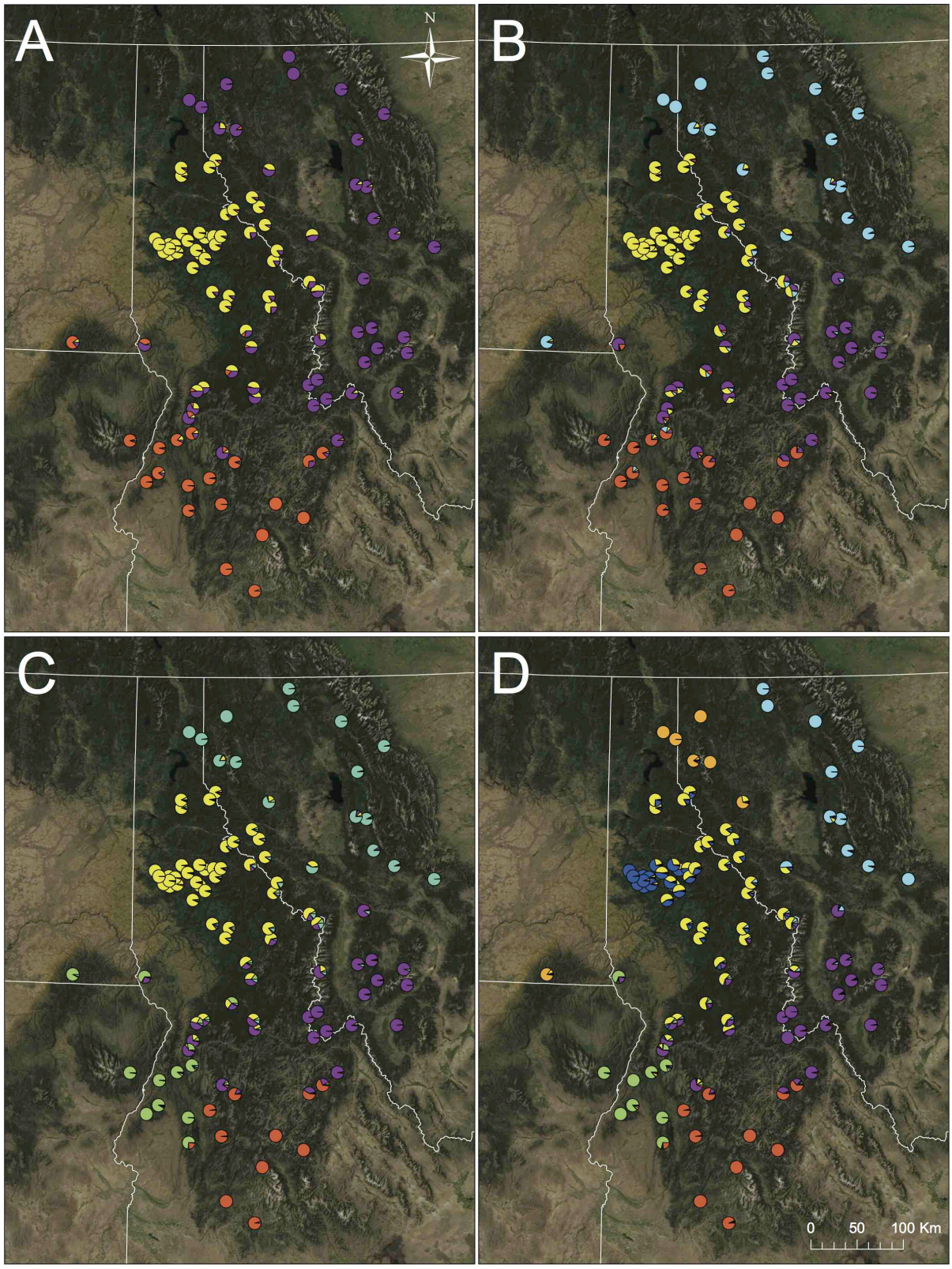

FIGURE A.4: Hierarchical pattern of clusters from Structure analyses with A. $K = 3$, B. $K = 4$, C. $K = 5$, and D. $K = 7$ under a co-dominant admixture model.

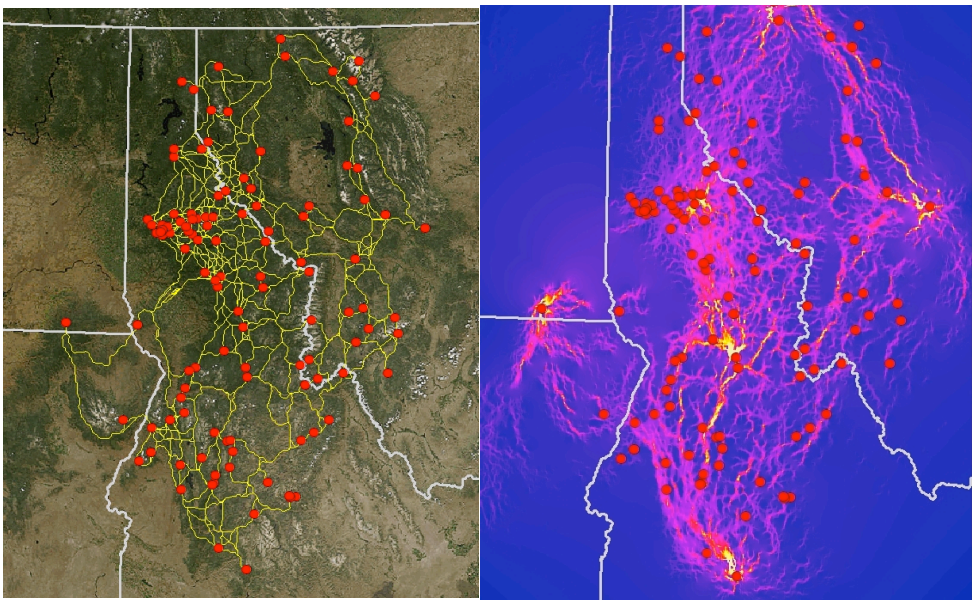

FIGURE A.5: Visualization of (left) Least-Cost Paths, and (right) CircuitScape resistance at 10X suitabilities.

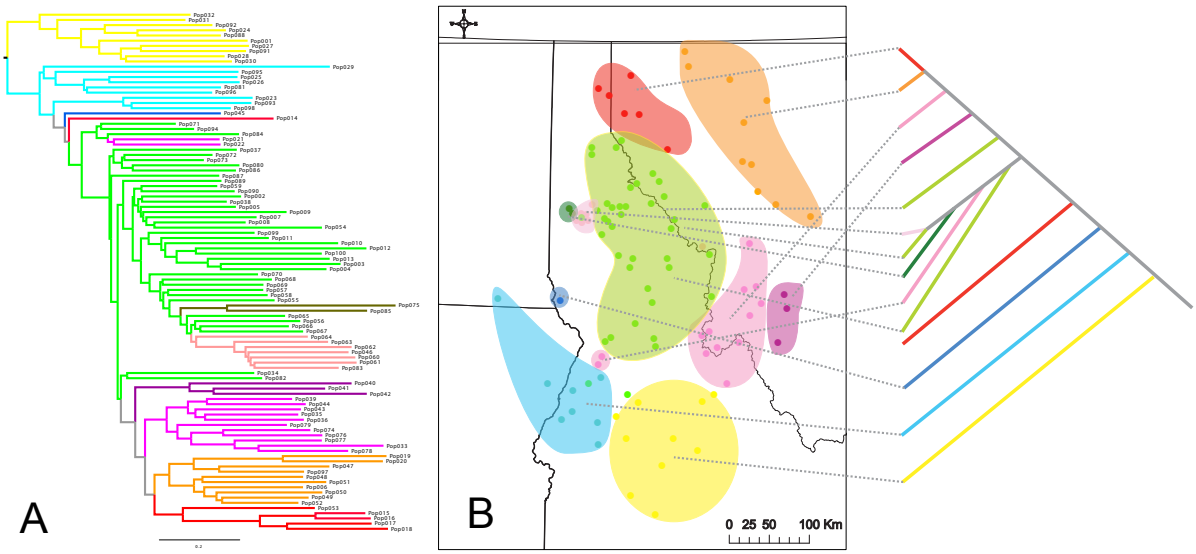

FIGURE A.6: A, neighbor-joining tree calculated from cord distances. Colors correspond to the result of Structure analysis with K=10 (B), with major lineages reduced and mapped geographically.

APPENDIX B

Supplementary Information to Chapter 3

B.1 Construction of ancestral *Acinetobacter baumannii* host-plasmid pair

To allow the host to adapt to the culture environment prior to experimental evolution with plasmid pB10, the archived Acinetobacter baumannii strain ATCC17978 was grown for 10 days with daily passage of 0.1% of the volume in fresh mineral medium (MBM) supplemented with 18.5 mM succinate as the main carbon source and 2 g/L casamino acids, hereafter named MBMS. MBM consists of 1X M9 salts (Sambrook and Russell, 2001), amended with 10 mL of a vitamin stock solution, and 10 mL trace element stock solution per liter (Wolin et al., 1963).

Following the medium adaptation phase, plasmid pB10 was introduced into *A. baumannii* by electroporation with selection on tetracycline (tet, 10 μg/ml). After electroporation 10 colonies were selected and grown overnight in MBMS with tetracycline (10 Îijg/ml), hereafter called MBMS-tet; plasmid DNA extractions were performed to determine the presence of pB10. Clones that appeared to contain intact pB10 were digested with the restriction enzymes PstII and HindIII and the patterns were compared to a control of pB10 DNA extracted from *Escherichia coli* to confirm the presence of intact pB10. The presence of full size pB10 was confirmed in one clone, which was then selected as the ancestor for all experiments.

To obtain a large liquid culture that contained as little genetic diversity among cells as possible, an extinction-dilution procedure was performed. A dilution series was made from 10^{-1} to 10^{-10} and used to inoculate large volumes of liquid media. These cultures were then allowed to grow until approximately 8 hours after the cultures inoculated from the 10^{-8} dilution showed turbidity. Because none of the more dilute cultures showed growth in that time period, the 10-8 culture was selected as the ancestral stock and 0.65 mL aliquots were archived at -70°C with 30% glycerol.

B.2 EXPERIMENTAL EVOLUTION

Biofilms were grown in flow cells that were set up in a manner similar to that of Ponciano et al. (2009), with some modifications described here and in the main text of this manuscript. The flow cell chambers and top and bottom slides were made of polycarbonate plastic and the chambers were sealed using silicone adhesive. The flow cells were inoculated with 200 μL of our archived ancestral culture, initially clamped for 24 hours to allow the cells to settle, and then subsequently unclamped to initiate a flow of MBMS-tet at a rate of 5.4 mL/hr. The flow cells, media, and waste bottles were contained within an incubator that was maintained at 37°C. Media and waste bottles were changed approximately every 2-3 days as needed. All flow cells were checked daily for overgrowth into the tubing supplying the media and for leaks. Leaking flow cells were removed and tubing and filters with substantial overgrowth were replaced as needed.

To harvest the biofilms the top plate of the flow cell was removed with a sterile scalpel blade. The media supernatant was pipetted off and the biofilm population was resuspended in 2 mL of phosphate buffer saline (PBS). A portion of the suspension was diluted in PBS and plated onto LB agar supplemented with tet (LB-tet). The remainder of the harvested cell suspension was combined with glycerol and stored in 0.65 mL aliquots in the -70°C freezer.

On day 4, 4.9 μL portions of each of the three t_0 biofilm cell suspensions were used to inoculate each of three test tubes containing 5 mL of MBMS-tet (that is, each t_0 biofilm was used to inoculate one of the liquid culture replicates); this represented the starting point of the liquid batch cultures. These liquid cultures were then grown at 37°C in a shaking incubator in serial batch cultures: 4.9μL was transferred into 5 mL of MBMS-tet every 24 ± 1 hours. Of each liquid cultures a 1 mL sample was archived every 5 days, as well as at 14 days (t_{14}) and 28 days (t_{28}) after inoculation (t_0). Here we only report on the populations harvested at t_{28}.

Six colonies were randomly chosen from the LB-tet agar plates for each biofilm or liquid culture harvested at both the t_0 and t_{28} time points. They were grown overnight in liquid MBMS-tet before being archived at -70°C with glycerol. The presence of pB10 was confirmed in these clones using a plasmid extraction protocol, followed by restriction enzyme digestion and gel electrophoresis. These clones were then tested for plasmid persistence.

B.3 PLASMID EXTRACTION PROCEDURE

We used the plasmid extraction procedure described in Selim and Hagag (2013) with the following modifications: 1) the volume of TE, lysis solution, and 2M Tris-HCl was doubled, 2) RNAse was added along with the 100 μL of TE before the lysis step, and 3) centrifugation at 20,000xg and 4°C was done for 45 minutes after adding Tris-HCl, followed by transferring the supernatant into phenol-chloroform.

B.4 DNA PURIFICATION AND QPCR CONDITIONS

To purify DNA for qPCR we used the DNA extraction protocol provided with QIAsymphony DSP DNA Mini kits. Briefly, 5 μL of liquid overnight cultures of *A. baumannii* was added to 220 μL of buffer ATL and transferred to 2 mL barcoded micro tubes. Next, 20 μL of proteinase K was added to the samples, followed by a brief vortexing and 1 hr of shaking at 900 rpm in a shaker-incubator (at 56°C). Later, the samples were briefly spun down to remove the condensation from tube caps. Four microliters of RNAse A (100 mg/ml) was then added and followed by a 2 min incubation period at room temperature. All tubes with cellular lysate were then transferred to the sample carrier trays for the QIAsymphony SP instrument set to run a DNA Blood & Tissue LC200 protocol (default parameters) with a 50 μL elution volume.

The amplification parameters for all qPCR reactions were 94°C for 10 min, 40 cycles of 94°C for 15 s, and 60°C for 60 s. Parameters for melting curves were 94°C for 15 s and 60°C for 30 s, followed by a temperature increase to 94°C with 0.1% ramp rate. The fluorescent signal was acquired after each 60°C amplification step and collected continuously during the melting curve analysis. 0.2 ng of template DNA was used per qPCR reaction, with every sample assayed in triplicate. On plasmid pB10 gene *trfA*, encoding the replication initiation protein, was amplified using primers 5'-GAACAGCACCACGATTTCGG-3' (forward) and 5'-TACTACACAAG GGCCGAGGA-3' (reverse). The *trfA* copies were compared to the *16S* rRNA genes in the *A. baumannii* chromosome, amplified using primers 1080γF (5'-TCGTCAGCTCGTGTYGTGA-3') and γ1202R (5'-CGTAAGGGCCATGATG-3') described previously by De Gregoris et al. (2011).

B.5 DATA ANALYSIS

Analysis of the raw qPCR fluorescence data was done using a Bayesian hierarchical model. We implemented our model in the Stan programming language and the R statistical programming language (Stan is available in R via the package RStan; Hoffman and Gelman, 2014). Stan utilizes a Hamilontian Monte Carlo (HMC) algorithm to sample posterior distributions of model parameters. Conceptually, the model being used is based on the simple enzyme kinetic model $x + y \rightarrow 2x$ where x is the concentration of template DNA and y is concentration of PCR primer. This model has the time dependent solution

$$x(t) = \frac{x_0 + y_0}{1 + \left(\frac{x_0}{y_0}\right) e^{k(x_0+y_0)t}} \tag{B.1}$$

where x_0 is the starting DNA concentration, y_0 is the starting primer concentration, k is the rate at which Taq polymerase catalyzes the reaction, and t is time measured in PCR cycles. We estimate the quantity x_0 because it represents the initial concentration of either *trfA* or *16S* rRNA. A hierarchical model allowed us to model the error to account for the fact that the error of replicate retention assays should be correlated with the error of the qPCR replicates, thus resulting in more accurate parameter estimates. The HMC sampler was run for 4 independent chains consisting of 2000 generations, the first 1000 of which were discarded as a burn-in period; further details can be found in the documentation for the Stan language (http://mc-stan.org). Potential scale reduction factor (PSRF) values were checked to verify that convergence was obtained across the 4 chains.

The qPCR-based *trfA*/*16S* rRNA ratio estimates were used as a measure of the fraction of plasmid-bearing cells. Specifically, we utilized a log-linear model to determine the rate at which plasmids were lost from *A. baumannii* cultures and the variation in that rate. The simple linear model was

$$log\left(\frac{trfA}{16S\ rRNA}\right) = day + environment + day\ x\ environment + error \tag{B.2}$$

The log transformation of the ratio improved model diagnostics and corrected for heteroscedasticity in residuals. Significant day ÃŮ environment terms indicate differences in the rate of plasmid loss between growth environments. Furthermore, the residual variance for each environment ($\sigma_{error|environment}$) from this model is an estimate of the biodiversity of plasmid persistence phenotypes.

APPENDIX C

Supplementary Information to Chapter 4

The following pages provide supplemental data for Chapter 4.

Table C.1: Summary of mutations, affected genes, and proposed gene functions for mutant versions of pB10 found in all clones isolated from the Tet+ arm of our experiment. * = Plasmid tested in other species

Antibiotics	Environment	Time Point	Gene(s)	Mutations	Proposed Function
Tet+	Starting Point	T0	kfrA	Δ27 bp; Δ141 bp	KfrA protein
			trbL	Δ405 bp	putative DNA topoisomerase
	Biofilm	T14	trbL	Δ 511; 694	putative DNA topoisomerase
		T28	[trbB]–[traN]	Δ32,730 bp	34 genes
			[trbD]–[kfrA]	Δ33,164 bp	34 genes
			[trbE]–[traN]	Δ 30395; *Delta*30,220 bp	31 genes
			trbB–[traL]	Δ31,764 bp	32 genes
			trbJ	Δ9 bp	conjugal transfer protein TrbJ
	Liquid	T14	[trbE]–[kfrA]	Δ30,679 bp (x6)*	33 genes
			[trbE]–[traM]	Δ28,988 bp (x5)	30 genes
			[trbF]–[kfrA]	Δ29,446 bp (x6); 30317	32 genes
			kfrA	Δ21 bp	KfrA protein
			xf2080 / tetR	SNP	hypothetical protein / tetracycline resistance repressor protein
		T28	[trbE]–[kfrA]	Δ30,679 bp (x6)	33 genes
			[trbE]–[traM]	Δ28,988 bp (x6); *Delta*28,549 bp (x6)*	30 genes
			kfrA	Δ84 bp; Δ178 bp	KfrA protein

Table C.2: Summary of mutations, affected genes, and proposed gene functions for mutant versions of pB10 found in all clones isolated from the Tet- arm of our experiment.

Antibiotics	Environment	Time Point	Gene(s)	Mutations	Proposed Function
Tet-	Starting point	T0	kfrA	Δ84 bp	KfrA protein
			klcB	Δ144 bp	KlcB protein
	Biofilm	T14	trbL	Δ219 bp, Δ306 bp	putative DNA topoisomerase
			trbI	Δ648 bp	conjugal transfer protein TrbI
			[trbL]–[trbN]	Δ1,881 bp	3 genes
			[trbB]–[traM]	Δ31,451 bp	33 genes
			[kfrA]–[korB]	Δ1,202 bp	2 genes
			trbB	SNP	conjugal transfer protein TrbB
	Liquid	T6	[trbE]–[incC2]	Δ34,594 bp (x4)	36 genes
		T10	[trbA]–[traD]	Δ23,369 bp (x4)	25 genes
			[trbE]–[kfrA]	Δ31,221 bp	33 genes
			[trbE]–[kfrA]	Δ31517	33 genes
		T14	[trbC]–[kfrA]	Δ32,964 bp	35 genes