5-2011

Molecular evolution and historical biogeography of new world birds

Brian T. Smith
University of Nevada, Las Vegas

MOLECULAR EVOLUTION AND HISTORICAL BIOGEOGRAPHY OF NEW WORLD BIRDS

by

Brian Tilston Smith

Bachelor of Science
University of Nevada, Las Vegas
2005

A dissertation submitted in partial fulfillment
of the requirements for the

Doctor of Philosophy in Biological Sciences
School of Life Sciences
College of Sciences

Graduate College
University of Nevada, Las Vegas
May 2011

ABSTRACT

Molecular Evolution and Historical Biogeography of New World Birds

by

Brian Tilston Smith

Dr. Javier Rodriguez, Examination Committee Co-chair
Professor of Biology
University of Nevada, Las Vegas

Dr. John Klicka Examination Committee Co-Chair
Adjunct Faculty of Biology
University of Nevada, Las Vegas

Deciphering the patterns of how biodiversity has evolved across time and space has remained a fundamental objective for biologists for the last 200 years. Researchers are faced with the challenge of interpreting the complexity of evolutionary patterns that have been generated over the deep history of the Earth. The advancement of DNA sequencing technology has yielded a new and powerful genetic toolkit that has allowed biologists to address novel evolutionary questions. For my dissertation research, I used molecular genetics and a statistical framework to study the evolution and historical biogeography of birds distributed in North and South America. My dissertation encompasses three separate research projects that examined the relationship between changes in the Earth's climate and geology and the evolution of avian taxa. I examined evolutionary processes that operate on three different biological scales: within populations, among populations, and species' assemblages. For my first research project, I used a bird species as a model to determine the processes responsible for generating genetic diversity and structure in populations (Chapter 2). I showed that genetic diversity is not just affected by processes that operate over long evolutionary periods, but

additionally genetic diversity is impacted by on-going processes. In my second research project, I evaluated the relationship between population size and the rate at which populations evolve (Chapter 3). For this project, I developed a new analytical framework using population genetic theory to show that small populations evolve faster than large populations. For my final research project, I showed how the formation of a land bridge between North and South America was critical for birds to disperse between continents (Chapter 4). Overall, my dissertation provides new insight into avian evolution by identifying processes across time and space linked to the genetic patterns observable in the birds of North and South America.

ACKNOWLEDGMENTS

I thank my dissertation committee, John Klicka, Javier Rodriguez, Jef Jaeger, Brett Riddle, and Amei Amei, for all their support, encouragement, and input during my tenure as a PhD student. The UNLV systematics group, SABR, and the Klicka lab were instrumental in my development as a student, and I am extremely grateful to them. I would especially like to thank Dawn Fletcher, Garth Spellman, Jeff DaCosta, Jeremey Batten, Rob Bryson Jr., Mavina Lim, Lia Africa, Tereza Jezkova, Derek Houston, Mallory Eckstut, Matt Graham, and Rebecca Carson. I would like to thank Nicholle Booker and Andy Andres with helping me navigate through the graduation paperwork maze. My family has provided continued support throughout graduate school and I thank them.

The following individuals and institutions provided critical samples for research: Patricia Escalante (Instituto de Biología, Universidad Nacional Autónoma de México), Blanca Hernández Baños and Adolfo G. Navarro-Sigüenza (Museo de Zoología, Facultad de Ciencias, Universidad Nacional Autónoma de México) Paul Sweet and George Barrowclough (American Museum of Natural History), Christopher Witt and Andrew Jones (Museum of Southwest Biology, University of New Mexico), Robert Zink (Bell Museum of Natural History), Sievert Rohwer and Sharon Birks (U of Washington, Burke Museum of Natural History), and Mark Robbins (University of Kansas). The following institutions that funded my research: a UNLV International Studies Scholarship, a UNLV GPSA grant, a UNLV President's Fellowship, and an American Museum of Natural History Chapman Grant.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The field of biogeography seeks to understand how species have evolved across time and space. A consensus of research has indicated that the evolutionary history of species is strongly linked to the Earth's geologic and climatic history. However, biogeographers are not able to directly observe how Earth history has impacted the evolutionary processes of speciation and extinction. Instead, researchers rely on deciphering patterns from living plants, animals, microbes, and fungi to reconstruct and test hypotheses on how species have evolved. There have been a number of tools developed that can be used to infer biogeographic and evolutionary patterns, which include measuring different aspects of species biology. Traditionally, researchers have used tools such as, morphological measurements, species geographic ranges, and the fossil record to to examine how species have evolved. More recently the development and widespread availability of DNA sequencing technology has provided a new and powerful resource for evolutionary studies. Changes to the DNA sequence known as mutations are recorded in species as they descend through time. The pattern of mutations recorded in DNA can be used to address various evolutionary questions at different spatial and temporal scales.

Biogeographers study evolution from micro-scales that look at the distribution of genetic variation within populations, to macro-scales that examine entire assemblages of species. For my dissertation research, I looked at the molecular evolution and historical biogeography of birds at three biological scales: within biological lineages or populations, among biological lineages or species, and between/within assemblages of

biological lineages or biotas. My research foscused on bird evolution patterns across the last ten million years. Although, each dissertation chapter can stand alone as an independent body of work, together the three research projects paint a picture of the evolution of birds in North and South America over the last ten million years.

Evolution within populations

In Chapter 2, I examine how populations of a widely distributed songbird responded to the glacial-intergalcial cycles over the last two million years. Climate change over evolutionary time is recognized as an important factor in shaping patterns of genetic structure and diversity in species. However, there is a lag time between when climate change events occurred and the modern day collection of DNA samples that are used to reconstruct past evolutionary events. In this chapter, I integrate genetics, ecological niche modeling, and current-day demography to reconcile potentially conflicting patterns inferred from processes occurring across different time frames. Specifically, I tested the prediction that genetic diversity in cardinals was highest in areas that provided suitable habitat during the last ice age.

Evolutionary rates among populations

In Chapter 3, I test whether population size affects the rate at which populations evolve. Molecular dating has relied on the assumption that DNA mutations become fixed at a clock-like rate. However, the so-called “molecular clock” is often violated, because species can evolve at different rates, which causes considerable difficulties for researchers interested in dating evolutionary events. There have been a number of reasons

proposed to help explain why rates can vary across species, but no clear answer has emerged. One contentious issue is the impact of population size and molecular rates with recent studies have finding contradictory results. In this chapter, I took a new approach to examining why species do not evolve according to the molecular clock by proposing a new analytical framework. My approach used DNA data from the mitochondrial and nuclear genomes and a statistical framework that compared expected and observed rates of evolution in different populations of cardinals.

Evolution of biotas and species assemblages

In Chapter 4, I address a longstanding question on how bird species responded to the formation of a land bridge that connected the isolated continents of North and South America approximately 3 million years ago. The fossil record indicates that during this time period a rapid exchange and mixing of mammalian faunas occurred. This exchange of mammals, referred to as the "Great American Biotic Interchange" initiated an unprecedented evolutionary event that led to the reconfiguration of mammal diversity in Central and South America. However, for birds, which do not have a rich fossil record, it is unclear whether they responded in manner similar to mammals. To address the question of bird participation in the Great American Biotic Interchange, I compiled data from 64 avian phylogenetic studies and tested whether there was a positive correlation between the rate of ancient dispersal events between North and South America and the time period associated with the closure of the Panamanian land bridge. I, also, assessed the movement of birds between continents has impacted the present day diversity of birds in the North and South America.

CHAPTER 2

PHYLOGEOGRAPHY OF THE NORTHERN CARDINAL

Introduction

A multitude of studies of various North American taxa have shown that historical events, such as Pleistocene climatic changes coupled with topographic, hydrologic and ecological barriers, were instrumental in generating phylogeographic structure within species (Riddle et al. 2000, Zink et al. 2001, Swenson and Howard 2005, Soltis et al. 2006). From the extensive body of phylogeographic work on North American species, two generalizations have emerged 1) taxa were isolated into independently evolving lineages (i.e. species, subspecies, phylogroups) during the Pleistocene (Klicka and Zink 1997) and 2) genetic diversity within species is highest in areas that remained stable (refugia) through glacial cycles (Hewitt 2000). Ultimately, these hypothesized predictions seek to link the phylogeographic structure in species to historical events.

A recent and critical historical event was the Last Glacial Maximum, a time when ice sheets covered much of present-day temperate North America. Entire biotas were fragmented and displaced to more southern latitudes (Hewitt 2000). After the glaciers receded many taxa colonized these newly available areas by rapid long-distance dispersal, a phenomenon referred to as the leading edge or pioneer model (Hewitt 1996). A major prediction of the leading edge model is that genetic diversity will be lower in the recently colonized areas. This idea is supported by the observation that many species have higher genetic diversity in European refugial areas (Hewitt 2000). Additionally, genetic data indicate that historical population expansions occurred in many species, a finding consistent with species colonizing newly available areas with suitable habitat. A

goal of phylogeography is to recover evidence of population expansions; however, the lag between earlier population expansions and the present represents a time when ongoing demographic processes can either erase earlier genetic signatures of population structure or expansions or prevent such signatures from developing at all.

Contemporary processes can affect the spatial distribution of genetic diversity in various ways. High levels of gene flow may homogenize haplotype diversity throughout a population that once showed genetic structure. Conversely, nonrandom mating with individuals in close geographic proximity can generate genetic structuring within a continuous population (Slatkin 1993). Autecology also plays a critical role in how species maintain genetic connectivity across the landscape (Burridge et al. 2008, Burney and Brumfield 2009). Vagile species with high ecological thresholds will be able spread haplotypes across an area more readily than poor dispersers with narrow ecological requirements. Interpretations of genetic diversity patterns are further confounded by changes in gene flow due to oscillations in population densities and by shifts in the connectivity of habitats. These points highlight the complexity of population-level processes and emphasize that genetic patterns of diversity may become so tightly coupled to on-going processes that signatures of historical differences or similarities may be erased.

To examine the impact of historical and contemporary processes on genetic pattern and diversity, I investigated a broadly distributed North American songbird, the Northern Cardinal (*Cardinalis cardinalis*). Cardinals are distributed from southeastern Canada, throughout the eastern United States, and in the southwestern states south through the Mexican lowlands to northern Central America (Halkin and Linville 1999).

Cardinalis cardinalis is currently divided into 18 subspecies representing four subspecies groups (Halkin and Linville 1999). The species inhabits multiple biomes which include deserts, dry and humid tropical forests, and deciduous temperate forests. In all regions it prefers similar habitats of shrubs, small trees, forest-edge, or secondary growth (Halkin and Linville 1999).

In this study I used mitochondrial DNA (mtDNA) sequences, ecological niche models, and demographic data to test the impacts of historical and contemporary processes on pattern of present-day genetic differentiation and diversity in *C. cardinalis*. First, I evaluated whether historical events had generated deep phylogeographic structure or monophyletic groups in mtDNA across the range of cardinals. Second, I tested for the signature of historical demographic expansions that should have been generated in populations that expanded following glacial retreat. Third, I used present-day and paleoecological niche models to explore if clades had geographically stable ranges since the Last Glacial Maximum. Finally, I evaluated if genetic diversity patterns are consistent with contemporary demographic processes.

Methods

I obtained samples from 163 individuals of *Cardinalis cardinalis* that includes representatives from 14 of the 18 recognized subspecies and the four subspecies groups: *carneus*, *igneus*, *coccineus*, and c*ardinalis* (Table 1). I included six closely related species (Klicka et al. 2007) in the dataset; *C. sinuatus*, *C. phoeniceus*, *Rhodothraupis celaeno, Periporphyrus eyrthromelas*, *Caryothraustes poliogaster*, and *C. canadensis*) and the more distantly related *Molothrus ater* as the outgroup. I extracted total genomic

DNA from tissues and toe-pads from voucher specimens using the DNeasy tissue extraction kit (Qiagen, Valencia, CA). I amplified the mitochondrial gene NADH dehydrogenase subunit II (ND2) via polymerase chain reaction (PCR) using the primers L5215 (Hackett et al. 1996) and HTrpc (STRI). I used 12.5 μl reactions using the following protocol: denaturation at 94° C for 10 min, 40 cycles of 94° C for 30 s, 54° C for 45 s, and 72° C for 2 min, followed by a final 10 min elongation at 72 °C. PCR products were sent to High-Throughput Genomics Unit (University of Washington, Seattle) for all subsequent sequencing steps. PCR products were purified using ExoSAP-IT (USB Corporation, Cambridge, MA), run through cycle-sequencing reactions and final products were sequenced using BigDye (Applied Biosystems, Foster City, CA) on a high-throughput capillary sequencer. I aligned chromatograms of the forward and reverse strands in Sequencher 4.9 (GeneCodes Corporation, Ann Arbor, MI) and sequences were translated into amino acids to check for premature stop codons.

Genetic pattern

I used several approaches to get an understanding of the distribution of genetic diversity within *C. cardinalis*. Aligned sequences were run through the program Mr.Modeltest v. 2.3 (Nylander 2004) to estimate an appropriate model of sequence evolution. I determined the sequence evolution model for the complete ND2 gene as well as data partitioned by codon position (1^{st} and 2^{nd} codon positions and 3^{rd} codon position). The best-fit sequence model was determined by use of the Akiake Information Criterion (AIC). I constructed a gene tree using the program, MrBayes v. 3.1.2 (Ronquist and Hulsenbeck 2003). I constructed gene trees with the non-partitioned data and partitioned

by codon positions. The branch length prior was set to unconstrained with an exponential distribution with the parameter set to 100.0 to avoid artificially long branches (Brown et al. 2010). All priors Ire unlinked for the partitioned data and I ran this analysis for seven million generations and sampled every 1000. I performed diagnostic tests to evaluate mixing and convergence of MCMC chains and the burn-in was determined from visual inspection of the likelihood plots in the program Tracer v 1.5 (Rambaut and Drummond 2010). To determine which approach was the best fit to the data, I used Bayes factors, computed from the harmonic mean of the non-partitioned and partitioned posterior likelihoods.

Genetic diversity

Clades identified from our gene tree were used to estimate the genetic distance (internode to clade node distance) using an uncorrected p-distance for each clade in the program MEGA 4 (Tamura et al. 2007). Two of the clades, *igneus* and *cardinalis* have relatively large ranges and are comprised of several subspecies, therefore I examined the amount of genetic structure within and among subpopulations for each of these clades by performing an AMOVA (Analysis of Molecular Variance) in the program Arlequin v. 3.11 (Excoffier and Schneider 2005). For the group designation for the AMOVA, I used U. S. states for *cardinalis* and for *igneus* I used the regions Baja California Sur, Tiburón Island, Arizona/New Mexico and Sinaloa. I calculated genetic diversity (the average number of nucleotide differences per site betIen two sequences), haplotype diversity (Nei 1987), and the frequency of private haplotypes in the program DNAsp v. 5.10.01 (Librado and Rozas 2009) for each clade and for subpopulations that had adequate

sample sizes ($n \geq 8$). Sample sizes among clades were skewed; therefore I estimated nucleotide diversity with all available samples in a clade and with equal samples (n=6).

Historical demography

To perform hypothesis tests on cardinal historical demography, I re-estimated sequence models for each individual clade that had adequate sampling (*cardinalis*, *coccineus*, and *igneus*). I evaluated demographic history in the program BEAST v. 1. 5.4 (Drummond and Rambaut 2007) by comparing the likelihood of three different models for the coalescent tree prior – constant population size, expansion growth, and Bayesian skyline. This approach allowed us to use Bayes factors to test whether a population expansion was more likely than a constant population size (Suchard et al. 2001). Although the methodology employed to estimate Bayes factors in Tracer has been shown to be yield unstable results (Lartillot and Philippe 2006), the results of demographic tests can be evaluated against the visual representation of population change over time in the Bayesian skyline plots. Again, I partitioned ND2 by codon position and allowed each partition to have separate sequence models. I ran the analyses with a fixed substitution rate of 1.23×10^{-2} subs/site/lineage/million years, a rate estimated relative to the avian "2% rule" (Lovette 2004) used for the gene cytochrome-b (see Smith and Klicka 2010. The posterior output was examined in Tracer v 1.5 (Rambaut and Drummond 2010) to assess mixing and convergence of MCMC chains. I ran the analysis for 20 million generations. Bayes factors were calculated by manually summing the tree likelihood and coalescent/skyline columns in the BEAST log file.

Ecological niche models

I downloaded locality records from two online databases, ORNIS (museum records only; http://ornisnet.org) and observation records from the Avian Knowledge Network (http://www.avianknowledge.net/content) and plotted all records to get an estimate of the *C. cardinalis* distribution. To build an ecological niche model for *C. cardinalis*, I used only records from observations and specimens collected for this study. I assumed that these locality records are the most precise because they have been collected in the last ten years. The number of observation records for *C. cardinalis* was heavily skewed towards records in eastern North America where it is a common "backyard" bird. Therefore, I randomly sampled a subset of the records in order to have a more even sampling scheme. In total I compiled 664 records for *C. cardinalis*.

Beginning in the late 1800s a well-documented northward range expansion has been recorded for *C. cardinalis* in the northern U. S. A. and the Colorado River Basin (Halkin and Linville 1999, Monson and Phillips 1981). I built additional ecological niche models that attempted to approximate the pre-expansion distribution by making an estimate of their pre-19th century range. The specific limits of the species' range prior to expansion are not precisely known. Therefore, based on distribution records (Halkin and Linville 1999), I used 40° N as a conservative and rough estimate of cardinal's pre-19th century range (post-glacial) in eastern North America and 33° N in the southwest U. S. A. All records above these latitudes were not included in the analysis in the post-glacial ecological niche models.

I used the Bioclimatic variables from the WorldClim dataset (v. 1.4) with a resolution of 2.5 min (Hijmans et al. 2005). Eight of the variables were correlated with other variables (R> 0.90); therefore, we used 11of the 19 temperature and precipitation variables (BIO1, BIO2, BIO5, BIO6, BIO8, BIO9, BIO10, BIO12, BIO13, BIO15 and BIO18). I generated five replicate models for each separate clade identified from our phylogenetic analysis using the maximum entropy algorithm in MAXENT 3.3.3 (Phillips et al. 2006). The model was then applied to the Model for Interdisciplinary Research on Climate (MIROC) layers to estimate suitable climatic conditions for *C. cardinalis* during the Last Glacial Maximum. I visualized models in the program ArcGIS 9.3 (ESRI Inc., Redlands, CA). Because the distribution of *C. cardinalis* is well known I used a digital range map (Ridgely et al. 2007) to set the logistic threshold values for the climatic suitability for each clade and applied these same threshold values to the paleoecological niche models. To calculate range size of each clade, I converted the model raster files into polygons in ArcGIS 9.3 and recorded polygon area in square kilometers.

Contemporary demography

To evaluate the impact of contemporary demography on genetic diversity, I used data from the North American Breeding Bird Survey, a long-term demographic study on breeding birds in the USA and Canada (Sauer et al. 2008). Breeding Bird Survey data is based on mean counts of species along specified routes from these data density and population growth estimates are approximated. The datasets I used Ire population growth (% change over time) and population density estimates. I used population growth based on estimates of demographic trends of this species in a U. S. state over 40 years (1966-

2007). For population densities, I directly extracted values from the Summer Distribution Map 1994 -2003 using the average latitude from our genetic samples.

Results

Genetic diversity and pattern

The best-fit models of sequence evolution were GTR + I + G for the complete ND2 gene, GTR + I for the 1st and 2nd codons and GTR + I +G for the 3rd codon. Both the partitioned and unpartitioned analyses yielded trees with similar topologies but with different likelihoods. The partitioned dataset (lnL -5082.52) was a better fit to the data than the unpartitioned data (lnL -5354.79) with a ln Bayes factor of 272.27. Although the harmonic mean method can be sensitive to a high degree of variance (Lartillot and Philippe 2006), the Bayes factors reported here are informative because the likelihoods of the two mtDNA gene trees are well separated (Liu and Pearl 2007). The mtDNA tree strongly indicated deep genetic structure across the range of *C. cardinalis* with no mtDNA haplotype sharing among regions except potentially in the Mexican Gulf coast region (Fig. 1). Most major nodes in the tree had posterior probabilities greater than 0.95. Exceptions were the node for the *cardinalis* clade (pp <0.50) and the lack of reciprocal monophyly between the *coccineus* and *saturatus* clades. Uncorrected node to internode genetic distance for each clade ranged from 0.35 – 2.60 % (Table 2).

The genetic breaks are largely concordant with the four long-recognized morphological groups: *carneus*, *igneus*, *coccineus*, and c*ardinalis* (Halkin and Linville 1999) with the addition of two monophyletic island lineages, *saturatus* within c*occineus* and m*ariae* within *igneus* (Fig. 1, Fig. 2). The basal split in the mtDNA tree separates the

carneus clade from all other cardinals. The remaining cardinals are split western and eastern clades. The western clade is comprised of a monophyletic *igneus*, distributed throughout the Baja California Peninsula, Sonoran and southern Mojave deserts and a monophyletic *mariae* on the Tres Marías Islands. The eastern clade is comprised of three groups (Fig. 1). *Coccineus* occurs on the Yucatán Peninsula and is paraphyletic with respect to the monophyletic *saturatus* group that occurs only on Cozumel Island. Haplotypes in northeastern México belong to *cardinalis*, although traditional taxonomy had placed specimens from this region in *coccineus*. The *cardinalis* clade is geographically widespread, distributed throughout eastern North America and potentially as far south as Veracruz, México.

Population genetic analyses revealed differing levels of genetic structure for the two more widespread clades, *igneus* of northwest México and the southwestern US and *cardinalis* of eastern North America. In *igneus* 20.3 % of the genetic variation was among and 79.7% was within populations (AMOVA, Table 3). There is very little structure in *cardinalis* with 97.8% of variation within and 2.2% among populations (Table 3). Nucleotide diversity within clades ranged from 0.00024 – 0.00434 (complete clade sampling) and 0.00032 – 0.00397 (equal sampling; Table 2). Nucleotide diversity within the *cardinalis* clade ranged from 0.00278 – 0.00511 and haplotype diversity was high across all geographic localities 0.900 – 1.000 (Table 4). The *igneus* clade had lower nucleotide diversity (0.00017 – 0.00368) with the lowest diversity in northern part of the range (Arizona and New México). The pattern of haplotype distributions was low haplotype diversity (0.182) and no private haplotypes in the north (Table 4).

For the historical demographic tests, the best-fit substitution model was the same for all three clades (*cardinalis*, *igneus*, and *coccineus*); HKY for 1st and 2nd codon positions and GTR for 3rd codon position. Because *igneus* shows geographic structure (Table 3), the analysis was run without the Tiburon Island samples (*igneus* A, Table 5) and a second time without both the Tiburón Island and Baja California samples (*igneus* B, Table 5). Population structure within *igneus* affected the results. For *igneus* A there was some evidence in support of the expansion growth model (Table 5; log Bayes factor 2.445), but in *igneus* B there was no support for the expansion growth model (Table 4; log Bayes factor 0.505). For *coccineus* I found no support for the expansion growth model (Table 4; log Bayes factor -0.002). Bayesian skyline plots for *igneus* and *coccineus* (Fig. 3A and Fig. 3B) showed no evidence of population expansion over time. In contrast, the expansion growth model and Bayesian skyline plot were a much stronger fit than the constant population growth model for the *cardinalis* clade (Table 5; log Bayes factors 7.520 and 13.896) indicating the population expanded over time (Fig. 3C).

Ecological niche models

Locality records obtained from museum specimens (ORNIS) and field observations (Avian Knowledge Network) indicated that the range of *C. cardinalis* is not continuous (Fig. 4A). I constructed separate models from the four clades with the following number of records; *cardinalis* (n = 232), *carneus* (n= 86), *igneus* (n = 195), and *coccineus* (n = 151). The ecological niche models produced by MAXENT performed better than random predictions and the area under the receiver operating characteristic (ROC) curve was close to one (AUC > 0.91) for the four mainland clades. Ecological

niche models performed well for the two island populations based on the MAXENT diagnostics but the models did not accurately predict the small distributions of the island populations; therefore these models are not presented. The distributions generated from the present-day ecological niche models (Fig. 4B) were consistent with the genetic breaks observed in our mtDNA gene tree (Fig. 1). There was no suitable climatic conditions connecting *cardinalis* and *igneus* across west Texas and New Mexico, but the model predicted areas with suitable climatic conditions for *cardinalis* into the range of *igneus*. There was a predicted continuous range across the Gulf Coast connecting the *cardinalis* and *coccineus* groups, but only with the lowest suitable values projected onto the map (Fig. 4B). The northern extent of the pre-19th century ecological niche models for *igneus* and *cardinalis* tightly follows the northern limits of the locaily records specified in the model (Fig. 4C).

The paleoecological niche model showed a pronounced reduction in the current range of the *cardinalis* clade in the eastern half of the U. S. A., which suggested that our samples in Louisiana, the Southeast, eastern México and possibly Texas are from areas that were a putative refugium (Fig. 4D) during the Last Glacial Maximum. The Last Glacial Maximum model also indicated that there were suitable conditions for the *cardinalis* clade across the Chihuahuan Desert into the Sonoran Desert and east of the Sierra Madre Oriental (Fig. 4D). There appears to have been a reduction in suitable climatic conditions for cardinals in México, such as on Baja California and the Mexican Pacific coast, but there was not the extreme range reduction seen at higher latitudes in the range of cardinals in eastern North America.

Spatial and temporal correlations

I found evidence for a positive linear relationship between clade range size and nucleotide diversity ($R^2 = 0.83$; $p = 0.01$; Table 2). This relationship was not significant when I converted range size to log values there was still a high R^2. There was no linear relationship between levels of nucleotide diversity and genetic distance ($R^2 = 0.04$; $p = 0.69$; Table 2). When I used nucleotide diversity estimates corrected for equal sample size, I obtained qualitatively similar results for the tests on range size and stem distance. Within the *Cardinalis* clade neither nucleotide diversity ($R^2 = 0.0003$; $p = 0.96$; Fig. 5A) or haplotype diversity ($R^2 = 0.25$; $p = 0.17$; Fig. 5B) or frequency of private haplotypes ($R^2 = 0.097$; $p = 0.41$; Fig. 5C) exhibited a linear relationship to latitude. However, there was a positive correlation with population growth and latitude ($R^2 = 0.690$; $p = 0.02$; Fig. 6B). This pattern is also seen in contemporary population density estimates, the areas that were stable and colonized in the post-glacial period currently have a density of 30 – 100 individuals per area, while areas colonized in the 19th century have 10 – 30 individuals per area (Fig. 6A).

Discussion

Historical signal - phylogeographic structure

Recent historical events, particularly Pleistocene glacial-interglacial cycles, have had a critical impact on generating phylogeographic structure in species (Swenson and Howard 2005). Typically, phylogeographic studies have examined how species responded to these events within a single biome or biogeographic region (Spellman and Klicka 2006, Houston et al. 2010, Arbeláez-Cortés et al. 2010). Here, I report how a

species distributed across multiple biomes and biogeographic regions of North America responded to the climatic oscillations of the Pleistocene. In the Northern Cardinal, I found six lineages that likely began differentiating well before the Late Pleistocene (Klicka and Zink 1997). Each of the four continental clades is found in a well-established biogeographic area: the Mexican Pacific coast, the Sonoran and Peninsular deserts, eastern North America and the Yucatán Peninsula. Additionally, two insular lineages on Cozumel and the Tres Marías Islands were monophyletic.

Major geological and climatic barriers appear to define the ranges of the mainland clades. The distribution of *carneus* is constrained by the Sierra Madre del Sur to the east and habitat turnover to the north and south. *Igenus* shares a range with many co-distributed desert species that border the Sierra Madre Occidental to the east and extend to the northern and southern limits of the Sonoran Desert. This clade does not exhibit the deep genetic structure between the Sonoran and Peninsular desert populations found in other co-distributed species (Riddle et al. 2000b). However, low-levels of genetic structure were detected in *igneus*, which is likely due to recent differentiation of the putative subspecies found in the Sonoran Desert, Baja California Peninsula, and Tiburón Island (Ridgway 1901). Present-day gene flow among cardinals in these areas is likely limited by their geographic isolation or by ecological constraints (García-Trejo 2009) , but gene exchange among these groups could have been more regular during the Last Glacial Maximum when sea levels were lower (Clark and Mix 2002) and populations were in closer proximity.

Igneus is replaced in the Chihuahuan Desert by *cardinalis* and haplotypes of both clades are found east and west of the Chihuahuan-Sonoran transition zone. This region has been identified as an area where many species exhibit genetic breaks and phylogroups often come into secondary contact (Swenson and Howard 2005). Although our sampling was limited in the transition zone, I found no evidence of geographic overlap in these haplotypes and our genetic data and ecological niche model support an earlier conclusion that these forms do not hybridize (Ridgway et al. 1901). The present-day ecological niche model found the geographic area between these clades unsuitable for either group, but the paleoecological niche model indicated a more continuous distribution across this gap during the Last Glacial Maximum.

The evolution of the *cardinalis* clade in eastern North America has been a history of population connectivity. Remarkably, *cardinalis,* with a range of over 3,000,000 km^2, has no appreciable geographic structure in mtDNA, an unexpected pattern for a widely distributed clade that has several described subspecies. Comparative phylogeographic work on other widely distributed avian species has shown a similar pattern of unstructured mtDNA in eastern North American birds (Zink 1996). Alternatively, a diverse array of taxa including birds distributed in the southeast U.S.A. exhibit distinct phylogeographic structure that was shaped by a diverse array of historical factors (Soltis et al. 2006).

The range limits of *cardinalis* and *coccineus* in eastern México are not well understood, but the mtDNA break coincides with the biogeographic barrier separating the Yucatán Peninsula biotic province from surrounding regions (Vázquez-Domínguez and Arita 2010). Morphological work has identified an area of turnover where *cardinalis*

grades into *coccineus* in Tamaulipas (Parkes 1997). Based on genetic data presented, however, specimens from Tamaulipas and areas farther south in Querétaro and Veracruz, México all possessed *cardinalis* haplotypes. Much of the Gulf Coast along México is unsuitable for cardinals because it has been converted to agriculture or encompasses low-lying areas that periodically flood. The ecological niche model also showed that habitat along the Gulf Coast to be sub-optimal. Additionally, *coccineus* appears to be a recently diverged lineage and is paraphyletic with respect to the monophyletic *saturatus*, a pattern consistent with incomplete lineage sorting between young groups (Omland et al. 2006).

The two island lineages, *mariae* and *saturatus*, are currently classified as subspecies in the *igneus* and *coccineus* groups, respectively, and both are monophyletic, suggesting that they are independently evolving lineages. Despite Cozumel being only 20 kms from the mainland, its biota has become distinct in a relatively short period of time, with eight endemic species of birds and mammals found on the island. *Mariae* occurs on the Tres Marías Islands, a small island chain, 100 km off the Pacific coast of México with several endemic subspecies (Stager 1957). Although *saturatus* and *mariae* belong to relatively young island assemblages, both have become distinct in morphology and mtDNA.

Mitochondrial DNA as a phylogeographic marker

Mitochondrial DNA is a powerful marker for identifying independently evolving populations within species because of its fast mutation rate and maternal inheritance (Zink and Barrowclough 2008). However, incomplete lineage sorting (Maddison 1997), rate variation (Arbogast et al. 2002), and selection (Cheviron and Brumfield 2009) can limit the utility of mtDNA as an evolutionary marker. Overall, the phylogeographic

structure presented in this study is largely congruent with the morphological groups previously described for the species. These mtDNA lineages reflect a long history of isolation and independent evolutionary trajectories that merit further evaluation using multilocus coalescent methods (Leaché and Fujita 2010).

Genetic diversity estimated from mtDNA is impacted by a number of factors, such as population size, age, genetic drift and mutation rate (Avise 2000). By using range size and genetic distances as proxies for population size and time of isolation, I was able to show that nucleotide diversity in *C. cardinalis* was more strongly coupled to range size than genetic distance. This relationship between range size and genetic diversity was expected because larger ranges have the potential for larger populations and thus are able to accrue higher levels of genetic diversity (Frankham 1996). This result suggests that mtDNA genetic diversity is at least partially coupled with cardinal population size, and indicates that the estimates provided here have not been strongly affected by selection and linkage to the W sex chromosome (Berlin et al. 2007) or genetic draft (Bazin et al. 2006).

Post-glacial expansion: historical to contemporary demography

The leading edge model of post-glacial expansion predicts there will be lower genetic diversity in recently colonized regions (Hewitt 1997). Despite strong evidence for this model in some taxa and some geographic areas, there has not been compelling evidence for this pattern in North American birds (Fry and Zink 1998, Spellman and Klicka 2006, Milá et al. 2007, but see Milá et al. 2000). These results are often attributed to other factors such as mixing of separate refugial populations, un-sampled populations,

or rapid expansions. Both our genetic data and our ecological niche modeling suggest a population expansion for the *cardinalis* and population stability for *igneus* and *coccineus*. Based on our Bayesian skyline plot (Fig. 3C) the expansion began before the Last Glacial Maximum, but this result may be attributable to the interspecific mutation rate I used (Ho et al. 2005).

Given that the *cardinalis* clade likely expanded out of a glacial refugium, it is surprising that I found nucleotide diversity, haplotype diversity and the frequency of private haplotypes to be uniform across *cardinalis* sampling localities. Intuitively, homogeneous genetic diversity would seem attributable to birds being more vagile than other vertebrates, but this is unlikely the case here. Many birds are constrained by the same ecological barriers as non-volant organisms (Smith and Klicka 2010). This ecological limitation of dispersal is evident in *igneus*, which has also undergone a recorded northern range expansion. My *igneus* samples from Arizona/New Mexico had lower haplotype and genetic diversity than samples from Sinaloa. A critical difference between *cardinalis* and *igneus* is the amount of habitat connectivity across their ranges. There are few major dispersal barriers for cardinals in eastern North America, allowing the species to freely disperse in any direction. On the other hand, in the areas of the arid Sonoran Desert where cardinals have recently colonized, they are more restricted to habitats with suitable vegetation such as riparian zones.

The high genetic diversity in recently established populations of the eastern *cardinalis* contradicts the leading edge model and begs explanation. One suggested alternative is the Phalanx model, which posits that large populations expanded slowly into newly available habitats with no loss of genetic diversity in newly colonized areas

(Hewitt 1996). Although the Phalanx predicts homogenous genetic diversity as seen in *cardinalis*, I see no reason for cardinals to have expanded their range slowly. Indeed, the rapid northward range expansion documented in the last 150 years suggests cardinals have high dispersal rates and can rapidly populate suitable habitat. Why may this be the case?

I suggest that cardinals in eastern North America were held at low numbers by a shortage of suitable habitat prior to the European settlement of North America. Throughout their range cardinals inhabit brushy open habitat, and prior to European settlement this ecotonal species lived in a sea of eastern hardwood forest where suitable habitat was patchy and ephemeral, created by water courses, tree falls, and fires. Throughout most of the evolutionary history of eastern cardinals, suitable habitat must have been ephemeral, leading to selection for high dispersal ability. Once European settlement converted most of Eastern North America into a vast ecotone of excellent cardinal habitat, their evolved propensity for high dispersal would homogenize genetic diversity across vast areas.

The hypothesis that cardinals exhibit high dispersal abilities is borne out by an analysis of banding records for cardinals (Dow and Scott 1971). Although conducted before root mean square dispersal was being estimated, Dow and Scott (1971) found that 190 of 1523 recovered cardinals were found outside the 10-minute block of latitude and longitude where they were banded; among those that moved beyond their block, first years moved an average distance of 60 km and adults moved an average of 130 km (Halkin and Linville 1999). These are large distances for a resident bird (Paradis et al. 2002), and comparing them with results for North American blackbirds analyzed in

modern ways (Moore and Dolbeer 1989) suggests that root mean square dispersal for Northern Cardinals is very high, possibly exceeding 100 km. Dispersal distances of this magnitude would be sufficient to homogenize genetic variation in a short period of time.

What these data suggest is that during the lag between historical expansion and the present-day, contemporary demography has played a role in shaping the observed genetic diversity pattern. Population densities are lowest at higher latitudes and population growth increases with latitude. I suggest that this population growth is not just from offspring staying in one area and increasing the density of individuals. Instead, populations are also growing, as new individuals with different haplotypes are moving into new areas. Given that genetic diversity is uniform even in the youngest cardinal populations, at the highest latitudes, the lag between expansion and genetic homogenization may be on the order of decades instead of hundreds or thousands of years. As new areas became inhabitable, dispersal would likely not have been just in a northern and linear fashion, but east and west. If cardinals expanded out of southern glacial refuges rapidly and in large numbers, the signature of decreased genetic diversity may have been rapidly erased or alternatively, potentially never existed at all.

Conclusions

In this study I used a widely distributed songbird, the Northern Cardinal, *Cardinalis cardinalis*, as a model to test historical predictions for how species evolved in response to Pleistocene glacial-interglacial cycles. I found evidence in the mitochondrial DNA and from our modeled distribution of the species that there is deep geographic structure in *C. cardinalis*, which is consistent with fragmentation caused by historical

climate change. This structure likely began developing well before the Late Pleistocene and is supported by the results of both contemporary and paleoecological niche models. A modeled paleodistribution along with historical demographic hypothesis tests indicated that cardinals expanded out of refugia in eastern North America since the Last Glacial Maximum. However, there is no signature of decreased genetic diversity in areas colonized after the expansion. I suggest on-going gene flow across eastern North America has likely homogenized genetic diversity across the region. These results demonstrate that both Earth history events and cotemporary processes are important in determining the geography of genetic diversity observed within species.

Table 1. Taxon List

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	BMNH	AWH349	USA	Arizona	Greenlee Co.
Cardinalis cardinalis	MBM	BTS07071	USA	Arizona	Verde River
Cardinalis cardinalis	MBM	BTS07072	USA	Arizona	Verde River
Cardinalis cardinalis	MBM	BTS07073	USA	Arizona	Verde River
Cardinalis cardinalis	MBM	BTS07074	USA	Arizona	Verde River
Cardinalis cardinalis	MBM	BTS07075	USA	Arizona	Verde River
Cardinalis cardinalis	MBM	DHB2691	USA	Arizona	Santa Cruz Co.
Cardinalis cardinalis	UWBM	VGR708	USA	Arizona	Gila River
Cardinalis cardinalis	UWBM	VGR712	USA	Arizona	Gila River
Cardinalis cardinalis	UNAM	CONA613	México	Baja California Sur	Rancho Monte Alto
Cardinalis cardinalis	UNAM	CONA667	México	Baja California Sur	Rancho Monte Alto
Cardinalis cardinalis	UNAM	CONA672	México	Baja California Sur	Rancho Monte Alto
Cardinalis cardinalis	UNAM	CONA715	México	Baja California Sur	San Dionisio
Cardinalis cardinalis	UNAM	CONA716	México	Baja California Sur	San Dionisio

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	AMNH	GFB061	México	Baja California Sur	Ciudad Insurgentes
Cardinalis cardinalis	AMNH	GFB3060	México	Baja California Sur	Ciudad Insurgentes
Cardinalis cardinalis	AMNH	GFB3062	México	Baja California Sur	Ciudad Insurgentes
Cardinalis cardinalis	UWBM	MFOR130	México	Baja California Sur	Rancho La Ballena
Cardinalis cardinalis	UNAM	PEP1669	México	Baja California Sur	San Jose del Cabo
Cardinalis cardinalis	UNAM	P010053	México	Baja California Sur	Isla San Jose
Cardinalis cardinalis	UNAM	P010054	México	Baja California Sur	Isla Santa Catalina
Cardinalis cardinalis	UWBM	VGR1082	México	Baja California Sur	Rancho La Ballena
Cardinalis cardinalis	MBM	BTS07165	México	Campeche	Municipo de la Candelaria
Cardinalis cardinalis	MBM	BTS0896	México	Coahuila	Municipo de Guerrero
Cardinalis cardinalis	MBM	BTS0897	México	Coahuila	Municipo de Guerrero
Cardinalis cardinalis	MBM	BTS0898	México	Coahuila	Municipo Ramos Arizipe
Cardinalis cardinalis	UWBM	MG08103	México	Coahuila	Alto de Norias
Cardinalis cardinalis	UWBM	MG08104	México	Coahuila	Alto de Norias

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	MSB	142220	USA	Florida	Lee County
Cardinalis cardinalis	MBM	DCO135	USA	Florida	Alachua Co.
Cardinalis cardinalis	MBM	DHB4137	USA	Florida	Alachua Co
Cardinalis cardinalis	MBM	DHB5730A	USA	Florida	Osceola Co.
Cardinalis cardinalis	MBM	GMS219	USA	Florida	Alachua Co.
Cardinalis cardinalis	MBM	GMS220	USA	Florida	Putnam County, Melrose
Cardinalis cardinalis	MBM	JMD343	USA	Georgia	Bryan Co.
Cardinalis cardinalis	MBM	JMD347	USA	Georgia	Bryan Co.
Cardinalis cardinalis	MBM	JMD355	USA	Georgia	Bryan Co.
Cardinalis cardinalis	UNAM	TECO08003	México	Guerrero	Tecomate, San Marcos
Cardinalis cardinalis	KU	3532	USA	Kansas	Douglas County
Cardinalis cardinalis	KU	6273	USA	Kansas	Lawrence County
Cardinalis cardinalis	MBM	BTS05049	USA	Kansas	Labette Co.
Cardinalis cardinalis	MBM	BTS05051	USA	Kansas	Montgomery Co.

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	MBM	BTS05050	USA	Louisiana	Tensas Parish, Buckhorn
Cardinalis cardinalis	MBM	BTS05057	USA	Louisiana	Ouachita Parish
Cardinalis cardinalis	MBM	DHB2768	USA	Louisiana	Cameron Parish
Cardinalis cardinalis	MBM	DHB2815	USA	Louisiana	Cameron Parish
Cardinalis cardinalis	MBM	DHB2991	USA	Louisiana	Cameron Parish
Cardinalis cardinalis	MBM	DHB5691	USA	Louisiana	Concordia Parish
Cardinalis cardinalis	MBM	DHB5692	USA	Louisiana	Concordia Parish
Cardinalis cardinalis	MBM	JMD319	USA	Louisiana	Ouachita Parish
Cardinalis cardinalis	MBM	TKA120	USA	Louisiana	St. Landry Parish
Cardinalis cardinalis	MBM	BTS08192	México	Michoacán	Lazaro Cardenas
Cardinalis cardinalis	MBM	BTS08217	México	Michoacán	Lazaro Cardenas
Cardinalis cardinalis	MBM	TKA135	USA	Louisiana	Cameron Parish
Cardinalis cardinalis	MBM	BTS08219	México	Michoacán	Lazaro Cardenas
Cardinalis cardinalis	MBM	BTS08220	México	Michoacán	Lazaro Cardenas

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	UNAM	CONA1057	México	Michoacán	Presa Infiernillo
Cardinalis cardinalis	BMNH	AEK489	USA	Minnesota	Chisago Co.
Cardinalis cardinalis	BMNH	AEK509	USA	Minnesota	Anoka Co
Cardinalis cardinalis	BMNH	X7320	USA	Minnesota	
Cardinalis cardinalis	BMNH	X8235	USA	Minnesota	Hennepin Co.
Cardinalis cardinalis	BMNH	X8252	USA	Minnesota	Washington Co.
Cardinalis cardinalis	BMNH	X8371	USA	Minnesota	Ramsey Co.
Cardinalis cardinalis	BMNH	X8403	USA	Minnesota	Hennepin Co.
Cardinalis cardinalis	BMNH	X8772	USA	Minnesota	Hennepin Co.
Cardinalis cardinalis	BMNH	X8772	USA	Minnesota	
Cardinalis cardinalis	UNAM	ITM049	México	Nayarit	Islas Marías
Cardinalis cardinalis	UNAM	ITM050	México	Nayarit	Islas Marías
Cardinalis cardinalis	UNAM	ITM086	México	Nayarit	Islas Marías
Cardinalis cardinalis	UNAM	ITM016	México	Nayarit	Islas Marías

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	UNAM	P010075	México	Nayarit	Islas Marías
Cardinalis cardinalis	MSB	3986	USA	New Mexico	Grant County, Lower Gila Box
Cardinalis cardinalis	MSB	14861	USA	New Mexico	Grant County
Cardinalis cardinalis	MSB	130538	USA	New Mexico	Eddy County, Carsbad
Cardinalis cardinalis	MSB	130539	USA	New Mexico	Eddy County, Carsbad
Cardinalis cardinalis	AMNH	PAC1245	USA	New York	Suffolk Co.
Cardinalis cardinalis	AMNH	PAC1330	USA	New York	New York Co.
Cardinalis cardinalis	AMNH	PAC1336	USA	New York	Suffolk Co.
Cardinalis cardinalis	AMNH	PAC565	USA	New York	Suffolk County; Huntington.
Cardinalis cardinalis	AMNH	PAC617	USA	New York	Suffolk County; Northport
Cardinalis cardinalis	AMNH	SPK136	USA	New York	Suffolk Co.; Melville
Cardinalis cardinalis	AMNH	SPK22	USA	New York	Nassau Co; Melville
Cardinalis cardinalis	AMNH	SPK23	USA	New York	Suffolk Co.; Smithtown
Cardinalis cardinalis	AMNH	SPK24	USA	New York	Nassau Co.; Locust Valley

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	MBM	RC1	Mexico	Nuevo Leon	Presa Cerralvo, Cerralvo
Cardinalis cardinalis	MBM	JMD396	USA	Oklahoma	Caddo Co
Cardinalis cardinalis	UNAM	P013211	México	Oaxaca	Mpio. San Pedro Huamelula
Cardinalis cardinalis	MBM	DHB5696	USA	Oklahoma	Caddo Co.
Cardinalis cardinalis	MBM	JMD327	USA	Oklahoma	Caddo Co.
Cardinalis cardinalis	MBM	JMD328	USA	Oklahoma	Caddo Co.
Cardinalis cardinalis	MBM	JMD339	USA	Oklahoma	Caddo Co.
Cardinalis cardinalis	MBM	JMD395	USA	Oklahoma	Caddo Co.
Cardinalis cardinalis	MBM	JSB116	USA	Oklahoma	Caddo Co.
Cardinalis cardinalis	MBM	JSB119	USA	Oklahoma	Caddo Co.
Cardinalis cardinalis	MBM	JSB120	USA	Oklahoma	Caddo Co.
Cardinalis cardinalis	MBM	MM149	USA	Oklahoma	Caddo Co
Cardinalis cardinalis	UNAM	QRO408	México	Queretaro	Higuerillas
Cardinalis cardinalis	UNAM	QRO510	México	Queretaro	Laguna de la Cruz

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	BMNH	Y408254	México	Yucatan	Dzilam de Bravo
Cardinalis cardinalis	UNAM	GLS179	México	Quintana Roo	Isla Cozumel El Cedral
Cardinalis cardinalis	UNAM	GLS180	México	Quintana Roo	Isla Cozumel El Cedral
Cardinalis cardinalis	UANM	MGL17	México	Quintana Roo	Isla Cozumel El Cedral
Cardinalis cardinalis	UNAM	Pe023292	México	Quintana Roo	Isla Cozumel
Cardinalis cardinalis	UNAM	PEP2979	México	Quintana Roo	Isla Cozumel El Cedral
Cardinalis cardinalis	UNAM	PEP3009	México	Quintana Roo	Isla Cozumel El Cedral
Cardinalis cardinalis	UWBM	MFOR139	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	MFOR151	México	Sinaloa	
Cardinalis cardinalis	UWBM	MFOR42	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	MFOR45	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	MFOR49	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	MFOR56	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	MFOR61	México	Sinaloa	Ejido Tesila

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	UWBM	MFOR82	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	SAR7793	México	Sinaloa	Tehueco
Cardinalis cardinalis	UWBM	MFOR144	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	SAR7798	México	Sinaloa	Tehueco
Cardinalis cardinalis	UWBM	SAR7950	México	Sinaloa	
Cardinalis cardinalis	UWBM	SEZ095	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	SEZ115	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	VGR1151	México	Sinaloa	lower side of irrigation canal
Cardinalis cardinalis	UWBM	VGR959	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UWBM	VGR984	México	Sinaloa	Ejido Tesila
Cardinalis cardinalis	UNAM	CONA170	México	Sonora	Isla Tiburón,
Cardinalis cardinalis	UNAM	CONA202	México	Sonora	Isla Tiburón, El Caracol
Cardinalis cardinalis	UNAM	CONA205	México	Sonora	Isla Tiburón, El Caracol
Cardinalis cardinalis	UNAM	CONA220	México	Sonora	Isla Tiburón, El Caracol

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	UNAM	CONA1160	México	Tamaulipas	Rancho el Mogote
Cardinalis cardinalis	UNAM	CONA1172	México	Tamaulipas	Rancho el Mogote
Cardinalis cardinalis	UNAM	CONA1177	México	Tamaulipas	Rancho el Mogote
Cardinalis cardinalis	UNAM	MAGH0968	México	Tamaulipas	Soto la Marina
Cardinalis cardinalis	UWBM	VGR960	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	UNAM	MAGHO969	México	Tamaulipas	Soto la Marina
Cardinalis cardinalis	MSB	100282	USA	Texas	Brewster County
Cardinalis cardinalis	MSB	116306	USA	Texas	Starr County
Cardinalis cardinalis	MBM	CAH110	USA	Texas	Coryell Co., Jonesboro
Cardinalis cardinalis	MBM	DHB5698	USA	Texas	Coryell Co., Jonesboro
Cardinalis cardinalis	MBM	DHB5699	USA	Texas	Coryell Co., Jonesboro
Cardinalis cardinalis	MBM	DHB5700	USA	Texas	Coryell Co., Jonesboro
Cardinalis cardinalis	MBM	DHB5753	USA	Texas	Coryell Co., Jonesboro
Cardinalis cardinalis	MBM	JSB117	USA	Texas	Coryell Co., Jonesboro

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Cardinalis cardinalis	MBM	PCPR004	USA	Texas	Coryell Co., Jonesboro
Cardinalis cardinalis	UNAM	Pe024539	México	Veracruz	Los Tuxtlas
Cardinalis cardinalis	BMNH	X8603	USA	Wisconsin	Pierce Co
Cardinalis cardinalis	KU	B1866	Mexico	Yucatan	Dzilam de Bravo
Cardinalis cardinalis	KU	B1868	Mexico	Yucatan	Dzilam de Bravo
Cardinalis cardinalis	MBM	BRB835	Mexico	Yucatan	El Cuyo
Cardinalis cardinalis	MBM	DHB5697	USA	Texas	Coryell Co., Jonesboro
Cardinalis cardinalis	MBM	BTS07190	México	Yucatan	Municipo de Uman
Cardinalis cardinalis	MBM	BTS07191	México	Yucatan	Municipo de Uman
Cardinalis cardinalis	MBM	BTS07192	México	Yucatan	Municipo de Uman
Cardinalis cardinalis	BMNH	Y408257	México	Yucatan	Dzilam de Bravo
Cardinalis cardinalis	BMNH	Y408284	México	Yucatan	Dzilam de Bravo
Cardinalis phoeniceus	STRI	CPH1	Venezuela	Guarapo	
Cardinalis sinuatus	MBM	DHB5750	USA	Minnesota	

Table 1. Taxon List continued.

Taxon	Museum	Collector No.	Country	State	County/Locality
Caryothraustes canadensis	LSUMNH	B1414	Panama	Darien	
Caryothraustes poligaster	MBM	JK01044	Honduras	Atlantida	
Periporphyrus erythromelas	ANSP	187491	Guyana	Potaro-Siparuni	
Rhodothraupis celaeno	BMNH	X17	México	Tamaulipas	
Molothrus ater	MBM	JK96016	USA	Minnesota	

Table 2. Clade genetic and range characteristics

Clade	Range Size (sq km)	n	Hap	Hap Diversity	Nuc Diversity	Nuc Diversity*	Genetic Dist
cardinalis	3,189,130	82	48	0.9530	0.00434	0.00397	0.75
carneus	35,165	8	3	0.4643	0.00065	0.00083	2.60
coccineus	204,229	11	3	0.6364	0.00175	0.00186	0.35
igneus	444,297	47	21	0.8608	0.00197	0.00179	1.65
mariae	245	6	2	0.7333	0.00173	0.00173	1.65
saturatus	647	8	2	0.2500	0.00024	0.00032	0.35

Summary of genetic diversity indices and range sizes for the six *Cardinalis cardinalis* clades. Number of individuals (n). Number of unique haplotypes (Hap). Metric on the proportion of unique haplotypes in a clade (Hap diversity). Nucleotide diversity with complete clade sampling (Nuc Diversity). Nucleotide diversity with equal sampling per clade (Nuc Diversity*). Node to internode genetic distance based uncorrected p-distances (Genetic Dist).

Table 3. AMOVA summary for the *cardinalis* and *igneus* clades

Source of Variation	**df**	**ss**	**Variance of Components**	**% of Variation**
cardinalis				
Among Populations	8	4.537	0.011	2.19
Within Populations	68	32.372	0.476	97.81
Total	76	36.909	0.487	
Fixation Index	FST:	0.022	p-value	0.028
igneus				
Among Populations	3	4.198	0.093	20.26
Within Populations	44	16.031	0.364	79.74
Total	47	20.229	0.457	
Fixation Index	FST:	0.203	p-value	<0.000

Table 4. Within clade genetic diversity

Locality	# of Ind	# of Hap	Nuc Div	Hap Div	Priv Hap
cardinalis					
Coahuila	8	6	0.00278	0.929	0.50
Florida/Georgia	9	7	0.00406	0.944	0.56
Kansas	5	4	0.00462	0.900	0.40
Louisiana	9	8	0.00443	0.972	0.67
Minnesota/Wisconsin	9	7	0.00406	0.917	0.56
New York	9	8	0.00422	0.972	0.67
Oklahoma	10	7	0.00312	0.933	0.40
Tamaulipas/Nuevo Leon	8	8	0.00511	1.000	0.88
Texas/New Mexico	12	10	0.00371	0.955	0.40
igneus					
Arizona/New Mexico	11	2	0.00017	0.182	0.00
Baja California Sur	13	9	0.00177	0.872	0.56
Sinaloa	19	13	0.00211	0.906	0.69
Tiburón Island	4	3	0.00368	0.833	0.67

Summary of genetic diversity within the *cardinalis* and *igneus* clades. Number of individuals (# of Ind). Number of haplotypes (# of Hap). Nucleotide diversity (Nuc Div). Haplotype diversity (Hap Div). Frequency of private haplotypes (Priv Hap).

Table 5. Hypothesis testing of historical demography for Northern Cardinal clades using different demographic models

Model	Marginal Likelihood (S.E.)	log Bayes Factors
cardinalis		
Constant Size	-2187.800 (+/- 0.336)	-
Expansion Growth	-2170.486 (+/- 0.354)	7.520
Bayesian Skyline	-2155.803 (+/- 0.411)	13.896
***igneus* A**		
Constant Size	-1585.983 (+/- 0.316)	-
Expansion Growth	-1580.354 (+/- 0.350)	2.445
Bayesian Skyline	-1565.140 (+/- 0.364)	9.052
***igneus* B**		
Constant Size	-1513.415 (+/- 0.311)	-
Expansion Growth	-1512.251 (+/- 0.254)	0.505
Bayesian Skyline	-1503.067 (+/- 0.303)	4.494
coccineus		
Constant Size	-1380.049 (+/- 0.174)	-
Expansion growth	-1380.054 (+/- 0.166)	-0.002
Bayesian Skyline	-1379.877 (+/- 0.170)	0.075

igneus A – Tiburón Island samples were removed.

igneus B – Baja California and Tiburón Island samples removed

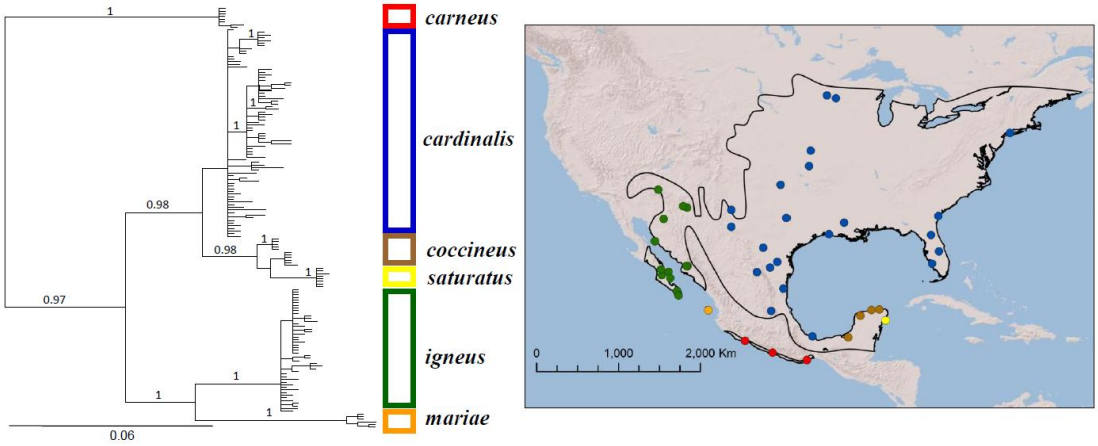


Figure 1. Bayesian mitochondrial gene tree for *Cardinalis cardinalis* with posterior probabilities. Localities for each clade are plotted on the accompanying map of North America. *carneus* (red); c*ardinalis* (blue); *coccineus* (brown); *saturatus* (yellow); *igneus* (green); *mariae* (orange).

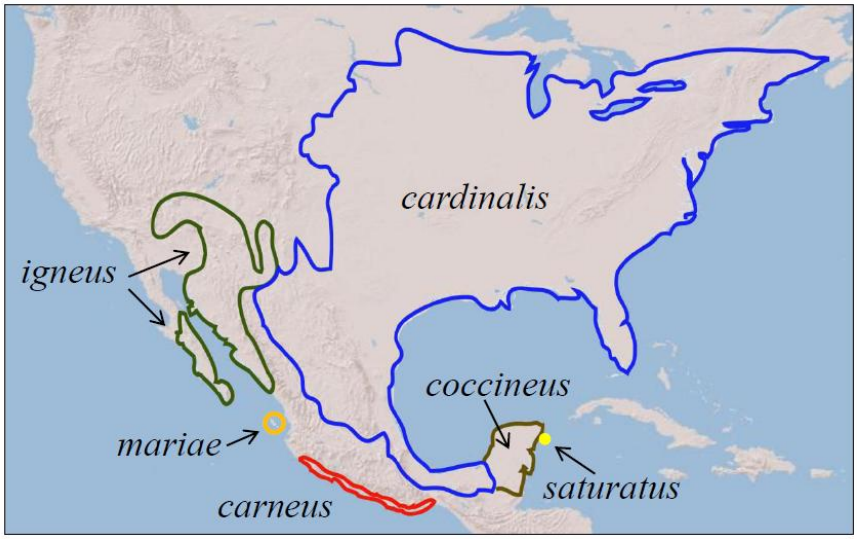


Figure 2. Approximated range limits of the six *Cardinalis cardinalis* clades based on mtDNA. *carneus* (red); c*ardinalis* (blue); *coccineus* (brown); *igneus* (green); *saturatus* (yellow); *mariae* (orange).

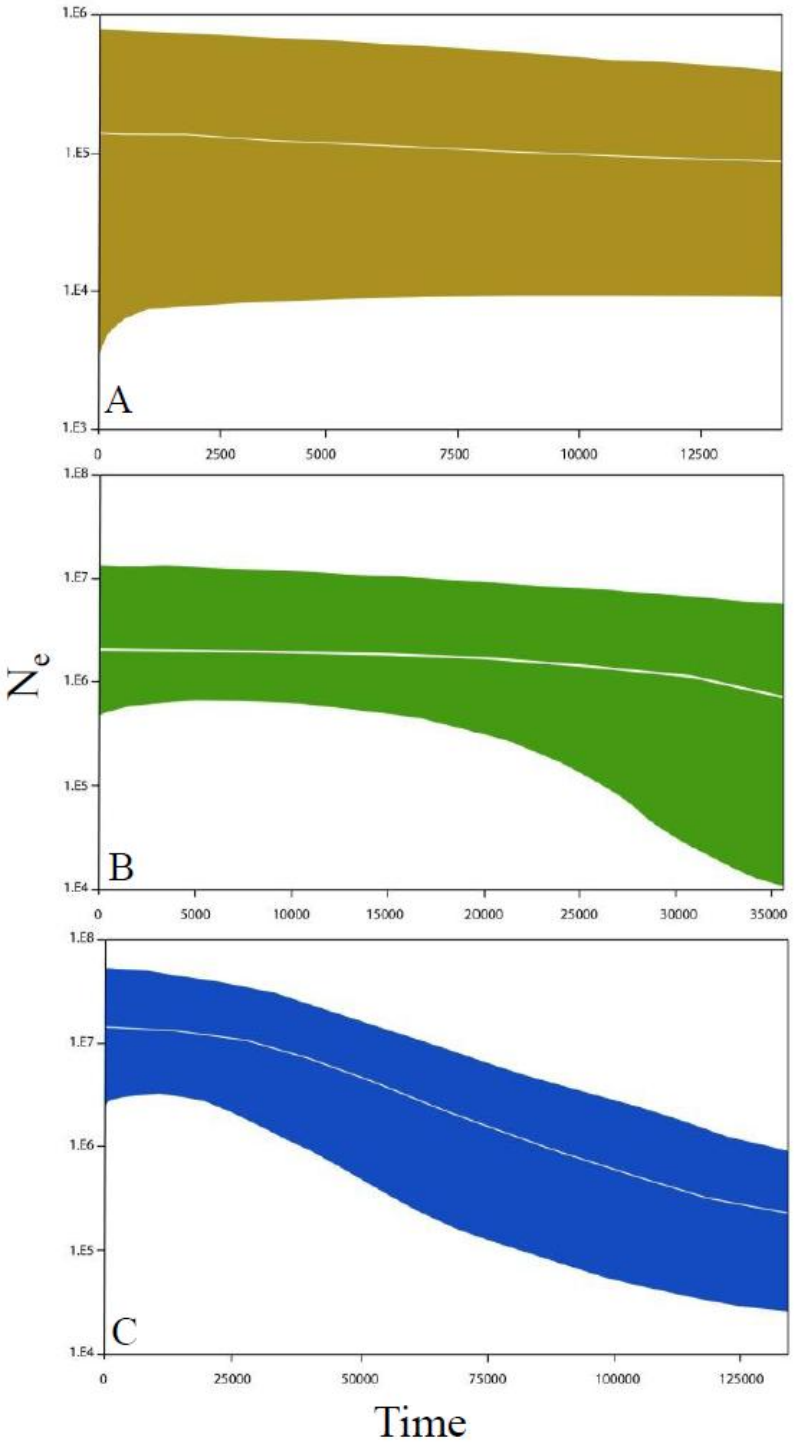


Figure 3. Bayesian Skyline Plots showing change in N_e (effective population size) across time. A) *coccineus* (brown); B) *igneus* (green); C) *cardinalis* (blue).

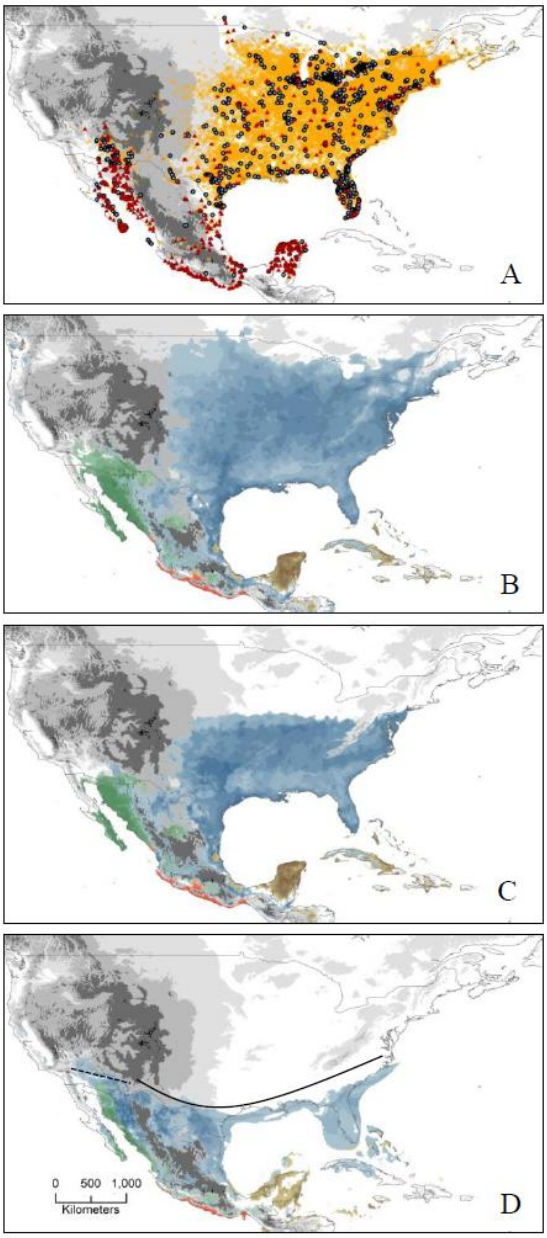


Figure 4. Ecological niche models for *Cardinalis cardinalis.* A) Locality Records: field observations (Avian Knowledge Network, orange hatch marks); museum specimens (ORNIS, blue circles), records used to build ecological niche models (red triangles) B) Current-day C) Post-glacial D) Last Glacial Maximum. Dotted and solid lines show northern extent of *igenus* and *cardinalis* during the Last Glacial Maximum.

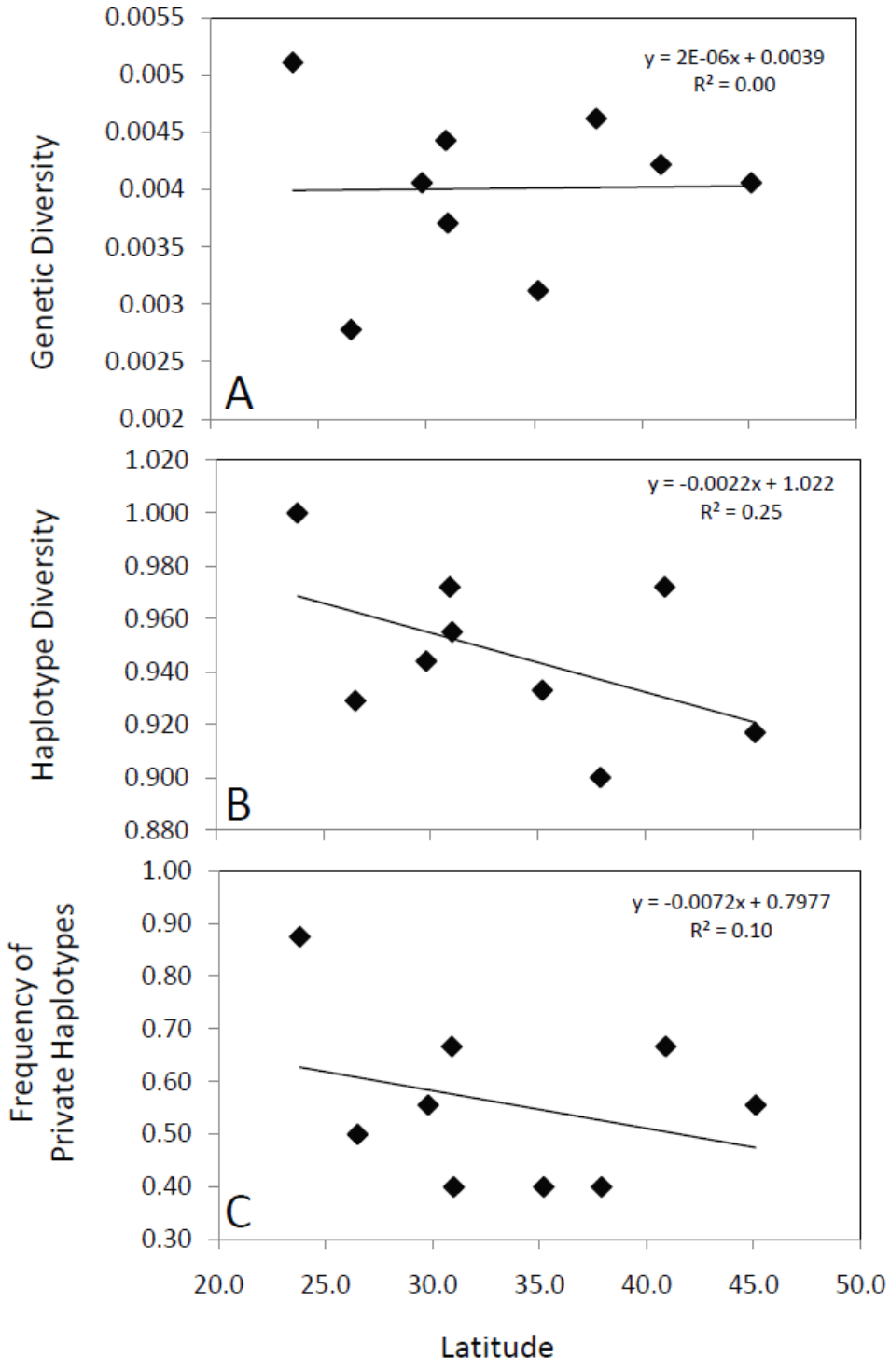


Figure 5. Population-level comparison between latitude and genetic diversity metrics estimated for geographic populations. A) Nucleotide diversity; B) Haplotype diversity; C) Frequency of private haplotypes.

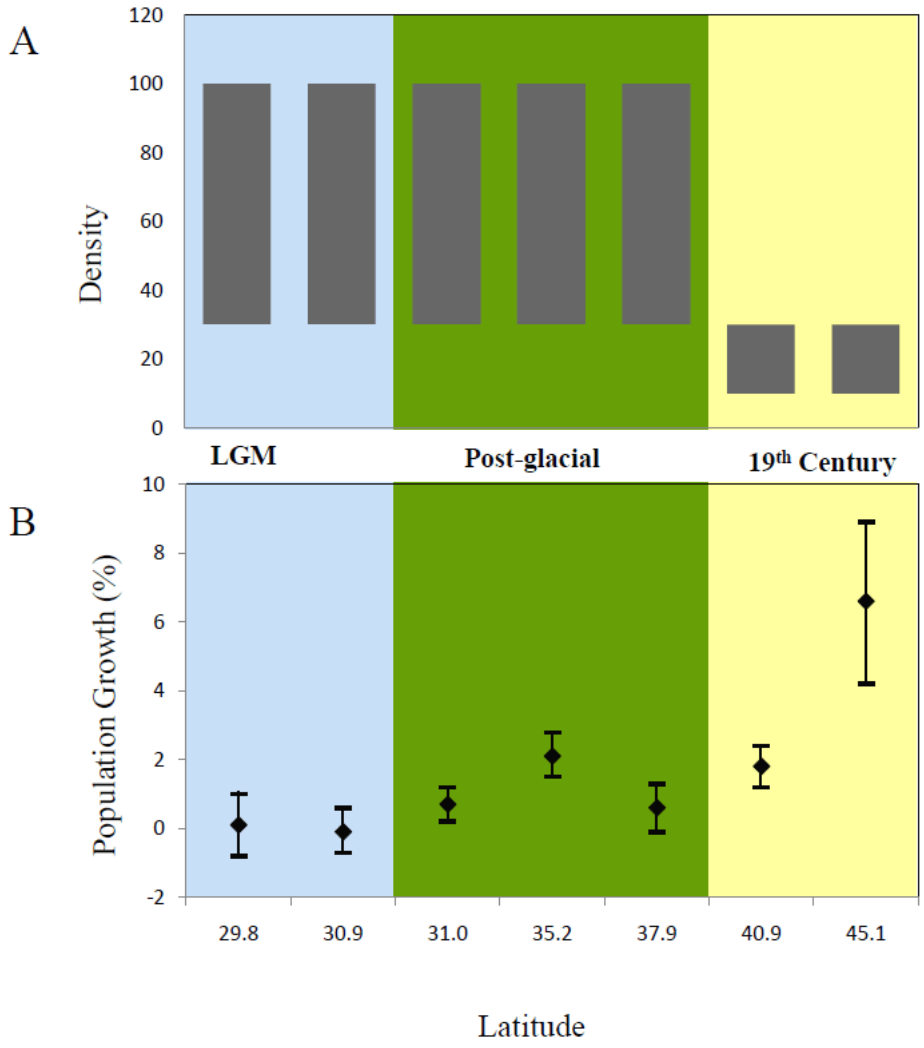


Figure 6. Comparison of contemporary demographic patterns (North American Breeding Bird Survey) across the distribution of the *cardinalis* clade. Demographic data were obtained for latitudes that had genetic samples included in the study. The colored portions represent latitudes that were inferred from the Ecological Niche Models to be stable over time (light blue), colonized in the post-glacial period (green), and colonized since the 19th century (yellow). A) Population density per latitude; B) Population growth (1966 -2007 with 95% CI) per latitude.

CHAPTER 3

EFFECTIVE POPULATION SIZE AND RATE OF EVOLUTION

Introduction

The assumption that DNA mutations become fixed at a nearly constant or "clock-like" rate is often violated (Duffy et al. 2008, Nabholz et al. 2008, Smith and Donoghue 2008, Brandley et al. 2011). Rates of DNA substitutions vary dramatically across the Tree of Life, causing considerable difficulties in inferring the temporal patterns of evolution (Lanfear et al. 2010). The mechanisms that generate evolutionary rate shifts are still not clear, but among lineage rate changes have been attributed to such factors as differences in generation times, body sizes, and metabolic rates (Martin and Palumbi 1993). Conversely, rates can also vary among closely related species or even within species that do not have substantial metabolic differences. At this scale, rate variation is often attributed to differences in effective population sizes (N_e) among taxa (Woolfit 2009).

The size of N_e will dictate the relative strength of selection or drift on the fixation of mutations in a population (Kimura 1962). The nearly neutral model of evolution predicts that slightly deleterious mutations will become fixed at a higher rate in small populations (Ohta 1972). In larger populations, these slightly deleterious mutations are presumed to be removed by purifying selection. The interplay between selection and drift can be seen in the accumulation of different types of mutations in protein coding regions of DNA. These DNA substitutions result in either a change in the amino acid (nonsynonymous change) that is subject to selection or a redundant change in the codon sequence (synonymous change) that presumably does not affect fitness. One approach

that attempts to explain the variation observed in amino acid changes is the estimation of the nonsynonymous/synonymous substitution rate ratio (dN/dS) between lineages with different sized N_e.

A consistent pattern on the relationship between N_e and the dN/dS ratio has not been shown (Woolfit and Bromham 2005; Wright et al. 2009). A range of studies have compared rates across mammalian orders (Easteal and Collet 1994), endosymbiotic and free-living bacteria and fungi (Woolfit and Bromham 2003), island-mainland birds (Johnson and Seger 2001), and insects (DeSalle and Templeton 1988, Bromham and Leys 2005). A general consensus from these studies is that small populations have elevated substitution rates and high dN/dS ratios. However, evolutionary rate patterns may be skewed when data is not phylogenetic independent and comparisons are made between evolutionary distant taxa with differing ecologies, selection regimes, and metabolic rates (Wright et al. 2009). A more recent study that presumably addressed many of these limitations found the opposite pattern of increased substitution rates in large mainland populations (Wright et al. 2009). These contradictory results are surprising and may indicate that there are additional yet unrecognized factors that may bias findings. These factors may not be due to species biology, but rather, to the methodological approaches used to infer evolutionary patterns.

There are two major analytical issues that have not been incorporated into previous comparative studies on molecular rates: phylogenetic error and coalescent or stochastic variance. Earlier work has relied on topologies inferred from either a single gene or concatenated multilocus data to identify sister lineages and estimate substitution rates. Both approaches have been shown to be subject to error and may be inadequate for

inferring phylogenetic relationships among species (Degnan and Rosenberg 2009). Furthermore, a single gene tree does not provide insight into the prevalence of rate variation across the genome, nor does it indicate how much of the variance across branch lengths can be attributed to stochastic processes. Even when multiple genes are used, the data are often combined into a single sequence fragment and patterns of rate variation occurring across genes and lineages may not be directly apparent. The recent development of coalescent methods that estimate species trees from multilocus data (Liu and Pearl 2007, Heled and Drummond 2010) can be used to infer a more accurate estimate of sister lineages, as well as, serve as a comparative framework for understanding the impact of random processes on individual gene tree branch lengths.

The stochastic sorting of ancestral polymorphisms can give the appearance that substitution rates have changed across a phylogeny (Peterson and Masel 2009). When branch lengths are converted to absolute time, the number of substitutions along the branch is supposed to represent lineage age. However, if the branch length is estimated from a single gene, the gene may have diverged before the species diverged from its last common ancestor (Edwards and Beerli 2000). Therefore, it is necessary to understand how both nucleotide substitutions and random lineage sorting or coalescent variance impacts branch lengths (Huang et al. 2009). The degree of gene-species divergence time discordance is positively correlated with size of the lineage's ancestral N_e, lineages with large ancestral N_e will be subject to wider coalescent variance. Multilocus coalescent methods are capable of correcting for this discrepancy by estimating ancestral N_e along with species divergence times (Rannala and Yang 2003). By using a coalescent

framework, I can determine if the apparent rate shifts between small and large populations can be attributed to randomness.

In this study, I used a multilocus coalescent approach to test the hypothesis that population size impacts the rate of DNA substitutions. As a model, I used a widespread songbird, the Northern Cardinal, *Cardinalis cardinalis* that consists of multiple island and mainland lineages (Smith Chapter 2). Further, I evaluated whether substitution rate shifts across lineages were present in both the mitochondrial and nuclear genomes. To complete these objectives, I performed a series of tests that began by constructing gene and species trees and then comparing mtDNA and multilocus divergence time estimates. Next, I used coalescent simulations to evaluate whether the difference between mtDNA and multilocus divergence time estimates could be attributed to due randomness associated with the coalescent process. Finally, we examined the correlation between the dN/dS ratio and multilocus estimates of N_e to test whether accelerated evolution occurs in small or large populations.

Methods

I selected 4-6 individuals from each of the six *C. cardinalis* lineages (Smith Chapter 2) and included multiple individuals of six other closely related species (*C. sinuatus*, *C. phoeniceus*, *Rhodothraupis celaeno*, *Periporphyrus eyrthromelas*, *Caryothraustes poliogaster*, and *C. Canadensis*; Klicka et al. 2007) and the outgroup *Molothrus ater* (Supplementary Data). Total genomic DNA was extracted from tissue using the DNeasy tissue extraction kit (Qiagen, Valenica, CA). I amplified nine nuclear markers (Table 6) via polymerase chain reaction (PCR) in 12.5 μl reactions using the

following protocol: denaturation at 94 °C for 10 min, 40 cycles of 94 °C for 30 s, 54 °C for 45 s, and 72 °C for 2 min, followed by 10 min elongation at 72 °C and 4 °C soak. I adjusted annealing temperatures for each intron: ACA, ACO1, βact3, βFib5, MYOI-2 60 °C; EEF 58 °C; HMG 56 °C; ODC 65 °C; Rho1 61 °C (Table 1). PCR products were sent to High-Throughput Genomics Unit (University of Washington) for all subsequent steps. Products were purified using ExoSAP-IT (USB Corporation, Cambridge, MA). PCR products were sent to High-Throughput Genomics Unit (University of Washington, Seattle) for all subsequent sequencing steps. PCR products were purified using ExoSAP-IT (USB Corporation, Cambridge, MA), run through cycle-sequencing reactions and final products were sequenced using BigDye (Applied Biosystems, Foster City, CA) on a high-throughput capillary sequencer. I aligned chromatograms of the forward and reverse strands in Sequencher 4.9 (GeneCodes Corporation, Ann Arbor, MI) and sequences were translated into amino acids to check for premature stop codons.

To resolve introns that had insertions/deletion events between homologous nuclear alleles, I used the program Indelligent (Dmitriev and Rakitov 2008). To phase heterozygous sites in the nuclear introns, I performed three separate runs for each marker in the program PHASE v. 1.2 (Stephens et al. 2001). Haplotypes that had posterior probabilities of < 0.90 were not included in subsequent analyses. I tested for recombination using the six different recombination tests in the program RDP3 (Martin et al. 2010).

Gene trees, species tree, and divergence times

I determined the best-fit sequence model for each gene based on the Akiake Information Criterion (AIC) scores from Mr.Modeltest v. 2.3 (Nylander 2004). To construct gene trees I used MrBayes v. 3.1.2 (Ronquist and Huelsenbeck 2003) and ran this analysis for 10^7 generations with four runs. To determine whether loci were evolving at a constant rate, I performed likelihood-ratio tests in PAUP* 4.0b10 (Swofford 2002) using both the Jukes Cantor and best-fit substitution models. I estimated mean relative substitution rate for each marker (Smith and Klicka 2010) by estimating uncorrected genetic distances from a common outgroup (*C. sinuatus*) and scaled each rate relative to the mtDNA cytochrome-b rate 1×10^{-8} (Lovette 2004).

I constructed a species tree using two recently developed approaches using the programs BEST (Liu and Pearl 2007) and *BEAST (Heled and Drummond 2010). Both programs require *a priori* designation of species or populations, so I used the six *C. cardinalis* lineages that are supported by both mtDNA (Smith Chapter 2) and morphology (Halkin and Linville 1999). These methods assume no gene flow after divergence, an assumption that is valid for *C. cardinalis* (Smith Chapter 2). For the BEST analysis, I performed extensive testing of priors, chain length and sampling output to achieve MCMC convergence and were only able achieve convergence with a single haplotype per lineage. I ran the analysis with two chains and two runs for 10^8 generations and sampled every 10^5. To account for differences between gene trees, I adjusted the default theta and mutation priors [thetapr=invgamma (2.3, 0.06); GeneMuPr=uniform (0.5, 3.5)]. I performed runs with three different sets of randomly selected loci and averaged the results.

For the *BEAST analysis, I used the complete dataset and performed several permutations to look at the impact of partitioning the mtDNA by codon position and the relative effects of strict and relaxed clocks. Additionally, I used *BEAST to estimate divergence times using all the markers (introns and ND2), only introns, and only mtDNA. For the mtDNA I used a relative rate for ND2 and for the nuclear markers, I used published avian intron rates (Axelsson et al. 2004). I used lognormal distributions for the relaxed uncorrelated ND2 rate (mean = 0.0123; SD 0.45), the autosomal intron rate (mean = 0.0018; SD = 0.45), and the sex-linked intron rate (mean = 0.00195; SD = 0.45) and ran the analysis for 20^9 generations sampling every 10^3 generations. All analyses implemented a Yule process on the tree priors and lognormal distributions on sequence model prior distributions. For gene and species trees, I assessed MCMC convergence and determined burn-ins by examining ESS values, likelihood plots, standard deviation of the split frequencies among trees in the programs Tracer v. 1.5 (Rambaut and Drummond 2010) .

I performed an additional divergence time analysis in the Bayesian coalescent program MCMCcoal (Rannala and Yang 2003). The method implements the Jukes Cantor sequence model and assumes no gene flow among species. MCMCcoal uses a gamma (α, β) distribution to specify the prior distribution for the parameters θ, the population mutation rate parameter ($\theta = 4N_e\mu$ for a diploid locus) and τ, the species divergence time parameter), ($\tau = T\mu$; T = species divergence time in millions of years). The shape parameter (α) and scale parameter (β) has a mean α/β and variance $s^2 = \alpha/\beta^2$. I selected the value 1 for all α values because this increased the variance around the mean allowing the prior to account for a wider range of values. All priors were scaled to the

ND2 mutation rate and I set θ (1, 100) to have a mean N_e of 150,000. I used the date of the re-emergence of Cozumel Island 121,000 years ago (Spaw 1979) to set the prior distribution (1, 670) for the *coccineus*/*saturatus* node. For the other nodes, I did not have strong prior knowledge, so I set the τ prior (1, 40) to have a mean divergence in the Early Pleistocene and an interval that covers much of the Pleistocene and Late Pliocene, a period often linked to the onset of intraspecific diversification in birds (Avise et al. 1998). I performed several preliminary runs with other reasonable τ prior distributions that indicated are choice of a vague Early Pleistocene prior was not strongly influencing our results. I ran analyses for 10^6 generations with a burn-in of 10^4 generations.

Additionally, I looked at the effect of locus number on divergence time by performing runs with just mtDNA, then 10 loci, and then five randomly selected loci datasets that included runs from 1 – 9 genes. For each run, I calculated the probability that the values estimated from each reduced dataset was within the 95% credible interval of the 10 loci divergence time estimate.

Coalescent simulations

To evaluate the impact of coalescent variance on divergence times estimated from the protein coding ND2 gene, I performed coalescent simulations using SIMCOAL2 (Laval and Excoffier 2004). I simulated a two population model that coalesced into a single ancestral population under constant growth using the following parameters: five individuals per population, mutation rate = 1.23×10^{-8}, locus length =1041 bp, transition bias = 0.8, gamma distribution shape = 0.5, and four discrete rate categories. I performed five different series of simulations based on the mean multilocus divergence time for

each of the five nodes in the species tree. Each series had independent simulations with the following population effective sizes: 1×10^3, 1×10^4, 2.5×10^4, 5×10^4, 1×10^5, 1.5×10^5, 2×10^5, 2.5×10^5, 3×10^5, 3.5×10^5, 4×10^5, 4.5×10^5, 5×10^5, 1×10^6, and 1.5×10^6. I compared the mean coalescent time of each simulation and identified how different $\Delta\theta_{\text{Node X}}$ ($\Delta\theta_{\text{Node X}} = \left| \theta_{\text{Node X multilocus}} - \theta_{\text{Node X mtDNA}} \right|$) would need to be to get the observed $\Delta T_{\text{Node X}}$ ($\Delta T_{\text{Node X}} = \left| T_{\text{Node X mtDNA}} - T_{\text{Node X multilocus}} \right| / T_{\text{Node X mtDNA}}$). This value was used as the critical value such that if 95% of the observed values fell below critical value, ΔT could not be attributed to coalescent variance.

Substitution rates

I used the program Crann 1.04 (Creevy and McInerney 2003) to estimate the rate of nonsynonymous (dN) and synonymous (dS) amino acid substitutions for the *C. cardinalis* lineages. The dN/dS ratio for each island lineage was compared its sister mainland lineage to evaluate the relative tempo of evolution between lineages with small Ne versus lineages with large Ne. I compared the ratios between island-mainland sister lineages ($[dN_{island}/dS_{island}]/[dN_{mainland}/dS_{mainland}]$, values above 1 are indicators of faster island evolution.

Results

Locus characteristics (chromosome, sequence length, sequence model, polymorphisms, GC content, and mean relative substitution rates) are listed in Table 7. There was considerable variation in the number of variable and parsimony informative sites across loci. Mean relative substitution rates ranged from 3.97×10^{-9} to 6.4×10^{-10} for

the autosomal markers and 1.02 x 10^{-9} for the z-linked marker ACO1. These rates are of similar magnitude to the autosomal (1.8 x 10^{-9}) and z-chromosome (1.95 x 10^{-9}) rates estimated from the turkey and chicken genomes (Axelsson et al. 2004). The six independent recombination tests detected recombination at only one sample in *C. sinuatus* for the HMG locus. Individual intron gene trees showed little resolution for the lineages, whereas the ND2 gene tree was well resolved (Fig. 7). Moelcular clock test using the Jukes Cantor (used in MCMCcoal) and best-fit sequence model (used in *BEAST) allowed us to use the appropriate model for each of our divergence time analysis. Most loci did not evolve in clock-like fashion when the best-fit sequence models were used, but only HMG and ACA failed the clock-like substitution rate test using the Jukes Cantor model (Table 10).

The BEST and *BEAST analyses produced the same topology as the mtDNA gene tree, but node support varied among the analyses (Fig. 8). All analyses indicated that the genus *Cardinalis* is sister to a resolved clade containing the genera *Rhodothraupis*, *Periporphyrus*, and *Caryothraustes*. There was no support (BEST Posterior Probability (PP) = 0.51, *BEAST PP = 0.47) for any preferred relationships among the three *Cardinalis* species (*C. cardinalis*, *C. phoeniceus*, and *C. sinuatus*) in the species tree, a result consistent with the mtDNA tree. Within *C. cardinalis*, *C. c. carneus* is sister (Node A; BEST/*BEAST PP = 1.0) to all other *C. cardinalis* lineages. The next break in the tree (Node B) separates the western clade (*C. c. igneus* and *C. c. mariae*) from the eastern clade (*C. c. cardinalis*, *C. c. coccineus*, and *C. c. saturatus*) and support for this node differed considerably across analyses (BEST PP = 0.59; *BEAST PP = 0.92). In the eastern clade, there is high support (Node C; BEST PP = 0.94, *BEAST PP

= 1.0) for *C. c. cardinalis* (*cardinalis*) being sister to *C. c. coccineus*/*C. c. saturatus*, and high support (Node D; BEST PP = 0.88, *BEAST PP = 0.98) for *C. c. coccineus* and *C. c. saturatus'* sister relationship. Within the western clade there is strong support for the *C. c. igneus* and *C. c. mariae* sister relationship (Node E; BEST PP = 0.98, *BEAST PP = 1.0).

Comparative divergence times

Mitochondrial time estimates (Table 8; 1A, 4A, 4B) were generally older than the multilocus and intron only estimates. Partitioning ND2 by codon positions and had a minimal impact on divergence time estimation (Table 8; 2A-5B). The multilocus MCMCcoal (Table 8; 1B) and *BEAST estimates yielded similar divergence times estimates except for the *saturatus*/*coccineus* node (Table 8). The removal of ND2 from multilocus divergence time estimates did not impact results (Table 8; 4A & 4B). The basal split, separating *carneus* from all other Northern Cardinals and the next divergence separating the eastern and western groups subsequently both occurred in the Early Pleistocene, (Table 8; Nodes A and B). In the eastern group, the divergence across the Isthmus of Tehuantepec occurred in the Early to Middle Pleistocene, separating *cardinalis* from *coccienus* and *saturatus* (Table 8; Node C). Surprisingly the age of both island lineages exhibited an incongruent mean time estimates across approaches. The divergence of Cozumel Island cardinals, s*aturatus*, from mainland Yucatán cardinals, *coccineus*, occurred from the Holocene to Middle Pleistocene, (Table 8; Node D). The timing of the founding of the Tres Marías Islands, represented by the split of *mariae* and i*gneus*, differed substantially between mtDNA and multilocus estimates. The mtDNA

estimate indicated it occurred from the Middle to Early Pleistocene, whereas, the multilocus estimate was more recent in the Late-Early Pleistocene (Table 8; Node E). Multlilocus estimates of N_e indicated that island lineages had lower N_e than mainland taxa and not expectedly mainland taxa had a wide range of N_e (Table 9; θ = 0. 0153 – 0.0418).

The difference between mean mtDNA and multilocus estimates was variable across nodes (Fig. 10; $\Delta T_{\text{Node A}}$ = 13%; $\Delta T_{\text{Node B}}$= 30%; $\Delta T_{\text{Node C}}$ = 43%; $\Delta T_{\text{Node D}}$ = 60%; $\Delta T_{\text{Node E}}$ = 83%). The probability that the mtDNA times estimate was within the multilocus 95% C.I. ranged from 0.78 – 0.00 (Fig. 11). For Node E, the divergence of the Tres Marías island lineage, there was a zero probability that the mtDNA estimate was within the multilocus 95 % CI. There was considerable variance in divergence times estimated from different randomly selected loci. This variance was reduced when more than five loci were used, except for Node B (Fig. 11B). Even when nine loci were used, the variance for Node B ranged 0.20 – 0.94 (Fig. 11B). The smallest variance was seen at node D (Fig. 11), the node with a geologic constraint on the prior distribution.

Coalescent variance and substitution rates

The two islands lineages had the lowest θ values (Table 8) and the largest ΔT, the difference between the mtDNA and multilocus time estimates (Fig. 12). The N_e size that established the critical values for each node were: Node A = 3×10^5; Node B = 5×10^5; Node C = 3.5×10^5, Node D = 1×10^5; Node E = 1×10^6. The only node for which I could reject coalescent variance as an explanation for ΔT was Node E, at which 99% of

the observed $\Delta\theta$ points fell below the critical value. For the four other nodes the percentage of points under the critical value ranged from 3 – 75% (Fig. 12).

Mitochondrial DNA branch lengths for island taxa were longer than the mean branch length for mainland taxa, but this pattern was not observed across the introns. Disparate branch lengths appeared to be randomly distributed across intron trees (Fig. 13. The median dN/dS ratio was 0.04 and the island lineage *mariae* had the highest ratio 0.18 (Fig. 14; Table 11). The ratio of island/mainland dN/dS for *mariae*/*igneus* was 3.39, whereas it was only 0.31 for *coccineus*/*saturatus*.

Discussion

Comparative evolution of island and mainland lineages

Comparative phylogenetic approaches rely on thoroughly understood taxonomic systems in order to make appropriate comparisons. However, the proliferation of DNA data for many biological species has shown that traditionally recognized species are often paraphyletic (Funk and Omland 2003); therefore, rigorously sampled phylogeographic studies are a prerequisite before defining ingroups or sister species. Our study focused on a single biological species, *C. cardinalis* that exhibits both morphological and genetic geographic variation. On the mainland there are four lineages distributed throughout the Mexican lowlands and eastern North America and two distinct island lineages.

To estimate the divergence times of these lineages, I implemented several analytical approaches using mtDNA, multilocus, and nuclear only data with both strict and relaxed molecular clocks. Overall, inferred divergence times were qualitatively similar across all approaches. Even when I removed the mtDNA from our multilocus

runs, I obtained similar results. Additionally, relaxing the mtDNA substitution rate and partitioning by codon position generated older mean time estimates and shifted credible intervals, but did not change divergence time interpretations. The only major difference observed among our different approaches was between mtDNA and multilocus estimates of the divergence times of island lineages from the mainland. The congruence across our time estimates indicated that results were robust to model complexity and the loci used.

The relative tempo of evolution as inferred from mitochondrial and nuclear DNA was incongruent between island and mainland groups. Mainland divergence estimates were generally consistent between the mtDNA and multilocus times estimates. These findings suggest that on the mainland, nuclear introns are evolving at an approximately constant rate, relative to the rate of mtDNA. Conversely, mtDNA time estimates of the age of island lineages were considerably earlier than multilocus estimates even when the mtDNA molecular clock was relaxed. Time estimates with only a few nuclear loci were highly discordant with the mtDNA estimates of the divergence of island lineages from the mainland (Fig.10 and Fig. 11). This finding indicates that the nuclear introns are not evolving at an elevated rate but rather at a rate similar to that of the nuclear introns on the mainland. The larger N_e of nuclear loci and the slower mutation rate may buffer the nuclear genome against island effects caused by small population sizes.

The disparity between the mtDNA and multilocus time estimates for the divergence of the two island lineages from the mainland was substantial, 60% different for the Cozumel Island lineage and 83% different for the Tres Marías Islands lineage. On the other hand, the overlap of the mtDNA and multilocus posterior distributions were quite different. The probability that the mtDNA time estimate was within the 95% CI of

the multilocus data was 0.62 for the Cozumel Island/mainland spilt (Fig. 11D), and a 0.00 for the Tres Marías Island/mainland split (Fig. 11E). Despite the higher probability for the Cozumel Island lineage, the difference between mean estimates is meaningful because the mtDNA estimate is considerably predates the emergence of Cozumel Island 121,000 years ago.

Effective Population Size

Direct estimates of N_e are important because the assumption that island taxa have smaller N_e than mainland taxa may not be always the case. For example, the usage of the geographic area of a taxon (i.e. continent or island area) as a proxy for the size of N_e (Wright et al. 2009) may confound comparisons because there are many narrowly or linearly distributed mainland species. Additionally, island taxa have been shown to colonize the mainland (Filardi and Moyle 2005, Jønson et al. 2011), which may reverse the expected pattern of small and large populations. Despite the uncertainties around estimates of N_e, even with multilocus data, (Liu and Pearl 2007), relative estimates of Ne are informative. I was able to infer relative N_e for the six *C. cardinalis* lineages and show that the two island lineages had lower N_e than those found on the mainland.

Temporal congruence and molecular clocks

Estimating the age of island founder events is challenging because the timing and pattern of these events are often unclear and inferred founder events can be "older" than the island (for a review see Heads 2011). In this study, temporal incongruence with island geology and the age of island lineages was seen in our mtDNA estimates. Cozumel Island

was fully submerged approximately 121,000±6000 years ago (Spaw 1979) and this time period is inconsistent with the age of divergence time estimates based on mtDNA and *BEAST. Only the MCMCcoal multilocus estimate with a prior distribution based on Cozumel Island's geology, had a divergence time after the island remerged. The Tres Marías Islands were likely submerged in the Late Pleistocene although the precise timing of this event is unclear (McCloy et al. 1988). The inferred mtDNA time estimates are highly inconsistent with a Late Pleistocene timeframe, whereas, the multilocus estimates fall within a Middle to Late Pleistocene timeframe. These finding indicate that the problem of "old taxa on young islands" may be due to undetected elevated island rates of evolution instead of an unclear colonization history.

A potential limitation of our study is that I based our time estimates on a universal molecular clock, instead of using an internal fossil or geologic calibration. Despite the findings of a recent study that used multiple independent calibrations and found support for the avian "2% rule" (Weir and Schluter 2008), the implementation of clocklike rates for molecular dating remains controversial (Nabholz et al. 2009, McCormack et al. 2011). For our study, which focused on a limited group of taxa, there were no informative fossils. Additionally, I did not use a geologic calibration because of the potential bias that may be introduced because of the disparity between island and mainland substitution rates. Instead, by using a universal molecular clock (strict and relaxed), I could control for any bias caused by a calibration on an island or mainland lineage. Overall, our inferred time estimates may be improved with newly discovered fossil calibrations, but the findings shown here provide a robust estimate of the relative divergence times between cardinal lineages.

Coalescent variance and elevated substitution rates

The apparent discordance between mtDNA and multilocus time estimates of the age of island lineages are consistent with the conclusion that small populations evolve "faster" than large ones. However, I needed to evaluate whether this discordance was due to an actual increased substitution rate or if instead, the rate only appears to be accelerated due to randomness associated with the coalescent process (Peterson and Masel 2009). This was a potential issue in our study because the mtDNA gene tree had five monophyletic lineages, one paraphyletic lineage, and no haplotype sharing among any lineages, thus there was no information on ancestral N_e to be inferred from the mtDNA. However, the multilocus data had information on the ancestral N_e because the intron gene trees contained unsorted haplotypes and unresolved clades. Therefore, the observed discrepancy between the mtDNA and multilocus divergence time estimates may have been due to the mtDNA diverging prior to the actual divergence event.

I was able to directly evaluate the impact of coalescent variance on branch lengths by determining the expected divergence time of lineages based on simulations. These simulations allowed us to determine how large population sizes needed to be to observe the differences we obtained between our mtDNA and multilocus time estimates. Our results from the simulated distributions were consistent with previous work that has shown coalescent variance is substantial for a single genetic marker (Edwards and Beerli 2000). Despite this expected wide degree of variance, for the Tres Marías Islands group 99% of the observed $\Delta\theta$ values fell below the critical value. Thus, I could strongly reject that coalescent variance was driving the difference in time estimates at this node (Fig.7, Node E). Alternatively, for Cozumel Island lineage, only 30% of the observed $\Delta\theta$ values

fell below the critical value indicating that the discrepancy between times estimates could be attributed to coalescent variance.

By accounting for coalescent variance, I could directly evaluate if the discordance between the mtDNA and multilocus estimates of the Tres Marías founder event was due to an increased substitution rate. There was clear evidence of an accelerated substitution rate in the Tres Marías Island lineage and no indication of rate shifts in the mainland and Cozumel Island groups. *Mariae* had both higher than average nonsynonymous and synonymous substitution rate. The rate of nonysnonymous substitutions was two-fold higher than any other lineage and dN/dS ratio was almost three-fold higher. Our findings support previous work that has suggested that some small populations have elevated rates of substitution because slightly deleterious nonsynonymous mutations are fixed due to genetic drift (Woolfit 2009). The mtDNA in the *mariae* lineage is evolving some four times faster than on the mainland.

Nuclear gene flow among the island and mainland lineages could exhibit a pattern of incongruence between mtDNA and multilocus divergence times (McCormack et al. 2011). Several studies have shown that mtDNA introgression is limited across avian hybrid zones, whereas, introgression in autosomal genes is more extensive (Saetre et al. 1997, Brumfield et al. 2001, Carling and Brumfield 2008b). Hybrid females are believed to be less fit according to Haldane's rule (Haldane 1922) and under this phenomenon mtDNA will not be able to diffuse across a contact zone. Although, I cannot rule out nuclear gene flow, it is a less parsimonious explanation. There is no evidence of hybrid phenotypes, contact zones (Halkin and Linville 1999), or mtDNA haplotype sharing

(Smith Chapter 2). Moreover, the mainland *igneus*, sister to the *mariae* group, is not distributed on the adjacent mainland to the Tres Marías Islands (Fig. 8).

To further understand the underlying mechanism of substitution rate heterogeneity, it will be necessary to show that the mtDNA and multilocus discordance is observed in other taxa with small N_e. Our study, while making advancement in incorporating a species tree and multilocus data, was restricted to two island taxa nested within a single a biological species. The next step is to expand the multilocus coalescent framework to look at the comparative rates of evolution in multiple taxa. The approach described here can be applied to any system that has lineages with both large and small population sizes. A question that became apparent to us is whether or not there is a temporal component to elevated substitution rates in small populations. I observed accelerated evolution in the older of the two island lineages. There may be a lag time from island colonization and when substitution rates begin increasing. Moreover, it is unclear if really old island lineages continue to exhibit increased substitution rates. There may be a fitness cost to island taxa that accumulate slightly deleterious substitutions over long evolutionary periods of time, similar to the constraints assumed to impact asexual lineages (Neiman et al. 2010).

Conclusions

Identifying the underlying causes of substitution rate heterogeneity among taxa has been difficult and recent findings have been contradictory (Woolfit and Bromham 2005, Wright et al. 2009). In this study, I used a unique approach to understanding the tempo of evolution between lineages with small and large population sizes by employing

a multilocus coalescent framework. The approach allowed us to address possible analytical limitations caused by phylogenetic uncertainty and differences in gene tree branch lengths caused by random coalescent processes. Here, I show that the small island populations can evolve faster than in large mainland populations. Multilocus time estimates were not consistent with the timing of island lineage divergence based mtDNA, even when the molecular clock was relaxed and coalescent variance was accounted for. These findings have important implications for the implementation of relaxed molecular clocks. If an extreme shift to a more rapid rate occurs on a phylogenetic tree, as the one recovered for the *mariae* lineage, then the relaxing of the molecular clock may be insufficient for accurate estimation of divergence times. A more robust approach is to incorporate multilocus data because the apparent accelerated evolution in taxa with small populations was not directly apparent in the nuclear genome.

Table 6. Taxon List

Taxon	Collector No.	Code	Country	State	Locality
Cardinalis cardinalis	3986	i11/i12	USA	New Mexico	Grant County
Cardinalis cardinalis	130539	c11/c12	USA	New Mexico	Eddy County
Cardinalis cardinalis	B1868	y1/y2	México	Yucatán	Dzilam de Bravo
Cardinalis cardinalis	BTS05051	c9/c10	USA	Kansas	Montgomery Co.
Cardinalis cardinalis	BTS07071	i3/i4	USA	Arizona	Verde River
Cardinalis cardinalis	BTS07165	y3/y4	México	Campeche	Municipo de la Candelaria
Cardinalis cardinalis	BTS07191	y5/y6	México	Yucatán	Municipo de Uman
Cardinalis cardinalis	BTS08164	a5/a6	México	Michoacán	Lazaro Cardenas
Cardinalis cardinalis	BTS08192	a7/a8	México	Michoacán	Lazaro Cardenas
Cardinalis cardinalis	BTS08217	a9/a10	México	Michoacán	Lazaro Cardenas
Cardinalis cardinalis	BTS08218	a11/a12	México	Michoacán	Lazaro Cardenas
Cardinalis cardinalis	BTS08219	a13/a14	México	Michoacán	Lazaro Cardenas
Cardinalis cardinalis	BTS08220	a15/a16	México	Michoacán	Lazaro Cardenas
Cardinalis cardinalis	CONA202	i5/i6	México	Sonora	Isla Tiburon, El Caracol
Cardinalis cardinalis	CONA1057	a1/a2	México	Michoacán	Presa Infiernillo

Table 6. Taxon List continued.

Taxon	Collector No.	Code	Country	State	Locality
Cardinalis cardinalis	CONA672	i1/i2	México	Baja California Sur	Rancho Monte Alto
Cardinalis cardinalis	DHB2691	i7/i8	USA	Arizona	Santa Cruz Co., Arivaca,
Cardinalis cardinalis	GLS163	s1/s2	México	Quintana Roo	Isla Cozumel El Cedral
Cardinalis cardinalis	QRO510	c5/c6	México	Queretaro	Laguna de la Cruz
Cardinalis cardinalis	GMS220	c1/c2	USA	Florida	Putnam County
Cardinalis cardinalis	ITM016	m1/m2	México	Nayarit	Islas Marías
Cardinalis cardinalis	ITM049	m3/m4	México	Nayarit	Islas Marías
Cardinalis cardinalis	ITM050	m5/m6	México	Nayarit	Islas Marías
Cardinalis cardinalis	ITM086	m7/m8	México	Nayarit	Islas Marías
Cardinalis cardinalis	ITM101	m9/m10	México	Nayarit	Islas Marías
Cardinalis cardinalis	JSB120	c3/c4	USA	Oklahoma	Caddo Co.
Cardinalis cardinalis	MGL17	s3/s4	México	Quintana Roo	Isla Cozumel
Cardinalis cardinalis	PAC 1336	c13/c14	USA	New York	Suffolk Co.
Cardinalis cardinalis	PEP2979	s5/s6	México	Quintana Roo	Isla Cozumel
Cardinalis cardinalis	PEP3009	s7/s8	México	Quintana Roo	Isla Cozumel

Table 6. Taxon List continued.

Taxon	Collector No.	Code	Country	State	Locality
Cardinalis cardinalis	SEZ095	i13/i14	México	Sinaloa	Ejido El Naranjo
Cardinalis cardinalis	TECO08003	a3/a4	México	Guerrero	Tecomate, San Marcos
Cardinalis cardinalis	X8403	c7/c8	USA	Minnesota	Hennepin Co.
Cardinalis cardinalis	Y408254	y7/y8	México	Yucatán	Dzilam de Bravo
Cardinalis cardinalis	Y408257	y9/y10	México	Yucatán	Dzilam de Bravo
Cardinalis cardinalis	Y408284	y11/y12	México	Yucatán	Dzilam de Bravo
Cardinalis phoeniceus	2238	p1/p2	Venezuela	Anzoategui	Piritu
Cardinalis phoeniceus	2239	p3/p4	Venezuela	Anzoategui	Piritu
Cardinalis sinuatus	BTS05044	Cs3/Cs4	USA	Texas	Brewster Co.
Cardinalis sinuatus	JSB222	cs1/cs2	USA		
Caryothraustes canadensis	427231	cc3/cc4	Brazil	Alagoas	
Caryothraustes canadensis	427232	cc5/cc6	Brazil	Alagoas	
Caryothraustes canadensis	KU1273	cc1/cc2	Guyana		
Caryothraustes poliogaster	KU2080	cp1/cp2	México		
Caryothraustes poliogaster	JK01044	cp5/cp6	Honduras		

Table 6. Taxon List continued.

Taxon	Collector No.	Code	Country	State	Locality
Caryothraustes poliogaster	JK01045	cp7/cp8	Honduras		
Caryothraustes poliogaster	KU2097	cp3/cp4	Mexico		
Periporphyrus erythromelas	KU1191	pe2/pe3	Guyana		
Periporphyrus erythromelas	KU1201	pe4/pe5	Guyana		
Rhodothraupis celaeno	CONA 1215	r2/r3	Mexico	Tamaulipas	
Rhodothraupis celaeno	HGOSLP431	r4/r5	Mexico	Tamaulipas	
Rhodothraupis celaeno	X17	r6/r7	Mexico	Tamaulipas	
Molothrus ater	JK03146	MA	USA	Maryland	Maryland Co.

Table 7. Locus Information.

Locus	Chrom.	Base Pairs	Sub. Model	Var. Sites	Pars. Inform. Sites	GC Content (%)	Mean Sub. Rate
ACA [a]	1	1160	HKY+I+G	166	104	54.0	1.29×10^{-9}
ACO1 [a]	Z	947	GTR+G	124	84	38.6	1.02×10^{-9}
βact3 [b]	10	414	HKY+G	70	49	49.0	2.18×10^{-9}
EEF2 [a]	28	729	HKY+I+G	101	63	54.1	2.05×10^{-9}
βFibI5 [a]	4	559	HKY+G	56	35	38.3	1.02×10^{-9}
HMG [a]	23	788	GTR+I+G	83	43	43.6	3.97×10^{-9}
MYOI-2 [c,d]	1	724	HKY+G	59	36	46.0	7.68×10^{-10}
ND2 [e,f]	mito	1041	GTR+I+G	385	349	45.7	1.23×10^{-8}
ODC [g]	3	676	HKY+G	78	40	38.7	6.4×10^{-10}
RHO1 [a]	12	280	HKY+I	34	15	59.0	1.54×10^{-9}

Chromosome (Chrom.); Substitution model (Sub. Model); Variable Sites (Var. Sites); Parsimony Informative Sites (Pars. Inform. Sites); Mean Substitution Rate (Mean Sub. Rate)

[a] Kimball et al. 2009

[b] Carling and Brumfield 2008

[c] Slade et al. 1993

[d] Helsewood et al. 1998

[e] Hackett et al. 1996

[f] Smithsonian Tropical Research Institute

[g] Primmer et al. 2002

Table 8. Mitochondrial DNA and multilocus divergence time estimates.

	Node	Mean (mya)	95 % CI		Mean (mya)	95% CI
1A	A	2.35	(1.81 - 2.93)	**1B**	2.03	(1.52 - 2.54)
	B	1.75	(1.27 - 2.24)		1.23	(0.85 - 1.63)
	C	0.70	(0.33 - 1.05)		0.40	(0.23 - 0.62)
	D	0.18	(0.02 - 0.38)		0.07	(0.01 - 0.20)
	E	1.15	(0.58 - 1.58)		0.19	(0.07 - 0.36)
2A	A	1.82	(1.38 – 2.28)	**2B**	1.85	(1.02 - 2.85)
	B	1.52	(1.09 – 1.96)		1.47	(0.76 - 2.27)
	C	0.44	(0.26 – 0.66)		0.46	(0.20 - 0.78)
	D	0.22	(0.09 - 0.35)		0.23	(0.07 - 0.42)
	E	0.26	(0.09 – 0.46)		0.26	(0.07 – 0.48)
3A	A	1.78	(1.26 - 2.36)	**3B**	1.74	(1.18 - 2.34)
	B	1.43	(0.97 - 1.87)		1.37	(0.94 - 1.79)
	C	0.41	(0.24 - 0.59)		0.41	(0.24 - 0.59)
	D	0.17	(0.06 - 0.29)		0.17	(0.07 - 0.28)
	E	0.23	(0.09 - 0.40)		0.23	(0.07 - 0.39)
4A	A	2.41	(1.31 - 3.77)	**4B**	2.56	(1.38 - 4.01)
	B	1.87	(1.05 - 2.80)		1.89	(1.10 - 2.85)
	C	1.00	(0.46 - 1.66)		1.02	(0.45 - 1.69)
	D	0.33	(0.12 - 0.59)		0.32	(0.12 - 0.56)
	E	1.07	(0.45 - 1.76)		1.08	(0.46 - 1.76)
5A	A	1.88	(1.08 - 2.79)	**5B**	1.93	(1.10 - 2.89)
	B	1.54	(0.86 - 2.31)		1.55	(0.86 - 2.33)
	C	0.48	(0.23 - 0.78)		0.48	(0.22 - 0.78)
	D	0.21	(0.06 - 0.37)		0.20	(0.07 - 0.34)
	E	0.28	(0.11 - 0.49)		0.28	(0.10 - 0.49)

1 MCMCcoal estimate: A) mtDNA only B) multilocus (mtDNA and introns)

2 *BEAST only introns: A) strict clock B) relaxed clock

3 *BEAST multilocus, strict clock mtDNA, strict clock introns: A) ND2 unpartitioned B) ND2 partitioned

4 *BEAST mtDNA only, relaxed clock: A) ND2 unpartitioned B) ND2 partitioned

5 *BEAST multilocus relaxed clock mtDNA, relaxed clock introns: A) ND2 unpartitioned B) ND2 partitioned

Table 9. MCMCcoal θ estimates inferred from mtDNA and multilocus data.

Theta	Mean (mtDNA)	95% CI	Mean (multilocus)	95 % CI
θ_{C}	0.0332	(0.0158 - 0.0608)	0.0418	(0.0253 - 0.0649)
θ_{A}	0.0051	(0.0008 - 0.0147)	0.0287	(0.0208 - 0.0383)
θ_{I}	0.0091	(0.0023 - 0.0229)	0.0292	(0.0143 - 0.0513)
θ_{M}	0.0098	(0.0025 - 0.0254)	0.0097	(0.0047 - 0.0170)
θ_{S}	0.0024	(0.0001 - 0.0107)	0.0065	(0.0012 - 0.0169)
θ_{Y}	0.0070	(0.0013 - 0.0203)	0.0153	(0.0034 - 0.0351)
θ_{SY}	0.0107	(0.0005 - 0.0344)	0.0474	(0.0215 - 0.0786)
θ_{SYC}	0.0093	(0.0002 - 0.0330)	0.0627	(0.0368 - 0.0943)
θ_{MI}	0.0102	(0.0002 - 0.0373)	0.0370	(0.0227 - 0.0556)
θ_{SYCMI}	0.0098	(0.0003 - 0.0354)	0.0446	(0.0158 - 0.0806)
θ_{SYCMIA}	0.0095	(0.0002 - 0.0345)	0.0150	(0.0028 - 0.0337)

Cardinalis cardinalis extant and ancestral lineages: *cardinalis* (C); *carneus* (A); *igneus* (I); *mariae* (M); *saturatus* (S); *coccineus* (Y).

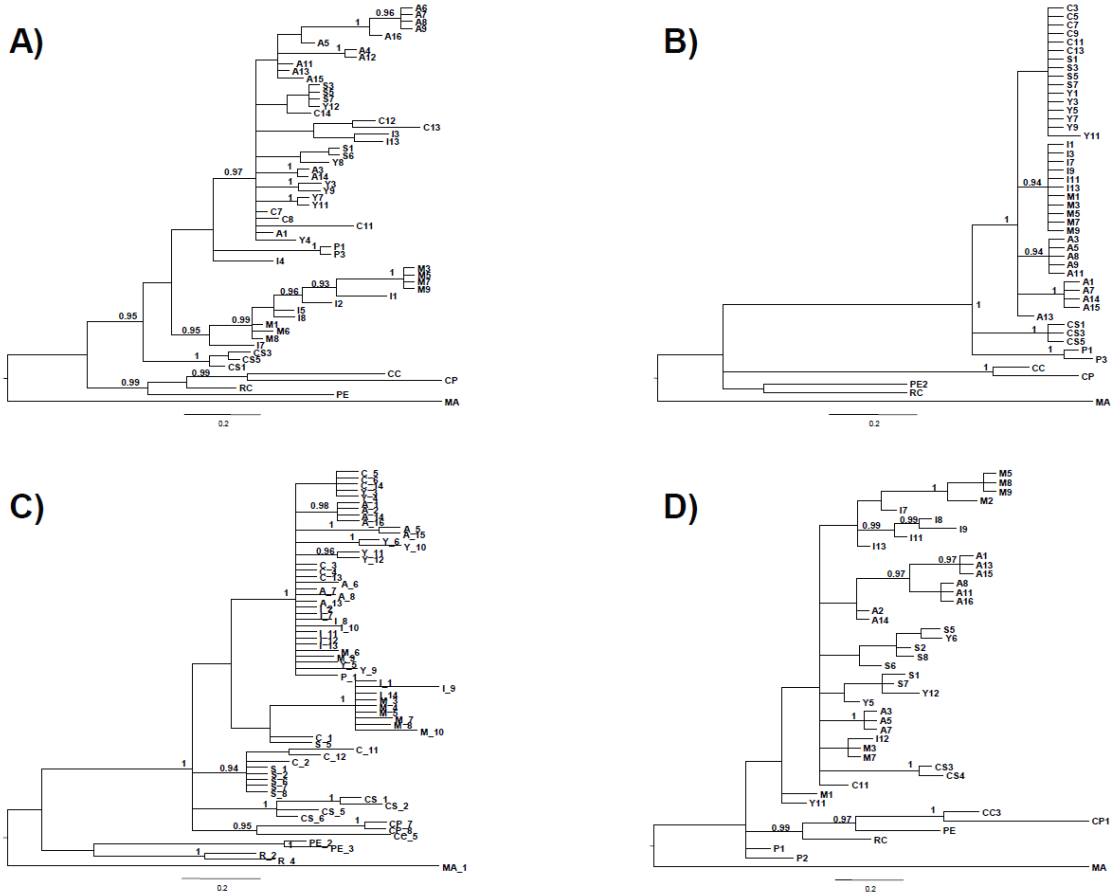
Table 10. Molecular Clock Tests.

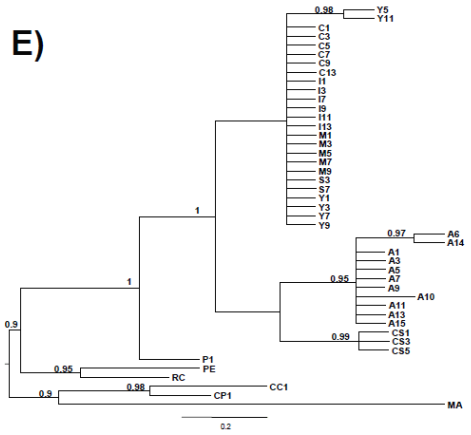
Locus	Substitution Model	Unconstrained Lk	Clock Constrained Lk	DF	2Δ	0.05
ACA	JC	-2328.19	-2378.24	55	**100.11**	73.31
ACO1	JC	-1331.75	-1336.06	70	8.62	90.53
Bact3	JC	-870.59	-900.72	52	60.25	69.83
EEF2	JC	-1402.89	-1433.58	44	61.38	61.66
Fib5	JC	-816.64	-824.57	66	15.86	85.97
HMG	JC	-1765.68	-1829.29	45	**127.21**	61.66
MYO-I2	JC	-1047.12	-1056.98	66	19.73	85.97
ND2	JC	-2341.79	-2363.28	35	42.98	49.80
ODC	JC	-1046.05	-1059.34	66	26.59	85.97
Rho1	JC	-422.90	-429.94	68	14.08	88.25
ACA	GTR + I + G	-2354.13	-2442.20	55	**176.14**	73.31
ACO1	HKY	-1483.47	-1605.84	70	**244.73**	90.53
Bact3	HKY + I + G	-899.20	-937.90	52	**77.41**	69.83
EEF2	HKY + I + G	-1422.00	-1481.02	44	**118.05**	61.66
Fib5	HKY	-855.89	-908.62	66	**105.46**	85.97
HMG	GTR + I + G	-1832.29	-1961.58	45	**258.57**	61.66
MYO-I2	K80	-1108.34	-1152.74	66	**88.79**	85.97
ND2	GTR + G	-2653.56	-2817.04	35	**326.96**	49.80
ODC	HKY	-1065.78	-1103.44	66	75.32	85.97
Rho1	HKY + I	-451.03	-486.02	68	69.99	88.25

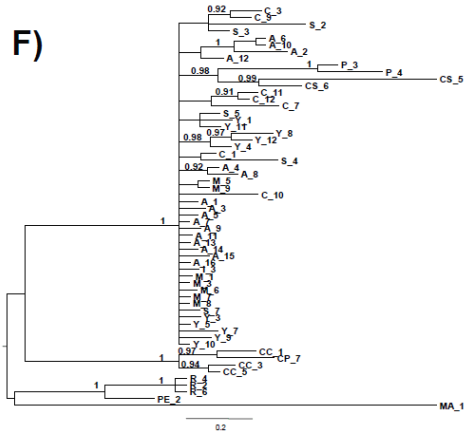
Table 11. Lineage ND2 substitution rates and branch lengths.

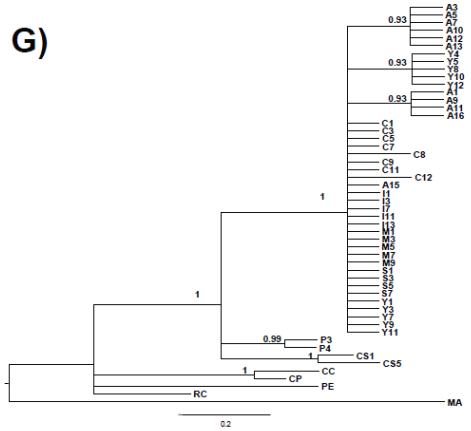
Haplotype	dN	dS	dN/dS	BL
A1	0.004587	0.108657	0.042217	2.886574
A3	0.004587	0.135714	0.033800	2.886070
A5	0.004587	0.103458	0.044338	2.886005
A7	0.004587	0.128692	0.035645	2.886005
C1	0.001524	0.026407	0.057727	3.216227
C3	0.000000	0.013337	0.000000	3.020157
C5	0.000000	0.035139	0.000000	3.103538
C7	0.000000	0.021523	0.000000	3.151217
I1	0.001524	0.024199	0.062994	3.202938
I3	0.001524	0.032541	0.046845	3.203484
I9	0.001524	0.024199	0.062994	3.202552
I7	0.001524	0.035373	0.043095	3.203127
M1	0.009231	0.051334	0.179823	3.549841
M3	0.009231	0.045660	0.202170	3.549306
M5	0.009231	0.054315	0.169952	3.548721
M7	0.009231	0.051334	0.179823	3.555508
S1	0.001524	0.026492	0.057541	3.321141
S3	0.000000	0.015564	0.000000	3.321154
S5	0.000000	0.015598	0.000000	3.321952
S7	0.000000	0.015564	0.000000	3.321677
Y1	0.001203	0.032736	0.036754	3.157226
Y3	0.001203	0.016003	0.075184	3.158116
Y5	0.000000	0.005301	0.000000	3.208751
Y7	0.001203	0.016003	0.075184	3.208537

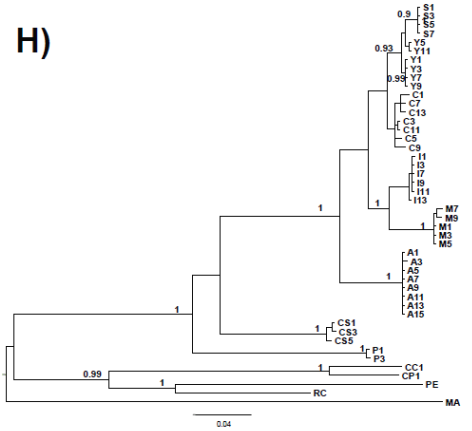
Cardinalis cardinalis lineages: *cardinalis* (C); *carneus* (A); *igneus* (I); *mariae* (M); *saturatus* (S); *coccineus* (Y). Nonsynonymous amino acid substitution rate dN; synonymous amino-acid substitution rate (dS); ND2 branch length estimated from a common outgroup (BL)


A)
B)
C)
D)


E)
0.98
Y5
Y11
C1
C3
C5
C7
C9
C13
I1
I3
I7
I9
I11
I13
M1
M3
M5
M7
M9
S3
S7
Y1
Y3
Y7
Y9
1
0.97
A6
A14
A1
A3
A5
A7
A9
1
0.95
A10
A11
A13
A15
0.9
0.99
CS1
CS3
CS5
P1
0.95
PE
RC
0.98
CC1
0.9
CP1
MA
0.2


F)
0.92
C_3
C_9
S_2
S_3
A_6
A_10
A_2
A_12
1
P_3
0.98
P_4
0.99
CS_5
CS_6
0.91
C_11
C_12
C_7
S_5
Y_11
0.97
Y_8
0.98
Y_12
Y_4
C_1
S_4
0.92
A_4
A_8
M_5
M_9
C_10
A_1
A_3
A_5
1
A_7
A_9
A_13
A_14
A_15
A_16
M_1
M_2
M_6
M_8
S_6
S_8
S_3
Y_5
Y_7
Y_9
Y_10
0.97
CC_1
CP_7
0.94
CC_3
CC_9
1
R_4
R_5
R_6
1
PE_2
MA_1
0.2


G)
A3
A5
0.93
A7
A10
A12
A13
Y4
0.93
Y5
Y8
Y10
Y12
A1
0.93
A9
A11
A16
C1
C3
C5
C7
C8
C9
C11
C12
A15
1
I1
I3
I7
I11
I13
M1
M3
M5
M7
M9
S1
S3
S5
S7
Y1
Y3
Y7
Y9
Y11
1
0.99
P3
P4
1
CS1
CS5
1
CC
CP
PE
RC
MA
0.2


H)
0.9
S1
S3
S5
S7
0.93
Y5
Y11
Y1
0.99
Y3
Y7
Y9
C1
C7
C13
C3
C11
C5
C9
I1
I3
I7
I9
I11
I13
1
1
M7
M9
M1
M3
M5
A1
A3
A5
1
A7
A9
A11
A13
A15
1
CS1
CS3
CS5
1
P1
P3
0.99
1
CC1
CP1
PE
1
RC
MA
0.04

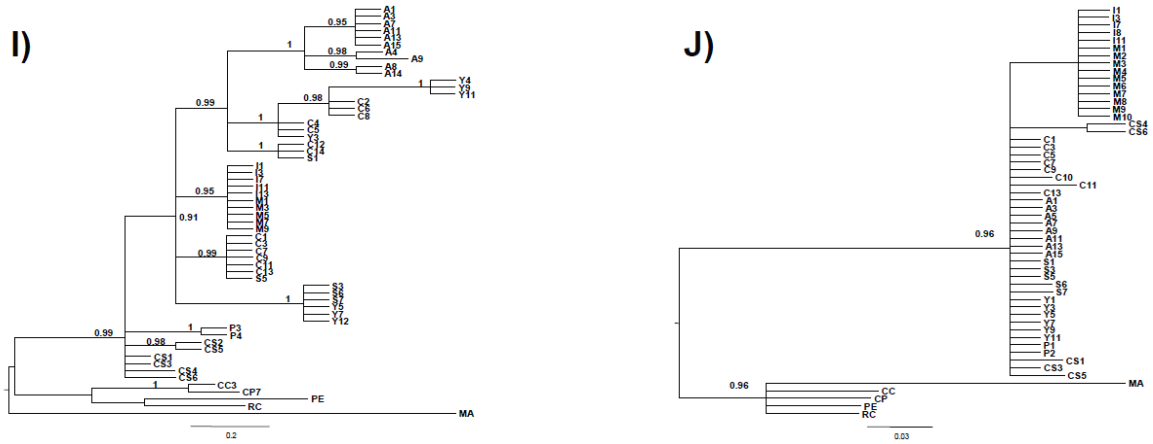


Figure 7. Gene trees. A) ACA; B) ACO-1; C) βact3 D) EEF2; E) βFibI5; F) HMG; G) MYOI-2; H) ND2; I) ODC; J) RHO1.

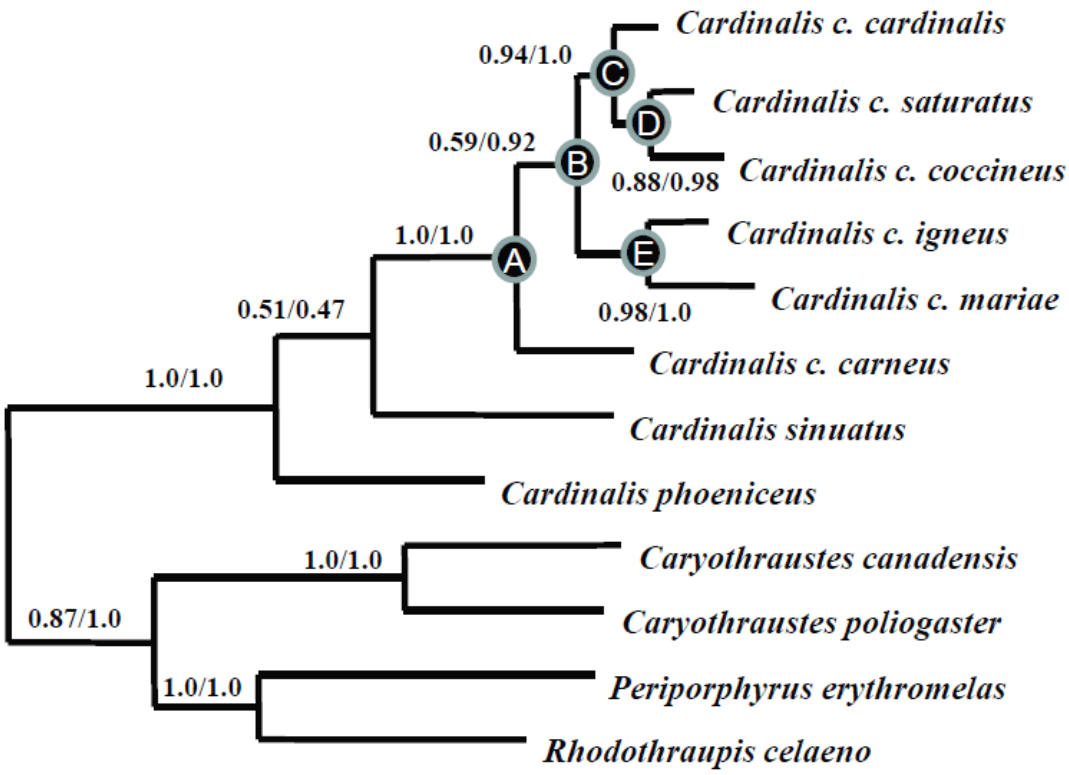


Figure 8. Species Tree for *Cardinalis cardinalis* and allies. Posterior Probabilities averaged across multiple runs are shown: BEST/*BEAST. Letters correspond to nodes in *C. cardinalis* species tree.

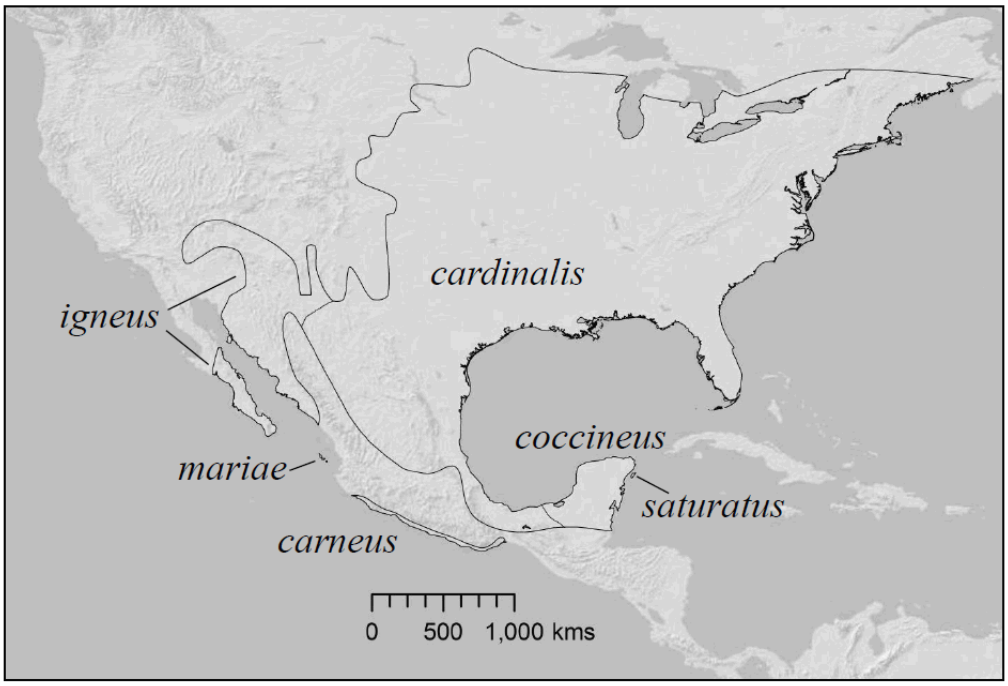


Figure 9. Range map of *Cardinalis cardinalis* lineages.

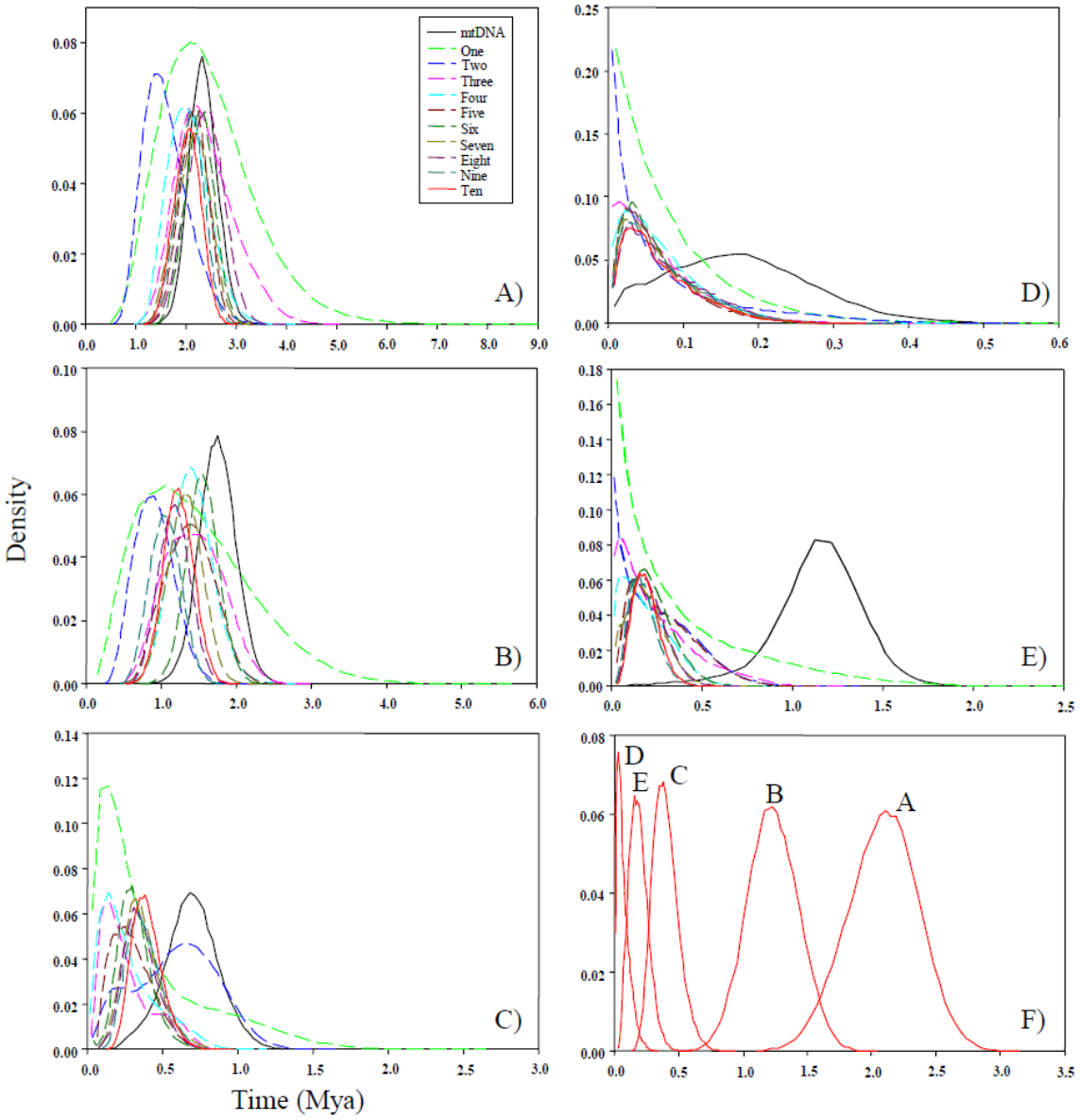


Figure 10. Posterior Distributions of MCMCcoal divergence time estimates. Each graph represents a species tree node and the colors correspond the number of loci used to estimate divergence time, million years ago (mya). A) *carneus* – all other lineages; B) *igneus*/*mariae* – *cardinalis*/*coccineus*/*saturatus*; C) *cardinalis* – *coccineus*/*saturatus*; D) *coccineus* – *saturatus*. E) i*gneus* – *mariae*; F) multilocus estimate for all five nodes

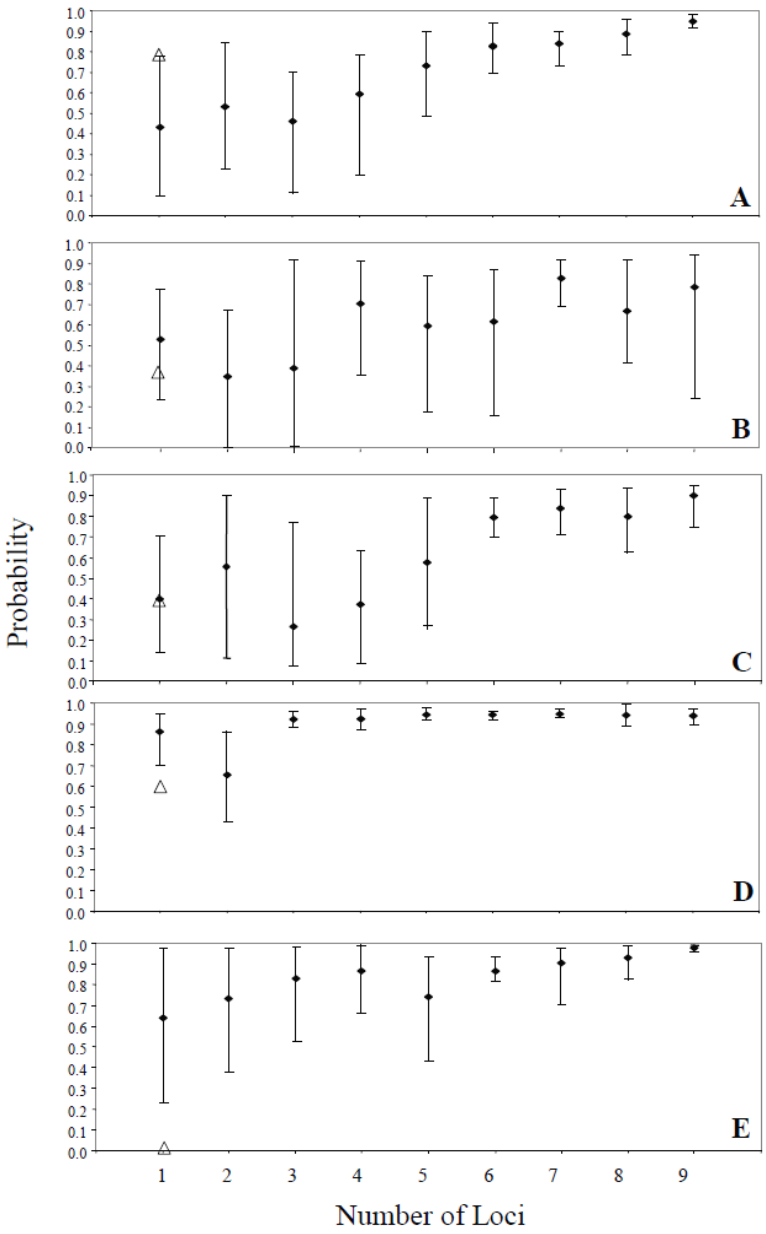


Figure 11. Probability of locus partition estimates falling within the 10 loci time estimate 95% CI. Each node on the species tree is shown. Five independent runs with randomly selected loci for each locus number partition, mean and variance of each partition are shown based on. Triangles show probability of mtDNA time estimates.

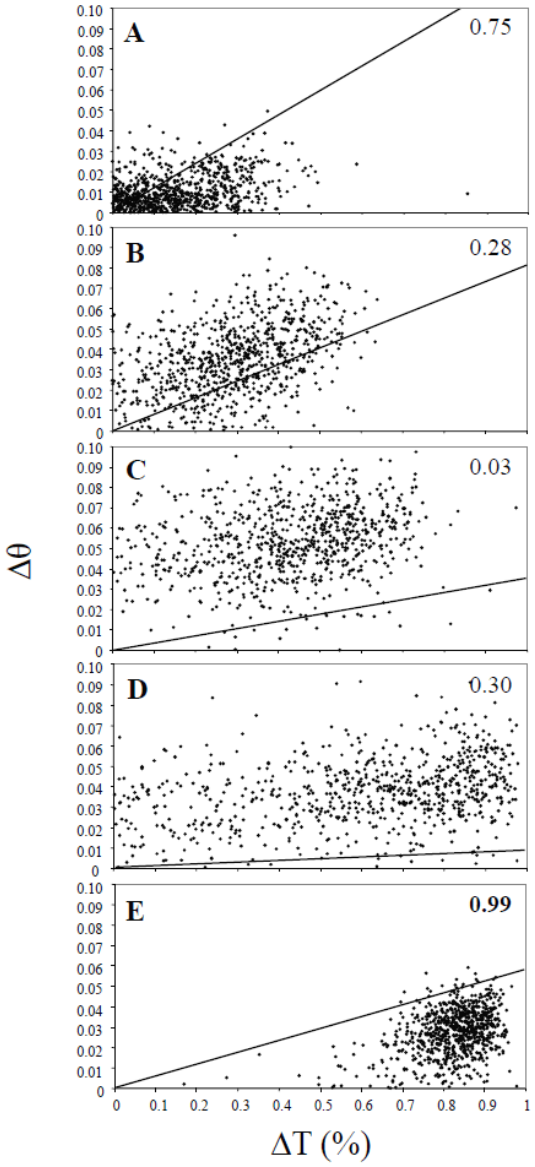


Figure 12. Coalescent variance test. Each graph represents $\Delta\theta_{\text{Node X}} = |\theta_{\text{Node X multilocus}} - \theta_{\text{Node X mtDNA}}|$ plotted against $\Delta T_{\text{Node X}} = |T_{\text{Node X mtDNA}} - T_{\text{Node X multilocus}}|$ for each species tree node. The lines represent the critical points at which the ΔT is not attributable to Δθ. Coalescent variance was rejected if 95% of the observed data fell below the critical value. Observed frequencies under the critical value are plotted on graph.

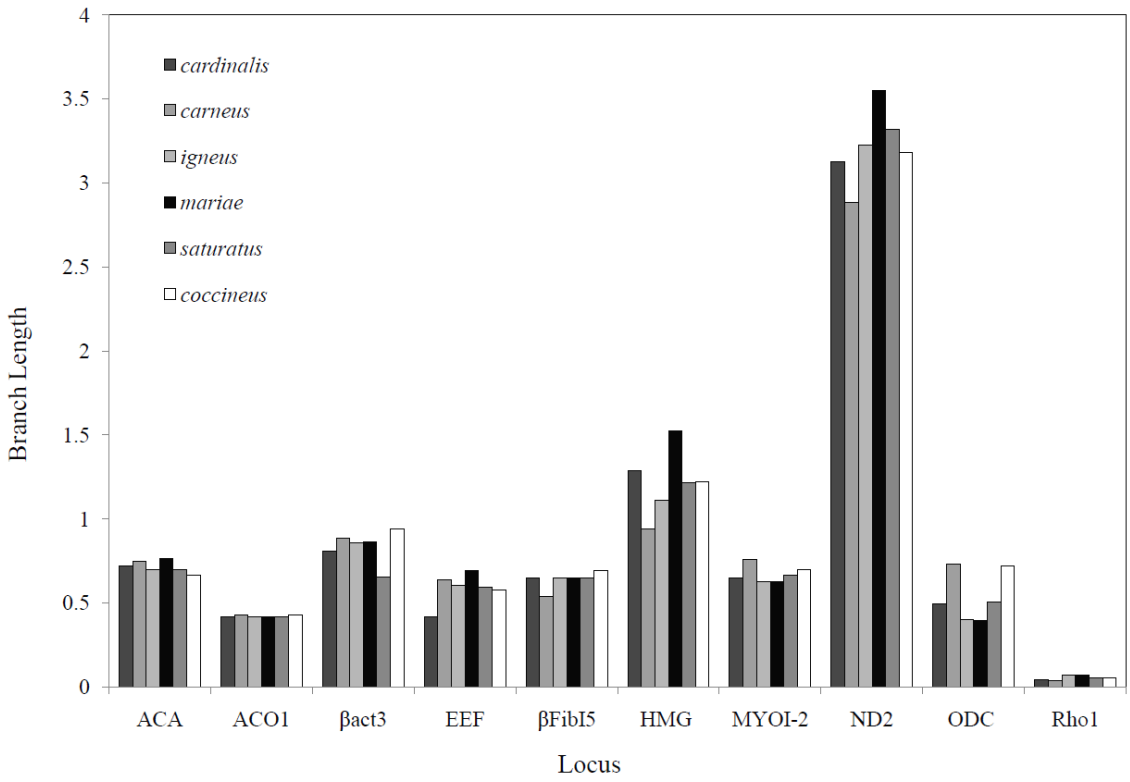


Figure 13. Mean branch length of lineage from common outgroup for each locus.

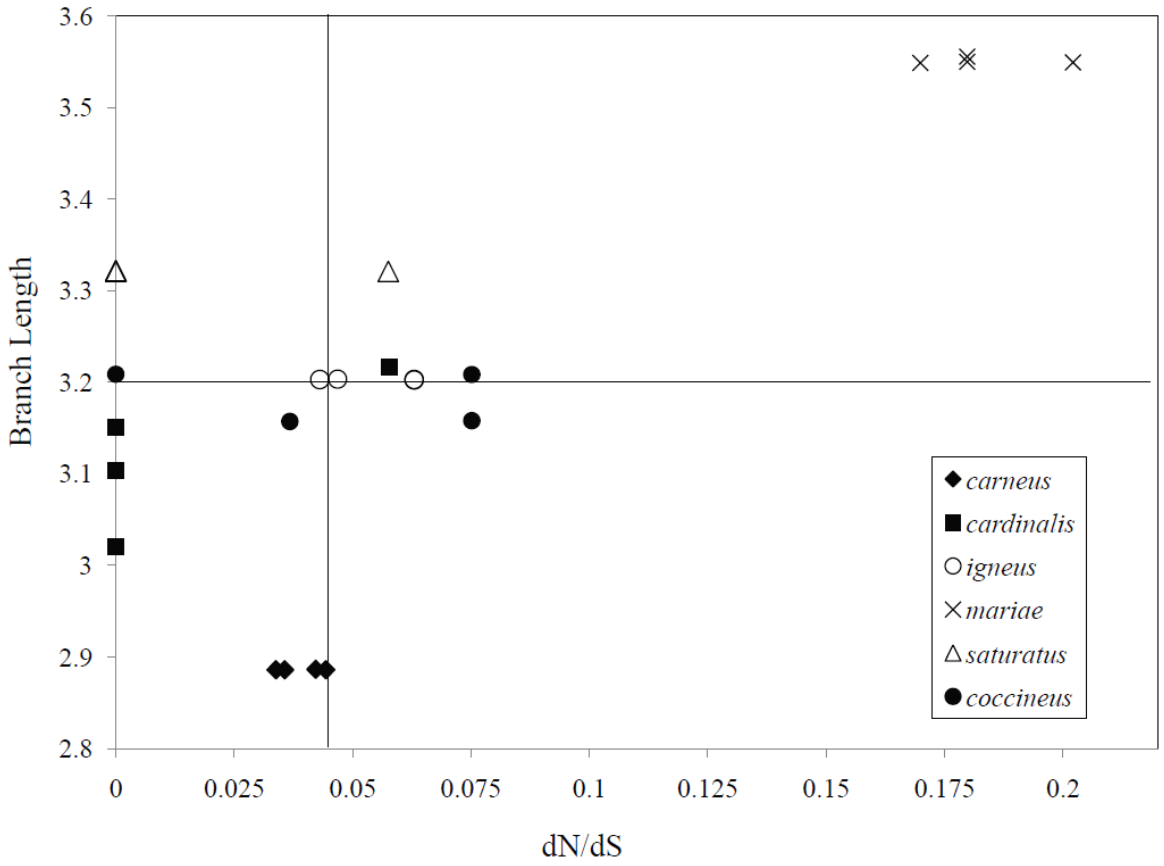


Figure 14. Mitochondrial DNA nonsynonymous/synonymous substitution rate ratio plotted against lineage branch length to a common outgroup. Results for multiple haplotypes per lineage are shown. Median branch length (horizontal line) and median dN/dS (vertical line).

CHAPTER 4

THE GREAT AMERICAN BIOTIC INTERCHANGE IN BIRDS

Introduction

The continents of North and South America have been isolated from one another throughout most of their histories. The fossil record indicates that a transient Beringian connection between North America and eastern Asia existed throughout much of the Cenozoic (Simpson 1947, Tiffney and Manchester 2001). Unlike North America, South America was an island continent during this time, allowing for the evolution of an ancient and largely endemic biota (Patteron and Pascual 1972, Simpson 1980, Vuilleumier 1984, Ricklefs 2002). Species assemblages unique to each continent met abruptly some 3 million years ago (Coates and Obando 1996) (4.0 - 2.5 million years ago [mya]) upon the final uplift of southern Central America and the formation of a new terrestrial corridor, the Panamanian land bridge. This uplift initiated a process that led to unprecedented ecological and evolutionary consequences for previously isolated biotas (Simpson 1980), an event now referred to as the Great American Biotic Interchange (GABI).

The mammalian response to the land bridge formation was written from the rich fossil deposits on both continents. Among the many evolutionary inferences gleaned from the mammalian fossil record are: 1) immigrant taxa appeared on each continent soon after the formation of the Panamanian Isthmus (Webb 1976); 2) an early wave of xeric-adapted species was followed by a second wave of mesic-adapted species (Webb and Rancy 1996); 3) there was an equal exchange of taxa between continents (Marshall et al. 1982, Webb and Marshall 1982); and 4) most exchange had ceased by the onset of the mid Pleistocene (Webb 1976). Over time, lineages moving into South America were

more “successful” than those moving in the opposite direction. Today, more than 50% of modern South American genera have North American origins whereas only about 10% of North American mammal genera are derived from southern immigrants (Webb and Marshall 1982).

Avian bones are thin and light relative to those of mammals and as a consequence fossilize poorly. The depauperate nature of the avian fossil record precludes a direct fossil-based comparison of the responses of birds and mammals to the land-bridge formation (Mayr 1964, Olson 1985, Vuilleumier 1985). The few avian fossils available do suggest some interchange during the late Cenozoic, although the timing is unclear (Vuilleumier 1985). Furthermore, there is the perception that because birds can fly, their dispersal abilities differ from non-volant organisms such as terrestrial mammals (Lomolino et al. 2006). The colonization of distant islands by birds provides ample evidence that they can cross water barriers. The over-water dispersal abilities of some groups of birds are well known (see “tramps” *sensu* Mayr and Diamond 2001; *Turdus* thrushes, Voelker 2009). However, other species of birds, such as those found in the Neotropical understory, do not readily disperse across water gaps (Hayes and Sewlal 2004) and the propensity for such dispersal is simply unknown for most birds. Thus, one cannot determine from the evidence at hand whether avian lineages regularly crossed between North and South America prior to the land bridge completion.

Intercontinental dispersal may be approximated by identifying trans-isthmus diversification events in well-resolved molecular phylogenies and then applying a molecular clock. Although, molecular clocks are not without controversy (see Methods) they provide a previously unavailable perspective on how birds responded to the presence

of a terrestrial corridor connecting these previously isolated continents. I approach this question in two ways. First, I used molecular data and a phylogenetic tree-based approach to summarize the timing and patterns of recent avian trans-isthmus exchange. I then compared discernible patterns to those inferred from the fossil record for mammals and I evaluated dispersal direction, and the influence of elevation and habitat preference on exchange. Second, I used a taxonomic approach in which I summarized and contrasted the historical and present-day distributions of Nearctic and Neotropical families to understand how the deep history of avian groups has affected present day diversity patterns of New World birds. When combined, my analyses allowed us to determine the relative role of the Panama land bridge formation on avifaunal exchange at different temporal and taxonomic scales.

Methods

I performed a literature search for molecular phylogenies on New World avifauna. Studies were selected if they had complete or near complete taxon sampling and if the constituent taxa included both North and South American lineages. Suitable datasets were downloaded from Genbank using the program Geneious v. 2.5.4 (Drummond et al. 2006). In all, I compiled 64 molecular phylogenetic studies of New World birds, summarizing those events in which sister lineages have diverged after intercontinental dispersal. The birds that I studied are an ecologically and taxonomically diverse group that includes representation from 11 orders, 34 families, and over 100 genera (complete electronic taxon list provided on http://www.oikos.ekol.lu.se/app.html).

Range evolution

In order to identify trans-isthmus diversification events I reconstructed phylogenetic trees for each data set in order to perform range evolution analyses. To select a model of sequence evolution for each dataset I used the program MrModeltest v. 2.3 (Nylander 2004) and the Akaike Information Criterion (AIC). I constructed Maximum Likelihood phylogenetic trees in the program PAUP* 4.0b10 (Swofford 2002) using the heuristic tree search, (TBR branch swapping, and 10 random addition replicates). Tree topology was verified by comparison to the original published tree. I then performed a likelihood ratio test for clocklike evolution. For trees that were not clock-like, I used the program BEAST v. 1.4.8 (Drummond and Rambaut 2007) and applied a relaxed molecular clock (molecular rates are discussed below) with a Yule process speciation prior. Each analysis was initially run for 10,000,000 generations and sampled every 1000. I used the program Tracer v1.4 (Rambaut and Drummond 2007) to assess the mixing of MCMC chains in the analyses and to determine the burn-in. If the runs did not converge or I did not achieve reliable ESS values (> 200) I re-ran the analysis for more generations. All trees generated prior to the point of stationarity Ire discarded. In cases where the greater topology was unresolved, I reduced the dataset to the clade of interest.

To estimate the direction of dispersal I employed the dispersal-extinction-cladogenesis (DEC) model (Ree and Smith 2008) in the program Lagrange 2.0.1 (Ree et al. 2005). I defined six geographically relevant New World regions; North America (from Panama, north of the Darien, to Alaska), South America, the Greater Antilles, the Lesser Antilles, the Hawaiian Archipelago, and the Galapagos Archipelago. For each dataset, I set the basal divergence time to 1.0 because I was only interested in the relative timing of

events for this particular analysis. I chose not to place dispersal time constraints on the analyses in order to optimize the data and allow all practical range evolution scenarios to occur. On trees that had an unresolved pre-isthmus ancestral area, I placed an ancestral area constraint to reduce uncertainty concerning the geographic origin of the clade. To justify the use of such a constraint, I performed ancestral state reconstruction using parsimony in Mesquite v. 2.6 (Maddison and Maddison 2009). Constraints were applied only if ancestral state reconstruction inferred that the basal node in a phylogeny was unambiguously assigned to one continent, and that node was older than 4 million years. This procedure eliminated a discontiguous ancestral area in both North and South America prior to the isthmus uplift.

Diversification times

For each phylogeny, I identified diversification (cladogenetic) events between North and South American taxa that represent a dispersal event across the isthmus. I recognize that a lag time between dispersal and diversification events exists, and that the duration of this time varies across species depending on such factors as ecology and dispersal ability. For this study, I assumed that the lag time between dispersal and lineage diversification was negligible with respect to the evolutionary time scale of the interchange. I reasoned that although that some intercontinental gene flow likely persisted for a short time after dispersal occurred, the relatively linear configuration of southern Central America, the geographic bottleneck at its contact point with South America, and the climatic and geologic turmoil that occurred at the time of the isthmus uplift led to relatively rapid genetic isolation. To improve precision, mtDNA sequences of all the

descendant lineages arising from each dispersal/diversification event (called diversification events hereafter) were re-run through Mr.ModelTest v. 2.3 (Nylander 2004) to obtain a sequence model that was unaffected by the phylogeny it was nested in. I re-ran each individual dataset through BEAST following the same methodology as described above. To evaluate the robustness of our results, I additionally performed analyses with alternative clock calibration times, a uniform clock, strict clocks, as well as uncorrected genetic distances.

Four different mitochondrial genes were used in the trans-isthmus dataset: cytochrome *b* (cyt-*b*), cytochrome oxidase I (COI), NADH dehydrogenase subunit II (ND2) and ATP-synthase 6 and 8 (ATPase6&8). I assessed the variability of rates among mitochondrial DNA (mtDNA) loci by using pairwise comparisons of average intrageneric genetic distances for each gene used (see Fig. 15). These analyses indicated that cyt-*b* and COI evolve at similar rates and for these I applied the widely used cyt-*b* rate of 2% sequence divergence per million years (0.010 substitutions/site/lineage/million years, [s/s/l/m]) (Lovette 2004). Our comparisons indicated that ND2 and ATPase6&8 evolve at approximately 1.25 times the rate of cyt-*b* (slope of regression line, Fig. 14), thus I applied a rate of 2.5% per million years (0.0125 s/s/l/m) for these genes.

Rate heterogeneity among taxa and across genes has led some researchers to question the general utility of a molecular clock (Pereira and Baker 2006, Nabholz et al. 2009) however, other studies have shown that when analyzed carefully (as I have done here) divergence time estimates can be robust and informative (e.g. Weir and Schluter 2008).

I acknowledge other sources of bias in estimates based on a single locus, such as mtDNA. For example, I did not correct our data for ancestral polymorphism (Arbogast and Slowinski 1998); nor, did I address time dependent rate variation (Ho et al. 2005). Both issues potentially affect a small proportion of our estimates in the tail of our distribution. Had I addressed these potential errors, a subset of points would be moved toward the present, strengthening our interpretation of the results. I did not account for coalescent variance (Edwards and Beerli 2000). Simulation studies have shown that a single vicariant event can generate a wide range of genetic distances among geminate species pairs (Hickerson et al. 2006). Discrete diversification pulses that Ire undetected with our analyses may have occurred during the GABI. Finally, any lag time between gene and population divergence (Edwards and Beerli 2000) would cause some of our estimates, particularly those associated with more recent diversification events, to be younger than I estimated.

Isthmus of Panama completion time estimates

The estimated time period of the final isthmus uplift and the subsequent formation of a terrestrial corridor have been inferred from multiple lines of evidence. Although most researchers agree that the Panamanian land bridge was present by the mid Pliocene, there is no consensus on precisely when this event occurred. Evidence based on marine deposits suggests that the connection between the Caribbean and Pacific was cut off by approximately 3.1 mya (Coates and Obando 1996). Radio-isotopic dating from mammalian fossil beds indicates an approximate Isthmus closure at 3 mya (Marshall et al. 1979); however, these same mammalian data have been interpreted as evidence for a

more recent Isthmus closure at 2.5 mya (Stehli & Webb 1985). The most recent geological estimate indicated that North and South America may have been connected as early as 4 mya (Kirby et al. 2008). To account for the uncertainty in isthmus closure times, I used two different estimates for the beginning of the post-land bridge period, the widely used date of 3.1 mya (Coates and Obando 1996) and the most recently published estimate of 4 mya (Kirby et al. 2008).

Phylogenetic tree-based comparative diversification analyses

To test the hypothesis that the land bridge was critical for avian exchange I compared observed diversification time estimates with appropriate null models (Zink et al. 2004). I used two complementary approaches. First I simulated distributions of diversification events using the program PHYL-O-GEN v. 1.2 (Rambaut 2002). These distributions were generated with parameters that would reflect a range of possible diversification scenarios. I used two different constant growth scenarios moderate speciation (birth rate = 0.2, extinction rate = 0.2); and high speciation (birth rate = 0.8, extinction rate = 0.2). I also generated distributions affected by extinction pulses, using a birth rate of 0.2 and episodic extinction events that removed 20% (low extinction pulse) and 80% (high extinction pulse) of the lineages in a simulation. I converted branching events in simulated phylogenies into time using the program LASER (Rabosky 2006) by setting the basal divergence of each simulation to 12 million years. Diversification events were periodically sampled to generate a simulated distribution. To make the distribution comparable to the distribution of empirically derived diversification events, all points less than 0.04 million years were omitted.

I realized that simulated distributions may not capture the complexity of diversification occurring within and among multiple avian lineages. To address this issue, I employed a second approach that used the results from our BEAST analysis to generate a distribution of all diversifications occurring within 28 completely sampled avian genera that are represented in both North and South America. Temporal estimates of all diversifications (n = 413) occurring within the 28 selected genera were then compiled into a single distribution that represents an empirically derived estimate of background diversification over time. To allow direct comparison with our distribution of isthmus related diversifications (n = 135), I generated a distribution that represents a temporal baseline of cladogenetic events occurring over the last 8 million years. From this distribution I took 100 random draws of 135 diversification events. To compare the simulated and empirical distributions of diversification events with distribution of trans-isthmus events, I used the nonparametric two-sample Kolmolgorov – Smirnov test (K-S test) across the entire temporal distribution (0 – 8 mya), as well as to pre-land bridge (4 – 8 and 3.1 – 8 mya) and post-land bridge (0 – 4 and 0 – 3.1 mya) time intervals.

Historical and ecological diversification patterns

I further explored the timing and pattern of dispersal by performing additional K-S tests on direction of dispersal and the potential role of elevation and habitat preference. Following the designations of Stotz et al (1996), I separated lineages into two broad elevation categories, highland and lowland. The transitional zone between montane and lowland habitats varies across mountain ranges and regions in the Neotropical region (Stotz et al. 1996). To account for this complexity, I defined lowland birds as those never

occurring above 2000 m and highland birds as those never occurring lower than 1000 m (*fide* Stotz et al. 2006). This analysis was used to test the hypothesis that birds occupying montane habitat "islands" may be predisposed to disperse with greater regularity than lowland forest birds. Lineages were also divided into two broad ecological categories based on predefined habitat and foraging strata categories (Stotz et al. 1996). These included birds that are obligately dependent on the vertical structure occurring in true forest habitats (hereafter: forest birds) and "non-obligates," birds occurring predominantly in successional, edge, and scrub habitats (hereafter: edge/scrub birds). This approach allowed us to use birds to evaluate the ecogeographic model (e.g. alternating periods of rainforest and drier, more open habitat corridors) that has been proposed to explain the mammalian interchange (Webb 1991, Vrba 1992).

Taxonomy–based comparative biogeographic analysis

The timing of avian exchange is not only important for discerning the impact of the Panamanian land bridge but it also provides critical insight into how the ancestral origin of avian groups affected the present-day distribution of New World avian diversity. To examine the contribution of historical assemblages to present-day patterns of diversity we identified New World avian families for which an ancestral origin had been inferred using molecular data. For each family, I summarized the number of species now inhabiting either Nearctic or Neotropical regions (source: Sibley and Monroe 1990). To test whether families with northern ancestral origins have a greater number of species in the Neotropical region than families with southern ancestral origins have in the

Nearctic (see Mayr 1964), I used a generalized linear model with a quasibinomial error term to account for over dispersion.

Results

I identified 135 trans-isthmus diversification events that occurred at multiple taxonomic scales including intra-specific, inter-specific, inter-generic, and inter-familial branching events (http://www.oikos.ekol.lu.se/app.html). Overall, our results indicate that the land bridge was critical in facilitating intercontinental exchange and subsequent diversification; although, a number of geologic, geographic, and climatic factors that are known to have occurred during this time period certainly played a role. I obtained qualitatively similar results irrespective of the clock approach used (results not shown), suggesting that considerable signal is present in these data. Diversification between North and South America occurred from the mid-Miocene through the Late Pleistocene, with 76% of the trans-isthmus diversification events I identified occurring within the last 4 million years. A histogram of trans-isthmus diversification times (Fig. 16) suggests a pulse from 4-2 mya, coincident with the land bridge formation. This pulse was followed by a pronounced decrease in the frequency of diversification events during the Pleistocene.

The distribution of trans-isthmus diversifications differed significantly from all simulations (Fig. 17; Table 11) but I could not reject the background distribution (i.e. all diversifications accruing within 28 genera). However, when our data are partitioned into biologically relevant time intervals, more precise patterns of diversification are recovered. All comparisons for the earliest time interval (4 – 8 or 3.1 – 8 mya) indicate

that early diversification events between North and South America followed the expected null diversification rate (Table 12). Importantly, this trend strongly shifts in the recent interval (regardless of which isthmus closure estimate is used), a time attributed to the GABI. During this period, the frequency of trans-isthmus diversifications exceeds those derived from both the simulated and background distributions (Table 12). Within the post-land bridge period, a decline in diversification is evident in the isthmus-related and background distributions, a phenomenon shown to be common in birds (Weir 2006, Phillimore & Price 2008). I suggest that the notably steeper decline apparent in the trans-isthmus distribution reflects a density dependent effect.

Within the tropics, the direction of exchange (Fig. 18a) was symmetrical and reciprocal (North to South, n = 52; South to North, n = 50; K-S test, D = 0.195, p = 0.259) but different trends emerge following the onset of the Pleistocene. Dispersals into South America slowed during this time while dispersal into North America reached its highest frequency. The role of elevation on trans-isthmus dispersal appears to be negligible. Our comparison between lowland (n = 72) and highland (n = 21) birds (Fig. 18b) indicates no difference between these groups (K-S test, D = 0.2480, p = 0.233). I did, however, detect potentially different temporal trends for forest and scrub/edge birds (Fig. 18c). The mean diversification time of scrub/edge forest birds (n = 25; 3.33 mya, 2.55 - 4.11) was older than the obligate forest birds (n = 49; 2.67 mya, 2.07 - 3.27) although the distributions were not significantly different (K-S test, D = 0.248, p = 0.224).

When family-level biogeography was examined, I detected contrasting evolutionary histories (Table 13). Of the 12 families I identified as having northern

origins, three are constrained today to the tropical region while the remaining nine occur widely across both Nearctic and Neotropical regions. Conversely, of the 17 families I identified as having southern (i.e. South American) origins, most invaded tropical Central America but only two were able to push further north to colonize Nearctic regions. The families with northern origins had a significantly greater proportion of species in the Neotropical region (83.4% of species from northern families now occur in the Neotropical region) than families with southern origins in the Nearctic (2.7% of species from southern families occur in the Nearctic; $2*P\ [t1>-9.627] <0.001$).

Discussion

Birds and mammals, a shared history

Molecular and fossil based evolutionary reconstructions can be affected by forms of bias unique to each approach, making direct comparisons between such studies difficult. Nevertheless, the avian response to the Isthmus of Panama formation that I describe shares a close resemblance to the late Pliocene interchange described for mammals. The temporal consistency between the avian and mammalian histories underscores the dramatic impact that the newly formed terrestrial corridor had on the two previously isolated biotas. For both birds and mammals, some lineages were able to disperse across the water gap prior to isthmus completion; although, in both groups intercontinental exchange for the vast majority was facilitated by the presence of the land bridge. While birds can fly, our result suggests that most birds are subject to the same dispersal constraints as their mammalian counterparts.

The similarities between the avian and mammalian responses to the Isthmus completion are also seen in the patterns of interchange. The shared decline in trans-isthmus dispersal over time likely reflects a real pattern that may be driven by the unavailability of ecological space. I suggest that this pattern is the result of a density dependent effect (the priority effect, MacArthur 1972), such that exchange across the isthmus would have decreased over time due to competition with congeners (or diversifying conspecifics) that had already crossed the isthmus. Other previously identified factors that may lead to a genetic signature of density-dependent diversification, such as extinction and inadequate taxon sampling (Rabosky and Lovette 2008), are not easily explained by the data. An increase in extinction rates that would cause a decline in diversification is not likely. There is no a priori prediction that would allow extinction to differentially affect recent colonists (last 2 million years) at a much higher rate than that experienced by earlier colonists. Inadequate taxon sampling is also not likely, as 35% of the points in the data are intra-specific genetic breaks identified by widespread geographic sampling.

Our data suggest that by the onset of the Pleistocene, the majority of species in geographic proximity to the land bridge had already crossed, even though source pools continued to receive new immigrant species from the interchange. Once these source taxa immigrated, their descendants only rarely dispersed back into their continent of origin. This is evident by the infrequency of backcrosses between continents, with 85% of the points represented by single dispersal events. Nearly half of the lineages that did backcross represented older events at higher taxonomic levels. It is unlikely that our results are confounded by sampling effects such as over-emphasizing highland taxa (at

the expense of more continuously distributed lowland species) or by selective study of only those taxa presumed a priori to possess complex population structure. Our data include intraspecific divergences for 34 lowland taxa that are presumed to be continuously distributed across the Isthmus of Panama. In the course of summarizing the relevant available data I encountered only one instance of haplotype sharing across the isthmus, suggesting that most lowland taxa do differ across this biogeographic break and that diversification occurs soon after trans-isthmus dispersal.

The similarity between the highland and lowland distributions is somewhat surprising given that the distributions of many highland birds are fragmented by intervening lowland valleys. Our result may indicate that highland birds are able to readily disperse through lowland corridors during interglacial habitat shifts. Regardless, additional data may be required to further clarify the role of elevation on trans-isthmus dispersal. I identified only 21 highland diversification events compared to seventy-two for lowland birds. The disparity in sample size between highland and lowland birds suggests that the land bridge corridor had a lesser impact in facilitating dispersal for highland species because the corridor only directly connected lowland habitats.

The mammalian GABI is known for the exchange of savanna species despite the present-day lack of a dry habitat corridor between the Americas. Our dataset contains several non-rainforest species (xeric, scrub, and edge taxa) that would have difficulty dispersing through the extensive rainforest on the isthmus today. The existence of historically more continuous dry habitats within the Neotropical region (Pennington et al. 2000) and a drier isthmus habitat corridor (Webb 1991) coincide with a known Pliocene warming period (Zachos et al. 2001). Thus, it seems reasonable that large mammalian

herbivores and birds adapted to more xeric habitats may have dispersed at this time with true forest birds crossing sometime later.

Although our analysis lacked statistical significance, differing temporal trends for these two ecological groups are apparent (Fig. 18c). Scrub/edge birds have an older mean diversification time and the frequency of dispersal declines in the Pleistocene, which is consistent with the early wave of large mammalian herbivores. In contrast, forest birds show a pattern similar to that seen for a second wave of mesic phase mammals (Webb and Rancy 1996), with both groups reaching their highest frequency in the Pleistocene. As rainforest continued to expand on the land bridge, many potential immigrants (xeric, scrub, and edge species) from the North would have been cut off, likely contributing to the slowdown in birds dispersing southward. Assuming our interpretations are correct, the decline in the rate of diversification events was driven in part by density dependent effects and also by habitat changes that may have prevented edge/scrub taxa from continuing to participate in the GABI. Despite the overall slowdown in avian exchange our results indicate that for tropical South American birds, the GABI does not appear to be over.

Intercontinental avian exchange and the latitudinal diversity gradient

Overall, the exchange of lineages in response to the Isthmus of Panama closure played a major role in assembling the present-day Neotropical avifauna. The molecular record shows that within the tropics, the number of birds dispersing between the continents in each direction was similar; however, from a broader geographic perspective this symmetry disappears. For both birds and mammals, the GABI had relatively little

effect on diversity in temperate North America. I suggest that this phenomenon is best explained by an asymmetry in niche conservatism and this asymmetry has contributed to the long-recognized latitudinal diversity gradient (Mittelbach et al. 2007).

Avian families with northern origins occur widely in both the Nearctic and Neotropical regions, thus northern lineages have successfully diversified throughout South America (Table 13). Conversely, the southern families that participated in the GABI remain almost entirely restricted to the Neotropical region, as they have been for tens of millions of years. As a specific example, the oscine and suboscine passerines are a large and diverse group representing over one half of all extant avian species (Sibley and Monroe 1990). Based on molecular dating, the oscines arrived in the New World between 34 – 14 mya, via multiple independent dispersal events, likely through Beringia (Barker et al. 2004). The suboscines have occupied South America some 65 million years, prior to its drifting from Antarctica and Australia (Barker et al. 2004). Despite having northern ancestral origins, 55% of New World oscine species now breed in South America, many of them in tropical habitats. In contrast, only 2.4% of suboscines have secondarily adapted to North American temperate zone habitats (Table 13).

The difficulty tropical organisms have in colonizing seasonal temperate zone environments is widely recognized. Niche conservatism at this scale, is an important component of the Tropical Niche Conservatism hypothesis, a model invoked to help explain the latitudinal diversity gradient (Wiens and Donoghue 2004). The observed asymmetry is also consistent with another component of the Tropical Niche Conservatism model, the "time for speciation" effect (Stephens and Wiens 2003). Most ancestrally southern bird lineages now occurring at the northern limits of the Neotropical

region, have arrived since the time of isthmus uplift and may not have had sufficient time to adapt to temperate conditions. Conversely, the ability of the northern birds to repeatedly invade the tropics may be linked to their long exposure to tropical conditions in Central America well before the uplift of the isthmus.

One part of the Tropical Niche Conservatism hypothesis that does not entirely explain the distribution of New World avian biodiversity is the role of shifting tropical habitats. During the early Tertiary, tropical forests were once much more widespread across the globe and have since contracted to equatorial latitudes (Behrensmeyer et al. 1992). This contraction along with the inability of tropical organisms to adapt to extratropical habitats has been argued be to a major driver of the latitudinal diversity gradient in New World birds (Hawkins et al. 2006). Support for this pattern in the Old World is seen in fossil evidence that indicates relatives of the present-day Ethiopian (Afrotropical) avifauna were once distributed in northern Europe and the United States (reviewed in Feduccia 1999, Ksepka and Clarke 2009). But, in the New World proper, neither fossil evidence (Vuilleumier 1985) nor molecular data support a unified contraction of the Neotropical avifauna, because the North and South American avifauna remained largely isolated from one another until the Panamanian land bridge formation (Fig. 16). The early Cenozoic North American tropical avifauna is largely extinct and has contributed very little to the extant Neotropical avifauna. I was able to identify, only two Neotropical families (Trogonidae and Motmotidae) which are presumed to have northern tropical origins (Feduccia 1999, Witt 2004, DaCosta and Klicka 2008). The contrasting biotic histories of the Old and New World are reflected in their geology. The Old World

continents have a long history of connectivity, which is in sharp contrast to the prolonged isolation of South America from North America.

Table 12. Results of comparisons between the distribution of trans-isthmus diversifications and alternative diversification scenarios. To compare the simulated and empirical distributions of diversification events with the trans-isthmus distribution, I used the nonparametric two-sample Kolmolgorov – Smirnov test (K-S test). The test statistic shown is the D statistic, a measure of difference between two distributions.

Distribution modeled	0 – 8 (mya)	4 – 8 (mya)	3.1 – 8 (mya)	0 – 4 (mya)	0 – 3.1 (mya)
Moderate speciation	0.6166 †	0.2575	0.1772	0.5840 †	0.5445 †
High speciation	0.3623 †	0.2256	0.1200	0.3626 †	0.3078 †
Low extinction pulse	0.5824 †	0.2363	0.1596	0.5533 †	0.5078 †
High extinction pulse	0.3333 †	0.2280	0.1348	0.3132 †	0.2560 *
"Background"	0.1495	0.2336	0.2272	0.2129 *	0.2449 *

* $P < 0.05$.

† $P < 0.001$.

Table 13. Breeding distributions and ancestral origins of New World avian families.

	No. of species	
Ancestral North	**Neotropical**	**Nearctic**
Trogonidae *	25	0
Momotidae	9	0
Mimidae	34	8
Vireonidae	38	13
Corvidae *	18	17
Polioptilidae	11	4
Troglodytidae	66	9
Thraupidae †	402	0
Parulidae	65	49
Icteridae	76	20
Cardinalidae	33	15
Emberizidae *	81	35

* Old World members of family are excluded from this total.

Table 13. Breeding distributions and ancestral origins of New World avian families continued.

	No. of species	
Ancestral South	**Neotropical**	**Nearctic**
Psittacidae *	150	0
Trochillidae	331	18
Ramphastidae	35	0
Semnornithidae	2	0
Capitonidae	11	0
Contingidae	60	0
Bucconidae	32	0
Galbulidae	17	0
Furnariidae	235	0
Thamnophilidae	209	0
Tityridae	31	0
Tyrannidae	400	27
Pipridae	51	0
Melanopareidae	4	0
Rhinocryptidae	44	0
Conopophagidae	10	0
Grallaridae	49	0

* Old World members of family are excluded from this total.

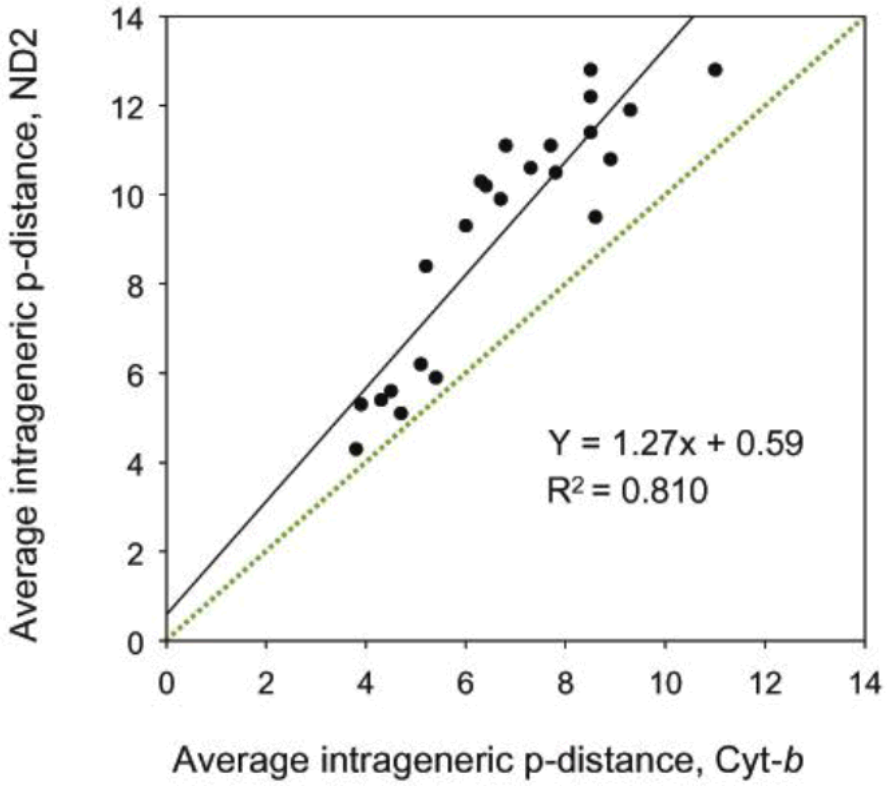


Figure 15. Plot of relative rates of cyt-*b* and ND2 genes for 23 genera of New World birds. The dotted green line indicates equal rates.

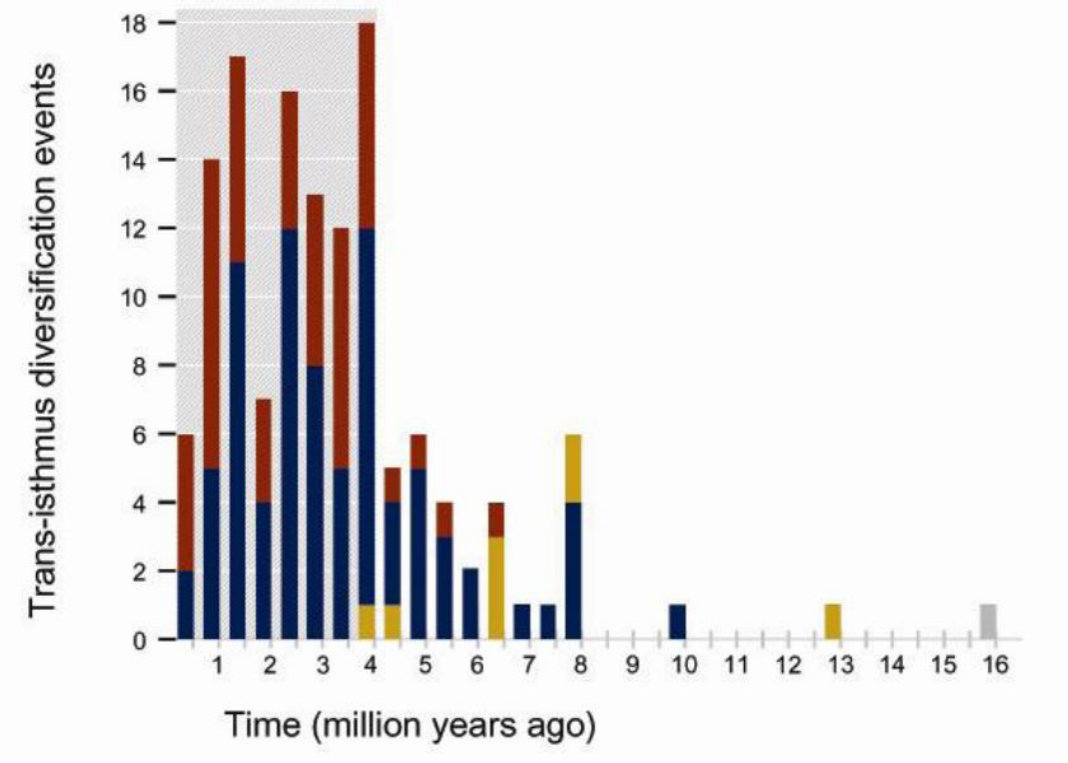


Figure 16. Histogram of trans-isthmus diversification events**.** Diversification times were estimated by employing a gene-specific relaxed molecular clock. The shaded area corresponds to the post-land bridge time period. Taxonomic rank is indicated by color: Intra-specific (red); inter-specific (blue); inter-generic (gold); and inter-familial (grey).

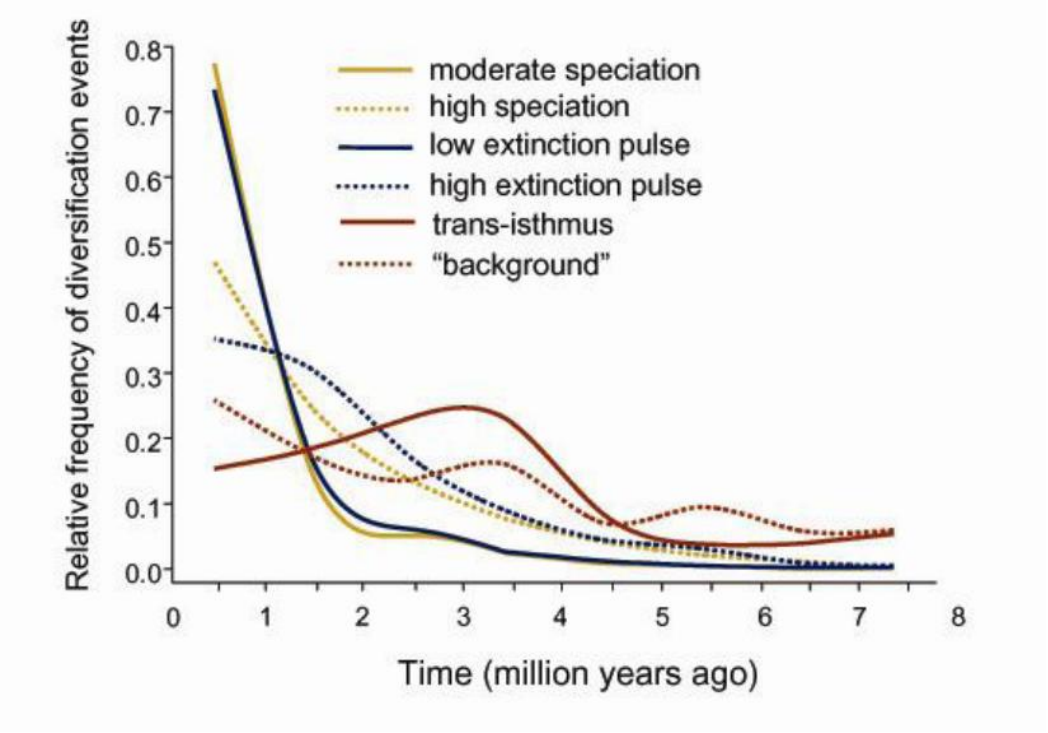


Figure 17. Relative frequency of simulated, empirical and trans-isthmus diversification events**.** Moderate speciation (birth rate = 0.2, extinction rate = 0.2); high speciation (birth rate = 0.8, extinction rate = 0.2); low extinction pulse (birth rate = 0.2) with episodic extinction events that removed 20% of lineages; high extinction pulse (birth rate = 0.2) with episodic extinction events that removed 80% of the lineages. "Background" diversification times were estimated for all cladogenetic events occurring within 28 widely distributed (i.e. with elements occurring on both continents) New World avian genera.

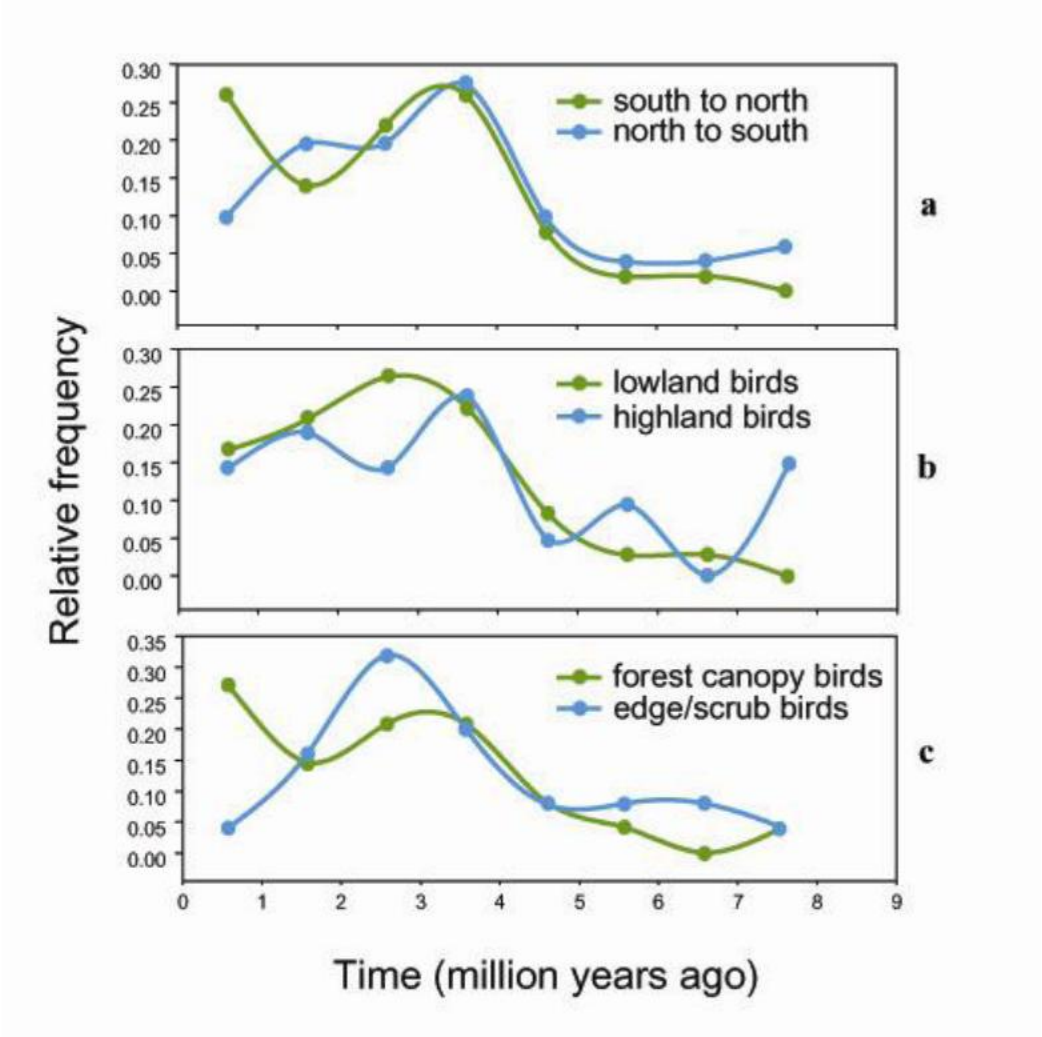


Figure 18. The frequency of diversification events relative to historical and ecological variables. (**a**) dispersal direction (K-S statistic: D = 0.1946, p = 0.259); (**b**) elevation (D = 0.2480, p = 0.233); and (**c**) habitat group (D = 0.2482, p = 0.224).

CHAPTER 5

CONCLUSION

Researchers have long been interested in the relationship between the evolution of biodiversity and changes to the Earth's geology and climate. The interaction between Earth history and biotic evolution operates on multiple scales and can be studied as a hierarchical system ranging from populations to biotic assemblages. The resolution at which researchers can interpret evolutionary and biogeographic patterns is time dependent. Population level study can provide insight into how population sizes and allele frequencies have changed over more recent time periods. Alternatively, the investigation of biotic assembly focuses on understanding the pattern of how entire communities of species have evolved across deep time scales. Although these processes operate at different time scales and are best interpreted independently they occur concomitantly. By examining processes that operate on different spatial and temporal scales I was able to infer a holistic view of the biogeography and evolution of avian biodiversity in the New World.

Evolutionary patterns within cardinal populations

In Chapter 2, I used a widely distributed songbird, the Northern Cardinal, *Cardinalis cardinalis*, as a model to test historical predictions for how species evolved in response to Pleistocene glacial-interglacial cycles. I found evidence in the mitochondrial DNA and from my modeled distribution of the species that there is deep geographic structure in *C. cardinalis*. This genetic pattern is consistent with fragmentation caused by historical climate change and likely began developing well before the Late Pleistocene.

These findings were corroborated by the results of both contemporary and paleo ecological niche models. A modeled paleo-distribution along with historical demographic hypothesis tests indicated that cardinals expanded out of refugia in eastern North America since the Last Glacial Maximum. However, there is no signature of decreased genetic diversity in areas colonized after the expansion. I suggested that this unexpected result of homogenous genetic diversity was due to on-going gene flow across eastern North America. These results demonstrate that both Earth history events and cotemporary processes are important in determining the geography of genetic diversity observed within species.

The tempo of evolution among cardinal populations

Determining the mechanism of why the molecular clock evolves at different rates has been an analytical and biological problem of interest for biologists. In Chapter 3, I employed a novel approach to understanding the tempo of evolution in island and mainland taxa by employing a multilocus coalescent framework. My approach allowed me to address possible analytical limitations caused by phylogenetic uncertainty and differences in gene tree branch lengths caused by random coalescent processes. I showed that multilocus time estimates were not consistent with mtDNA estimates of an island founder event even when the molecular clock was relaxed and coalescent variance was accounted for. These findings have important implications for the implementation of relaxed molecular clocks. If rate shifts are extremely rapid, as the one recovered in my study, then the relaxing of the molecular clock may be insufficient for accurately estimating divergence times. A more robust approach is to incorporate multilocus data

because the apparent accelerated evolution in taxa with small populations may be restricted to the mitochondrial genome.

The Great American Biotic Interchange in birds

Chapter 4 suggests the formation of the Isthmus of Panama was critical for facilitating avian exchange in a manner similar to that known for mammals. This exchange, contributed to a reconfiguration of the taxonomic composition of New World bird communities. The Great American Biotic Interchange was the process that united the isolated North and South American tropical bird faunas into a single biogeographic unit. However, the Great American Biotic Interchange had little impact on bird diversity in the temperate North America. This asymmetric exchange between the temperate and tropical regions provides an additional insight into why species diversity is higher towards the equator. The interchange of birds from ancestrally northern families increased bird diversity in the tropics, whereas, birds from the ancient South America bird fauna contributed little to extant diversity in temperate North America. I suggested that the ability of organisms with northern origins, to colonize and radiate within the tropics may represent an important and underappreciated contributor to the latitudinal diversity gradient.

Conclusion

In conclusion, I was able to show a baseline pattern that the evolution of New World avifaunas was strongly impacted by historical events across the last ten million years. These findings serve as hypotheses that can now be tested and expanded with new

data. As pointed out in Chapter 3, the first step to expanding phylogenetic and biogeographic studies is to include multiple unlinked genetic markers that can reduce errors associated with genetic parameters. The results in Chapter 2 and 4 were based on only a single genetic marker mitochondrial DNA and are subject to a wide degree of variance. Given the relative ease of generating multilocus sequence data, the robustness of my dissertation findings can now be used evaluated with more genetic data. Although, it is unlikely that the overall pattern inferred from mitochondrial DNA will change, it is reasonable to assume that there are more fine scale genetic patterns that were undetected in my studies. The second major issue, which is relevant for all three research chapters, is that my findings need to be expanded to other species in order to determine how these results apply to other bird species or other species of animals and plants. Despite these limitations and the uncertainty around the patterns I inferred, the broader findings of research will be able to influence future studies on the molecular evolution and historical biogeography of New World birds.

BIBLIOGRAPHY

Arbeláez-Cortés E, Nyári ÁS, Navarro-Sigüenza AG. 2010. The differential effect of lowlands on the phylogeographic pattern of a Mesoamerican montane species (*Lepidocolaptes affinis,* Aves: Furnariidae). Mol. Phyl. Evol. 57: 658-668.

Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Annu. Rev. Ecol. Syst. 33: 707-740.

Arbogast BS, Slowinski JB. 1998. Pleistocene speciation and the mitochondrial DNA clock. Science 282:1955a.

Avise JC. 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge.

Avise JC, Walker D, Johns GC. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. Proc. Biol. Sci. 265:1707-1712.

Axelsson E, Smith NGC, Sundstrom H, Berlin S, Ellegren H. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. Mol. Biol. Evol. 21:1538–1547.

Barker FK. et al. 2004. Phylogeny and diversification of the largest avian radiation. Proc. Nat. Acad. Sci. 101: 11040-11045.

Bazin E, Glémin S, Galtier N. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312: 570-572.

Behrensmeyer AK et al. 1992. Terrestrial ecosystems through time. Volutionary paleoecology of terrestrial plants and animals. University of Chicago Press.

Berlin S, Tomaras D, Charlesworth B. 2007. Low mitochondrial variability in birds may indicate Hill-Robertson effects on the W chromosome. Heredity 99:389-396.

Brandley MC, Wang Y, Guo X, Montes de Oca AN, Fería-Ortíz M, Hikida T, Ota H. 2011. Accommodating heterogenous rates of evolution in molecular divergence dating methods: an example using intercontinental dispersal of *Plestiodon* (*Eumeces*) lizards. Syst. Biol. 60: 3-15.

Bromham L, Leys R. 2005. Sociality and the rate of molecular evolution. Mol. Biol. Evol. 22: 1392-1402.

Brown JM, Hedtke SM, Lemmon AR, Lemmon EM. 2010. When trees grow too long: investigating the causes of highly inaccurate Bayesian branch length estimates. Syst. Biol. 59: 145-161.

Brumfield RT, Jernigan RW, McDonald DB, Braun MJ. 2001. Evolutionary implications of divergent clines in an avian (Manacus: Aves) hybrid zone. Evolution 55:2070-2087.

Burney CW, Brumfield RT. 2009. Ecology predicts levels of genetic differentiation in Neotropical birds. Am. Nat. 174:358-368.

Burridge CP, Craw D, Jack DC, King TM, Waters JM. 2008. Does fish ecology predict dispersal across a river drainage divide? Evolution 62:1484-1499.

Carling, MD RT Brumfield RT. 2008. Integrating phylogenetic and population genetic analyses of multiple loci to test species divergence hypotheses in *Passerina* buntings. Genetics 178: 363-377.

Carling, MD RT Brumfield RT. 2008b. Haldane's rule in an avian system: using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the *Passerina* bunting hybrid zone. Evolution 62:2600-2615.

Cheviron ZA, Brumfield RT. 2009. Migration-selection balance and local adaptation of mitochondrial haplotypes in Rufous-collared Sparrows (*Zonotrichia capensis*) along an elevational gradient. Evolution 63:1593-1605.

Clark PU, Mix AC. 2002. Ice sheets and sea level of the Last Glacial Maximum. Quat. Sci. Rev. 21: 1-7.

Coates A, Obando J. 1996. The geologic evolution of the Central American Isthmus. In: Jackson, J. et al. (eds), Evolution and Environment in Tropical America. The University of Chicago Press, pp. 21-56.

Creevey C, McInerney, JO. 2003. CRANN: Detecting adaptive evolution in protein coding DNA sequences Bioinformatics 19: 1726.

DaCosta JM, Klicka J. 2008. The Great American Interchange in birds: a phylogenetic perspective with the genus *Trogon*. Mol. Ecol. 17: 1328-1343.

DeSalle R, Templeton AR. 1988. Founder effects and the rate of mitochondrial DNA evolution in Hawaiian *Drosophila*. Evolution 42:1076-1084.

Dmitriev DA, Rakitov RA. 2008. Decoding of superimposed traces produced by direct sequencing of heterozygous indels. PLoS Comput. Biol. 4(7): e1000113. doi:10.1371/journal.pcbi.1000113

Dow DD, Scott DM. 1971. Dispersal and range expansion by the cardinal: an analysis of banding records. Can. J. Zool. 49: 185-198.

Drummond AJ et al. 2006. Geneious v.2.5.4.

Drummond A J, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis sampling trees. BMC Evol. Biol. 7: 214.

Duffy S. et al. 2008 Rates of evolutionary change in viruses: patterns and determinants. Nat. Rev. Genet. 9: 267-276.

Edwards SV, Beerli P. 2000. Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution 54: 1839-1854.

Easteal S, Collet C. 1994. Consistent variation in amino-acid substitution rate, despite uniformity of mutation rate: Protein evolution in mammals is not neutral. Mol. Biol. Evol. 11: 643-647.

Excoffier LGL, Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online. 1:47-50.

Feduccia A. 1999. The origin and evolution of birds. Yale University.

Filardi CE, Moyle RG. 2005. Single origin of a pan-Pacific bird group and upstream colonization of Australasia. Nature 438:216-219.

Frankham R. 1996. Relationships of genetic variation to population size in wildlife. Cons. Biol. 10: 1500-1508.

Fry AJ, Zink RM. 1998. Geographic analysis of nucleotide diversity and song sparrow population history. Mol. Ecol. 7:1303-1313.

Funk D, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34:397-423.

Hackett SJ. 1996. Molecular phylogenetics and biogeography of tanagers in the genus Ramphocelus (Aves). Mol. Phyl. Evol. 5: 368–382.

Haldane JBS. 1922. Sex ratio and unisex sterility in hybrid animals. J. Genetics 12: 101-109.

Halkin SL, Linville SU. 1999. Northern Cardinal (Cardinalis cardinalis) The Birds of North America No 44 Edited by: Poole A, Gill F. The Birds of North America Inc, Philadelphia.

Hawkins BA. et al. 2006. Post-Eocene climate change, niche conservatism, and the latitudinal diversity gradient of New World birds. J. Biogeogr. 33: 770-780.

Hayes FE, Sewlal JN. 2004. The Amazon River as a dispersal barrier to passerine birds: effects of river width, habitat and taxonomy. J. Biogeogr. 31: 1809-1818.

Heads M. 2011. Old taxa on young islands: A critique of the use of island age to date island endemic clades and calibrate phylogenies. Syst. Biol. 60: 204-218.

Heslewood MM, Elphinstone MS, Tidemann SC, Baverstock PR. 1998 Myoglobin intron variation in the Gouldian Finch *Erythrura gouldiae* assessed by temperature gradient gel electrophoreses. Electrophoresis 19: 142-151.

Hewitt GM. 1996. Some genetic consequences of ice ages, and their role, in divergence and speciation. Biol. J. Linn. Soc. 58: 247-276.

Hewitt GM. 2000. The genetic legacy of the Quaternary ice ages. Nature 405: 907-913.

Hickerson M J et al. 2006. Comparative phylogeographic summary statistics for testing simultaneous vicariance across taxon-pairs. Mol. Ecol. 15: 209 - 224.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Clim. 25: 1965-1978.

Ho SYW, Phillips MJ, Cooper A, Drummond AJ. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Mol. Biol. Evol. 22: 1561-1568.

Huang H, He Q, Kubatko LS, Knowles LL. 2010. Sources of error inherent in species tree estimation: impact of mutational and coalescent effects on accuracy and implications for choosing among different methods. Syst. Biol. 59: 573-583.

Houston D, Shiozawa DK, Riddle BR. 2010. Phylogenetic relationships of the western North American cyprinid genus Richardsonius, with an overview of phylogeographic structure. Mol. Phyl. Evol. 55: 259-273.

Johnson KP, Seger J. 2001. Elevated rates of nonsynonymous substitution in island birds. Mol. Biol. Evol. 18: 874-881.

Jønson KA, Fabre PH, Ricklefs R, Fjeldså. 2011. Major global radiation of corvoid birds originated in the proto-Papuan archipelago. Proc. Nat. Acad. Sci. 108:2328-2333.

Kirby MX et al. 2008. Lower Miocene Stratigraphy along the Panama Canal and its Bearing on the Central American Peninsula. PloS ONE, 3: e2791.

Klicka J, Burns K, Spellman GM. 2007. Defining a monophyletic Cardinalini: A molecular perspective. Mol. Phyl. Evol. 45: 1014-1032.

Klicka J, Zink RM. 1997. The importance of recent ice ages in speciation: a failed paradigm. Science 277: 1666-1669.

Kimball RT, et al. 2009. A well-tested set of primers to amplify regions spread across the avian genome. Mol. Phyl. Evol. 50: 654-660.

Kimura M. 1962. On the probability of mutant genes in a population. Genetics. 47: 641-783.

Ksepka D T, Clarke J A. 2009. Affinities of *Palaeospiza bella* and the phylogeny and biogeography of Mousebirds (Coliiformes). The Auk 126: 245-259.

Lartillot N, Philippe H. 2006. Computing Bayes factors using thermodynamic integration. Syst. Biol. 55: 195-207.

Laval G, Excoffier L. 2004. SIMCOAL 2.0: a program to simulate genomic diversity over large recombining regions in a subdivided population with a complex history. Bioinformatics Applications Note 20: 2485-2487.

Leaché AD, Fujita MK. 2010. Bayesian species delimitation in west African forest geckos (*Hemidactylus fasciatus*). Proc. Roy. Soc. B 277: 3071-3077.

Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25: 1451-1452.

Liu L, Pearl DK. 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions Syst. Biol. 56: 504-514.

Lomolino MV. et al. 2006. Biogeography. Sinauer Associates Inc.

Lovette IJ. 2004. Mitochondrial dating and mixed-support for the "2% rule" in birds. Auk 121: 1-6.

Mittelbach GG et al. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecology Letters 10: 315-331.

MacArthur RH. 1972. Geographical ecology: Patterns in the distribution of species. Princeton University Press.

Maddison WP. 1997. Gene trees in species trees. Syst. Biol. 46: 523-536.

Maddison WP, Maddison D. 2009. MESQUITE: A Modular System for Evolutionary Analysis v 2.6.

Marshall L G. et al. 1979. Calibration of the Great American Interchange. Science 204: 272-279.

Marshall L G et al.1982. Mammalian evolution and the Great American Interchange. Science 215: 1351-1357.

Martin AP, Palumbi SR. 1993. Body size, metabolic rate, generation time, and the molecular clock. Proc. Natl. Acad. Sci. 90: 4087-4091.

Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P. 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26: 2462-2463.

Mayr E. 1964. Inferences concerning the Tertiary American bird faunas. Proc. Natl. Acad. Sci. 51: 280-288.

Mayr E, Diamond J. 2001. The birds of northern Melanesia: Speciation, ecology, and biogeography. Oxford University Press.

McCloy C, Ingle JC, Barron JA. 1988. Neogene stratigraphy, foraminifera, diatoms, and depostional history of Maria Madre Island, Mexico: Evidence of Early Neogene marine conditions in the southern Gulf of California. Marine Micropaleontology. 13:193-212.

McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL. 2011. Calibrating divergence times on species trees versus gene trees: implications for speciation history of *Aphelocoma* jays. Evolution 65: 184–202.

Milá B, Girman DJ, Kimura M, Smith TB. 2000. Genetic evidence for the effect of a post glacial population expansion on the phylogeography of a North American songbird. Proc . Roy. Soc. B 267:1033-1040.

Milá B, Smith TB, Wayne RK. 2007. Speciation and rapid phenotypic differentiation in the yellow-rumped warbler (*Dendroica coronata*) complex. Mol. Ecol. 16: 159-173.

Monson G, Phillips AR. 1981. Annotated checklist of birds of Arizona. 2nd edn. Univ. of Arizona Press, Tucson.

Moore RP et al. 2008. Experimental evidence for extreme dispersal limitation in tropical forest birds. Ecol. Lett. 11: 960-968.

Moore WS, Dolbeer RA. 1989. The use of banding recovery data to estimate dispersal rates and gene flow in avian species: case studies in red-winged blackbird and common grackle. The Condor 91: 242-253.

Nabholz B, Glémin S, Galtier N. 2008.Strong variations of mitochondrial mutation rate across mammals - the longevity hypothesis. Mol. Biol. Evol. 25:120-130.

Nabholz B et al. 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. BMC Evol. Biol. 9: 54.

Neiman M, Hehman G, Miller JT, Logsdon Jr JM, Taylor DR. 2010. Accelerated mutation accumulation in asexual lineages of a freshwater snail. Mol. Biol. Evol. 27: 954-963.

Nylander J. 2004. MrModeltest v 2.3. Evolutionary Biology Centre, Uppsala University.

Ohta T. 1972. Population size and rate of evolution. J. Mol. Evol. 1: 305-314.

Omland KE, Baker J.M, Peters JL. 2006. Genetic signatures of intermediate divergence: population history of Old and New World Holarctic ravens (Corvus corax). Mol. Ecol. 15:795-808.

Olson, SL 1985. The fossil record of birds. In: Farner, D. S. et al. (eds) Avian Biology. Academic Press, pp. 179-238.

Paradis E, Baillie SR, Sutherland WJ, Gregory RD. 2002. Patterns of natal and breeding dispersal in birds. J. Animal Ecol. 67: 518-536.

Parkes KC. 1997. The Northern Cardinals of the Caribbean slope of Mexico, with the description of an additional subspecies from Yucatan. The era of Alan R. Phillips: A festschrift. Compiler: Dickerman RW Horizon Communications, Albuquegue.

Patterson B, Pascual R. 1972. The fossil mammal fauna of South America. In: Keast, A. et al. (eds), Evolution, Mammals, and Southern Continents. State University of New York Press.

Pereira SL, Baker AJ. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. Mol. Biol. Evol. 23: 1731-1740.

Pennington RT et al. 2000. Neotropical seasonally dry forests and Quaternary vegetation changes. J. Biogeogr. 27: 261-273.

Peterson GI, Masel J. 2009. Quantitative prediction of molecular clock and Ka/Ks at short time scales. Mol. Biol. Evol. 26:2595-2603.

Phillimore AB, Price T D. 2008. Density-dependent cladogenesis in birds. PloS Biol. 6: e71.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. Ecol. Mod. 190: 231-259.

Primmer CR, Borge T, Saetre GP. 2002. Single-nucleotide polymorphism characterization in species with limited availablesequence information: high nucleotide diversity revealed in the avian genome. Mol. Ecol. 11: 603-612.

Rabosky DL. 2006. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates. Evol. Bioinf. 2: 257-260.

Rabosky DL, Lovette I J. 2008. Density-dependent diversification in North American wood warblers. Proc. Roy. Soc. B 275: 2363-2371.

Rambaut A. 2002. PHYL-O-GEN: phylogenetic tree simulator package v 1.1.

Rambaut A, Drummond AJ. 2007. Tracer v 1.4.

Rambaut A, Drummond A J. 2010. Tracer v1.5.

Ree R H, Smith SA. 2008. Maximum Likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57: 4 - 14.

Ree RH. et al. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution 59: 2299-2311.

Ricklefs RE. 2002. Splendid isolation: historical ecology of the South American passerine fauna. J. Avian Biol. 33, 207-211.

Riddle BR, Hafner DJ, Alexander LF. 2000. Phylogeography and systematics of the *Peromyscus eremicus* species group and the historical biogeography of North American warm regional deserts. Mol. Phyl. Evol. 17:145-160.

Riddle BR, Hafner DJ, Alexander LF, Jaeger JR. 2000. Cryptic vicariance in the historical assembly of a Baja California peninsular desert biota. Proc. Natl. Acad. Sci. 97:14438-14443.

Ridgely RS, Allnutt TF, Brooks T, McNicol DK, Mehlman DW, Young BE, Zook, JR. 2007. Digital Distribution Maps of the Birds of the Western Hemisphere, version 3.0. NatureServe, Arlington.

Ridgway R. 1901. The birds of Middle and North America: a descriptive catalogue of the higher groups, genera, species, subspecies of birds known to occur in North America. Vol. 1 U.S. Nat. Mus. Bull. 50.

Rojas-Soto OR, Westberg M, Navarro-Sigüenza AG, Zink RM. 2010. Genetic and ecological differentiation in the endemic avifauna of Tiburón Island. J. Avian Biol. 41:398-406.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19:1572-1574.

Saetre GP, Moum T, Bures S, Kral M, Adamjan M, Moreno J. 1997. A sexually selected character displacement in flycatchers reinforces premating isolation. Nature 387: 589-592.

Sauer JR, Hines JE, Fallon J. 2008. The North American Breeding Bird Survey, results and analysis 1966 - 2007. Version 5.15.2008. USGS Patuxent Wildlife Research Center, Laurel, MD.

Sibley CG, Monroe BLJ. 1990. Distribution and taxonomy of birds of the world. Yale University Press.

Simpson GG. 1947. Holarctic mammalian faunas and continental relationships during the Cenozoic. Bull. Geo. Soc. Am. 58: 613-688.

Simpson GG. 1980. Splendid Isolation: The curious history of South American mammals. Yale University Press.

Slade RW, Moritz C, Heideman A, Hale PT. 1993. Rapid assessment of single-copy nuclear DNA variation in diverse species. Mol. Ecol. 2: 359-373.

Slatkin M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution. 47: 264-279.

Smith SA, Donoghue MJ. 2008. Rates of molecular evolution are linked to life history in flowering plants. Science 322: 86-89.

Smith BT, Klicka J. 2010. The profound influence of the Late Pliocene Panamanian uplift on the exchange, diversification, and distribution of New World birds. Ecography.33: 333-342.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS. 2006. Comparative phylogeography of unglaciated eastern North America. Mol. Ecol. 15: 4261-4293.

Spaw RH. 1978. Late Pleistocene carbonate bank deposition; Cozumel Island, Quintana Roo, Mexico. Gulf Coast Assoc. Geol. Soc .Trans. 28: 601-619.

Spellman GM, Klicka J. 2006. Testing hypotheses of Pleistocene population history using coalescent simulations: phylogeography of the pygmy nuthatch (Sitta pygmaea). Proc. Roy. Soc. B 273: 3057-3063.

Stager KE. 1957. The avifauna of the Tres Marias Islands, Mexico. The Auk. 74: 413-432.

Stehli F, Webb S. 1985. The Great American Biotic Interchange. Plenum Press.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Amer. J. Hum. Gen. 68: 978-989.

Stephens PR, Wiens JJ. 2003. Explaining species richness from continents to communities: the time-for-speciation effect in Emydid turtles. Am. Nat. 161: 112-128.

Stotz, D. F. et al. 1996. Neotropical birds: ecology and conservation. The University of Chicago Press.

Suchard MA, Weiss RE, Sinsheimer JS. 2001. Bayesian selection of continuous-time Markov chain evolutionary models. Mol. Biol. Evol. 18:1001-1013.

Swenson NG, Howard DJ. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. Am. Nat. 166: 581-91.

Swofford DL. 2002. PAUP* v. 4.0b10: phylogenetic analysis using parsimony (and other methods). Sinauer Associates.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24: 1596-1599.

Tiffney BH, Manchester S R. 2001. The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the northern hemisphere Tertiary. Intl. J. Plant Sci. 162: 3-17.

Voelker G et al. 2009. Repeated trans-Atlantic dispersal catalysed a global songbird radiation. Glob. Ecol. Biogo. 18: 41-49.

Vrba ES. 1992. Mammals as a key to evolutionary theory. J. Mammal.73: 1-28.

Vuilleumier F. 1984. Faunal turnover and development of fossil avifaunas in South America. Evolution 38: 1384-1396.

Vuilleumier F. 1985. Fossil and recent avifaunas and the Interamerican Interchange. In: Stehli F Webb S, (eds), The Great American Biotic Interchange. Plenum Press, pp. 387-424.

Webb SD. 1976. Mammalian faunal dynamics of the Great American Interchange. Paleobiology 2: 220-234.

Webb SD, Marshall LG. 1982. Historical biogeography of recent South American land mammals. In: Mares MA, Genoways HH. (eds), Evolution of Neotropical Mammals. Pymatuning Laboratory of Ecology, pp. 39-54.

Webb SD. 1991. Ecogeography and the Great American Interchange. Paleobiology 17: 266-280.

Webb SD, Rancy A. 1996. Late Cenozoic evolution of the Neotropical mammal fauna. In: Jackson, J. et al. (eds), Evolution and Environment in Tropical America. The University of Chicago Press, pp. 335-358.

Weir JT, 2006. Divergent timing and patterns of species accumulation in lowland and highland Neotropical birds. Evolution 60 842-855.

Weir JT, Schluter D. 2008. Calibrating the avian molecular clock. Mol. Ecol. 17: 232 -2328.

Wiens JJ, Donoghue MJ. 2004. Historical biogeography, ecology and species richness. Trends Ecol. Evol. 19: 639-644.

Witt CC. 2004. PhD. dissertation. Rates of molecular evolution and their application to Neotropical avian biogeography. Louisiana State University.

Woofit, M. 2009. Effective population size and the rate and pattern of nucleotide substitutions. Biol. Lett. 5:417-420.

Woolfit M, Bromham L. 2003. Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population sizes. Mol. Biol. Evol. 20: 1545-1555.

Woolfit M, Bromham L. 2005. Population size and the rate of molecular evolution on islands. Proc. Roy. Soc. B 272:2277-2282.

Wright SD, Gillman LN, Ross HA, Keeling DJ. 2009. Slower tempo of microevolution in island birds: implications for conservation. Evolution 63:2275-2287.

Zachos J et al. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292: 686-693.

Zink RM. 1996. Comparative phylogeography in North American birds. Evolution 50: 308-317.

Zink RM, Kessen AE, Line TV, Blackwell-Rago RC. 2001. Comparative phylogeography of some aridland bird species. The Condor. 103:1-10.

Zink RM. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. Proc. Roy. Soc. B 271: 561-564.

Zink RM, Klicka J, Barber B. 2004. The tempo of avian diversification during the Quaternary. Proc. R. Soc. Lond. B 359: 215-220.

Zink RM, Barrowclough GF. 2008. Mitochondrial DNA under siege in avian phylogeography. Mol. Ecol. 17: 2107-2121.

VITA

Graduate College
University of Nevada Las Vegas

Brian Tilston Smith

Home Address:
8201 Omni Court
Las Vegas, NV 89154

Degrees:
Bachelor of Science, 2005
University of Nevada, Las Vegas

Dissertation Title: Molecular Evolution and Historical Biogeography of New World Birds

Thesis Examination Committee
Co-Chair, Dr. Javier Rodriguez
Co-Chair, Dr. John Klicka
Committee Member, Dr. Brett Riddle
Committee Member, Dr. Jef Jaeger
Graduate Faculty Representative, Dr. Amei Amei