From species to populations: Genetic and eco-physiological divergence across Puerto Rican *Anolis* lizards

Matthew T. McElroy

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Adam D. Leaché, Chair

Raymond B. Huey

Harvey D. Bradshaw

Program Authorized to Offer Degree:

Biology

University of Washington

Abstract

From species to populations: Genetic and eco-physiological divergence across Puerto Rican *Anolis* lizards

Matthew T. McElroy

Chair of the Supervisory Committee:

Adam D. Leaché, PhD

Department of Biology

Species pairs of Puerto Rican anoles inhabit distinct thermal environments and are locally adapted to those environments with regard to physiological traits. In chapter 1, I analyzed multilocus nuclear genetic data and genome-wide SNP data to test for simultaneous divergence and gene flow between physiologically divergent species pairs. I utilized new model-based methods that rigorously test for shared evolutionary history by jointly estimating divergence times for multiple species while implementing models of divergence with and without gene flow, and compared results from Sanger and SNP datasets to determine which dataset is best suited for comparative phylogeographic analyses. These analyses indicate that 'xeric' adapted lizards may have evolved synchronously, but in general, recent speciation of Puerto Rican anoles proceeded idiosyncratically.

Recent models show that tropical ectotherms may respond to increasing temperatures by undergoing range shifts, acclimatizing, or becoming locally extinct. These models often overlook standing genetic variation and assume that tropical ectotherms will not be able to adapt rapidly enough to track global climate change. In chapter 2, I used multilocus nuclear genetic data and genome-wide SNP data to investigate phylogeographic patterns in *A. cristatellus* across Puerto Rico, and to test whether genetic variation correlates with environmental climate layers. These analyses identified multiple genetic populations of *A. cristatellus* that correlate with thermal environment indicating that selection on physiological traits may drive population divergence. Surprisingly, I found asymmetrical gene flow from xeric to mesic populations. Genes flowing 'out' of xeric habitats fits a pattern in which selection against 'warm-adapted' genotypes in cool habitats is weak, whereas selection against 'cool-adaptive' genotypes in warm habitats is relatively strong.

Phylogeographic approaches identify geographic and ecological features that generate and maintain genetic diversity in co-distributed taxa. I analyzed genome-wide SNP data for three species of widespread Puerto Rican anoles to test the relative contributions of geography and ecology in driving population divergence. While only one species (*A. cristatellus*) had phylogeographic structure associated with ecological transitions, all three (*A. cristatellus, A. pulchellus,* and *A. stratulus*) shared concordant population breaks and simultaneous divergence across the San Loiza Basin in eastern Puerto Rico. These results are consistent with phylogeographic work on other Puerto Rican vertebrates including frogs, indicating the San Loiza Basin as an important geographic feature driving population divergence across diverse ecological groups on Puerto Rico.

Divergent selection and local adaptation are recognized as primary drivers of reproductive isolation. However, thermoregulatory behavior may buffer divergent selection acting on thermal physiology, resulting in high gene flow across thermal gradients. I tested this prediction by analyzing genome-wide SNP data to investigate genetic clines along elevation gradient for populations of *A. cristatellus* that differ in operative temperatures and body temperatures. I also ran ecological association analysis of SNPs to test for signatures of selection resulting from temperature and precipitation. I found sharp genetic clines, despite thermoregulatory behavior that should buffer natural selection, indicating that precipitation may be the main agent of selection along the elevational gradient in southern Puerto Rico.

Acknowledgments.

I thank Paul Hertz, Luiso Otiero, Efren Ortiz, and Sofia Prado-Irwin for all their hard work during numerous field trips to Puerto Rico to collect body temperature data. I thank Jose L Herrera and JP Zegarra for welcoming me to Puerto Rico during my first field trip in 2011. I thank Sandra Moya Guzman for her involvement with overnight field trips with UPR Ponce undergraduates. I also thank the Department of Natural Resources in Puerto Rico and the US Forest Service at El Yunque for permits and permission. I thank the game wardens at Bosque Estatal de Guanica, Carite, Maricao, Toro Negro. I used field stations at Mata de Platano and Jobos Bay.

I was funded by the National Science Foundation via the Graduate Research Fellowship Program and the Doctoral Dissertation Improvement Grant. Thanks to Karen Bergeron for her patience in helping me apply to the NSF DDIF. I also received funding from the Department of Biology at University of Washington via the Snyder Award and the Field Experimental Ecology Award.

Special thanks for the Leaché and Buckley/Huey lab groups for the invigorating scientific environment. Thank you to Andreas Chavez for conceptual feedback and discussions. Thank you to Rebecca Harris, Jared Grummer, Sima Bouzid, Barb Banbury, Jamie Oaks, and Charles Linkem for help with analysis, methods development, and computer programming. Thanks to Duncan Reid, Tony Gill, and Kevin Epperly for maintaining the genetic lab. Thank you to Sharon Birks, Peter Miller, and Shane Armstrong for accessions at the Burke Museum of Natural History and Culture. Thank you to the Museum of Vertebrate Zoology at UC Berkeley for tissue loans.

Chapter 1: Testing for synchronous speciation of physiologically divergent Puerto Rican *Anolis* lizards.

Matthew T. McElroy[12], Jamie R. Oaks[3], Adam D. Leaché[12]

[1]Department of Biology, University of Washington, Seattle, WA 98195

[2]Burke Museum of Natural History and Culture, Seattle, WA 98195

[3]Department of Biological Sciences, Auburn University, Auburn, AL 36849

Abstract

Adaptive radiations are responsible for some of the most iconic examples of evolution in nature. Biologists have long proposed that adaptive radiations proceed in predictable stages. For instance, the *Anolis* radiation is thought to proceed first by changes in body size, then in morphology, and finally in thermal biology. To test whether a final stage of evolution in the *Anolis* radiation occurs in a predictable manner, we investigated sister-species pairs of Puerto Rican anoles that inhabit distinct thermal environments and are physiologically adapted to those environments. We analyzed multilocus nuclear sequence data and genome-wide SNP data to test for synchronous divergence and gene flow between these physiologically divergent species pairs. We utilized new model-based methods that rigorously test for shared evolutionary history by jointly estimating divergence times for multiple species while implementing models of divergence with and without gene flow, and compared results from nuclear sequence and SNP datasets to determine which dataset is best suited for comparative phylogeographic analyses. Our analysis of SNP data outperformed the analysis of nuclear loci. Despite evidence for a final stage

of evolution in *Anolis* lizards, we find that speciation events in Puerto Rican anoles are idiosyncratic and complex.

Introduction

Adaptive radiations result from phenotypic differentiation and speciation due to divergent natural selection (Lack 1947; Dobzhanksy 1951; Simpson 1955); they are responsible for some of the most iconic examples of evolution in nature. Defined as the evolution of ecological and phenotypic diversity in rapidly multiplying lineages, adaptive radiations occur when a common ancestor diversifies into species that inhabit a variety of habitats and that have suites of phenotypic traits to exploit those habitats (Simpson 1955; Schluter 2000). Early developers of the theory of adaptive radiation postulated that adaptive radiations occur in predictable fashion. Grant (1949) proposed that floral differentiation precedes vegetative differentiation in animal-dispersed angiosperms, but that the reverse is true in angiosperms that are dispersed by abiotic vectors, such as wind. Many lineages of birds undergo differentiation in resource use pertaining to habitat, then food size, and finally in food type (Diamond 1973; Terborgh and Diamond 1970). Trophic diversification is perhaps best represented by the evolution of beak morphology in Hawaiian honeycreepers and in Galapagos finches and in the evolution of specialized pharyngeal jaws of East African cichlids. In response to accumulating empirical studies, Streelman and Danley (2003) developed a generalized model for the stages of vertebrate adaptive radiations whereby species diverge first in habitat, then morphology, and finally in traits relating to sexual-selection and communication. This ordering of the stages of adaptive radiations indicate the relative importance of drivers of diversification and speciation, and even addresses the longstanding question of whether evolution is predictable (Schluter 2000).

Independent adaptive radiations that are replicated in closely related groups represent robust tests for predictable evolutionary outcomes. For instance, as glacial sheets retreated in North America during the Pleistocene, stickleback fish independently infiltrated post-glacial lakes from marine habitats and diverged in predictable phenotypic traits that are adaptive for freshwater habitats (McKinnon and Rundle 2002). Independent radiations of anoles on Cuba, Hispañola, Puerto Rico, and Jamaica represent another example of replicate adaptive radiation (Williams 1983). On each island, anoles are the dominant vertebrate group and have diversified in morphology and behavior to form ecomorphs that are named after the vegetative micro-habitats that they utilize (e.g. grass-bush, trunk-ground, trunk-crown) (Losos 2009). Anoles have also diversified along a climatic or physiological axis. Ruibal (1961) on Cuba, and Rand (1964) and Heatwole et al. (1969) on Puerto Rico, observed that closely-related anoles are similar in morphology but utilize different thermal habitat and showed that species differences in thermal biology explain differences in habitat use. Williams incorporated the morphological analysis of Schoener (1971) as well as physiological work of Ruibal and Rand to synthesize a hypothesis for the stages of the adaptive radiation of Greater Antillean *Anolis* (Williams 1972, 1983). He proposed that anoles first diversify in body size, then in morphological traits relating to arboreal habitat, and finally in thermal biology. William's hypothesis for the stages of anole evolution has withstood the test of time. Molecular phylogenetics and comparative methods confirm that morphological adaptations occurred early in anole evolution, and that thermal physiology is labile (Hertz et al. 2013); closely related anoles often share morphology, but exhibit differences in thermal biology.

The *cristatellus* group on mainland Puerto Rico is comprised of four pairs of sister-species (Brandley and de Queiroz 2004; Jackman et al. 1999); aside from *Anolis gundlachi* and

A. pulchellus, species within each pair are similar in morphology but divergent in thermal biology (Huey and Webster 1976; Hertz 1992a; Hertz, Huey, and Stevenson 1993; Rivero 1998; Henderson and Powell 2009), indicating that thermal adaptation is an important and repeated component of recent speciation. Anoles in the *cristatellus-group* inhabit one of three generalized distributions across Puerto Rico: central and eastern montane mesic forests, southwestern xeric scrub forest, or widespread inhabiting both mesic and xeric forests. Physiological traits such as preferred body temperature, critical minimum and maximum temperatures, and rates of evaporative water loss indicate that species have thermal biologies that are adapted to the thermal and hydric environments that they inhabit (Huey and Webster 1976; Hillman, Gorman, and Thomas 1979; Hertz 1992b; Losos 2009). Reconstructing the demographic history of divergence in the sister species pairs is a starting point for testing whether there are predictable similarities in how this final stage of physiological divergence in the anole adaptive radiation happened, and if diversification during "stages" ever proceeds in concordant fashion. William's posited that similar ecological species may share similar evolutionary histories - though this prediction remains unfounded (e.g. Jezkova et al., 2009). Testing Williams' prediction in the context of divergence events (i.e. pairs of species with similar ecological differences) can provide insight into the climatic or geological factors - as well as the geographic and demographic context of speciation - that lead thermal adaptation and speciation in anoles.

To understand the evolutionary factors that contribute to the final stage of anole evolution, we investigated the demographic history of *Anolis* species-pairs on Puerto Rico. The Puerto Rican radiation has been well studied, and was the initial testing grounds for William's synthesis, in part because it is a relatively simple radiation with fewer species and has lower morphological diversity than do Cuba and Hispaniola. The evolution of thermal physiology

is best understood within the context of an accurate phylogenetic framework, and the phylogenetic placement of *Anolis poncensis* and *A. gundlachi* within the clade differ among datasets and phylogenetic approaches (Brandley and de Queiroz 2004; Nicholson et al. 2005; Mahler et al. 2010), and recent phylogenetic studies of broad scale diversification of anoles did not have complete genetic sampling for Puerto Rican species (Alföldi et al. 2011; Poe et al. 2017) We used species tree analysis of nuclear sequences in *BEAST and of SNP data in SNAPP to infer phylogenetic relationships for the *cristatellus* group. In order to estimate divergence dates and to test for migration between sister-species pairs we used IMa2 to analyze nuclear sequence data. Finally, to explicitly test for shared demographic histories among pairs of sister-species, we implemented *ecoevolity*, a novel full-likelihood approach that infers the number of divergence events from SNP data. We compared results from *ecoevolity* with analysis of nuclear sequences in the approximate-Bayesian approach, PyMsbayes, to determine the utility of novel methods and datatypes in comparative phylogeographic inference.

Materials and Methods

Sampling

We collected seven of the eight *Anolis* species in the *cristatellus* series from across mainland Puerto Rico (2011-IC-044). We excluded species that are within the *cristatellus* series but not found on mainland Puerto Rico (e.g. *Anolis monensis*, *A. desechensis*) because these species likely result from dispersal and allopatric speciation (Rodríguez-Robles, Jezkova, and García 2007). We froze liver tissue in liquid nitrogen and used 10% buffered formalin to prepare whole body specimens, which were accessioned at the Burke Museum of Natural History and Culture

(UWBM) at the University of Washington in Seattle (Table 1). We obtained six tissue samples for *Anolis cooki* from the Museum of Vertebrate Zoology at University of California, Berkeley.

Sanger Data Collection

We collected eight nuclear gene sequences (*ALMS1*, *EXPH5*, KIAA2018, *KIF24*, *NOS1*, *PPP2R5C*, *RAPGEF2*, *R35*) and one mitochondrial gene sequence (*NADH2*) for 189 individuals from the *cristatellus* series (Table 2). We used Qiagen DNeasy spin columns to extract DNA from frozen liver tissue and followed standard PCR protocols to amplify gene regions. PCR primer citations and characteristics of gene sequences are noted in Table 3. We sequenced loci in both directions on an ABI3730 capillary sequencer, cleaned and assembled DNA sequences in Geneious v6.1.8, and generated sequence alignments in MUSCLE v3.8.1 (Edgar 2004). We resolved phased alleles in the program PHASE v2.1 (Stephens, Smith, and Donnelly 2001) with the recombination (-MR) and parent-independent (-d) models invoked. We used TOPALi v2 (Milne et al. 2009) and the difference of sums of squares (DSS) to test for intragenic recombination. We used JModelTest v2.1.10 (Posada 2008) and AIC to estimate and rank nucleotide substitution models for each locus.

ddRADseq Data Collection

We used the double-digest RADseq protocol (Peterson et al. 2012) to generate genome-wide SNP data for three to seven individuals per species in the *cristatellus* series (Table 1). Tissues were digested overnight at 37°C with 10% SDS and Proteinase K. Following digestion, we used a 2x Serapure bead clean-up protocol to isolate and purify DNA. We used 20 units each of the common cutter MspI (restriction site 5'-CCGG-3') and the rare cutter SbfI (restriction site 5'-

CCTGCAGG-3') to digest 500 ng of DNA at 37°C for 10 hours. We purified the samples using Serapure speedbeads before ligating Illumina barcoded adapters onto the DNA fragments. Following ligation, we selected DNA fragments between 415-515 bp using a Pippin Prep (Sage Science) and then used proofreading Taq and Illumina specific primers to amplify size-selected libraries. We used an Agillent 2200 TapeStation to quantify and assess the distribution of size fragments of the final libraries. We sent the final libraries to the QB3 Vincent Coates Genomics Sequencing Laboratory at UC Berkeley where they were qPCRed to determine the sequence-able library and pooled in equimolar ratios for one lane of single-end 50-bp Illumina sequencing on a HiSeq4000.

ddRADseq Bioinformatics

We processed Illumina data within the software pipeline iPyrad version 0.7.15 (Eaton 2014). iPyrad assembles *de novo* clusters of putative loci and allows for insertions and deletions in the final alignments, making it suitable for producing genetic alignments across species. We demultiplexed the samples from their unique barcode and adapter sequences. After filtering reads with more than five low-quality bases and removing the sequences of the barcode adaptors and the restriction-site overhangs we retained 39 bp sequences. iPyrad implements VSEARCH (Rognes et al. 2016) to make stacks of similar reads within individuals, MUSCLE v3.8.1 (Edgar 2004) for aligning putative loci based on a clustering threshold, and maximum-likelihood joint estimation of mean sequencing error-rate and heterozygosity from base frequencies (Lynch 2008). *De novo* clustering of short-read genetic data produces two types of errors: the splitting of alleles into different clusters artificially increases homozygosity, and clustering paralogs together creates false heterozygosity (Ilut, Nydam, and Hare 2014; Harvey et al. 2015). We found the

optimal parameter settings to minimize over-splitting and formation of paralogs by testing a series of clustering thresholds from 86-98% (Ilut, Nydam, and Hare 2014; Nieto-Montes de Oca et al. 2017). We identified the best clustering threshold by plotting the percentage of resulting clusters that were homozygous, heterozygous, or paralogous – defined as having more than two alleles per individual per cluster. We consider the optimal clustering threshold to be the highest threshold at which the proportions of homozygous, heterozygous, and paralogous loci equilibrate. After identifying the best clustering threshold, we iteratively tested the following parameters: the maximum number of SNPs per locus (from 1-15), the maximum number of indels per locus (from 1-5), and the maximum proportion of shared heterozygous sites across individuals (from 5-50%). RADseq experiments that include divergent species can suffer from allelic drop-out due to mutations in the restriction sites, resulting in large amounts of missing data (Arnold et al. 2013). To reduce the impact of allelic-drop out and increase the number of loci for our study, we reran the iPyrad pipeline and clustering series analysis for each sister-species pair. We utilized the same pipeline and threshold procedure that we used for the *cristatellus*-series. After identifying the best clustering threshold for each species pair, we tested the maximum number of SNPs per locus (from 1-10), the maximum number of indels per locus (from 1-5), and the maximum proportion of shared heterozygous sites across individuals (from 10-100%). Finally, we generated final datasets for each assembly by retaining all loci that were represented in at least 50% of individuals.

Species Tree Analysis

We implemented the multi-species coalescent model in *BEAST v1.7.5 (Drummond et al. 2012) and inferred a species tree for the *cristatellus* group from the nuclear sequence data. To reduce

computation time, we randomly selected one allele per individual. We used the uncorrelated relaxed clock model and set the nucleotide partition scheme for each gene to the top model from JModelTest (Posada 2008) (Table 3). We ran *BEAST for 100 million iterations and sampled every 10,000 steps to record 10,000 trees. We removed the first 25% as burn-in and combined the results from two independent runs for a total of 15,000 trees. We used TreeAnnotator to visualize the maximum-clade credibility tree with median root height and plotted the 95% highest posterior density on the nodes of sister-species. We also used SNAPP (Bryant et al. 2012), implemented in BEAST v 2.4.7 (Bouckaert et al. 2014), to estimate a species tree from SNP data. We used the SNP alignment from the *cristatellus*-group pipeline because SNAPP requires loci to be shared across all species. We used the R-package *phrynomics* (https://github.com/bbanbury/phrynomics) to reduce the dataset to bi-allelic SNPs. We used BEAUTI v 2.4.7 to create input files for SNAPP. We set mutation rates $u = v = 1$ and estimated the coalescent rate from the data. We set the prior for population size to a gamma distribution with shape = 1.0 and scale = 4.0. We set the snap.prior to alpha = 100 and beta to 10^6 and ran SNAPP for 1,000,000 iterations, sampling every 1,000 steps for a total of 1,000 trees.

Demographic modeling of species pairs

To estimate divergence times and infer migration between species pairs, we used the phased nuclear gene sequences and modeled population divergence in the program IMa2 (Hey, 2010). IMa2 uses multi-locus genetic data and implements an isolation-migration model to jointly estimate population divergent parameters such as population sizes, ancestral population size, divergence time, and bi-directional migration rates for multiple populations (Hey, 2010). We set the nucleotide substitution model for each nuclear gene to HKY and scaled the results to

demographic parameters by including the mitochondrial *ND2* gene with a substitution rate of 0.65% (Macey et al. 1998). We used likelihood ratio testing to rank the top demographic models for each species pair (Carstens et al. 2013).

Testing for co-evolutionary events

A rigorous test of simultaneous divergence would jointly estimate divergence times for multiple species, consider models of divergence with and without gene flow, and would calculate the probability of models with different number of divergence events. We used two approaches to test for evidence of co-evolutionary responses of sister-species: *dpp-msbayes* (implemented in *PyMsBayes*) and *ecoevolity*. *Dpp-msbayes* is an approximate Bayesian computation method that estimates clusters of divergence times of co-distributed taxa from nuclear sequence data (Huang et al. 2011; Oaks 2014). *Dpp-msbayes* utilizes flexible prior probability distributions for divergence parameters to reduce biases towards inferring highly clustered divergences (Oaks et al. 2013), and estimates posterior probabilities for the number of divergence events as well as whether divergence occurs with or without gene flow. We used *dpp*-msbayed, implemented in the wrapper PyMsBayes, to estimate the number of divergence events from nuclear sequences for pairs of sister-species in the *cristatellus* group. We set the shape and scale of priors as follows: concentration: 0, 0; theta: 1, 0.002; ancestral theta: 0, 0; and tau: 1.5, 20.0. *Ecoevolity* is a full-likelihood method for estimating population divergence parameters and the number of divergence events from SNP data for pairs of species or populations (Oaks, 2017). Unlike *dpp-msbayes*, *ecoevolity* assumes that there is no gene flow between sister-taxa. Because *ecoevolity* does not required shared loci among species comparisons, we used the SNP assemblies generated from sister-pair iPyrad pipelines to maximize the number of loci for analysis. We set the

concentration prior for the number of divergent events to 1.8 in order to match the prior probability distribution for the *dpp-msbayes* analysis. We used an exponential distribution with rate of 10 (mean = 0.1) for the divergence time (i.e. expected number of substitutions per site) and set the population size prior to a gamma distribution with shape = 4.0 and scale = 0.001. Finally, we ran the analysis twice from unique starting seeds, sampling every 50 steps for 75,000 iterations. We assessed convergence in Tracer, discarded the first 50% of samples, and combined the two runs before using *pycoevolity* scripts to plot the results. We used Bayes factors to determine support for different models of divergence.

Results

Sanger dataset

The final nuclear sequence dataset comprised 692,679 bp with 9% missing data. All loci tested negative for recombination. The top nucleotide substitution model for each locus is listed in Table 3.

ddRADseq SNP assemblies

We generated approximately 110 million raw reads with an average of 2.4 million reads per individual after filtering for read quality and barcode match (Table 4). The optimal clustering threshold for the *cristatellus*-group pipeline was 90% (Figure 1). The optimal parameter value for the number of SNPs per locus was 11, the number of indels per locus was three, and the proportion of shared heterozygous sites was 25% (Figure 1). These parameter values were chosen because loci that could be recovered from higher parameter values were likely paralogs. All four species-pairs had similar optimal parameter values: the clustering threshold was 92%,

the number of SNPs per loci was 6, the number of indels per locus was three, and the maximum proportion of shared heterozygous sites was 50% (Figure 2). On average, we recovered between 4000-6500 loci per species from the species-pairs pipelines; approximately three times as many loci as from the *cristatellus*-group pipeline (Table 5).

Species trees

Species trees estimated from *BEAST and from SNAPP were congruent with each other. Sister-species relationships are all highly supported (>0.99) in both analysis. Unlike previous mitochondrial analyses (Brandley and de Queiroz 2004; Nicholson et al. 2005; Mahler et al. 2010), our results support *A. gundlachi – A. poncensis* being closely related to *A. krugi – A. pulchellus* (>0.95). However, our species tree analysis indicates uncertainty in the relationships among the remaining sister-species pairs, which may imply rapid diversification early in the radiation.

Isolation-migration

Gene flow rates for three sister-species pairs were zero, but IMA2 estimated gene flow between the sister-species *A. krugi* and *A. pulchellus* (Fig. 3F; Table 6), indicating that this species pair either diverged with gene flow, or experienced gene flow at secondary contact. This result complements a previous study — which documented uni-directional mtDNA gene flow from *A. krugi* to *A. pulchellus* in this species pair (Jezkova et al., 2013) — but shows bi-directional nuclear gene flow between the species pair. Divergence dates for Puerto Rican sister-species estimated from IMA2 range between 5-15 mya, and the 95% CI all overlap, providing anecdotal and inconclusive support for synchronous divergence (Fig 3E).

Co-evolutionary divergence

Our results for the number of divergences in Puerto Rican anoles depends on the datasets and analytical methods (Figure S2). The approximate Bayesian approach (*dpp-msbayes*) with the nuclear sequence dataset favors no gene flow (PP no-migration: 0.82 vs. PP migration: 0.18) and a single divergence event for all Puerto Rican sister-species (BF = 1.82); however any given model has little support (all BF <2; Table 7) indicating that information in the dataset is insufficient to estimate precise divergence times. The full likelihood method (*ecoevolity*) with the SNP dataset supports idiosyncratic divergence times among the species pairs (Figure 4). The best model has unique divergence parameters for each species pair (PP = 0.56, BF= 10.16; Table 7). The second-best model is a three-divergence model in which *A. cooki – A. cristatellus* and *A. gundlachi – A. poncensis* diverge synchronously (PP = 0.372, BF = 8.69; Table 7)).

Discussion

We confirmed that the *cristatellus* group is represented by well-supported pairs of sister-species (Figure 3; Nicholson et al., 2005; Mahler et al., 2010). Prior phylogenetic estimates of the *cristatellus* group differ in their placement of *Anolis gundlachi*. Morphological studies placed *A. gundlachi* with the other trunk-ground anoles, *A. cooki* and *A. cristatellus*, and placed *A. poncensis* with the other twig anoles, *A. krugi* and *A. pulchellus*. Mitochondrial studies place *A. gundlachi* within the twig anoles, either as sister to *A. krugi* + *A. pulchellus* (Brandley and de Queiroz 2004) or as sister to *A. poncensis* (Nicholson et al. 2005; Mahler et al. 2010). However, studies that placed *A. gundlachi* and *A. poncensis* as sister taxa were unable to resolve the

phylogenetic placement of the two species within the *cristatellus* group. Subsequent studies based on nuclear data either did not incorporate all the species of the *cristatellus* group (Alföldi et al. 2011) or they also included morphological data (Poe et al. 2017), which may have biased the phylogenetic placement of *A. poncensis* and *A. gundlachi*. Our species-tree analyses of nuclear sequences and of SNP datasets both support the sister-species relationship of *A. gundlachi* and *A. poncensis*, and place the species pair as sister to the other grass-bush species, *A. krugi* and *A. pulchellus*. Our inability to confidently place the *cristatellus-cooki* pair in the tree likely stems from deep coalescence due to large population sizes and from rapid diversification early in the Puerto Rican radiation.

Species-pair divergence in Puerto Rican anoles is represented by two general scenarios – deep divergence without gene flow, and recent divergence with gene flow. The confidence intervals around nodes are wide, and divergence time estimates from nuclear sequences analyzed in IMa2 for three species-pairs all overlap. These analyses suggest that three species pairs diverged at roughly similar times, but with *A. krugi – A. pulchellus* diverging much more recently. IMa2 results for *A. krugi* and *A. pulchellus* support a model of divergence with gene flow, which compliments previous reports of mtDNA introgression from *A. krugi* to *A. pulchellus* (T. Jezkova, Leal, and Rodríguez-Robles 2013). The species tree methods *BEAST and SNAPP both assume that no gene flow, which may lead to artificially short branch lengths for *A. krugi* and *A. pulchellus* (Leaché et al. 2014) (Figure 3).

Species trees and population divergence analysis indicated synchronous divergence events, but we wanted to follow up and use methods (https://github.com/phyletica/ecoevolity; Oaks, 2014) that are explicitly designed to estimate co-evolutionary diversification events and to

resolve idiosyncrasies in the timing of diversification. Species pairs exhibiting concordant divergence times may have been influenced by a shared geological event (Brown et al. 2013; Oaks et al. 2013). Conversely, idiosyncratic or multiple waves of divergences of species pairs likely result from incomplete formation of historical barriers, unequal selective pressure of climatic gradients, and differing organismal responses to both due to variation in the phenotypic traits, life history, or physiology (Leache, Crews, and Hickerson 2007; Bell et al. 2010; Zamudio, Bell, and Mason 2016). Our results favor the later explanation for *in situ* speciation of Puerto Rican anoles. Inundation is a driver of allopatric divergence of anoles on Hispaniola and on Cuba, but the center of Puerto Rico has not been inundated (MacPHEE, ITURRALDE-VINENT, and Gaffney 2003). In the top model, sister-species divergence occurs idiosyncratically, and in the second best model there are three events, with *A. gundlachi – A. poncensis* and *A. cooki – A. cristatellus* exhibiting simultaneous divergence.

The potential synchronous divergence of *A. gundlachi – A. poncensis* and *A. cooki – A. cristatellus* roughly 10 mya provides insight into xeric adaptation on Puerto Rico. First, the divergence of the two xeric-specialists, *A. cooki* and *A. poncensis*, may have coincided with adaptation to novel or expanded arid habitats in the Southwest. The divergence of *A. gundlachi – A. poncensis* also indicates that Puerto Rico can only sustain three physiological forms (xeric, semi-xeric, mesic) of a given ecomorph. If we assume that the common ancestor to *A. gundlachi* (trunk-ground) – *A. poncensis* (grass-bush) was a trunk-ground anole, then the simultaneous divergence of *A. gundlachi – A. poncensis* and *A. cooki – A. cristatellus* would have led to four trunk ground anoles. Character displacement likely drove physiological adaptation of *A. cooki* to the xeric habitat, *A. cristatellus* to the semi-xeric environment, and *A. gundlachi* to the mesic habitat. The xeric grass-bush habitat may have been relatively unoccupied, however, because *A.*

pulchellus and *A. krugi* had not yet diverged. In this case, all the trunk-ground thermal niches would have been occupied, and *A. poncensis* would have independently evolved the grass-bush morphology to exploit newly available habitat in the Southwest. This scenario holds true even all four divergences occur idiosyncratically, because the *A. gundlachi – A. poncensis* divergence is estimated to occur after the *A. cooki – A. cristatellus* divergence and before the *A. krugi – A. pulchellus* divergence (Figures 3,4).

Our results indicate that the evolution of thermal physiology on Puerto Rico, while predictable, is perhaps more idiosyncratic than initially anticipated by Williams. First, our understanding of the phylogenetic relationships of Puerto Rican anoles have changed. Williams' initial synthesis was based on the hypothesis that trunk-ground and grass-bush ecomorphology each evolved once, with subsequent divergence into the three physiological forms of xeric, semi-xeric, and mesic. Our current understanding of the phylogenetic relationships requires a messier explanation, including two evolutions of either the grass-bush or trunk-ground morphologies. While the evolution of thermal biology occurring does appear to represent a "final stage" of the Puerto Rican *Anolis* adaptive radiation, the outcome of physiological adaptation on Puerto Rica is not predictable. The sister-pairs represent all possible combinations of physiological outcomes: xeric and semi-xeric; semi-xeric and mesic; as well as xeric and mesic. Finally, in the case of *A. gundlachi* and *A. poncensis*, the evolution of thermal physiology and morphology can be dynamic and complex even in the later stages of the *Anolis* radiation.

References

Alföldi, J, FD Palma, M Grabherr, C Williams, L Kong, E Mauceli, P Russell, et al. 2011. "The Genome of the Green Anole Lizard and a Comparative Analysis with Birds and Mammals." *Nature* 477 (7366): 587–91. https://doi.org/10.1038/nature10390.

Arnold, B, RB Corbett-Detig, D Hartl, and K Bomblies. 2013. "RADseq Underestimates Diversity and Introduces Genealogical Biases Due to Nonrandom Haplotype Sampling." *Molecular Ecology* 22 (11): 3179–90. https://doi.org/10.1111/mec.12276.

Bell, RC, JL Parra, M Tonione, CJ Hoskin, JB Mackenzie, SE Williams, and C Moritz. 2010. "Patterns of Persistence and Isolation Indicate Resilience to Climate Change in Montane Rainforest Lizards" *Molecular Ecology* 19: 2531-2544. https://doi.org/10.1111/j.1365-294X.2010.04676.x.

Bouckaert, R, J Heled, D Kühnert, T Vaughan, CH Wu, D Xie, MA Suchard, A Rambaut, and AJ Drummond. 2014. "BEAST 2: A Software Platform for Bayesian Evolutionary Analysis." *PLoS Computational Biology* 10 (4): e1003537. https://doi.org/10.1371/journal.pcbi.1003537.

Brandley, MC, and K Queiroz. 2004. "Phylogeny, Ecomorphological Evolution, and Historical Biogeography of the *Anolis cristatellus* Series." *Herpetological Monographs* 18 (1): 90–126.

Brown, RM, CD Siler, CH Oliveros, JA Esselstyn, AC Diesmos, PA Hosner, CW Linkem, et al. 2013. "Evolutionary Processes of Diversification in a Model Island Archipelago." *Annual Review of Ecology, Evolution, and Systematics* 44 (1): 411–35. https://doi.org/10.1146/annurev-ecolsys-110411-160323.

Bryant, D, R Bouckaert, J Felsenstein, NA Rosenberg, and A RoyChoudhury. 2012. "Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis." *Molecular Biology and Evolution* 29 (8): 1917–32. https://doi.org/10.1093/molbev/mss086.

Carstens, BC, RS Brennan, V Chua, CV Duffie, MG Harvey, RA Koch, CD McMahan, et al. 2013. "Model Selection as a Tool for Phylogeographic Inference: An Example from the Willow Salix Melanopsis." *Molecular Ecology* 22 (15): 4014–28. https://doi.org/10.1111/mec.12347.

Diamond, JM 1973. "Distributional Ecology of New Guinea Birds." *Science* 179 (4075): 759–69.

Dobzhanksy, T. 1951. *Genetics and the Origin of Species. 3rd Ed.* New York, NY: Columbia University Press.

Drummond, AJ, MA Suchard, D Xie, and A Rambaut. 2012. “Bayesian Phylogenetics with BEAUti and the BEAST 1.7.” *Molecular Biology and Evolution* 29 (8): 1969–73. https://doi.org/10.1093/molbev/mss075.

Eaton, DA R. 2014. “PyRAD: Assembly of de Novo RADseq Loci for Phylogenetic Analyses.” *Bioinformatics* 30 (13): 1844–49. https://doi.org/10.1093/bioinformatics/btu121.

Edgar, RC. 2004. “MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput.” *Nucleic Acids Research* 32 (5): 1792–97. https://doi.org/10.1093/nar/gkh340.

Grant, V. 1949. “Pollination Systems as Isolating Mechanisms in Angiosperms.” *Evolution* 3 (1): 82. https://doi.org/10.2307/2405454.

Grummer, JA, Calderón-Espinosa, ML, Nieto-Montes de Oca, A, Smith, EN, Méndez-de la Cruz, FR, & Leaché, AD. 2015. “Estimating the Temporal and Spatial Extent of Gene Flow among Sympatric Lizard Populations (genus *Sceloporus*) in the Southern Mexican Highlands.” *Molecular ecology*, *24*(7), 1523-1542.

Harvey, MG, CD Judy, GF Seeholzer, JM Maley, GR Graves, and RT Brumfield. 2015. “Similarity Thresholds Used in DNA Sequence Assembly from Short Reads Can Reduce the Comparability of Population Histories across Species.” *PeerJ* 3 (April): e895. https://doi.org/10.7717/peerj.895.

Heatwole, H, TH Lin, E Villalon, A Muniz, and A Matta. 1969. “Some Aspects of the Thermal Ecology of Puerto Rican Anoline Lizards.” *Journal of Herpetology* 3 (1/2): 65. https://doi.org/10.2307/1563225.

Henderson, RW, and R Powell. 2009. *Natural History of West Indian Reptiles and Amphibians*. University Press of Florida.

Hertz, PE 1992a. “Evaluating Thermal Resource Partitioning.” *Oecologia*. 90: 127-136.

———. 1992b. “Temperature Regulation in Puerto Rican Anolis Lizards: A Field Test Using Null Hypotheses.” *Ecology* 73 (4): 1405–17. https://doi.org/10.2307/1940686.

Hertz, PE, Y Arima, A Harrison, RB Huey, JB Losos, and RE Glor. 2013. “Asynchronous Evolution of Physiology and Morphology in *Anolis* Lizards.” *Evolution* 67 (7): 2101–13. https://doi.org/10.1111/evo.12072.

Hertz, PE, RB Huey, and RD Stevenson. 1993. “Evaluating Temperature Regulation by Field-Active Ectotherms: The Fallacy of the Inappropriate Question.” *American Naturalist*, 142 (5): 796–818.

Hillman, S, GC Gorman, and Rd Thomas. 1979. “Water Loss in Anolis Lizards Evidence for Acclimation and Intraspecific Differences along a Habitat Gradient.” *Comparative Biochemistry and Physiology*. 62: 491-494.

Huang, W, N Takebayashi, Y Qi, and MJ Hickerson. 2011. “MTML-MsBayes: Approximate Bayesian Comparative Phylogeographic Inference from Multiple Taxa and Multiple Loci with Rate Heterogeneity.” *BMC Bioinformatics* 12 (1): 14.

Huey, RB, and TP Webster. 1976. “Thermal Biology of Anolis Lizards in a Complex Fauna: The *Cristatellus* Group on Puerto Rico.” *Ecology* 57 (5): 985–94. https://doi.org/10.2307/1941063.

Ilut, DC, ML Nydam, and MP Hare. 2014. “Defining Loci in Restriction-Based Reduced Representation Genomic Data from Nonmodel Species: Sources of Bias and Diagnostics for Optimal Clustering.” *BioMed Research International* 2014: 1–9. https://doi.org/10.1155/2014/675158.

Jackman, TR., A Larson, K deQueiroz, and JB Losos. 1999. “Phylogenetic Relationships and Tempo of Early Diversification in *Anolis* Lizards.” *Systematic Biology* 48 (2): 254–285.

Jezkova, T, M Leal, and JA Rodríguez-Robles. 2013. “Genetic Drift or Natural Selection? Hybridization and Asymmetric Mitochondrial Introgression in Two Caribbean Lizards

(*Anolis pulchellus* and *Anolis krugi*)." *Journal of Evolutionary Biology* 26 (7): 1458–71. https://doi.org/10.1111/jeb.12149.

Jezkova, T, M Leal, and JA Rodríguez-Robles. 2009. "Living Together but Remaining Apart: Comparative Phylogeography of *Anolis poncensis* and *A. cooki,* Two Lizards Endemic to the Aridlands of Puerto Rico." *Biological Journal of the Linnean Society* 96 (3): 617–634.

Lack, D 1947. *Darwin's Finches*. Cambridge, UK: Cambridge University Press.

Leaché, AD, SC Crews, and MJ Hickerson. 2007. "Two Waves of Diversification in Mammals and Reptiles of Baja California Revealed by Hierarchical Bayesian Analysis." *Biology Letters* 3 (6): 646–50. https://doi.org/10.1098/rsbl.2007.0368.

Leaché, AD, RB Harris, B Rannala, and Z Yang. 2014. "The Influence of Gene Flow on Species Tree Estimation: A Simulation Study." *Systematic Biology* 63 (1): 17–30. https://doi.org/10.1093/sysbio/syt049.

Losos, JB 2009. *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*. Vol. 10. Berkeley, CA: University of California Press.

Lynch, M. 2008. “Estimation of Nucleotide Diversity, Disequilibrium Coefficients, and Mutation Rates from High-Coverage Genome-Sequencing Projects.” *Molecular Biology and Evolution* 25 (11): 2409–19. https://doi.org/10.1093/molbev/msn185.

Macey, RJ, JA Schulte II, NB Ananjeva, A Larson, N Rastegar-Pouyani, SM Shammakov, and TJ Papenfuss. 1998. “Phylogenetic Relationships among Agamid Lizards of the *Laudakia caucasia* Species Group: Testing Hypotheses of Biogeographic Fragmentation and an Area Cladogram for the Iranian Plateau.” *Molecular Phylogenetics and Evolution* 10 (1): 118–31.

MacPhee, RDE, MA Iturralde-Vinent, and ES Gaffney. 2003. “Domo de Zaza, an Early Miocene Vertebrate Locality in South-Central Cuba, with Notes on the Tectonic Evolution of Puerto Rico and the Mona Passage.” *American Museum Novitates*, 1–42.

Mahler, DL, LJ Revell, RE Glor, and JB Losos. 2010. “Ecological Opportunity and the Rate of Morphological Evolution in the Diversification of Greater Antillean Anoles: Opportunity and Rate in *Anolis* Lizards.” *Evolution* 64 (9): 2731–45. https://doi.org/10.1111/j.1558-5646.2010.01026.x.

McKinnon, JS, and HD Rundle. 2002. “Speciation in Nature: The Threespine Stickleback Model Systems.” *Trends in Ecology & Evolution* 17 (10): 480–88. https://doi.org/10.1016/S0169-5347(02)02579-X.

Milne, I, D Lindner, M Bayer, D Husmeier, G McGuire, DF Marshall, and F Wright. 2009. "TOPALi v2: A Rich Graphical Interface for Evolutionary Analyses of Multiple Alignments on HPC Clusters and Multi-Core Desktops." *Bioinformatics* 25 (1): 126–27. https://doi.org/10.1093/bioinformatics/btn575.

Nicholson, KE, RE Glor, JJ Kolbe, A Larson, SB Hedges, and JB Losos. 2005. "Mainland Colonization by Island Lizards: Mainland Colonization by *Anolis*." *Journal of Biogeography* 32 (6): 929–38. https://doi.org/10.1111/j.1365-2699.2004.01222.x.

Nieto-Montes de Oca, A, AJ Barley, RN Meza-Lázaro, UO García-Vázquez, JG Zamora-Abrego, RC Thomson, and AD Leaché. 2017. "Phylogenomics and Species Delimitation in the Knob-Scaled Lizards of the Genus *Xenosaurus* (Squamata: Xenosauridae) Using DdRADseq Data Reveal a Substantial Underestimation of Diversity." *Molecular Phylogenetics and Evolution* 106 (January): 241–53. https://doi.org/10.1016/j.ympev.2016.09.001.

Oaks, JR 2014. "An Improved Approximate-Bayesian Model-Choice Method for Estimating Shared Evolutionary History." *BMC Evolutionary Biology* 14 (1): 150.

Oaks, JR, J Sukumaran, JA Esselstyn, CW Linkem, CD Siler, MT Holder, and RM Brown. 2013. "Evidence For Climate-Driven Diversification? A Caution For Interpreting ABC Inferences of Simultaneous Historical Events." *Evolution* 67 (4): 991–1010. https://doi.org/10.1111/j.1558-5646.2012.01840.x.

Peterson, BK, JN Weber, EH Kay, HS Fisher, and HE Hoekstra. 2012. "Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species." *PLoS ONE* 7 (5): e37135. https://doi.org/10.1371/journal.pone.0037135.

Poe, S, A Nieto-montes de Oca, O Torres-carvajal, K deQueiroz, JA Velasco, B Truett, LN Gray, et al. 2017. "A Phylogenetic, Biogeographic, and Taxonomic Study of All Extant Species of *Anolis* (Squamata; Iguanidae)." *Systematic Biology* 66 (5): 663–97. https://doi.org/10.1093/sysbio/syx029.

Portik, DM, PL Wood Jr., JL Grismer, EL Stanley, and TR Jackman. 2012. "Identification of 104 Rapidly-Evolving Nuclear Protein-Coding Markers for Amplification across Scaled Reptiles using Genomic Resources." *Conservation Genetic Resources* 4 (1): 1–10.

Posada, D, 2008. "JModelTest: Phylogenetic Model Averaging." *Molecular Biology and Evolution* 25 (7): 1253–56. https://doi.org/10.1093/molbev/msn083.

Rand, AS. 1964. "Ecological Distribution in Anoline Lizards of Puerto Rico." *Ecology* 45 (4): 745–52. https://doi.org/10.2307/1934922.

Rivero, JA 1998. *Amphibians and Reptiles of Puerto Rico*. Puerto Rico: La Editorial de la Universidad de Puerto Rico.

Rodríguez-Robles, JA, T Jezkova, and MA García. 2007. "Evolutionary Relationships and Historical Biogeography of *Anolis desechensis* and *Anolis monensis*, Two Lizards Endemic to Small Islands in the Eastern Caribbean Sea." *Journal of Biogeography* 34 (9): 1546–58. https://doi.org/10.1111/j.1365-2699.2007.01718.x.

Rognes, T, T Flouri, B Nichols, C Quince, and F Mahé. 2016. "VSEARCH: A Versatile Open Source Tool for Metagenomics." *PeerJ* 4 (October): e2584. https://doi.org/10.7717/peerj.2584.

Ruibal, R. 1961. "Thermal Relations of Five Species of Tropical Lizards." *Evolution* 15 (1): 98. https://doi.org/10.2307/2405846.

Schluter, D. 2000. *The Ecology of Adaptive Radiation*. Oxford, UK: Oxford University Press.

Schoener, TW. 1971. "Structural Habitats of West Indian Anolis Lizards. II. Puerto Rican Uplands." *Breviora*, no. 368: 1–39.

Simpson, GG. 1955. *The Major Features of Evolution*. New York, NY: Columbia University Press.

Stephens, M, NJ Smith, and P Donnelly. 2001. "A New Statistical Method for Haplotype Reconstruction from Population Data." *The American Journal of Human Genetics* 68 (4): 978–989.

Streelman, JT, and PD Danley. 2003. “The Stages of Vertebrate Evolutionary Radiation.” *Trends in Ecology & Evolution* 18 (3): 126–131.

Terborgh, J, and JM Diamond. 1970. “Niche Overlap in Feeding Assemblages of New Guinea Birds.” *The Wilson Bulletin*, 29–52.

Williams, EE. 1972. “The Origin of Faunas: Evolution of Lizard Congeners in a Complex Island Fauna - A Trial Analysis.” In *Evolutionary Biology,* 6:47–89. Boston, MA: Springer.

———. 1983. “Ecomorphs, Faunas, Island Size, and Diverse End Points in Island Radiations of *Anolis*.” In *Lizard Ecology: Studies of a Model Organism*, 326–70. Cambridge, MA: Harvard Univeristy Press.

Zamudio, KR, RC Bell, and NA Mason. 2016. “Phenotypes in Phylogeography: Species Traits, Environmental Variation, and Vertebrate Diversification.” *Proceedings of the National Academy of Sciences* 113 (29): 8041–48. https://doi.org/10.1073/pnas.1602237113.

Table 1. Catalogue numbers and localities of specimens.

Species	Locality	Lat	Long	Catalogue
A. cooki	Boqueron	17.93	-67.19	MVZ 235170, 235172
A. cooki	Bahia Ballena	17.96	-66.86	MVZ 226113-226117
A. cristatellus	Rio Piedras	18.40	-66.05	UWBM 6686-6688
A. cristatellus	Guanica	17.95	-66.85	UWBM 6689-6692
A. cristatellus	Maricao	18.15	-66.99	UWBM 6694-6697
A. cristatellus	Mata de Platano	18.41	-66.73	UWBM 6699-6702
A. cristatellus	Toro Negro	18.17	-66.49	UWBM 6703-6707
A. cristatellus	Carite	18.11	-66.07	UWBM 6709
A. cristatellus	Carite	18.09	-66.03	UWBM 6710-6712
A. cristatellus	El Yunque	18.30	-65.86	UWBM 6716-6717
A. cristatellus	El Yunque	18.33	-65.86	UWBM 6718
A. cristatellus	Guajataca	18.42	-66.97	UWBM 6721-6724
A. evermanni	Maricao	18.15	-66.99	UWBM 6734-6738
A. evermanni	Toro Negro	18.17	-66.49	UWBM 6739-6741
A. evermanni	Piedra Escrito	18.22	-66.57	UWBM 6742-6745
A. evermanni	Carite	18.09	-66.03	UWBM 6746-6749
A. evermanni	El Yunque	18.29	-65.85	UWBM 6750-6754
A. gundlachi	Maricao	18.15	-66.99	UWBM 6757-6760
A. gundlachi	Rio Abajo	18.32	-66.69	UWBM 6761-6765
A. gundlachi	Toro Negro	18.17	-66.49	UWBM 6766-6770
A. gundlachi	Carite	18.12	-66.08	UWBM 6771-6773
A. gundlachi	Carite	18.09	-66.03	UWBM 6774-6777
A. gundlachi	El Yunque	18.29	-65.85	UWBM 6778-6781
A. gundlachi	El Yunque	18.31	-65.78	UWBM 6782-6785
A. gundlachi	Guajataca	18.42	-66.96	UWBM 6786-6789
A. krugi	Maricao	18.15	-66.99	UWBM 6790-6893
A. krugi	Mata de Platano	18.41	-66.73	UWBM 6794-6798
A. krugi	Toro Negro	18.17	-66.49	UWBM 6799-6803
A. krugi	Jayuya	18.21	-66.57	UWBM 6804-6805
A. krugi	Carite	18.12	-66.08	UWBM 6806-6807
A. krugi	Carite	18.11	-66.07	UWBM 6808-6812
A. krugi	Yunque	18.34	-65.83	UWBM 6813-6816
A. krugi	Yunque	18.31	-65.78	UWBM 6817-6820
A. krugi	Guajataca	18.42	-66.96	UWBM 6821-6824
A. poncensis	Ponce	17.97	-66.67	UWBM 6831-6834
A. poncensis	Ponce	17.98	-66.67	UWBM 6835-6838
A. poncensis	Cabo Rojo	18.09	-67.15	UWBM 6839-6840
A. pulchellus	Guanica	17.95	-66.85	UWBM 6841-6842
A. pulchellus	Ponce	18.02	-66.57	UWBM 6843-6846
A. pulchellus	Mata da Platano	18.41	-66.73	UWBM 6847-6850
A. pulchellus	Carite	18.11	-66.07	UWBM 6851-6853
A. pulchellus	Palmer	18.37	-65.77	UWBM 6854-6857

Table 1. Continued…

Species	Locality	Lat	Long	Catalogue
A. pulchellus	Ponce	17.98	-66.67	UWBM 6858-6859
A. pulchellus	Palomas	18.01	-66.89	UWBM 6860-6863
A. pulchellus	Guajataca	18.44	-66.97	UWBM 6864-6866
A. stratulus	Maricao	18.15	-66.99	UWBM 6867-6868
A. stratulus	Maricao	18.16	-67.00	UWBM 6869-6870
A. stratulus	Carite	18.11	-66.07	UWBM 6871
A. stratulus	Yunque	18.33	-65.86	UWBM 6872-6875
A. stratulus	Yunque	18.34	-65.83	UWBM 6876-6880
A. stratulus	Carite	18.09	-66.04	UWBM 6881-6883
A. stratulus	Guajataca	18.42	-66.97	UWBM 6884-6887
A. stratulus	Mayaguez	18.22	-67.16	UWBM 6888-6889

Table 2. Number of individuals sequenced for each gene.

Species	N	*ALMS1*	*EXPH5*	*KIAA2018*	*KIF24*	*NOS1*	*PPP2R5C*	*RAPGEF2*	*R35*	*ND2*[a]
A. cooki	6	6	6	6	6	5	6	6	6	3
A. cristatellus	33	31	25	31	33	32	32	29	32	26
A. evermanni	23	23	19	21	23	18	22	20	21	21
A. stratulus	23	22	21	22	22	21	20	21	21	22
A. krugi	35	30	33	32	35	29	26	29	33	33
A. pulchellus	27	24	23	24	17	26	25	26	21	24
A. gundlachi	33	32	31	31	33	30	30	31	31	23
A. poncensis	9	8	9	9	9	9	8	8	9	9

[a]mitochondrial DNA sequence

Table 3. Characteristics of Sanger genes.

Gene	Sequences	Basepairs	Substitution Model
ALMS1	176	438	HKY+G
EXPH5	167	864	HKY+I+G
KIAA2018	176	617	GTR+G
KIF24	178	473	HKY+I+G
NOS1	170	673	HKY+G
PPP2R5C	169	569	GTR
RAPGEF2	170	231	HKY+I
R35	174	574	GTR+G
$ND2^{mtdna}$	161	1058	GTR+I+G

Table 4. Characteristics of ddRADseq datasets.

Species[a]	N	Reads[b]	Passed[c]	Clusters[d]	Depth[e]	Loci[f]	Het[g]
A. evermanni – A. stratulus	12	2,352,128	2,346,460	11,509	64.5	7,492	0.0092
A. cooki – A. cristatellus	12	2,615,284	2,603,006	12,304	67.7	7,024	0.0095
A. krugi – A. pulchellus	12	2,470,964	2,463,950	12,359	68.6	8,231	0.0099
A. gundlachi – A. poncensis	8	2,338,825	2,332,511	12,003	65.4	8,473	0.0097
cristatellus series	44	2,453,889	2,445,933	12,003	66.9	2,776	0.0098

[a]Species used for iPyrad assembly.
[b]Mean raw read count after sample demultiplexing.
[c]Mean number of reads passing quality filters.
[d]Mean number of clusters passing quality filters.
[e]Mean sequencing depth per cluster.
[f]Number of loci retained for datasets in which 50% of individuals are present.
[g]Mean heterozygosity

Table 5. Mean number of loci per individual (and range) for two iPyrad assembly procedures.

Species	N	Sister-pairs[abcd]	*Cristatellus* clade[e]
A. evermanni	5	4,040 (2,113 – 4,652)[a]	1,487.0 (820 – 1,735)
A. stratulus	7	6,110 (5,141 – 6,996)[a]	1,580.7 (1,214 – 2,028)
A. cooki	5	5,246 (5,056 – 5,495)[b]	2,064.2 (1,950 – 2,182)
A. cristatellus	7	4,826 (4,066 – 5,584)[b]	1,825.0 (1,477 – 2,185)
A. krugi	5	6,181 (5,636 – 6,477)[c]	2,203.4 (1,986 – 2,298)
A. pulchellus	7	5,996 (4,879 – 6,740)[c]	2,078.1 (1,645 – 2,396)
A. gundlachi	5	6,524 (6,120 – 7,015)[d]	2,078.0 (1,928 – 2,205)
A. poncensis	3	5,522 (5,318 – 5,782)[d]	2,000.0 (1,912 – 2,138)

[a]*A. evermanni – A. stratulus*.
[b]*A. cooki – A. cristatellus.*
[c]*A. krugi – A. pulchellus*.
[d]*A. gundlachi – A. poncensis*.
[e]Assembly with all eight species.

Table 6. Top model rankings for *A. krugi – A. pulchellus* Isolation-Migration analysis.

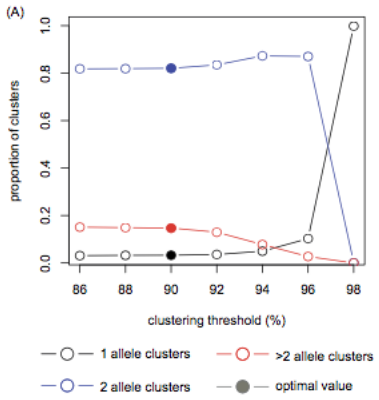
Model	Log(P)	K	kr (q)	pu (q)	anc. (q)	kr>pu (m)	pu>kr (m)	AIC	ΔAIC	ω	Cum.prob
m1=m2,p1,p2,p3	1.681	4	4.10	12.41	7.16	0.15	[0.15]	4.64	0.00	0.54	0.54
m1,m2,p1,p2,p3	2.009	5	3.98	12.75	7.01	0.19	0.12	5.98	-1.34	0.27	0.81
m1=m2,p1,p2=p3	-1.367	3	3.87	10.44	[10.44]	0.17	[0.17]	8.73	-4.10	0.07	0.88
m1=m2,p1=p3,p2	-1.467	3	4.91	12.75	[4.91]	0.13	[0.13]	8.93	-4.30	0.06	0.94
m1,m2,p1=p3,p2	-1.13	4	5.54	13.90	[5.54]	0.06	0.15	10.26	-5.62	0.03	0.97
m1,m2,p1,p2=p3	-1.367	4	3.87	10.44	[10.44]	0.17	0.17	10.73	-6.10	0.03	1.00

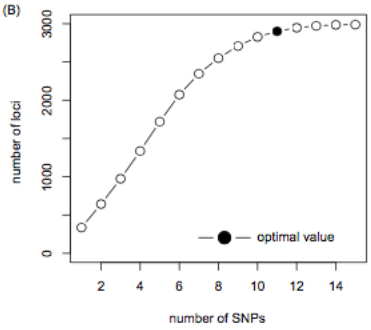
'Model' describes the migration/population model implemented in IMa2.
K = number of migration and population size parameters. Population size parameters were independent in each model.
ω = akiake weights. Can be interpreted as the probability of the model.
kr = *A. krugi*
pu = *A. pulchellus*
anc. = ancestral population
q = population size
m = migration

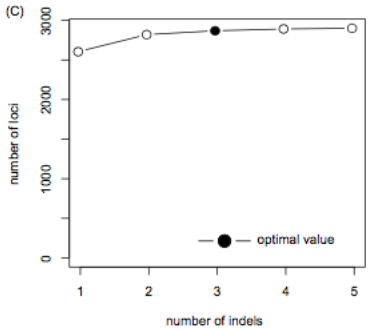
Table 7. Best divergence models ranked by Bayes factors.

dpp-msbayes			ecoevolity				
divs[a]	**model[b]**	**posterior**	**divs**	**model**	**posterior**	**prior**	**BF**
1	cocr=gupo=krpu=evst	0.284	4	cocr / gupo / krpu / evst	0.566	0.114	10.14
2	cocr= evst / gupo=krpu	0.077	3	cocr=gupo / krpu / evst	0.372	0.063	8.69
2	cocr / gupo=krpu=evst	0.071	3	evst =gupo / krpu / cocr	0.062	0.063	0.97
2	cocr=gupo=evst / krpu,	0.070	-	-	-	-	-

[a]number of divergence events
[b]description of divergence model
[c]Bayes Factor
cocr: *A. cooki – A. cristatellus*
evst: *A. evermanni – A. stratulus*
gupo: *A. gundlachi – A. poncensis*
krpu: *A. krugi – A. pulchellus*







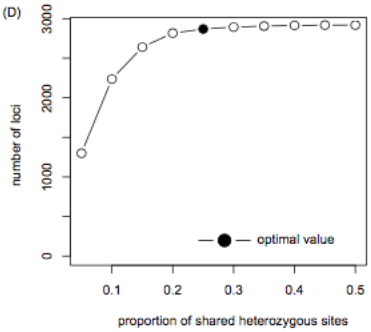


Figure 1. Summary plots of iPyRad output for the *cristatellus*-series. (A) Variation in the proportions of clusters with 1, 2, and >2 alleles per cluster. (B) Variation in the number of recovered loci with different maximum number of SNPs in each locus. (C) Variation in the number of recovered loci with different maximum numbers of insertions and deletions in the each locus. (D) Variation in the number of recovered loci with different maximum number of samples with a shared heterozygous site. Filled in circles represent the optimal parameter value.

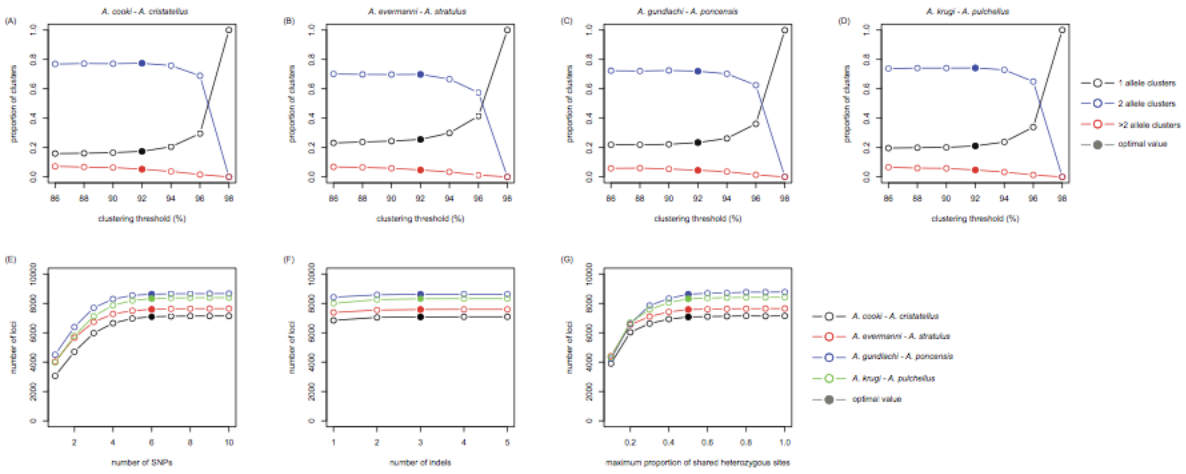


Figure 2. Summary plots of iPyRad output for each sister-pair. (A-D) Variation in the proportions of clusters with 1, 2, and >2 alleles per cluster. (E) Variation in the number of recovered loci with different maximum number of SNPs in each locus. (F) Variation in the number of recovered loci with different maximum numbers of insertions and deletions in the each locus. (G) Variation in the number of recovered loci with different maximum number of samples with a shared heterozygous site. Filled in circles represent the optimal parameter value. Different colored lines represent different assemblies; see key for details.

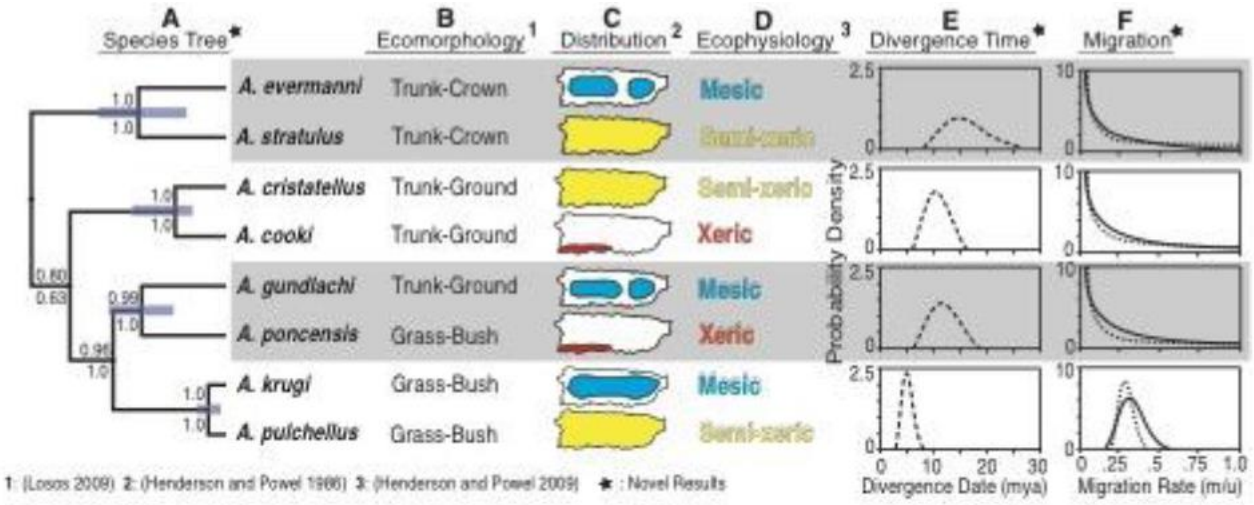


Figure 3. Ecomorphology (B), distribution (C), and ecophysiology (D), placed in a phylogenetic context (A). Divergence time (E) and migration rate (F) for sister-species pairs were inferred in the program IMa2. Posterior probabilities from *BEAST (above) and SNAPP (below) each node.

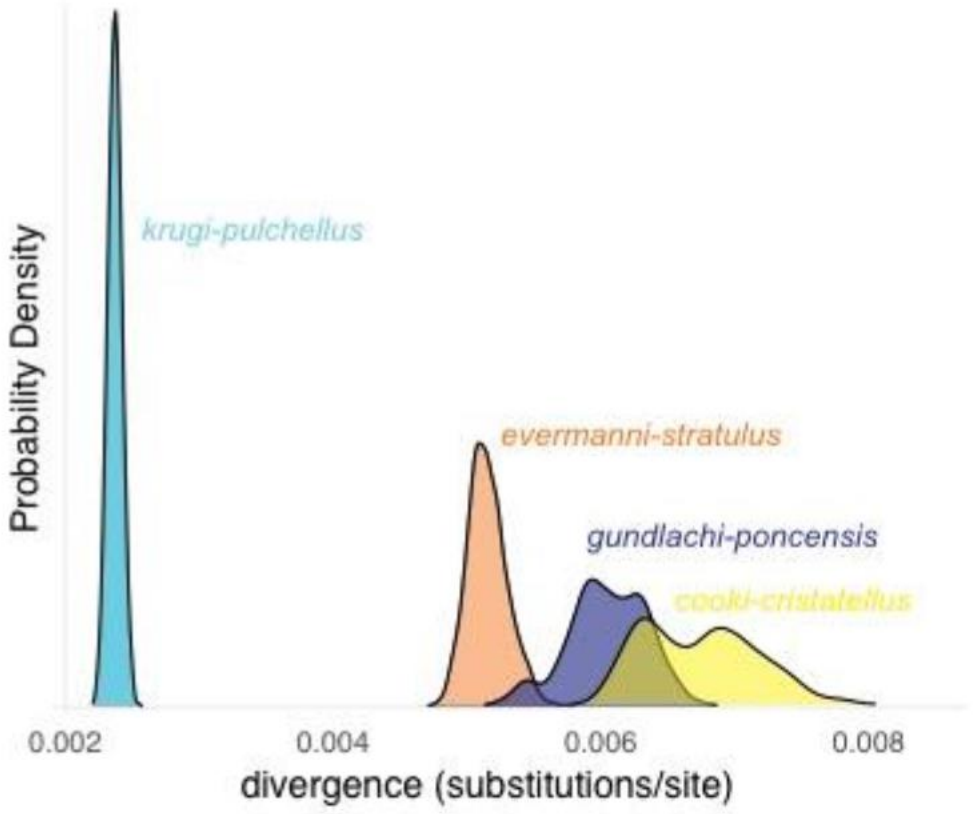


Figure 4. Probability densities of divergence estimates for species pairs across all ecoevolity divergence models.

Chapter 2: Precipitation mediates phylogeographic structure and gene flow in the Puerto Rican Crested Anole (*Anolis cristatellus*)

Matthew T. McElroy[12]

[1]Department of Biology, University of Washington, Seattle, WA 98195

[2]Burke Museum of Natural History and Culture, Seattle, WA 98195

Abstract

Recent models show that tropical ectotherms may respond to increasing temperatures by undergoing range shifts, acclimatizing, or becoming locally extinct. These models often overlook standing genetic variation and assume that tropical ectotherms will not be able to adapt rapidly enough to track global climate change. I used multilocus nuclear genetic data and genome-wide SNP data to investigate phylogeographic patterns in a terrestrial ectotherm, the Puerto Rican Crested Anole (*Anolis cristatellus*), *t*o test whether existing genetic variation correlates with environmental climate layers. Phylogenetic and genetic clustering analyses identified multiple genetic populations of *A. cristatellus*. Principle component analysis of bioclimatic variables demonstrate that genetic populations occupy distinct environments. Isolation-migration modeling of nuclear sequences identifies asymmetrical gene flow from xeric to mesic habitats. Genes flowing 'out' of xeric habitats fits a pattern in which selection against 'warm-adapted' genotypes in cool habitats is weak, whereas selection against 'cool-adaptive' genotypes in warm habitats is relatively strong. Therefore, the xeric populations of *A. cristatellus* may serve as a source for adaptive standing genetic variation during future climate change as Puerto Rico becomes warmer and drier.

Introduction

Environmental variation across habitats is ubiquitous in nature, imposes physiological challenges for organisms, and limits species ranges. For widely distributed species that span climatic gradients, populations may experience divergent selective pressures between environments. The balance between strength of selection and rates of gene flow may facilitate local adaptation, potentially leading to locally adapted populations (Feder et al., 2012; Yeaman and Otto, 2011; Yeaman and Whitlock, 2011) When genetic variation is correlated with environmental parameters like temperature and precipitation, the genomic underpinnings of physiological traits may be under selection (Bay and Palumbi, 2014; Hancock et al., 2011). However, teasing apart whether genetic divergence initially occurred in allopatry followed by secondary contact and physiological adaptation, or whether selection on physiological traits along an ecological gradient is sufficient to cause bimodal distributions of genetic variation, remains an outstanding question. Furthermore, whether reinforcement-like processes can potentially lead to thermal adaptation and ecological speciation is unclear (Keller and Seehausen, 2012). Phylogeography offers tractable approaches for studying how climatic variation mediates population genetic structure and gene flow (Marko and Hart, 2011; 2012)

Tropical ectotherms are predicted to have limited ability for physiological adaptation because they experience relatively stable climates relative to their temperate counterparts, and will be forced to migrate, acclimate, or go extinct due to climate change (Deutsch et al., 2008; Huey et al., 2009; Sinervo, 2010). This has inspired research on geographic variation in physiological traits (Gunderson and Leal, 2012; Gunderson et al., 2011; Moritz et al., 2012) and rapidly evolving physiological traits (Kolbe et al., 2012; Leal and Gunderson, 2012; Logan et al., 2014). Lower-latitude species are generally "older" than temperate species, and potentially have

been resilient to climate change in the past (Cattin et al., 2016). Therefore, understanding how climate-mediated population divergence has resulted in standing genetic variation across heterogeneous environments is necessary for understanding adaptation and refining models of species extinction, persistence, and adaptation due to climate and land-use changes.

Anolis lizards have radiated within Greater Antillean islands where they have repeatedly evolved "ecomorphs" which are characterized by a shared morphology, behavior, and arboreal niche (Losos, 2009). Ernest Williams (1972) suggested that the *Anolis* adaptive radiation in Puerto Rico proceeded in stages, whereby anoles first diverged in body size, then in morphology and arboreal niche, and finally in climatic niche. Hertz et al. (2012) found bigger changes in climatic niche than in morphology among closely related species, indicating that thermal physiology is labile and is important for understanding recent speciation in *Anolis*. Furthermore, intraspecific variation in physiological traits - such as rates of cutaneous water loss - has been documented along elevational gradients (Leal and Gunderson 2012; Kolbe et al. 2012; Logan, Cox, and Calsbeek 2014), leading to the hypothesis that intraspecific physiological differences are due to adaptive genetic differentiation between populations (Gunderson et al., 2011). Recent genetic studies on *Anolis* have investigated "isolation by ecology" (Wang et al., 2013) gene flow along ecological gradients (Ogden and Thorpe, 2002; Thorpe et al., 2008), and the potential for ecological speciation (Thorpe et al., 2010; Muñoz et al., 2014), but the relationship between phylogeographic structure and thermal and hydric ecology has remained largely unexplored.

Here, I assess how climate-mediated pressures influence demographic history and potential for local adaptation of *Anolis cristatellus* on Puerto Rico. Despite being a relatively small island Puerto Rico has substantial variation in temperature, rainfall, and elevation. *Anolis cristatellus* spans this environmental variation and inhabits a range of thermal and hydric environments,

including the southwestern xeric scrub forest. Specifically, I evaluate whether population structure correlates with climatic variation, and whether gene flow is reduced between populations inhabiting different climates.

Materials and Methods

Sampling.

I generated multi-locus Sanger and genome-wide SNP data for *Anolis cristatellus* from Puerto Rico. Between 2011 – 2015 I collected 157 specimens and tissues from 27 localities across Puerto Rico (2011-IC-044; 2015-IC-005) (Figure 1). I fixed specimens in 10% buffered formalin and preserved liver tissues in liquid nitrogen in the field. Samples are accessioned at the University of Washington Burke Museum of Natural History and Culture (UWBM) in Seattle, WA.

Sanger Sequencing.

To extract DNA from frozen liver tissue, I used Qiagen DNeasy spin columns. I followed standard PCR protocols and collected nine coding- and non-coding nuclear loci and the mtDNA NADH2 region (Jackman et al., 1999) for 64 individuals of *A. cristatellus* from 16 localities across Puerto Rico (Table 1). I sequenced all loci in both directions on an ABI 3730 capillary sequencer and used the program Geneious v6.1.8 to assemble DNA contigs and MUSCLE v3.8.1 (Edgar, 2004) to generate sequence alignments. I resolved length-polymorphisms by hand in Geneious. Finally, I picked the best of five runs of PHASE v2.1 (Stephens et al., 2001) with the recombination (-MR; Stephens and Donnelly, 2003) and parent-independent (-d) models to

resolve phased alleles. For phylogenetic analysis, ambiguity codes were retained for alleles with reconstruction probabilities lower than 90%. For population structure and isolation-migration analysis, I retained the highest probability alleles. I tested for intra-genic recombination using the difference of sums of squares (DSS) test in TOPALi v2 (Milne et al., 2009). I used JModelTest v2.1.10 (Posada, 2008) and AIC to estimate and rank nucleotide substitution models for each locus.

Double-digest RAD sequencing.

Tissues were digested overnight at 37°C with 10% SDS and Proteinase K. Following digestion, I isolated and purified DNA using a 2x serapure bead clean-up protocol. I collected restriction-site associated loci for 157 individuals following the double-digest RADseq protocol (Peterson et al., 2012). I used 20 units each of the common cutter MspI (restriction site 5'-CCGG-3') and the rare cutter SbfI (restriction site 5'-CCTGCAGG-3') to digest 500 ng of DNA at 37 ℃ for 10 hours. I used serapure beads to purify samples before ligating Illumina barcoded adapters onto the DNA fragments. Following ligation, I used a Blue Pippin Prep (Sage Science) to select DNA fragments between 415-515 bp. I then used proofreading Taq and Illumina specific primers to amplify size-selected libraries before using an Agillent 2200 TapeStation to quantify and assess the distribution of size fragments of the final libraries. I used RT-qPRC to determine the sequence-able library and pooled in equimolar ratios for single-end 50-bp Illumina sequencing on either the HiSeq2500 or the HiSeq4000. Final libraries were sequenced at the QB3 Vincent Coates Genomics Sequencing Laboratory at UC Berkeley.

Stacks Bioinformatics.

I processed Illumina data with the STACKS pipeline ver. 1.44, which identifies putative loci and infers haplotypes for individuals (Catchen et al., 2013, 2011). To create putative loci and call SNPs, I implemented *ustacks* (Hohenlohe et al., 2011) and grouped reads into loci that differed by a threshold of one bp mismatch with a minimum of 10x coverage. I ran *cstacks* with a threshold allowing two bp mismatches to create a catalogue of consensus loci across individuals. Finally, I ran *sstacks* to resolve the allelic state at each locus in individuals. I used the *populations* script to resolve individual haplotypes, to retain nucleotides with a minimum allele frequency of 5%, and to generate datasets for which loci are retained if they are present in 50, 65, or 80% of individuals (hereafter, "min.ind"). I generated final datasets by removing individuals with over 50% missing loci (Table 3).

Population Genetic Structure.

I analyzed the nuclear sequences using the program STRUCTURE v2.3 (Pritchard et al., 2000) to infer the number of genetic populations (*K*) and assign individuals to those populations. STRUCTURE uses allele frequencies from multi-locus data and estimates the number of populations by minimizing Hardy-Weinberg and linkage disequilibrium between putative genetic populations. I implemented the correlated-allele frequency model, which assigns individuals to multiple populations if they are admixed (Falush et al., 2003). I tested *K*-values 1-6 and sampled over 1 million generations with the first 100k samples discarded as burn-in for 10 replicate runs at each K-value. The program ADMIXTURE also estimates the number of genetic populations (*K*) and assigns individuals to those populations, but it implements a relaxed-block maximum-likelihood algorithm that is faster and better suited for genome-wide SNP datasets. I used ADMIXTURE v1.3 (Alexander et al., 2009) and analyzed the unlinked-SNP dataset to infer *K*

and to generate maximum-likelihood estimates of individual ancestry coefficients. I tested *K*-values from 1-10 and used cross-validation to determine the optimal number of populations.

Phylogenetic Inference.

I inferred phylogenetic relationships among *Anolis cristatellus* sampling localities from nuclear sequence data and from genome-wide SNPs. To estimate phylogenetic relationships from nuclear sequence data I used *BEAST version 1.8.6 (Drummond et al., 2012). I set assigned individuals to 'species' based on their sampling localities. I unlinked gene tree topologies and used a relaxed-lognormal clock calibration. I estimated phylogenetic relationships from SNP data by computing single value decomposition scores in the program SVDQuartes (Chifman and Kubatko, 2014). SVDQuartets is a coalescent-model based program that infers unrooted phylogenies for quartets of species before agglomerating quartets into a full species tree. Because SVDQuartets makes the assumption of no gene flow between lineages I removed samples if they exhibited K=2 ancestry coefficients that were less than 90% for either population. I analyzed 1917 unlinked SNPs from 143 individuals across 25 localities. I randomly sampled 500,000 species quartets (~5% of possible quartets) and ran 1000 bootstraps to assess node support for phylogenetic relationships.

Multivariate Analysis of Bioclimatic Variables.

I downloaded the set of 19 bioclimatic variables (~1km resolution; version 1.4) from Worldclim.org (Hijmans et al., 2005). I extracted the values of bioclimatic variable for each sampling locality and used this dataset for multivariate analysis of *Anolis cristatellus* environments. To ease interpretation of the analysis and to limit the impact of redundant

variables, I reduced the dataset by retaining only one environmental variable if multiple variables were >95% correlated with each other. This resulted in a final dataset of bioclimatic variables pertaining to temperature (annual mean temperature, mean diurnal range, isothermality, temperature seasonality) and precipitation (annual precipitation, precipitation of wettest month, precipitation of driest month, precipitation seasonality). To investigate environmental differences between genetic populations of *Anolis cristatellus*, I implemented Principle Component Analysis (PCA) on the remaining eight bioclimatic variables. I used the R-package *ggplot2* to plot the results. Because admixed localities could not be assigned to populations based on genetics alone, I conducted Linear Discriminant Analysis (LDA) to assign them to populations based on environment. I used the K=2 population assignments from ADMIXTURE for localities with pure ancestry. To assess the accuracy of LDA to discriminate between northern and southern environments I trained the model by randomly selecting 70% of the localities, and I tested the model with the remaining localities. I calculated the misclassification rate across 1,000 iterations. Finally, I used the model to calculate probabilities for the assignment of admixed localities to either northern or southern environments. I used the R-package *beeswarm* to plot the LDA results.

Estimated Effective Migration Surfaces

To visualize a spatial representation of genetic clusters and *Anolis cristatellus* migration on Puerto Rico, I used the estimated effective migration surfaces (EEMS) program (Petkova et al., 2015). Clustering-based analyses can over estimate the number of genetic clusters when genetic variation is continuously distributed, such as in cases of isolation by distance (IBD; Frantz et al., 2009; Meirmans, 2012). EEMS implements a stepping-stone model to model migration rates

across a landscape and identifies corridors and barriers to gene flow by highlighting areas that deviate from a pattern of isolation by distance (Petkova et al., 2015). I averaged two runs each of 6,000,000 MCMC iterations (sampled every 10,000 steps following a 2,000,000 step burn-in) with 600 demes. I chose to use 600-demes (~9 km^2 grid size) because reflects the dense sampling design of the study.

Isolation-migration Analysis

To estimate migration rates between xeric and mesic environments I analyzed the nuclear sequence data using the program IMa2, which uses multi-locus genetic data and jointly estimates population divergence parameters such as population sizes, divergence times, and bi-directional migration rates between populations (Hey, 2010). I assigned the "admixed" localities to the mesic environment, based on the LDA analysis. I used the HKY nucleotide substitution model for each locus and ran IMa2 for 4,000,000 iterations with 1,000,000 burn-in and assessed convergence by comparing the scatter plots and by comparing results between four independent runs. I conducted joint-estimates of parameter estimates in L-mode and compared model output with AIC (Carstens et al., 2013).

Results

Double-digest RAD sequencing.

I generated approximately 250 million raw reads with an average of ~1.2 million reads per individual after filtering based on read quality and barcode match. The most liberal dataset in

terms of missing data had 1917 loci for 157 individuals (min.ind = 50%; min.loci = 90%) and the most conservative dataset had 323 loci for 153 individuals (min.ind = 80%; min.loci = 50%) (Table 3). For some of the more conservative thresholds all individuals from Rio Piedras (RIOP) were removed from the dataset.

Population Genetic Structure and Assignment.

Estimates of population genetic structure are largely consistent between Sanger and ddRAD datasets. Both datasets recover a strong population boundary separating northern and southern *Anolis cristatellus* (Figure 2). The nuclear sequence dataset and the Evanno method (Evanno et al., 2005) support four populations, and the SNP dataset and ADMIXTURE's five-fold cross-validation support five populations. The four- and five-population models both result in northern and southern populations (i.e. $K = 2$) followed by east-west population breaks within the northern and southern populations.

Phylogenetic Inference.

The phylogenetic relationships largely reflect the hierarchical population structure inferred from STRUCTURE and ADMIXTURE (Figure 2). *BEAST and SVDQuartets both recovered two major clades that correspond to north and south of the Cordillera Central mountain range. Within the southern clade, *BEAST and SVDQuartets recovered reciprocally monophyletic groups in the southwest and southeast. The topology within the northern clade does not have strong posterior support in the SVDQuartets tree, but the clades do correspond to the northern populations from the population genetic cluster analysis. In contrast to SVDQuartets, *BEAST inferred only two clades in the North, which may have resulted from limited sampling in the

nuclear DNA dataset (Figure S1).

Estimates of Effective Migration Surfaces.

Migration surfaces indicate a strong gene-flow barrier running west to east, highlighting the Cordillera Central mountain range and southern rain shadow which represent potential geographic and ecological barriers to gene flow (Figure 3). This pattern best reflects the K=2 clustering analysis from ADMIXTURE. EEMS does not find barriers to gene flow between southwestern and southeastern populations or between northern populations, and instead indicates that these populations - which were recovered in cluster analysis - do not deviate from the expectations of IBD.

Multivariate Analysis of Environmental Data.

I ran PCA to characterize the environmental variation of *A. cristatellus* populations on Puerto Rico. The first PC describes 68% of the variation and can be mainly categorized as a precipitation gradient. The second PC described 17% of the variation and was negatively correlated with mean annual temperature and positively correlated with mean diurnal temperature range and isothermality (Fig 4). The southern populations overlap with each other in PCA space and are characterized by arid and hot conditions, as well as increased daily temperature ranges and increased seasonality of precipitation. In contrast, the northern populations overlap with each other and are characterized as being more stable environments that are wetter and cooler than southern environments. The admixed localities overlap in PCA space with the northern environments. I ran LDA to explicitly test whether admixed localities are more similar to northern or southern environments. The top three loadings in the LDA are precipitation

of the driest month, mean annual precipitation, and precipitation seasonality (Fig X). The LDA assigns the "test" localities to their correct population assignment 90.5% of the time, indicating that these environmental variables can accurately discriminate between northern and southern environments. LDA assigns the admixed localities to the northern population with high confidence (mean = 97.5%, range = 87.9 - 99.9%).

Isolation-migration modeling and asymmetrical gene flow.

The top two models from the joint estimates of demographic parameters for isolation-migration modeling combine for 98% of the Akaike weights (Table 4). The top model has 4 parameters and migration from "mesic" to "xeric" is set to zero. The second-best model is a full model, in which each parameter is freely estimated, and migration from "mesic" to "xeric" is estimated to be very low. Taken together, these models indicate that gene flow is asymmetrical, with gene flow from "xeric" to "mesic" being greater than gene flow from "mesic" to "xeric" environments.

Discussion

Anolis cristatellus is represented by genetic populations that inhabit distinct climatic habitats north and south of the Cordillera Central mountain range. The habitat of the southern population lies in the rain shadow of the Cordillera Central and is warmer and drier than the northern habitat. The populations represented by finer-scale genetic clustering within northern and southern populations does not correlate with obvious environmental variation and may result from isolation by distance or from range-shifts and allopatric divergence in forest refugia during glaciation-induced Pleistocene climate change. Admixed individuals occupy moderate-elevation

habitat in a contact zone between northern and southern populations in the Carite mountains. The admixture zone is best characterized as mesic habitat (i.e. northern), indicating weak selection against xeric genotypes in mesic habitats.

The influence of the Cordillera Central

The ridge of the Cordillera Central appears to represent a geographic barrier to gene flow between the north and south *A. cristatellus* populations (Fig 3). *Anolis cristatellus* reach elevations of ~900 meters; portions of the Cordillera Central exceeding 900 meters are therefore inhospitable which reduces the opportunity for north and southern populations to come in contact with each other. Northeasterly trade-winds bring maritime air across Puerto Rico, causing increased precipitation along the northern faces of mountains and creating a rain shadow south of the Cordillera Central. Because the Cordillera Central as a geographic barrier is therefore correlated with the ecological (i.e. temperature and precipitation) gradient on Puerto Rico, it is difficult to tease apart ecological and geographic drivers of population differentiation. Still, these populations inhabit starkly different habitats and are likely exposed to different selective pressures regardless of whether geography or climate were the initial drivers of differentiation. Principle component analysis of bioclimatic variables confirm that southern population sampling localities are warmer and dried (Fig 4). Furthermore, these localities are characterized by high seasonality of precipitation, indicating that for most of the year southern populations experience even lower levels of precipitation, with increases in precipitation during rainstorms from September – November. In contrast, the northern populations are largely characterized by increased precipitation, decreased temperatures, and more stable levels of rainfall throughout the year.

Fine scale genetic clustering

Population genetic differentiation within northern or southern clades are unlikely to be driven by climatic factors because these populations largely overlap in environmental space (Fig 4). While there is a slight gradient in both precipitation and forest types within the northern clade - from tropical rainforest (Northeast), to high-elevation mesic forest (Northcentral), to mid-elevation and slightly drier karst forest (Northwest) – these differences are slight compared to the north versus south habitats. Instead, population genetic structure within the northern clade might result from isolation by distance or from allopatric divergence during population expansions and contractions. The San Loiza River Basin, located in northeast Puerto Rico, is a relatively arid lowland basin that is hypothesized to represent a temporary geographic barrier. During glacial maxima when conditions were drier, increased aridity may have caused savannah-like habitats to expand into the San Loiza River Basin, potentially separating forested habitats in the Cordillera Central and Luquillo mountains. Repeated cycles of forest expansion and retraction may have promoted population differentiation in forest taxa (Barker et al., 2012; Rodríguez-Robles et al., 2010; Velo-Antón et al., 2007).

Admixture zone at moderate elevation

Individuals at moderate elevations between Toro Negro and the Carite mountains have mixed genetic ancestry despite the localities being relatively cool and mesic. Principle component analysis of bioclimatic variables cannot distinguish the admixture localities from northeastern or northcentral habitats and linear discriminant analysis assigns the admixed localities confidently to the northern "mesic" population, indicating that admixture is caused by lizards from the xeric

habitat migrating into the mesic habitat. This pattern is corroborated by demographic parameter estimates in IMa2; the top two models support asymmetrical gene flow from the xeric population into the mesic population (Table 4).

Implications for adaptation

A primary challenge facing evolutionary physiologists is understanding the environmental factors and demographic histories that enable physiological adaptation. Genetic variation in *A. cristatellus* correlates with temperature and precipitation on Puerto Rico and offers a tractable scenario for investigating the phylogeography of physiological adaptation. In a common garden experiment *A. cristatellus* in the South had lower cutaneous water-loss rates than *A. cristatellus* from the North (Gunderson et al., 2011). Provided that North and South populations are locally adapted to their respective environments, the genetic results fit a conceptual model in which there is weak selection acting against xeric-adapted *A. cristatellus* in mesic habitats, but relatively strong selection acting against mesic-adapted *A. cristatellus* in xeric habitats. Intuitively, mesic-adapted lizards in xeric habitats will face increased desiccation stress, whereas xeric-adapted lizards should be well equipped to maintain water-balance in mesic environments. Furthermore, habitats that are similar in climatic environment might be excepted to have relatively elevated levels of gene flow because of reduced selection between populations (Logan et al., 2016). Consistent with this, migration surfaces from EEMS show migration corridors within North and South populations that overlap in climatic PCA space (Fig 3). Understanding the relationship between phylogeographic patterns and environmental variation may help improve predictions of population and species responses to climate change.

References

Alexander, DH, J Novembre, and K Lange. 2009. "Fast Model-Based Estimation of Ancestry in Unrelated Individuals." *Genome Research* 19 (9): 1655–64. https://doi.org/10.1101/gr.094052.109.

Barker, BS, JA Rodríguez-Robles, VS Aran, A Montoya, RB Waide, and JA Cook. 2012. "Sea Level, Topography and Island Diversity: Phylogeography of the Puerto Rican Red-Eyed Coquí, *Eleutherodactylus Antillensis*." *Molecular Ecology* 21 (24): 6033–52. https://doi.org/10.1111/mec.12020.

Bay, RA, and SR Palumbi. 2014. "Multilocus Adaptation Associated with Heat Resistance in Reef-Building Corals." *Current Biology* 24 (24): 2952–56. https://doi.org/10.1016/j.cub.2014.10.044.

Carstens, BC, RS Brennan, V Chua, CV Duffie, MG Harvey, RA Koch, CD McMahan, et al. 2013. "Model Selection as a Tool for Phylogeographic Inference: An Example from the Willow *Salix melanopsis*." *Molecular Ecology* 22 (15): 4014–28. https://doi.org/10.1111/mec.12347.

Catchen, JM, A Amores, P Hohenlohe, W Cresko, JH Postlethwait, and DJD Koning. 2011. "Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences." *G3 - Genes|Genomes|Genetics* 1 (3): 171–82. https://doi.org/10.1534/g3.111.000240.

Catchen, J, PA Hohenlohe, S Bassham, A Amores, and WA Cresko. 2013. “Stacks: An Analysis Tool Set for Population Genomics.” *Molecular Ecology* 22 (11): 3124–40. https://doi.org/10.1111/mec.12354.

Cattin, L, J Schuerch, N Salamin, and S Dubey. 2016. “Why Are Some Species Older than Others? A Large-Scale Study of Vertebrates.” *BMC Evolutionary Biology* 16 (1). https://doi.org/10.1186/s12862-016-0646-8.

Chifman, J, and L Kubatko. 2014. “Quartet Inference from SNP Data Under the Coalescent Model.” *Bioinformatics* 30 (23): 3317–24. https://doi.org/10.1093/bioinformatics/btu530.

Deutsch, CA, JJ Tewksbury, RB Huey, KS Sheldon, CK Ghalambor, DC Haak, and PR Martin. 2008. “Impacts of Climate Warming on Terrestrial Ectotherms across Latitude.” *Proceedings of the National Academy of Sciences* 105 (18): 6668–6672.

Drummond, AJ, MA Suchard, D Xie, and A Rambaut. 2012. “Bayesian Phylogenetics with BEAUti and the BEAST 1.7.” *Molecular Biology and Evolution* 29 (8): 1969–73. https://doi.org/10.1093/molbev/mss075.

Edgar, RC. 2004. “MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput.” *Nucleic Acids Research* 32 (5): 1792–97. https://doi.org/10.1093/nar/gkh340.

Evanno, G, S Regnaut, and J Goudet. 2005. "Detecting the Number of Clusters of Individuals Using the Software Structure: A Simulation Study." *Molecular Ecology* 14 (8): 2611–20. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

Falush, D, M Stephens, and JK Pritchard. 2003. "Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies." *Genetics* 164 (4): 1567–1587.

Feder, JL, SP Egan, and P Nosil. 2012. "The Genomics of Speciation-with-Gene-Flow." *Trends in Genetics* 28 (7): 342–50. https://doi.org/10.1016/j.tig.2012.03.009.

Frantz, AC, S Cellina, A Krier, L Schley, and T Burke. 2009. "Using Spatial Bayesian Methods to Determine the Genetic Structure of a Continuously Distributed Population: Clusters or Isolation by Distance?" *Journal of Applied Ecology* 46 (2): 493–505. https://doi.org/10.1111/j.1365-2664.2008.01606.x.

Gunderson, AR, and M Leal. 2012. "Geographic Variation in Vulnerability to Climate Warming in a Tropical Caribbean Lizard: Vulnerability of a Tropical Lizard to Climate Warming." *Functional Ecology* 26 (4): 783–93. https://doi.org/10.1111/j.1365-2435.2012.01987.x.

Gunderson, AR, J Siegel, and M Leal. 2011. "Tests of the Contribution of Acclimation to Geographic Variation in Water Loss Rates of the West Indian Lizard *Anolis cristatellus*."

Journal of Comparative Physiology B 181 (7): 965–72. https://doi.org/10.1007/s00360-011-0576-0.

Hancock, AM, B Brachi, N Faure, MW Horton, LB Jarymowycz, FG Sperone, C Toomajian, F Roux, and J Bergelson. 2011. "Adaptation to Climate across the *Arabidopsis thaliana* Genome." *Science* 334 (6052): 83–86. https://doi.org/10.1126/science.1209997.

Hey, J. 2010. "Isolation with Migration Models for More Than Two Populations." *Molecular Biology and Evolution* 27 (4): 905–20. https://doi.org/10.1093/molbev/msp296.

Hijmans, RJ, SE Cameron, JL Parra, PG Jones, and A Jarvis. 2005. "Very High Resolution Interpolated Climate Surfaces for Global Land Areas." *International Journal of Climatology* 25 (15): 1965–78. https://doi.org/10.1002/joc.1276.

Hohenlohe, PA, SJ Amish, JM Catchen, FW Allendorf, and G Luikart. 2011. "Next-Generation RAD Sequencing Identifies Thousands of SNPs for Assessing Hybridization between Rainbow and Westslope Cutthroat Trout." *Molecular Ecology Resources* 11 (March): 117–22. https://doi.org/10.1111/j.1755-0998.2010.02967.x.

Huey, RB, CA Deutsch, JJ Tewksbury, LJ Vitt, PE Hertz, HJA Perez, and T Garland. 2009. "Why Tropical Forest Lizards Are Vulnerable to Climate Warming." *Proceedings of the Royal Society B: Biological Sciences* 276 (1664): 1939–48. https://doi.org/10.1098/rspb.2008.1957.

Jackman, TR, A Larson, K deQueiroz, and JB Losos. 1999. "Phylogenetic Relationships and Tempo of Early Diversification in *Anolis* Lizards." *Systematic Biology* 48 (2): 254–285.

Keller, I, and O Seehausen. 2012. "Thermal Adaptation and Ecological Speciation." *Molecular Ecology* 21 (4): 782–99. https://doi.org/10.1111/j.1365-294X.2011.05397.x.

Kolbe, JJ, PS VanMiddlesworth, N Losin, N Dappen, and JB Losos. 2012. "Climatic Niche Shift Predicts Thermal Trait Response in One but Not Both Introductions of the Puerto Rican Lizard *Anolis Cristatellus* to Miami, Florida, USA: Thermal Trait Evolution during a Lizard Introduction." *Ecology and Evolution* 2 (7): 1503–16. https://doi.org/10.1002/ece3.263.

Leal, M, and AR Gunderson. 2012. "Rapid Change in the Thermal Tolerance of a Tropical Lizard." *The American Naturalist* 180 (6): 815–22. https://doi.org/10.1086/668077.

Logan, ML, RM Cox, and R Calsbeek. 2014. "Natural Selection on Thermal Performance in a Novel Thermal Environment." *Proceedings of the National Academy of Sciences* 111 (39): 14165–69. https://doi.org/10.1073/pnas.1404885111.

Logan, ML, MC Duryea, OR Molnar, BJ Kessler, and R Calsbeek. 2016. "Spatial Variation in Climate Mediates Gene Flow across an Island Archipelago." *Evolution* 70 (10): 2395–2403. https://doi.org/10.1111/evo.13031.

Marko, PB, and MW Hart. 2011. "The Complex Analytical Landscape of Gene Flow Inference." *Trends in Ecology & Evolution* 26 (9): 448–56. https://doi.org/10.1016/j.tree.2011.05.007.

———. 2012. "Retrospective Coalescent Methods and the Reconstruction of Metapopulation Histories in the Sea." *Evolutionary Ecology* 26 (2): 291–315. https://doi.org/10.1007/s10682-011-9467-9.

Meirmans, PG. 2012. "The Trouble with Isolation by Distance." *Molecular Ecology* 21 (12): 2839–46. https://doi.org/10.1111/j.1365-294X.2012.05578.x.

Milne, I, D Lindner, M Bayer, D Husmeier, G McGuire, DF Marshall, and F Wright. 2009. "TOPALi v2: A Rich Graphical Interface for Evolutionary Analyses of Multiple Alignments on HPC Clusters and Multi-Core Desktops." *Bioinformatics* 25 (1): 126–27. https://doi.org/10.1093/bioinformatics/btn575.

Moritz, C, G Langham, M Kearney, A Krockenberger, J VanDerWal, and S Williams. 2012. "Integrating Phylogeography and Physiology Reveals Divergence of Thermal Traits between Central and Peripheral Lineages of Tropical Rainforest Lizards." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1596): 1680–87. https://doi.org/10.1098/rstb.2012.0018.

Munoz, MM, MA Stimola, AC Algar, A Conover, AJ Rodriguez, MA Landestoy, GS Bakken, and JB Losos. 2014. "Evolutionary Stasis and Lability in Thermal Physiology in a Group of Tropical Lizards." *Proceedings of the Royal Society B: Biological Sciences* 281 (1778): 20132433–20132433. https://doi.org/10.1098/rspb.2013.2433.

Ogden, R, and RS Thorpe. 2002. "Molecular Evidence for Ecological Speciation in Tropical Habitats." *Proceedings of the National Academy of Sciences* 99 (21): 13612–13615.

Peterson, BK, JN Weber, EH Kay, HS Fisher, and HE Hoekstra. 2012. "Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species." *PLoS ONE* 7 (5): e37135. https://doi.org/10.1371/journal.pone.0037135.

Petkova, D, J Novembre, and M Stephens. 2015. "Visualizing Spatial Population Structure with Estimated Effective Migration Surfaces." *Nature Genetics* 48 (1): 94–100. https://doi.org/10.1038/ng.3464.

Posada, D. 2008. "JModelTest: Phylogenetic Model Averaging." *Molecular Biology and Evolution* 25 (7): 1253–56. https://doi.org/10.1093/molbev/msn083.

Pritchard, JK, M Stephens, and P Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155 (2): 945–959.

Rodríguez-Robles, JA, T Jezkova, and M Leal. 2010. "Climatic Stability and Genetic Divergence in the Tropical Insular Lizard *Anolis krugi*, the Puerto Rican 'Lagartijo Jardinero de La Montaña.'" *Molecular Ecology* 19 (9): 1860–76. https://doi.org/10.1111/j.1365-294X.2010.04616.x.

Sinervo, B. 2010. "Erosion of Lizard Diversity by Climate Change and Altered Thermal Niches." *Science* 328 (5980): 892–94. https://doi.org/10.1126/science.1188013.

Stephens, M, and P Donnelly. 2003. "A Comparison of Bayesian Methods for Haplotype Reconstruction from Population Genotype Data." *The American Journal of Human Genetics* 73 (5): 1162–1169.

Stephens, M, NJ Smith, and P Donnelly. 2001. "A New Statistical Method for Haplotype Reconstruction from Population Data." *The American Journal of Human Genetics* 68 (4): 978–989.

Thorpe RS, Y Surget-Groba, and H Johansson. 2008. "The Relative Importance of Ecology and Geographic Isolation for Speciation in Anoles." *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1506): 3071–81. https://doi.org/10.1098/rstb.2008.0077.

Thorpe, RS, Y Surget-Groba, and H Johansson. 2010. "Genetic Tests for Ecological and Allopatric Speciation in Anoles on an Island Archipelago." *PLoS Genetics* 6 (4): e1000929. https://doi.org/10.1371/journal.pgen.1000929.

Velo-Antón, G, PA Burrowes, RL Joglar, I Martínez-Solano, KH Beard, and G Parra-Olea. 2007. "Phylogenetic Study of *Eleutherodactylus coqui* (Anura: Leptodactylidae) Reveals Deep Genetic Fragmentation in Puerto Rico and Pinpoints Origins of Hawaiian Populations." *Molecular Phylogenetics and Evolution* 45 (2): 716–28. https://doi.org/10.1016/j.ympev.2007.06.025.

Wang, IJ, RE Glor, and JB Losos. 2013. "Quantifying the Roles of Ecology and Geography in Spatial Genetic Divergence." *Ecology Letters* 16 (2): 175–82. https://doi.org/10.1111/ele.12025.

Yeaman, S, and SP Otto. 2011. "Establishment and Maintenance of Adaptive Genetic Divergence Under Migration, Selection, and Drift." *Evolution* 65 (7): 2123–29. https://doi.org/10.1111/j.1558-5646.2011.01277.x.

Yeaman, S, and MC Whitlock. 2011. "The Genetic Architecture of Adaptation Under Migration–Selection Balance." *Evolution* 65 (7): 1897–1911. https://doi.org/10.1111/j.1558-5646.2011.01269.x.

Table 1. Sampling localities and tissue samples that were sequenced for nuDNA and for ddRAD.

Locality	Latitude	Longitude	nuDNA	ddRAD
Aibonito	18.12634	-66.28806	UWBM 8432-8435(4)	UWBM 8432-8435(4)
Carite	18.091	-66.0348	UWBM 6710-6712 (3)	UWBM 6710-6715 (6)
Carite	18.1096	-66.0714	UWBM 6809 (1)	UWBM 6809 (1)
Ciales	18.31367	-66.57156	-	UWBM 8416-4820 (5)
Corozal	18.33128	-66.33289	-	UWBM8411-8415 (5)
Guajataca	18.4194	-66.9668	UWBM 6721-6724 (4)	UWBM 6721-6724 (4)
Guanica	17.9539	-66.8491	UWBM 6689-6692 (4)	UWBM 6689-6692 (4)
Guilarte	18.14622	-66.76758	-	UWBM 7903-7910 (8)
Jobos	17.9565	-66.22343	UWBM 8426-8429 (4)	UWBM 8426-8427, 8429-8430 (4)
JuanaDiaz	18.02766	-66.51318	UWBM 7718-7721 (4)	UWBM 7718-7721, 7723-7724 (7)
Lares	18.28992	-66.85603	-	UWBM 8437-8439 (3)
Margarita	18.11366	-66.53562	UWBM 7736-7739 (4)	UWBM 7736-7742 (7)
Maricao	18.1496	-66.9935	UWBM 6694-6696, 6697(4)	UWBM 6693-6696 (4)
MataPlatano	18.4144	-66.729	UWBM 6699-6702 (4)	UWBM 6698, 6699-6702 (5)
Naguabo	18.19811	-65.71193	UWBM 8397-8399 (3)	UWBM 8397 (1)
Parguera	17.97365	-67.03517	-	UWBM 7805-7811 (7)
Ponce	17.9796	-66.6709	UWBM 6719-6720 (2)	UWBM 6719-6720 (2)
Quebradas	18.08294	-66.80804	-	UWBM 7826-7833 (8)
Rincon	18.35046	-67.26315	-	UWBM 8421-8422, 8424-8425 (4)
RioPiedras	18.4028	-66.0494	UWBM 6686- 6688 (3)	UWBM 6686- 6688 (3)
SanGerman	18.06399	-66.97513	UWBM 7788-7791 (4)	UWBM 7788-7793, 7795 (7)
SanLorenzo	18.18216	-65.94117	-	UWBM 8401-8405 (5)
ToroNegro	18.1732	-66.4926	UWBM 6703-6707 (5)	UWBM 6703-6707 (5)
Yabucoa	18.01921	-65.85684	UWBM 8406, 8408-8410 (4)	UWBM 8406-8410 (5)
Yunque	18.2966	-65.8513	UWBM 6716-6717 (2)	UWBM 6716-6717 (2)
Yunque	18.3322	-65.8579	UWBM 6718 (1)	UWBM 6718 (1)

Table 2. Summary statistics for sanger loci.

Locus	Ind[1]	bp[2]	Indels[3]	Class[4]	Ref.[5]
ALMS1	64	438	5	P	P
EXPH5	57	861	13	P	P
KIAA1549	58	720	18	P	P
KIAA2018	64	617	12	P	P
KIF24	64	473	0	P	P
NOS1	59	705	8	P	G
PPP2R5C	64	568	2	I	A
R35	64	574	0	P	G
RAPGEF2	60	229	6	P	A

[1]number of individuals.
[2]basepairs.
[3]number of individuals with heterozygous length polymorphisms.
[4]P = protein coding; I = intron.
[5]P = Portik et al. 2012; A = Alfoldi et al., 2011; G = Grummer et al. 2015

Table 3. ddRADseq summary of number of loci and number of individuals in each dataset.

min.ind[1] (%)	min.loci[2] (%)	Inds (#)	Loci (#)
50	50	143	1917
50	70	153	1917
50	90	157	1917
65	50	154	916
65	70	156	916
65	90	157	916
80	50	153	323
80	70	157	323

[1]the percentage of individuals required to retain a locus in the dataset
[2]the percentage of loci required to retain an individual in the dataset

Table 4. Joint estimates of demographic parameters for migration-isolation model in IMa2.

Model	Log(P)	K	nAc(q)	sAc(q)	anc(q)	nAc>sAc	sAc>nAc	AIC	ΔAIC	ω
m1, m2=0	0.6424	4	19.68	7.43	4.12	0.43	[0.00]	6.71	0	0.62
m1, m2	1.115	5	20.06	7.03	3.78	0.37	0.046	7.77	-1.05	0.36
m1=m2	-2.853	4	21.51	5.94	2.84	0.25	[0.25]	13.70	-6.99	0.02
m1=0, m2	-871.2	4	19.63	6.64	3.80	[0.00]	0.063	1750.4	-1743.6	0.0
m1=m2=0	-994.8	3	19.63	6.64	3.80	[0.00]	[0.00]	1995.6	-1988.8	0.0

'Model' describes the migration/population model implemented in IMa2.

K = number of migration and population size parameters. Population size parameters were independent in each model.

ω = akiake weights. Can be interpreted as the probability of the model.

nAc = northern *A. cristatellus*

sAc = southern *A. cristatellus*

anc = ancestral population

q = population size

m = migration

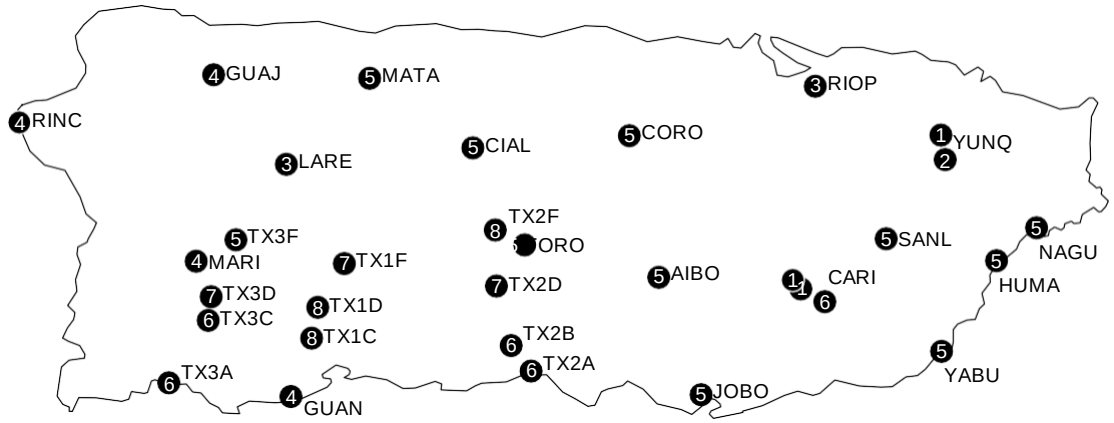


Figure 1. *Anolis cristatellus* sampling localities on Puerto Rico. Numbers represent the number of individuals sequenced by ddRADseq at each locality. Labels are as follows:
AIBO: Aibonito, CARI: Carite, CIAL: Ciales, CORO: Corozol, GUAJ: Guajataca, GUAN: Guanica, HUMA: Humacao, JOBO: Jobos, LARE: Lares, MATA: Mata de Platano, NAGU: Naguaba, RIOP: Rio Piedras, SANL: San Lorenzo, TORO: Toro Negro, YABU: Yabucoa, YUNQ: El Yunque,

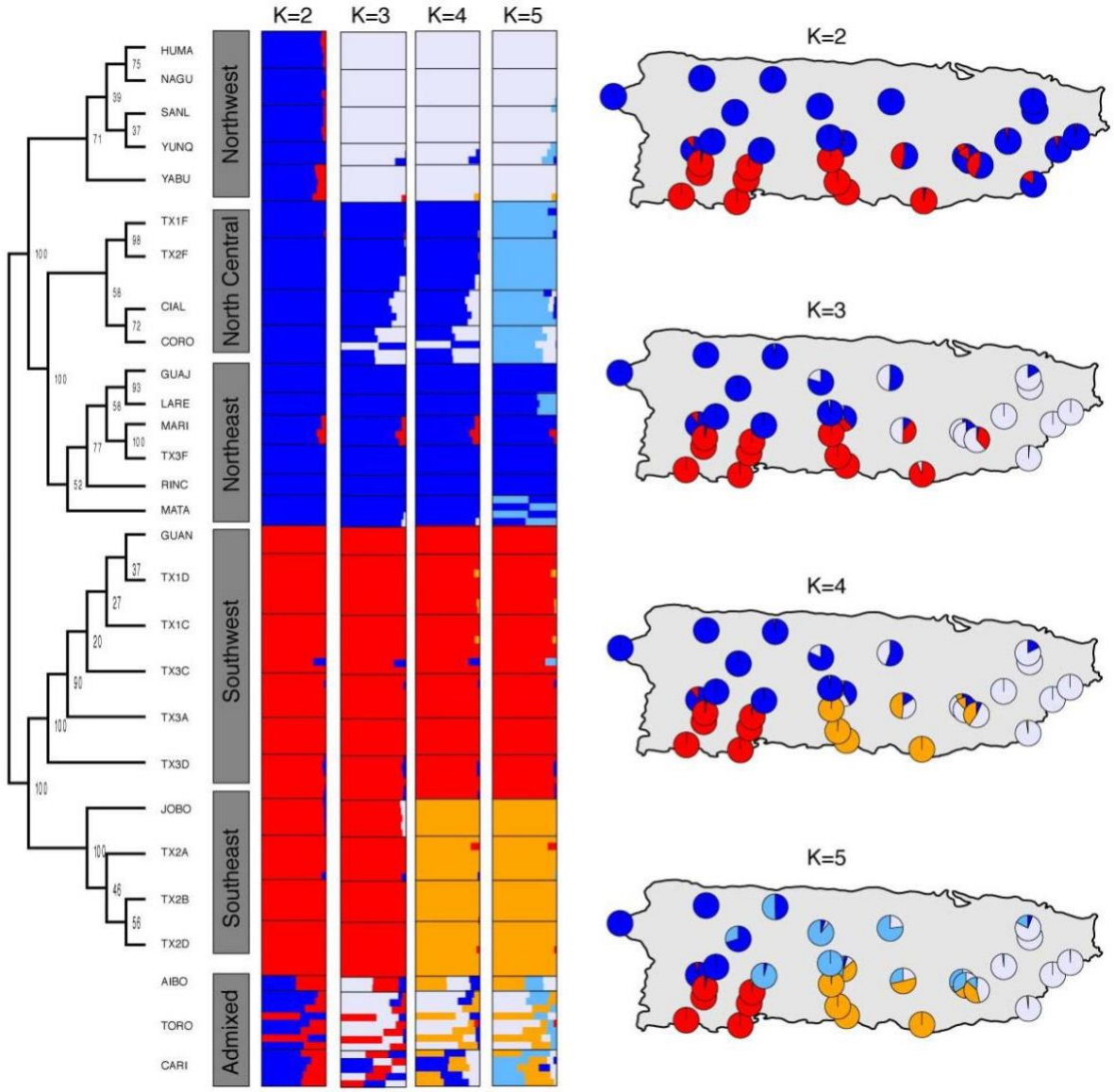


Figure 2. Phylogeographic structure in *Anolis cristatellus*. (A) Phylogenetic reconstruction from SVDQuartets. (B) Individual ancestry coefficients estimated from ADMIXTURE (K=2-5). (C) Ancestry coefficients for sampling localities mapped on Puerto Rico (K=2-5).

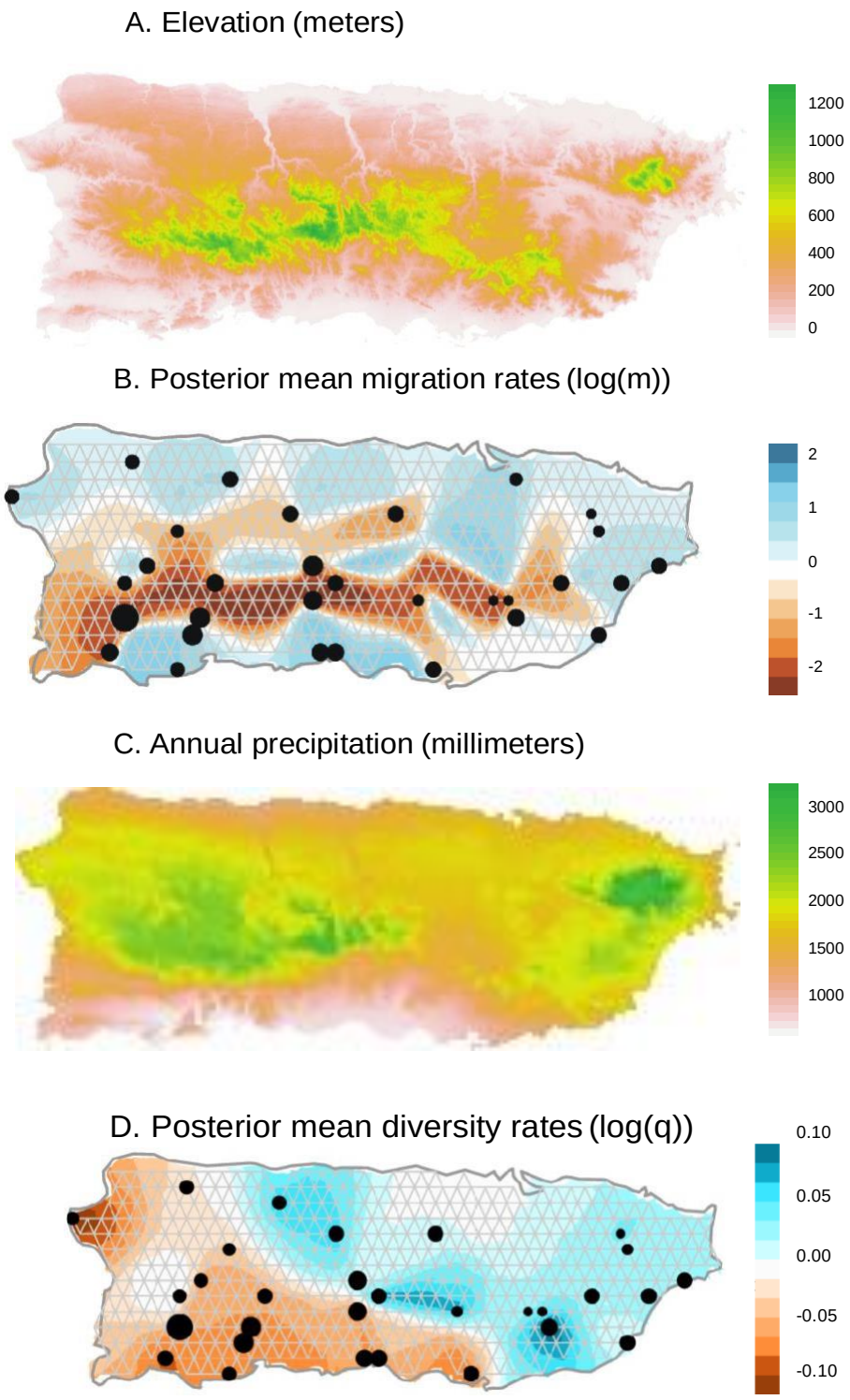


Figure 3. Effective migration surfaces. (A) Elevation on Puerto Rico. (B) Effective migration surfaces estimated in EEMS. (C) Annual precipitation on Puerto Rico. (D) Genetic diversity estimated in EEMS

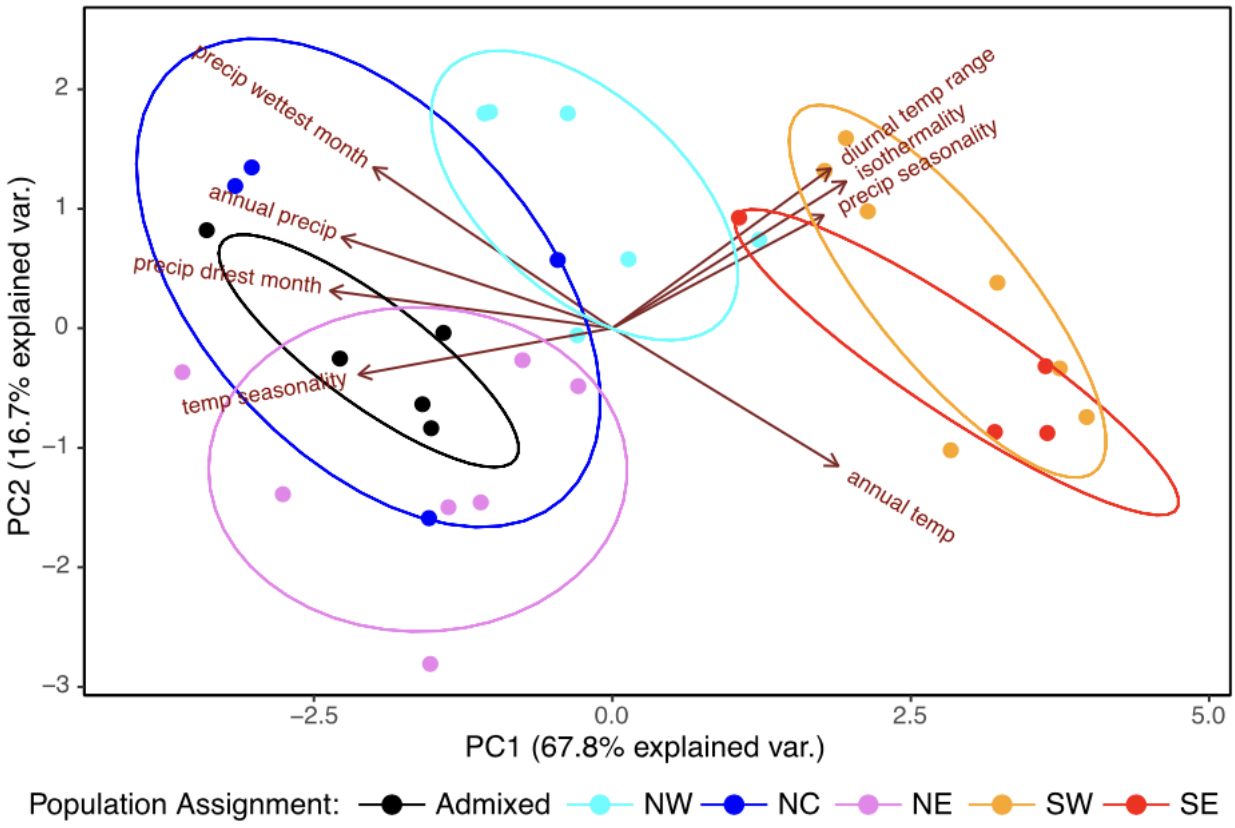


Figure 4. Principle Components of environmental variation. The legend indicates population assignment based on K=5 clustering (Fig 2). "Admixed" localities are those that are admixed between North and South. Arrows and labels indicate the direction and magnitude of the loadings for each bioclimatic variable.

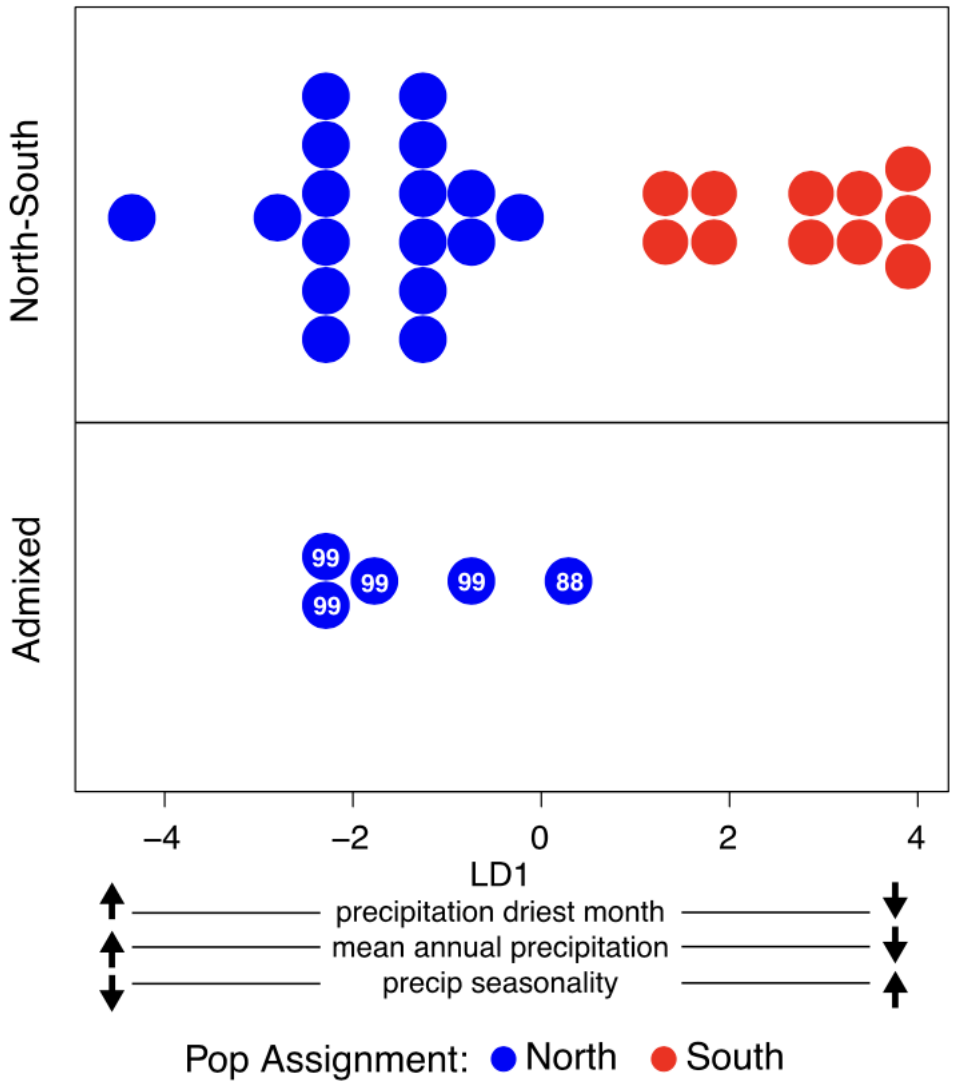


Figure 5. Linear Discriminant Analysis plot. The legend indicates population assignment based on K=2 clustering. Admixed localities are colored by their population assignment, with the number indicating the posterior probability of assignment. The three bioclimatic variables with the highest loadings on LD1 are listed at the bottom, with arrows indicating the direction of the gradient.

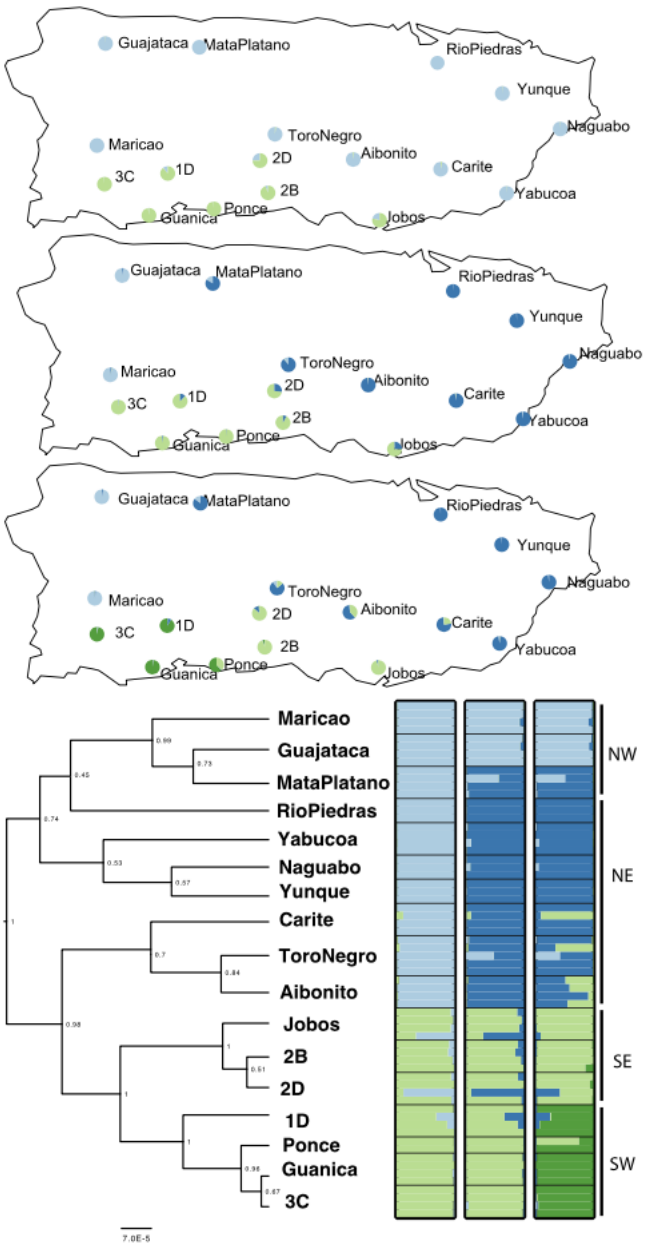


Figure S1. Phylogeography of nuclear sequences. Population genetic structure (K= 2-5) and species tree analysis of nuclear sequence data.

Chapter 3: Comparative phylogeography of co-distributed anoles on Puerto Rico

Matthew T. McElroy[12]

[1]Department of Biology, University of Washington, Seattle, WA 98195

[2]Burke Museum of Natural History and Culture, Seattle, WA 98195

Abstract

Comparative phylogeographic approaches identify geographic and ecological features that generate and maintain genetic diversity in co-distributed taxa. I analyzed genome-wide SNP data for three species of widespread Puerto Rican anoles to test the relative contributions of geography and ecology in driving population divergence. While only one species (*Anolis cristatellus*) had phylogeographic structure associated with the xeric-mesic transition in the Southwest, all three (*A. cristatellus*, *A. pulchellus*, and *A. stratulus*) exhibited concordance in geographic population breaks across the San Loiza Basin in eastern Puerto Rico. These results are consistent with phylogeographic work on other Puerto Rican vertebrates implicating the San Loiza Basin as an important geographic feature driving population divergence across diverse ecological groups on Puerto Rico.

Introduction

Understanding how geographic and ecological features influence genetic drift, natural selection, and ultimately population divergence is a primary goal of phylogeography. In a comparative framework, investigating co-distributed taxa illuminates to what extent communities of organisms experienced concerted or idiosyncratic responses to historical climatic events and

geographic features (Knowles, 2009; Hickerson et al., 2010; Avise et al., 2016). For instance, the closing of the Isthmus of Panama instigated allopatric divergence between Caribbean and Atlantic fauna (Hickerson et al., 2006), climatic fluctuations in the Australian Wet Tropics caused forests to retract and expand which created periods of allopatry and secondary contact for forest vertebrates (Schneider et al., 1998; Moritz et al. 2009; Bell et al. 2010), and cycles of sea-level inundation promoted vicariant divergence in the Philippines (Brown et al. 2013; Oaks et al. 2013).

An emerging pattern in phylogeography is that co-distributed taxa often have unique spatial genetic patterns (Moritz et al. 2009; Bagley & Johnson 2014), or, those taxa that do have concordant spatial genetic patterns have idiosyncratic demographic histories (Dolman & Joseph 2012; d'Horta et al. 2013; Demos et al. 2014). Species-specific evolutionary responses can be driven by phenotypic traits, plasticity, and differing dispersal abilities (Zamudio et al., 2016). In the tropics, idiosyncratic evolutionary histories of co-distributed taxa may be more prominent due to longer evolutionary time (Cattin et al. 2016) as well as increased biotic interactions. If idiosyncratic population and demographic histories in tropical communities are the norm, then more studies are needed to develop theories for species-specific responses to geological and climatic events (Prates et al. 2016; Zamudio et al., 2016).

The Greater Antilles – comprised of Cuba, Hispaniola, Puerto Rico, and Jamaica – have a dynamic paleo-geographic history that influenced the Caribbean biodiversity (Hedges et al., 1992; Nicholson et al. 2005; Tucker et al. 2017). Understanding how sea-level fluctuations influenced the paleo-geography and paleo-ecology of the Caribbean provides important context

for evaluating the processes that drive population differentiation and speciation. The paleo-geography of the Greater Antilles remained relatively dynamic until the mid-Miocene, at which point the islands were largely in their current configuration (Iturralde-Vinent & MacPhee 1999; Graham 2003; MacPhee et al., 2003). However, sea-level fluctuations due to glacial cycles continued to influence shorelines and island areas, exposing land-bridges between local offshore islands during glacial maxima, and causing marine incursions during inter-glacial periods. The locations of paleo-island sutures, as well as other lowland basins on Cuba and Hispaniola, represent biogeographical boundaries today in part because of repeated marine incursions and over-water barriers that generated and reinforced population divergence (Glor & Warren 2011; Glor et al. 2004; Gifford et al. 2004). In contrast to Cuba and Hispaniola, the interior of Puerto Rico has not been inundated since its emergence, which means that over-water barriers cannot explain intra-island population divergence (MacPhee et al., 2003). However, periods of low sea levels correlated with increased global aridity, which in the tropics caused savannah-like habitats to expand and forested habitats to retract. Consequently, Puerto Rican forests likely retreated up mountains and formed disjunct forest patches that were separated by increasingly arid habitat (Renken et al., 2002). For montane and mesic-adapted species, the expansion of arid habitats into the San Loiza River Basin may have resulted in a temporary dispersal barrier and caused population divergence between populations in the Cordillera Central Mountain Range and the Luquillo Mountains (Velo-Antón et al., 2007; Rodríguez-Robles et al., 2010; Barker et al., 2012).

Despite being the smallest island in the Greater Antilles, Puerto Rico is topographically and climatically diverse enabling researchers to test both geographic and ecological features that

generate phylogeographic diversity. The tropical climate has little seasonal variation in temperature, and the island is primarily blanketed with lush mesic and tropical rainforest. The Cordillera Central (1,338 m) represents the primary mountain range on Puerto Rico and bisects the island into northern upland plains and southern coastal lowlands (Figure 1). The Luquillo (1,070 m) and Carite (1,000 m) mountain ranges create topographic relief in the Northeast and Southeast, respectively. Northeasterly trade winds pass humid maritime air over the island releasing precipitation on the north and northeastern faces of the Luquillo Mountains and Cordillera Central, creating a rain shadow along the southern coast of Puerto Rico (Daly et al., 2003; Murphy et al. 2017). In stark contrast to the tropical rainforests of Luquillo and mesic forests of the Cordillera Central, the scrubby dry forest of southwest Puerto Rico supports endemic xerophilic flora and fauna (Genet et al., 2001) (Figure 1).

To test the hypothesis that the San Loiza River Basin and the southwest rain shadow represent ecological features that promote and maintain genetic divergence on Puerto Rico, I investigated the comparative phylogeography of three widespread and co-distributed *Anolis* lizards. *Anolis cristatellus*, *A. pulchellus*, and *A. stratulus* are all within the *cristatellus* group, but they are not sister species. They are all heliothermic and relatively warm adapted (Rand, 1964; Heatwole et al., 1969; Huey and Webster, 1976; Henderson and Powell, 2009; Gunderson and Leal, 2012) but they have distinct arboreal habitat preferences (Schoener, 1971). *Anolis cristatellus* uses tree-trunks, *A. stratulus* utilizes tree canopies, while *A. pulchellus* is found in bushes and high grasses (Losos, 2009). Each species is widely distributed across Puerto Rico, including both sides of the San Loiza River Basin as well as in the arid southwest coast and elevations from sea-level to approximately 900 m. I collected multiple individuals per species

from localities across their range and generated genome-wide SNP data. I then inferred phylogenies and assessed population genetic structure to determine whether phylogeographic patterns across species are concordant, or whether they reflect species-specific responses to ecological and geographic features of Puerto Rico.

Materials and Methods

Sampling & Labwork

I collected tissues and voucher specimens of *Anolis cristatellus* (n=157), *Anolis pulchellus* (n=105), and *Anolis stratulus* (n=94) during field work between 2011—2015 (2011-IC-044; 2015-IC-005). I fixed specimens in 10% buffered formalin and preserved liver tissues in liquid nitrogen in the field. All samples are accessioned at the University of Washington Burke Museum of Natural History and Culture (UWBM) in Seattle, WA.

ddRAD sequencing

I used the double-digest RADseq protocol (Peterson et al. 2012) to generate genome-wide SNP data. Tissues were digested overnight at 37°C with 10% SDS and Proteinase K. Following digestion, I used a 2x Serapure bead clean-up protocol to isolate and purify DNA. I collected restriction-site associated loci for 356 individuals following the double-digest RADseq protocol (Peterson et al. 2012). I used 20 units each of the common cutter MspI (restriction site 5'-CCGG-3') and the rare cutter SbfI (restriction site 5'-CCTGCAGG-3') to digest 500 ng of DNA at 37°C for 10 hours. I purified the samples using Serapure speedbeads before ligating Illumina barcoded adapters onto the DNA fragments. Following ligation, I selected DNA fragments

between 415-515 bp using a Pippin Prep (Sage Science) and then used proofreading Taq polymerase and Illumina specific primers to amplify size-selected libraries. I used an Agilent 2200 TapeStation to quantify and assess the distribution of size fragments of the final libraries. Final libraries were quantified using RT-qPCR at the QB3 Vincent Coates Genomics Sequencing Laboratory (UC Berkeley) and pooled in equimolar ratios for single-end 50-bp Illumina sequencing on two lanes of the HiSeq2500 and one lane of HiSeq4000.

Stacks Bioinformatics

I processed Illumina data with the STACKS pipeline ver. 1.44, which identifies putative loci and infers haplotypes for individuals (Catchen et al., 2011; Catchen et al., 2013). To create putative loci and call SNPs, I implemented *ustacks* (Hohenlohe et al., 2011) and grouped reads into loci that differed by a threshold of one basepair (bp) mismatch with a minimum of 10x coverage. I ran *cstacks* with a threshold allowing two bp mismatches to create a catalogue of consensus loci across individuals. Finally, I ran *sstacks* to resolve the allelic state at each locus in individuals. I ran the populations script to resolve individual haplotypes, to retain nucleotides with a minor allele frequency (maf) of 5%, and to generate datasets for which loci are retained if they are present in 50% of the individuals. I then generated final datasets by removing individuals that were missing 50% of the loci (Tables 2-4).

Population Genetic Structure

I used the program ADMIXTURE v1.3 to analyze the putative unlinked-SNP data and inferred the number of genetic populations (K) and generated maximum likelihood estimates of individual ancestry coefficients (Alexander et al., 2009). I tested multiple K-values (1-10) and

used cross validation to assess the best population model. I evaluated the sensitivity of population assignment to levels of missing data by running ADMIXTURE on datasets with different levels of missing data. Results were largely insensitive to levels of missing data (supplement), so I used the data set with the most SNPs and that had the most complete data per individual (min.ind=50, min.loci=90).

Phylogenetic Inference

I inferred phylogenetic relationships for sampling localities of each species from genome-wide SNPs using the program SVDQuartets (Chifman and Kubatko, 2014). SVDQuartets is a coalescent-model based program that infers unrooted phylogenies for quartets of samples before combining quartets into a species tree. I assigned individuals to their sampling locality, and estimated a phylogeny for each species from the putative unlinked-SNPs. I randomly sampled 500,000 species quartets and ran 1000 bootstraps to assess node support for phylogenetic relationships.

Estimated Effective Migration Surfaces

To visualize a spatial representation of genetic clusters and migration of the three widespread anoles on Puerto Rico, I used the estimated effective migration surfaces (EEMS) program (Petkova et al., 2015). Clustering-based analyses can over estimate the number of genetic clusters when genetic variation is continuously distributed, such as in cases of isolation by distance (Frantz et al., 2009; Meirmans, 2012). EEMS implements a stepping-stone model to model migration rates across a landscape and identifies corridors and barriers to gene flow by highlighting areas that deviate from a pattern of isolation by distance (Petkova et al., 2015).

Effective migration surfaces for each species to identify genetic corridors and barriers, and to determine whether cluster-based analyses over-estimate genetic clusters due to IBD. I averaged two runs each of 6,000,000 MCMC iterations (sampled every 10,000 steps following a 2,000,000 step burn-in) with 600 demes. I chose to use 600-demes (~9 km^2 grid size) to reflect the dense sampling design of the study.

Results

Labwork & Bioinformatics

I generated approximately 570 million sequence reads (~1.6 million reads per individual) after filtering based on read quality and barcode match. Datasets with higher levels of missing data contained more loci compared to datasets that allowed fewer missing data. Roughly 10% of individuals (14 A. cristatellus, 10 *A. pulchellus*, 15 *A. stratulus*) did not have data for at least half of the loci and were excluded from the dataset.

Population Genetic Structure

Species differ in both their number of inferred genetics clusters as well as in their distributions of genetic variation. ADMIXTURE cross-validation indicates that *A. cristatellus* is best represented by K=5, while *A. pulchellus* and *A. stratulus* are best represented by K=3 (Figure 2B,E,H). The most predominant population break in *A. cristatellus* delineates northern and southern populations and highlights the central mountains as a putative population barrier (Figures 1&2B). The northern population is subsequently represented by three smaller populations that

span Puerto Rico from east to west, and the southern population is represented by two smaller populations in southwest and in southcentral Puerto Rico. In contrast, *A. pulchellus* and *A. stratulus* are each represented by three genetic populations (Fig 2). The populations in eastern and central Puerto Rico span the San Loiza River Basin while the populations in central and western Puerto Rico appear to correlate with a transition to Karst forests in the West.

Phylogeographic Relationships

Phylogenetic trees estimated using SVDQuartets predominantly correspond to the population genetic groups inferred from ADMIXTURE (Figure 2). *Anolis cristatellus* is represented by two highly supported clades in the North and South. The three northern clades correspond to the three northern populations identified in ADMIXTURE but their relationships are not well supported in the phylogeny, likely due to gene flow between northern populations. Similarly, the two southern clades are well supported in the phylogeny and correspond to the populations inferred in ADMIXTURE. *Anolis pulchellus* and *A. stratulus* are both represented by well-supported east and west clades that correspond to the population genetic groups inferred from ADMIXTURE. In *A. pulchellus*, the central population forms a clade that is closely related to the eastern population, while in *A. stratulus* the central localities are more closely related to the western clade (Figure 2).

Effective Estimates of Migration

In general, estimates of migration barriers reflect the primary (i.e. K=2) population genetic structure for each species. However, mountain ranges also represent barriers to gene flow for each species (Figures 2C,F,I). This is especially evident in central Puerto Rico, where the Cordillera Central reaches elevations above 1,000 m (Figure 1). *Anolis pulchellus* and *A.*

stratulus localities in the West do not display sharp north-south barriers to gene flow like those seen in *A. cristatellus*.

Effective Estimates of Genetic Diversity

Despite exhibiting different patterns of population genetic structure and migration, estimates of genetic diversity for each species show similar patterns. Each species has increased genetic diversity in eastern Puerto Rico and reduced genetic diversity in the West (Figure 4). In *A. pulchellus*, low genetic diversity correlates with the transition to westernmost population. In *A. stratulus* populations west of the San Loiza River Basin have low genetic diversity. Populations with low genetic diversity are present across the deep north-south split in *A. cristatellus*, with populations in the Southeast, Southwest, and Northwest all being characterized by low genetic diversity.

Discussion

Co-distributed anoles on Puerto Rico have idiosyncratic phylogeographic histories. *Anolis cristatellus*, *A. pulchellus*, and *A. stratulus* are continuously distributed across Puerto Rico, yet each species has a phylogeographic break separating eastern and western populations. Shared allopatric population divergence across the San Loiza River Basin may be explained by periods of increased aridity during which forests likely retracted from the lowland valley. However, following population divergence and the re-forestation of the Puerto Rican lowlands populations expanded in species-specific manners. Furthermore, while *A. cristatellus* is represented by three genetic populations in the north that are analogous to the eastern, central, and western populations of *A. pulchellus* and *A. stratulus*, *A. cristatellus'* deepest phylogeographic split

distinguishes populations north and south of the Cordillera Central range. This phylogeographic break is almost certainly dictated by the ecological transition to the rain shadow habitat in the South, and yet it is neither represented in the phylogeography of *A. pulchellus* nor *A. stratulus*, perhaps because *A. cristatellus* is restricted to open habitats at elevation as it more sensitive to low temperatures. Finally, low genetic diversity in western Puerto Rico for each species may be explained by climate-driven population bottlenecks and expansions in western populations of anoles. Thus, paleoclimate cycles likely caused allopatric population divergence as well of population size fluctuations in *A. cristatellus*, *A. pulchellus*, and *A. stratulus*.

Population Divergence and the San Loiza River Basin

Glacial maxima during the Quaternary caused decreased sea-level, cooler temperatures, and increased global aridity. The forest refugia hypothesis is commonly invoked to explain tropical diversification, and states that increased aridity caused savannah-like habitats to expand and forested habitats to retract, which promoted allopatric divergence in forest refuges. The San Loiza River Basin (SLRB) in Puerto Rico is a relatively arid lowland basin that separates the Cordillera Central from the Luquillo and Carite Mountains. The SLRB is hypothesized to have caused population divergence in montane (*Eleutherodactylus coqui*) and widespread anurans (*E. antillensis*), as well as a montane lizard (*Anolis krugi*) on Puerto Rico (Velo-Antón et al., 2007; Rodríguez-Robles et al., 2010; Barker et al., 2012). The basin was likely more arid during glacial maxima, which would have caused populations to retract along with the forests into the mountains of the east and west (Renken et al., 2002). Our phylogenetic trees indicate two different models for how the SLRB may have caused population divergence: divergence and population expansion, and repeated cycles of refugia.

The deepest phylogenetic split in *A. stratulus* is between the eastern population and the central and western populations. The western population renders the central population paraphyletic (Figure 2), and appears to represent population expansion from central Puerto Rico, which is supported by the low genetic diversity in the west (Figure 3) and the hypothesis that the arid Southwest was even more xeric in the past. This pattern is explained by allopatric divergence in forest refugia on either side of the SLRB, followed by population expansion into the West following forest expansion and more mesic conditions.

An alternative scenario, exemplified by *A. cristatellus* and *A. pulchellus*, is characterized by repeated cycles of forest refugia and population divergence. Imagine a widespread population retracting into refugia on either side of a valley, diverging, and remaining distinct upon secondary contact. If this species experiences another refugia scenario, we would expect another bifurcation, but the topology would reflect how the populations inhabited the refugia (Figure 4). This is precisely what we see in *A. pulchellus* and in the northern group of *A. cristatellus* (i.e. second divergence – model 1; Figure 4). Both species exhibit a deep split between the western population and a shallower split between central and eastern populations that span the San Loiza River Basin.

Paleo-climate induced population bottleneck and expansion

Despite being relatively arid adapted, *A. cristatellus*, *A. pulchellus*, and *A. stratulus* have phylogeographic breaks at the SLRB, similar to species that are more sensitive to water loss (*A. krugi*, *E. coqui*, *E. antillensis*). *Anolis cristatellus*, *A. pulchellus*, and *A. stratulus* are all relatively arid-adapted and heat tolerant, and they are found in the xeric southwest of Puerto Rico. Conditions would have to have been significantly drier for the SLRB to represent an

effective dispersal barrier for these species. One of the best indications that conditions were substantially drier in the past come from the fossil record on Puerto Rico which includes numerous examples of widespread obligate xerophiles. For instance, burrowing owls, once found widespread are now extinct on Puerto Rico indicating a reduction in grassland and xeric habitats (Pregill and Olson, 1981). A similar case can be made for the Bahaman Mocking Bird. Currently the species lives on the Bahamas and isolated xeric habitats of Cuba and Jamaica, though it is known from the Puerto Rican fossil record (Pregill and Olson, 1981). In the Bahamas currently, the species is more abundant in the drier southern islands. Evidence of widespread xeric conditions are also inferred from the Pleistocene distributions of herpetofauna. The curly-tailed lizards of the genus *Leiocephalus* are found in the driest parts of Cuba, Hispaniola, and the Bahamas, yet fossils of extinct species from Puerto Rico and Jamaica indicate a much wider distribution, particularly in the Pleistocene (Pregill and Olson, 1981).

Despite being represented by longstanding phylogenetic lineages, populations in the West have low genetic diversity, providing further evidence that climate cycles influenced the genetic diversity of Puerto Rican anoles. Low genetic diversity in the West can be explained by western populations undergoing bottleneck and expansion and is supported by two lines of evidence. First, if the western population expanded from the eastern population we would expect the western localities to have short branch lengths and to render the eastern group paraphyletic. Instead we find that east and west populations are monophyletic and deeply divergent. Second, the distribution of low genetic diversity in the EEMS plots correlates with the spatial structure inferred from ADMIXTURE. If western populations diverged prior to going through a population bottleneck and expansion into their current range we would expect low genetic diversity to be correlated with spatial genetic structure. Finally, *A. cristatellus* has low genetic

diversity in the southern populations as well, which is difficult to reconcile with a model of recent colonization given the deep phylogenetic split of the southern lineage. During glacial maxima the sea-level was lower as well; the combination of increased aridity with increased habitat may have benefitted the xeric-adapted congeners *A. cooki* and *A. poncensis*, increasing competition, and causing population bottlenecks in *A. cristatellus*, *A. pulchellus*, and *A. stratulus*. Thus, climate cycles may have promoted allopatric divergence, as well as shaped genetic diversity following population divergence.

Species-specific evolutionary histories

The main phylogeographic split in *Anolis cristatellus* separates northern and southern populations. The population boundary in the South correlates with the xeric rain shadow on Puerto Rico, where *A. cristatellus* experiences drier and warmer conditions than elsewhere on the island (McElroy, Ch2). In contrast, *A. pulchellus* and *A. stratulus* do not demonstrate phylogeographic structure that relates to the rain shadow. Instead they have east-west population structure that may originate from forest refugia during Pleistocene glacial maxima. That *A. cristatellus* has substantially more phylogeographic structure than its co-distributed congeners may indicate that it has experienced more habitat fragmentation during climate change, or that its fragmented populations were more likely to persist than fragmented populations of *A. pulchellus* and *A. stratulus*. Given that it inhabits grasses and bushes, one might expect *A. pulchellus* to have different spatial and demographic patterns than *A. cristatellus* and *A. stratulus*, which are both tree inhabitants. However, the evolutionary histories of *A. cristatellus* and *A. stratulus* are surprisingly different. In conclusion, despite finding support that population divergence of co-

distributed anoles on Puerto Rico has been influenced by the San Loiza River Basin, each species has a unique evolutionary and demographic history.

References

Alexander, DH, J Novembre, and K Lange. 2009. "Fast Model-Based Estimation of Ancestry in Unrelated Individuals." *Genome Research* 19 (9): 1655–64. https://doi.org/10.1101/gr.094052.109.

Avise, JC, BW Bowen, and FJ Ayala. 2016. "In the Light of Evolution X: Comparative Phylogeography." *Proceedings of the National Academy of Sciences* 113 (29): 7957–61. https://doi.org/10.1073/pnas.1604338113.

Bagley, JC, and JB Johnson. 2014. "Testing for Shared Biogeographic History in the Lower Central American Freshwater Fish Assemblage Using Comparative Phylogeography: Concerted, Independent, or Multiple Evolutionary Responses?" *Ecology and Evolution* 4 (9): 1686–1705. https://doi.org/10.1002/ece3.1058.

Barker, BS, JA Rodríguez-Robles, VS Aran, A Montoya, RB Waide, and JA Cook. 2012. "Sea Level, Topography and Island Diversity: Phylogeography of the Puerto Rican Red-Eyed Coquí, *Eleutherodactylus antillensis*." *Molecular Ecology* 21 (24): 6033–52. https://doi.org/10.1111/mec.12020.

Bell, RC, JL Parra, M Tonione, CJ Hoskin, JB Mackenzie, SE Williams, and C Moritz. 2010. "Patterns of Persistence and Isolation Indicate Resilience to Climate Change in Montane Rainforest Lizards." *Molecular Ecology* 19: 2531-2544. https://doi.org/10.1111/j.1365-294X.2010.04676.x.

Brown, RM, CD Siler, CH Oliveros, JA Esselstyn, AC Diesmos, PA Hosner, CW Linkem, et al. 2013. "Evolutionary Processes of Diversification in a Model Island Archipelago." *Annual Review of*

Ecology, Evolution, and Systematics 44 (1): 411–35. https://doi.org/10.1146/annurev-ecolsys-110411-160323.

Catchen, JM, A Amores, PA Hohenlohe, W Cresko, JH Postlethwait, and DJ De Koning. 2011. "Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences." *G3 - Genes|Genomes|Genetics* 1 (3): 171–82. https://doi.org/10.1534/g3.111.000240.

Catchen, JM, PA Hohenlohe, S Bassham, A Amores, and WA Cresko. 2013. "Stacks: An Analysis Tool Set for Population Genomics." *Molecular Ecology* 22 (11): 3124–40. https://doi.org/10.1111/mec.12354.

Cattin, L, J Schuerch, N Salamin, and S Dubey. 2016. "Why Are Some Species Older than Others? A Large-Scale Study of Vertebrates." *BMC Evolutionary Biology* 16 (1). https://doi.org/10.1186/s12862-016-0646-8.

Chifman, J, and L Kubatko. 2014. "Quartet Inference from SNP Data Under the Coalescent Model." *Bioinformatics* 30 (23): 3317–24. https://doi.org/10.1093/bioinformatics/btu530.

Daly, C, EH Helmer, and M Quiñones. 2003. "Mapping the Climate of Puerto Rico, Vieques and Culebra." *International Journal of Climatology* 23 (11): 1359–81. https://doi.org/10.1002/joc.937.

Demos, TC, JCK Peterhans, B Agwanda, and MJ Hickerson. 2014. "Uncovering Cryptic Diversity and Refugial Persistence among Small Mammal Lineages across the Eastern Afromontane

Biodiversity Hotspot." *Molecular Phylogenetics and Evolution* 71 (February): 41–54. https://doi.org/10.1016/j.ympev.2013.10.014.

Dolman, G, and L Joseph. 2012. "A Species Assemblage Approach to Comparative Phylogeography of Birds in Southern Australia: Divergence in Southern Australian Birds." *Ecology and Evolution* 2 (2): 354–69. https://doi.org/10.1002/ece3.87.

Frantz, AC, S Cellina, A Krier, L Schley, and T Burke. 2009. "Using Spatial Bayesian Methods to Determine the Genetic Structure of a Continuously Distributed Population: Clusters or Isolation by Distance?" *Journal of Applied Ecology* 46 (2): 493–505. https://doi.org/10.1111/j.1365-2664.2008.01606.x.

Genet, KS, JA Genet, TM Burton, and PG Murphy. 2001. "The Lizard Community of a Subtropical Dry Forest: Guanica Forest, Puerto Rico." *Tropical Ecology* 42 (1): 97–109.

Gifford, ME, R Powell, A Larson, and RL Gutberlet. 2004. "Population Structure and History of a Phenotypically Variable Teiid Lizard (*Ameiva chrysolaema*) from Hispaniola: The Influence of a Geologically Complex Island." *Molecular Phylogenetics and Evolution* 32 (3): 735–48. https://doi.org/10.1016/j.ympev.2004.04.003.

Glor, RE, ME Gifford, A Larson, JB Losos, LR Schettino, ARC Lara, and TR Jackman. 2004. "Partial Island Submergence and Speciation in an Adaptive Radiation: A Multilocus Analysis of the Cuban Green Anoles." *Proceedings of the Royal Society B: Biological Sciences* 271 (1554): 2257–65. https://doi.org/10.1098/rspb.2004.2819.

Glor, RE, and D Warren. 2011. "Testing Ecological Explanations For Biogeographic Boundaries: Ecology and Biogeographic Boundaries." *Evolution* 65 (3): 673–83. https://doi.org/10.1111/j.1558-5646.2010.01177.x.

Graham, A. 2003. "Geohistory Models and Cenozoic Paleoenvironments of the Caribbean Region." *Systematic Botany* 28 (2): 378–86.

Gunderson, AR, and M Leal. 2012. "Geographic Variation in Vulnerability to Climate Warming in a Tropical Caribbean Lizard." *Functional Ecology* 26 (4): 783–93. https://doi.org/10.1111/j.1365-2435.2012.01987.x.

Heatwole, H, TH Lin, E Villalon, A Muniz, and A Matta. 1969. "Some Aspects of the Thermal Ecology of Puerto Rican Anoline Lizards." *Journal of Herpetology* 3 (1/2): 65. https://doi.org/10.2307/1563225.

Hedges, SB, CA Hass, and LR Maxson. 1992. "Caribbean Biogeography: Molecular Evidence for Dispersal in West Indian Terrestrial Vertebrates." *Proceedings of the National Academy of Sciences* 89 (5): 1909–1913.

Henderson, RW, and R Powell. 2009. *Natural History of West Indian Reptiles and Amphibians*. University Press of Florida.

Hickerson, MJ, EA Stahl, and HA Lessios. 2006. "Test for Simultaneous Divergence Using Approximate Bayesian Computation." *Evolution* 60 (12): 2435–2453.

Hickerson, MJ, BC Carstens, J Cavender-Bares, KA Crandall, CH Graham, JB Johnson, L Rissler, PF Victoriano, and AD Yoder. 2010. "Phylogeography's Past, Present, and Future: 10 Years after Avise, 2000." *Molecular Phylogenetics and Evolution* 54 (1): 291–301. https://doi.org/10.1016/j.ympev.2009.09.016.

Hohenlohe, PA, SJ Amish, JM Catchen, FW Allendorf, and G Luikart. 2011. "Next-Generation RAD Sequencing Identifies Thousands of SNPs for Assessing Hybridization between Rainbow and Westslope Cutthroat Trout." *Molecular Ecology Resources* 11 (March): 117–22. https://doi.org/10.1111/j.1755-0998.2010.02967.x.

Horta, FM, AM Cuervo, CC Ribas, RT Brumfield, and CY Miyaki. 2013. "Phylogeny and Comparative Phylogeography of *Sclerurus* (Aves: Furnariidae) Reveal Constant and Cryptic Diversification in an Old Radiation of Rain Forest Understorey Specialists." *Journal of Biogeography* 40 (1): 37–49. https://doi.org/10.1111/j.1365-2699.2012.02760.x.

Huey, RB, and TP Webster. 1976. "Thermal Biology of Anolis Lizards in a Complex Fauna: The *Cristatellus* Group on Puerto Rico." *Ecology* 57 (5): 985–94. https://doi.org/10.2307/1941063.

Iturralde-Vinent, M, and RDE MacPhee. 1999. "Paleogeography of the Caribbean Region: Implications for Cenozoic Biogeography." *Bulletin of the AMNH*: No. 238. http://digitallibrary.amnh.org/handle/2246/1642.

Knowles, L. 2009. "Statistical Phylogeography." *Annual Review of Ecology, Evolution, and Systematics* 40 (1): 593–612. https://doi.org/10.1146/annurev.ecolsys.38.091206.095702.

MacPhee, RDE, MA Iturralde-Vinent, and ES Gaffney. 2003. “Domo de Zaza, an Early Miocene Vertebrate Locality in South-Central Cuba, with Notes on the Tectonic Evolution of Puerto Rico and the Mona Passage.” *American Museum Novitates*, 1–42.

Meirmans, PG. 2012. “The Trouble with Isolation by Distance.” *Molecular Ecology* 21 (12): 2839–46. https://doi.org/10.1111/j.1365-294X.2012.05578.x.

Moritz, C, CJ Hoskin, JB MacKenzie, BL Phillips, M Tonione, N Silva, J VanDerWal, SE Williams, and CH Graham. 2009. “Identification and Dynamics of a Cryptic Suture Zone in Tropical Rainforest.” *Proceedings of the Royal Society B: Biological Sciences* 276 (1660): 1235–44. https://doi.org/10.1098/rspb.2008.1622.

Murphy, SF, RF Stallard, MA Scholl, G González, and AJ Torres-Sánchez. 2017. “Reassessing Rainfall in the Luquillo Mountains, Puerto Rico: Local and Global Ecohydrological Implications.” *PLOS ONE* 12 (7): e0180987.

Nicholson, KE, RE Glor, JJ Kolbe, A Larson, SB Hedges, and JB Losos. 2005. “Mainland Colonization by Island Lizards: Mainland Colonization by *Anolis*.” *Journal of Biogeography* 32 (6): 929–38. https://doi.org/10.1111/j.1365-2699.2004.01222.x.

Oaks, JR, J Sukumaran, JA Esselstyn, CW Linkem, CD Siler, MT Holder, and RM Brown. 2013. “Evidence For Climate-Driven Diversification? A Caution For Interpreting ABC Inferences Of Simultaneous Historical Events.” *Evolution* 67 (4): 991–1010. https://doi.org/10.1111/j.1558-5646.2012.01840.x.

Peterson, BK, JN Weber, EH Kay, HS Fisher, and HE Hoekstra. 2012. "Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species." *PLoS ONE* 7 (5): e37135. https://doi.org/10.1371/journal.pone.0037135.

Petkova, D, J Novembre, and M Stephens. 2015. "Visualizing Spatial Population Structure with Estimated Effective Migration Surfaces." *Nature Genetics* 48 (1): 94–100. https://doi.org/10.1038/ng.3464.

Prates, I, AT Xue, JL Brown, DF Alvarado-Serrano, MT Rodrigues, MJ Hickerson, and AC Carnaval. 2016. "Inferring Responses to Climate Dynamics from Historical Demography in Neotropical Forest Lizards." *Proceedings of the National Academy of Sciences* 113 (29): 7978–7985.

Pregill, GK, and SL Olson. 1981. "Zoogeography of West Indian Vertebrates in Relation to Pleistocene Climatic Cycles." *Annual Review of Ecology and Systematics* 12 (1): 75–98.

Rand, AS. 1964. "Ecological Distribution in Anoline Lizards of Puerto Rico." *Ecology* 45 (4): 745–52. https://doi.org/10.2307/1934922.

Renken, RA, WC Ward, IP Gill, and F Gómez-gómez. 2002. "Geology and Hydrogeology of the Caribbean Islands Aquifer System of the Commonwealth of Puerto Rico and the U.S. Virgin Islands." U.S. Geological Survey Professional Paper 1419. Regional Aquifer-System Analysis -- Caribbean Islands. Reston, VA: U.S. Dept. of the Interior, U.S. Geological Survey.

Rodríguez-Robles, JA, T Jezkova, and M Leal. 2010. "Climatic Stability and Genetic Divergence in the Tropical Insular Lizard *Anolis krugi* , the Puerto Rican 'Lagartijo Jardinero de La Montaña.'" *Molecular Ecology* 19 (9): 1860–76. https://doi.org/10.1111/j.1365-294X.2010.04616.x.

Schneider, CJ, M Cunningham, and C Moritz. 1998. "Comparative Phylogeography and the History of Endemic Vertebrates in the Wet Tropics Rainforests of Australia." *Molecular Ecology* 7 (4): 487–98. https://doi.org/10.1046/j.1365-294x.1998.00334.x.

Schoener, TW. 1971. "Structural Habitats of West Indian Anolis Lizards. II. Puerto Rican Uplands." *Breviora*, no. 368: 1–39.

Tucker, DB, SBHedges, GR Colli, RA Pyron, and JW Sites. 2017. "Genomic Timetree and Historical Biogeography of Caribbean Island *Ameiva* Lizards (Pholidoscelis : Teiidae)." *Ecology and Evolution* 7 (17): 7080–90. https://doi.org/10.1002/ece3.3157.

Velo-Antón, G, PA Burrowes, RL Joglar, I Martínez-Solano, KH Beard, and G Parra-Olea. 2007. "Phylogenetic Study of *Eleutherodactylus coqui* (Anura: Leptodactylidae) Reveals Deep Genetic Fragmentation in Puerto Rico and Pinpoints Origins of Hawaiian Populations." *Molecular Phylogenetics and Evolution* 45 (2): 716–28. https://doi.org/10.1016/j.ympev.2007.06.025.

Zamudio, KR, RC Bell, and NA Mason. 2016. "Phenotypes in Phylogeography: Species Traits, Environmental Variation, and Vertebrate Diversification." *Proceedings of the National Academy of Sciences* 113 (29): 8041–48. https://doi.org/10.1073/pnas.1602237113.

Table 1. Individuals and raw reads.

Species	n[1]	Unique Localities[2]	Raw Reads[3]
A. cristatellus	157	32	1,211,459
A. pulchellus	105	22	1,948,158
A. stratulus	94	23	1,711,729

[1]Number of individuals
[2]Number of sampling localities
[3]Average number of raw reads / individual

Table 2. Summary of number of loci and number of *A. cristatellus* individuals in each dataset.

min.ind (%)	min.loci (%)	N[1]	Loci[2]
50	50	143	1917
50	70	153	1917
50	90	157	1917
65	50	154	916
65	70	156	916
65	90	157	916
80	50	153	323
80	70	157	323
80	90	157	323

[1]Number of individuals

[2]Number of loci in the dataset

min.ind = the percentage of individuals required to retain a locus in the dataset

min.loci = the percentage of loci required to retain an individual in the dataset

Table 3. Summary of number of loci and number of *A. pulchellus* individuals in each dataset.

min.ind (%)	min.loci (%)	N[1]	Loci[2]
50	50	95	2685
50	70	105	2685
50	90	105	2685
65	50	103	1215
65	70	105	1215
65	90	105	1215
80	50	105	399
80	70	105	399
80	90	105	399

[1]Number of individuals

[2]Number of loci in the dataset

min.ind = the percentage of individuals required to retain a locus in the dataset

min.loci = the percentage of loci required to retain an individual in the dataset

Table 4. Summary of number of loci and number of *A. stratulus* individuals in each dataset.

min.ind (%)	min.loci (%)	N[1]	Loci[2]
50	50	79	1487
50	70	90	1487
50	90	94	1487
65	50	86	1044
65	70	90	1044
65	90	94	1044
80	50	90	294
80	70	90	294
80	90	94	294

[1]Number of individuals

[2]Number of loci in the dataset

min.ind = the percentage of individuals required to retain a locus in the dataset

min.loci = the percentage of loci required to retain an individual in the dataset

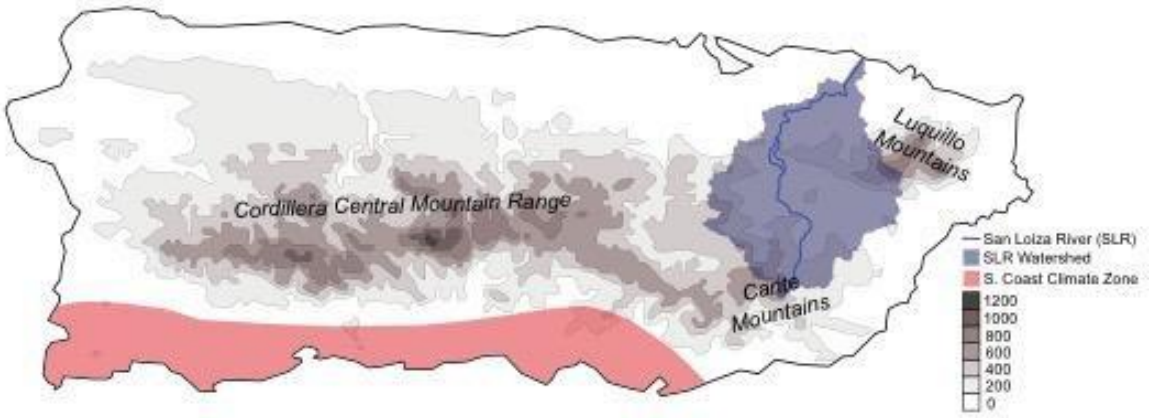


Figure 1. Map of Puerto Rico with elevation and location of the San Loiza River Basin and Southern Coast rain shadow.

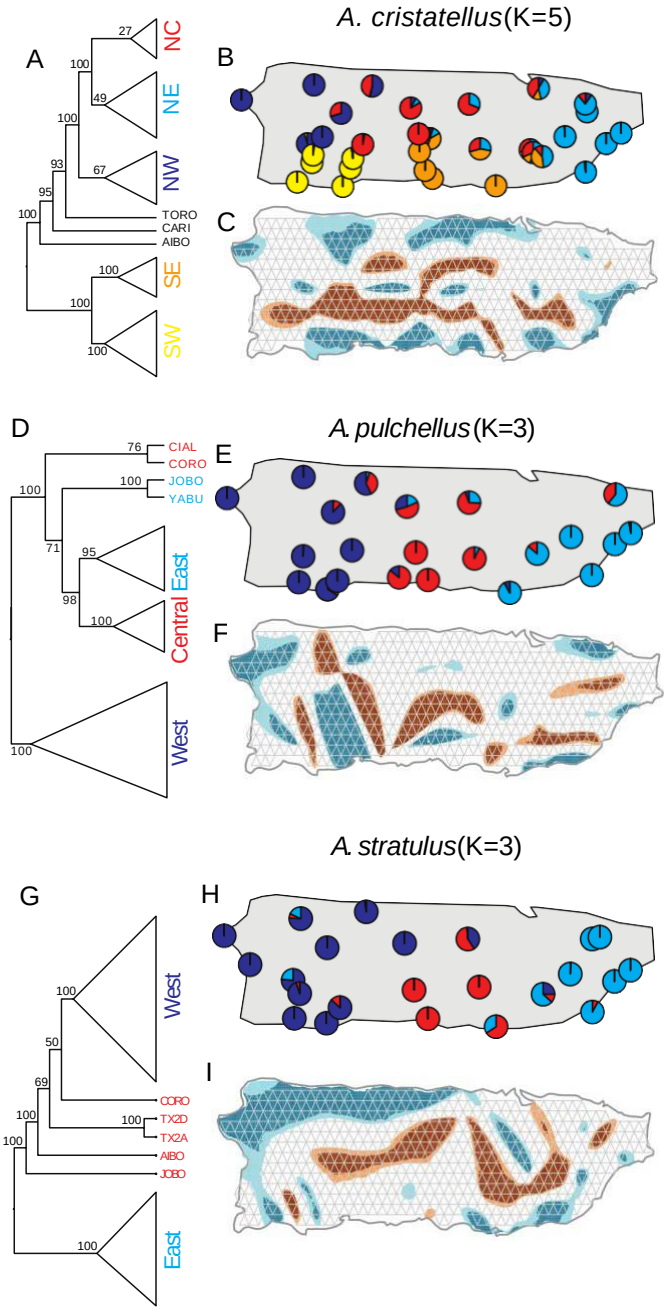


Figure 2. Phylogeographic patterns for *A. cristatellus* (A-C), *A. pulchellus* (D-F), and *A. stratulus* (G-I). (A,D,G) Phylogenetic reconstructions from SVDQuartets show major clades and localities with admixed ancestry. (B,E,F) Best population model based on cross-validation and ancestry coefficients for sampling localities. (C,F,I) Effective migration surfaces estimated in EEMS. Warm colors indicate population barriers, cool colors indicate corridors for gene flow.

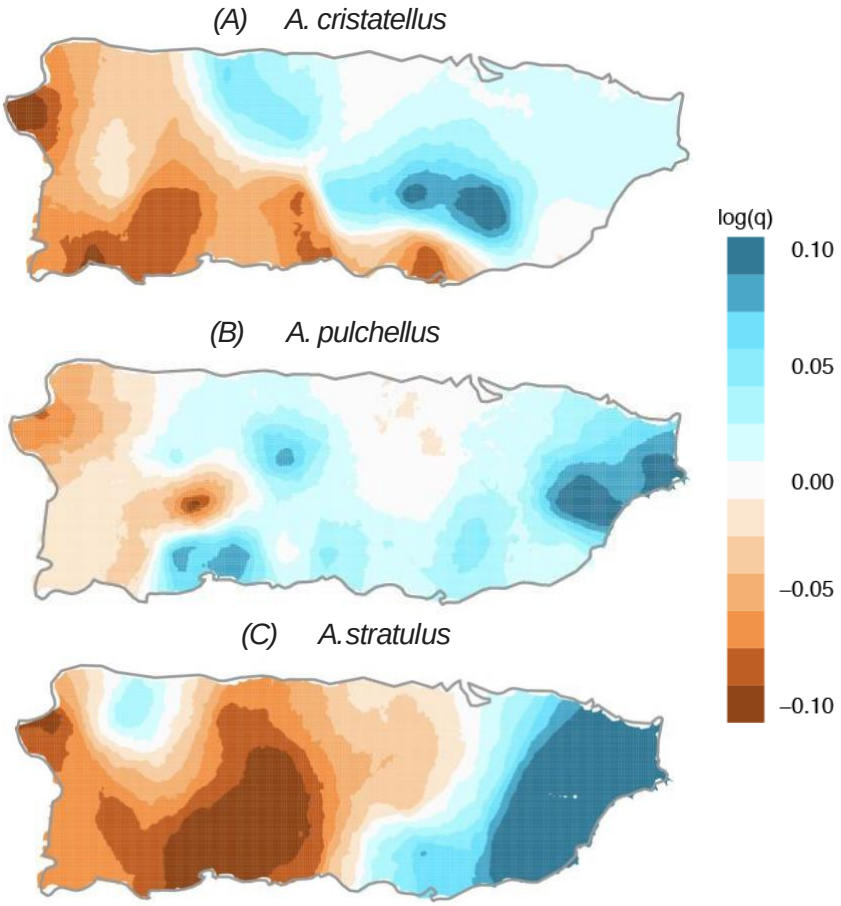


Figure 3. Distribution of genetic diversity on Puerto Rico estimated from EEMS for (A) *Anolis cristatellus*, (B) *A. pulchellus*, and (C) *A. stratulus*.

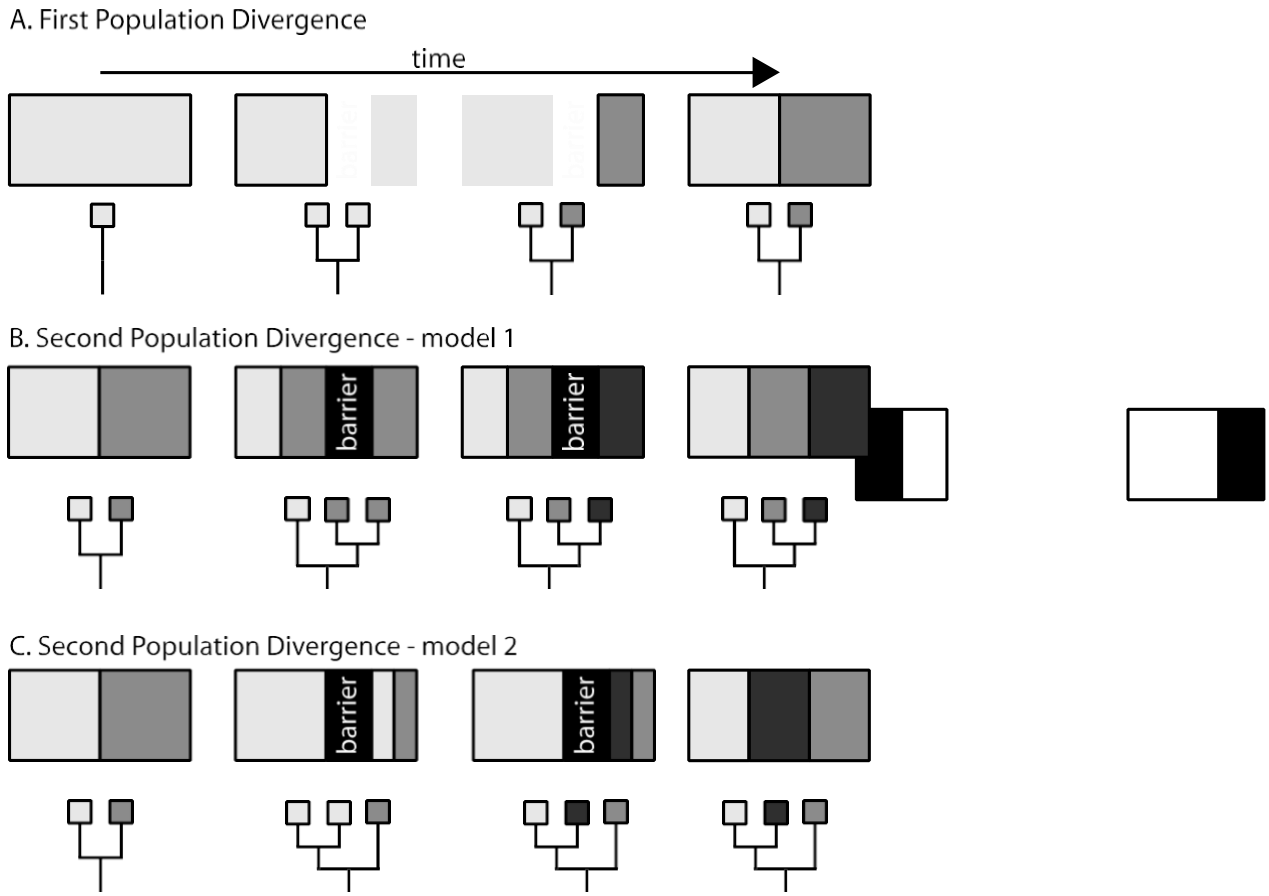


Figure 4. Alternative models depicting how repeating barriers can generate bi-furcating population trees.

Chapter 4: Testing the Bogert Effect with SNP data in *A. cristatellus*

Matthew T. McElroy[12], Raymond B. Huey[1], and Paul E. Hertz[3]

[1]Department of Biology, University of Washington, Seattle WA

[2]Burke Museum of Natural History and Culture, University of Washington, Seattle WA

[3]Department of Biology, Barnard College, New York NY

Abstract

Physiological adaptation will be essential for many species to persist in a changing climate. Counter-intuitively, ectotherms that are adept thermoregulators may be particularly challenged to adapt if their thermoregulatory behavior buffers natural selection. This idea, termed the Bogert Effect, posits that behavioral thermoregulation buffers natural selection by minimizing exposure to novel environments and selective pressures. We test the predictions of reduced natural selection and high gene flow across a thermal gradient in *Anolis cristatellus*, a thermoregulating lizard from Puerto Rico. We collected genome-wide SNPs for *A. cristatellus* that were sampled along elevational transects in southern Puerto Rico. We conducted cline analysis and latent-factor mixed-modeling of SNP data to determine whether elevation, temperature, and precipitation represent selective pressures for *A. cristatellus.* In contradiction with the Bogert Effect, we find sharp genomic clines along the transects despite active thermoregulation. We only found one SNP to be associated with mean annual temperature, whereas we found 15 candidate SNPs in association with precipitation. There results indicate the potential for an important role for adaptation in water loss physiology in thermoregulating species.

Introduction

Physiological adaptation will be essential for many species to persist in a changing climate. Understanding how thermal and hydric adaptation occur across heterogenous landscapes provides a foundation for predictions about species responses. Despite a resurgent focus on how divergent natural selection promotes adaptation despite gene flow across ecological gradients, physiological traits are often overlooked as important phenotypes for selection because they are difficult to observe, probably multi-genic, and potentially plastic (Angilletta, 2009; Keller and Seehausen, 2012). Furthermore, while organismal behavior typically promotes diversification by exposing species to new habitats and novel selective pressures, when it comes to physiology, organismal behavior may actually buffer natural selection (Bogert, 1949; Bartholomew, 1963; Huey et al., 2003). This idea is called the Bogert Effect, named after the herpetologist Charles Bogert who conducted pioneering work on reptile thermoregulation and who proposed that behavioral thermoregulation could buffer selection on thermal physiology (Huey et al., 2003).

The Bogert Effect posits that behavioral thermoregulation buffers natural selection by minimizing exposure to novel thermal environments and selective pressures. Ectotherms that are proficient behavioral thermoregulators will experience similar body temperatures largely independent of the thermal environment they inhabit (Huey and Slatkin, 1976). In contrast, ectotherms that do not behaviorally regulate their body temperatures (i.e. thermoconformers) will experience variable body temperatures across heterogenous landscapes, enabling selection to act on thermal physiology. One concern is that by buffering natural selection, the Bogert Effect will delay thermal adaptation in ectotherms that behaviorally thermoregulate and exacerbate the challenges for organisms experiencing rapid climate change (Huey et al., 2012; Buckley et al., 2015).

Two key predictions of the Bogert Effect are that relative to thermoconformers, thermoregulators should experience reduced natural selection on physiological traits along thermal gradients. Huey et al. (Huey et al., 2003) compared the thermoregulatory behavior of two Puerto Rican anoles, *Anolis cristatellus* and *A. gundlachi*, as a case study for demonstrating the Bogert Effect. *Anolis cristatellus* is a thermoregulating species that inhabits lowland xeric habitat as well as montane mesic forests. At higher elevations *A. cristatellus* utilizes mostly open canopy and sunny patches where it can maintain its body temperature near a preferred range (Huey and Webster, 1976; Hertz, 1992a; Huey et al., 2003. We revisit this system and use genetics to test the predictions of the Bogert Effect in *Anolis cristatellus*. We collected genome-wide SNPs for *A. cristatellus* that were sampled along elevational transects that span a thermal gradient in southern Puerto Rico. At the top and bottom of one transect, we collected operative and body temperatures to assess behavioral thermoregulation. We conducted cline analysis and latent-factor mixed-modeling of SNP data to determine whether elevation, temperature, and precipitation represent selective pressures for *A. cristatellus*.

Materials and Methods

Field-active body temperatures and operative temperatures

During August of 2011-2013 and January of 2013 and 2015, we conducted field experiments in the coastal dry forest of Guanica (elevation = 5 m) and in the montane mesic forest of Guilarte (elevation = 1004 m). We followed the methodology of Hertz et al. (1993) to collect operative temperatures (Te) and body temperatures (Tb) for *Anolis cristatellus*. Briefly, operative temperatures are estimated from lizard-shaped hollow copper-models are used to estimate the

distribution of body temperatures that would be expected in a given environment if lizards were not engaging in behavioral thermoregulation (Hertz, 1992). The temperatures measured from the copper-models therefore represent a null-hypothesis. The body temperatures of field-active lizards are then compared to the null-distribution to assess the degree that a species thermoregulates. We mounted copper-models throughout the habitat in randomized locations on tree trunks and tree branches to mimic *A. cristatellus* perch structures (Guanica: n=60; Guilarte: n=50). These models accurately represent the equilibrium body temperature of a live lizard within 1°C (Hertz, 1992). During each hour between 0700 and 1800, we recorded the Tb of approximately ten wild-caught *A. cristatellus* individuals as well as the Te of all copper-model.

Sampling Design

We collected 148 *A. cristatellus* along three 25 km transects that began at sea level in the southern dry forest and that ended at approximately 900 m in the central mesic mountains. We sampled eight individuals every 5 km, for a total of 48 individuals per transect. Transects were spaced 20-30 kms apart and ran approximately parallel to each other. We improvised a fourth transect in the East by capitalizing on island wide sampling for another study (McElroy, Ch2). We used this transect — with the caveat that it was not designed explicitly *a priori* as a transect — as a comparison to the steeper elevational gradients in the west. Specimens were fixed in 10% buffered formalin and liver samples were flash frozen in liquid nitrogen in the field. Specimens and tissues were accessioned at the Burke Museum of Natural History and Culture (UWBM) at the University of Washington in Seattle.

ddRAD sequencing

We used the double-digest RADseq protocol (Peterson et al., 2012) to generate genome-wide SNP data. Tissues were digested overnight at 37°C with 10% SDS and Proteinase K. Following digestion, I used a 2x Serapure bead clean-up protocol to isolate and purify DNA. I collected restriction-site associated loci for 356 individuals following the double-digest RADseq protocol (Peterson et al., 2012). I used 20 units each of the common cutter MspI (restriction site 5'-CCGG-3') and the rare cutter SbfI (restriction site 5'-CCTGCAGG-3') to digest 500 ng of DNA at 37 degrees for 10 hours. I purified the samples using Serapure speedbeads before ligating Illumina barcoded adapters onto the DNA fragments. Following ligation, I selected DNA fragments between 415-515 bp using a Pippin Prep (Sage Science) and then used proofreading Taq and Illumina specific primers to amplify size-selected libraries. I used an Agillent 2200 TapeStation to quantify and assess the distribution of size fragments of the final libraries. Final libraries were quantified using RT-qPCR at the QB3 Vincent Coates Genomics Sequencing Laboratory (UC Berkeley) and pooled in equimolar ratios for single-end 50-bp Illumina sequencing on two lanes of the HiSeq2500 and one lane of HiSeq4000.

Stacks Bioinformatics

I processed Illumina data with the STACKS pipeline ver. 1.44, which identifies putative loci and infers haplotypes for individuals (Catchen et al., 2013; 2011). To create putative loci and call SNPs, I implemented *ustacks* (Hohenlohe et al., 2011) and grouped reads into loci that differed by a threshold of one basepair (bp) mismatch with a minimum of 10x coverage. I ran *cstacks* with a threshold allowing two bp mismatches to create a catalogue of consensus loci across individuals. Finally, I ran *sstacks* to resolve the allelic state at each locus in individuals. I ran the

populations script to resolve individual haplotypes, to retain nucleotides with a minor allele frequency (maf) of 5%, and to generate datasets for which loci are retained if they are present in 50% of the individuals. I then generated final datasets by removing individuals that were missing 50% of the loci.

Cline fitting Analysis

We used the R package HZAR (Derryberry et al., 2014) to perform maximum-likelihood cline-fitting analyses of the SNP data along our four transects. We compared the null model to the cline model that assumes trait intervals are fixed at 0 and 1 without exponential trail fitting. We determined the best model from AICc scores, and then extracted and compared ML model estimates for cline centers and widths of each transect.

Latent-factor Mixed Modeling

Latent-factor mixed modeling (LFMM) explores associations between allele frequencies and environmental characteristics while accounting for background levels of population structure (Frichot et al., 2013). We conducted LFMM in the R package LEA (Frichot and François, 2015). Because the statistical power to detect outlier loci is greater with widespread sampling than it is from transects (Lotterhos and Whitlock, 2015), we analyzed the island-wide SNP dataset (from Chapter 2) for *Anolis cristatellus*. We tested for ecological associations between SNPs and two environmental variables: annual precipitation and mean annual temperature (Hijmans et al., 2005). We tuned to number of latent factors (K=10) to obtain a lambda value (~1.5) that generated an ideal distribution of p-values. Finally, we set the false discovery rate to 10% and generated a list of candidate loci.

Results

Lizard body and operative temperatures

Within the lowland xeric forest, *A. cristatellus* maintained body temperatures at or below the temperatures of copper models (Figure 1) indicating that they are behaviorally down-regulating their body temperature. At the montane forest site, *A. cristatellus* body temperatures were above those of copper model (Figure 1) indicating that they are upregulating their body temperatures. As a result of thermoregulatory behavior, the difference between low- and high-elevation operative temperatures was greater than the difference between low- and high-elevation body temperatures.

Cline fitting analysis

For each transect, the simple cline model outperformed the null model (deltaAICc scores ranged from 29.97 – 40.77; Table 2). The ranges for cline widths all overlapped, but the posterior probability distributions were different. The highest probability values for cline widths for the the Parguera transect was 4 km, the Jobos transect was 19 km, and the Guanica and Juana Diaz transects were zero (Table 2).

LFMM, precipitation, and temperature

15 SNP loci were associated with annual precipitation on Puerto Rico, and only one SNP locus was associated with mean annual temperature. The locus that was associated with temperature was also associated with precipitation.

Discussion

Body temperatures of xeric and montane *Anolis cristatellus* are more similar than xeric and mesic operative temperatures because montane lizards up-regulate and xeric lizards down-regulate their body temperatures. In support of the Bogert Effect, the body and operative temperatures highlight that the full magnitude of differences between thermal environments aren't necessarily experienced by individuals, and therefore temperature-mediated selection should be weaker than if thermoregulatory behavior was ignored. However, along the steep elevational gradient in Southwest Puerto Rico behavioral thermoregulation is not enough to completely overcome differences in thermal environment.

In our analysis, populations that are spanning the steepest elevational transects (i.e. Guanica and Juana Diaz) have the narrowest cline widths and display the least admixture, indicating that climate-mediated selection is stronger along steeper elevational profiles. To make a robust inference of climate-mediated selection, one would ideally identify candidate or outlier loci, and then verify the gene function matches the a priori hypothesis. Unfortunately, our 39 bp RAD loci are too short to confidently BLAST to available genomic resources and identify gene classes or functions. Despite this limitation in our SNP dataset, our LFMM analysis partly validates the SNP clines and the lizard body temperature data. We found that 0.7% of our RAD loci showed statistical associations with mean annual precipitation. In contrast, <0.1% loci had an association with temperature.

With respect to the Bogert Effect, the body temperatures and the SNP clines seem contradictory – why do we have steep genetic clines if thermoregulatory behavior partially buffers body temperature differences between at either end of our transects? Our LFMM analysis

indicates that the answer may relate to the hydric environment. Relative to body temperature, it may be harder for individuals to behaviorally-regulate their rates of water loss (e.g. through habitat selection). If this were the case, individuals would experience stronger selection for traits relating to water loss physiology than they would for thermal physiology and the predictions from the Bogert Effect might appear unmet.

References

Angilletta, MJ. 2009. *Thermal Adaptation: A Theorethical and Empirical Synthesis*. Oxford University Press.

Bartholomew, GA. 1963. "The Roles of Physiology and Behaviour in the Maintenance of Homeostasis in the Desert Environment." *Symposia of Hte Society for Experimental Biology* 18.

Bogert, CM. 1949. "Thermoregulation in Reptiles, a Factor in Evolution." *Evolution* 3 (3): 195–211.

Buckley, LB, JC Ehrenberger, and MJ Angilletta. 2015. "Thermoregulatory Behaviour Limits Local Adaptation of Thermal Niches and Confers Sensitivity to Climate Change." *Functional Ecology* 29 (8): 1038–47. https://doi.org/10.1111/1365-2435.12406.

Derryberry, EP, GE Derryberry, JM Maley, and RT Brumfield. 2014. "HZAR : Hybrid Zone Analysis Using an R Software Package." *Molecular Ecology Resources* 14 (3): 652–63. https://doi.org/10.1111/1755-0998.12209.

Frichot, E, and O François. 2015. "LEA : An R Package for Landscape and Ecological Association Studies." *Methods in Ecology and Evolution* 6 (8): 925–29. https://doi.org/10.1111/2041-210X.12382.

Frichot, E, SD Schoville, G Bouchard, and O François. 2013. "Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models." *Molecular Biology and Evolution* 30 (7): 1687–99. https://doi.org/10.1093/molbev/mst063.

Hertz, PE. 1992. "Temperature Regulation in Puerto Rican Anolis Lizards: A Field Test Using Null Hypotheses." *Ecology* 73 (4): 1405–17. https://doi.org/10.2307/1940686.

Hertz, PE, RB Huey, and RD Stevenson. 1993. "Evaluating Temperature Regulation by Field-Active Ectotherms: The Fallacy of the Inappropriate Question." *American Naturalist* 142 (5): 796–818.

Hijmans, RJ, SE Cameron, JL Parra, PG Jones, and A Jarvis. 2005. "Very High Resolution Interpolated Climate Surfaces for Global Land Areas." *International Journal of Climatology* 25 (15): 1965–78. https://doi.org/10.1002/joc.1276.

Hohenlohe, PA, SJ Amish, JM Catchen, FW Allendorf, and G Luikart. 2011. "Next-Generation RAD Sequencing Identifies Thousands of SNPs for Assessing Hybridization between Rainbow and Westslope Cutthroat Trout." *Molecular Ecology Resources* 11 (March): 117–22. https://doi.org/10.1111/j.1755-0998.2010.02967.x.

Huey, RB, MR Kearney, A Krockenberger, JAM Holtum, M Jess, and SE Williams. 2012. "Predicting Organismal Vulnerability to Climate Warming: Roles of Behaviour,

Physiology and Adaptation." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1596): 1665–79. https://doi.org/10.1098/rstb.2012.0005.

Huey, RB, PE Hertz, and B Sinervo. 2003. "Behavioral Drive versus Behavioral Inertia in Evolution: A Null Model Approach." *The American Naturalist* 161 (3): 357–366.

Huey, RB, and M Slatkin. 1976. "Costs and Benefits of Lizard Thermoregulation." *The Quarterly Review of Biology* 51 (3): 363–84.

Keller, I, and O Seehausen. 2012. "Thermal Adaptation and Ecological Speciation." *Molecular Ecology* 21 (4): 782–99. https://doi.org/10.1111/j.1365-294X.2011.05397.x.

Lotterhos, KE, and MC Whitlock. 2015. "The Relative Power of Genome Scans to Detect Local Adaptation Depends on Sampling Design and Statistical Method." *Molecular Ecology* 24 (5): 1031–46. https://doi.org/10.1111/mec.13100.

Peterson, BK, JN Weber, EH Kay, HS Fisher, and HE Hoekstra. 2012. "Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species." *PLoS ONE* 7 (5): e37135. https://doi.org/10.1371/journal.pone.0037135.

Table 1. Characteristics of transects and ddRADseq datasets.

Transect[a]	N[b]	Distance (km)	Loci[d]	Loci[e]
Parguera	29	25	1154.3	1690
Guanica	39	25	1453.2	2073
Juana Diaz	37	25	1455.2	2073
Jobos	27	60	1953.2	2982

[a]Name of town at start of transect.
[b]Number of individuals in final dataset.
[c]Length of transect in kilometers.
[d]Average number of loci per individual.

Table 2. Results from cline analysis for each transect.

Transect[a]	Cline width[b]	Width range[c]	AICc[d]	AICc-null[e]
Parguera	4	0.1 – 7.8	4.72	45.49
Guanica	0	0.0 - 7.2	11.12	51.32
Juana Diaz	0	0.0 – 7.7	11.20	49.32
Jobos	19	0.5 – 30.7	4.32	34.29

[a]Name of town at start of transect.
[b]Highest probability value for cline width (km).
[c]Range of 2LL values for cline width (km).
[d]AICc value for the "cline" model in HZAR.
[d]AICc value for the "null" model in HZAR.

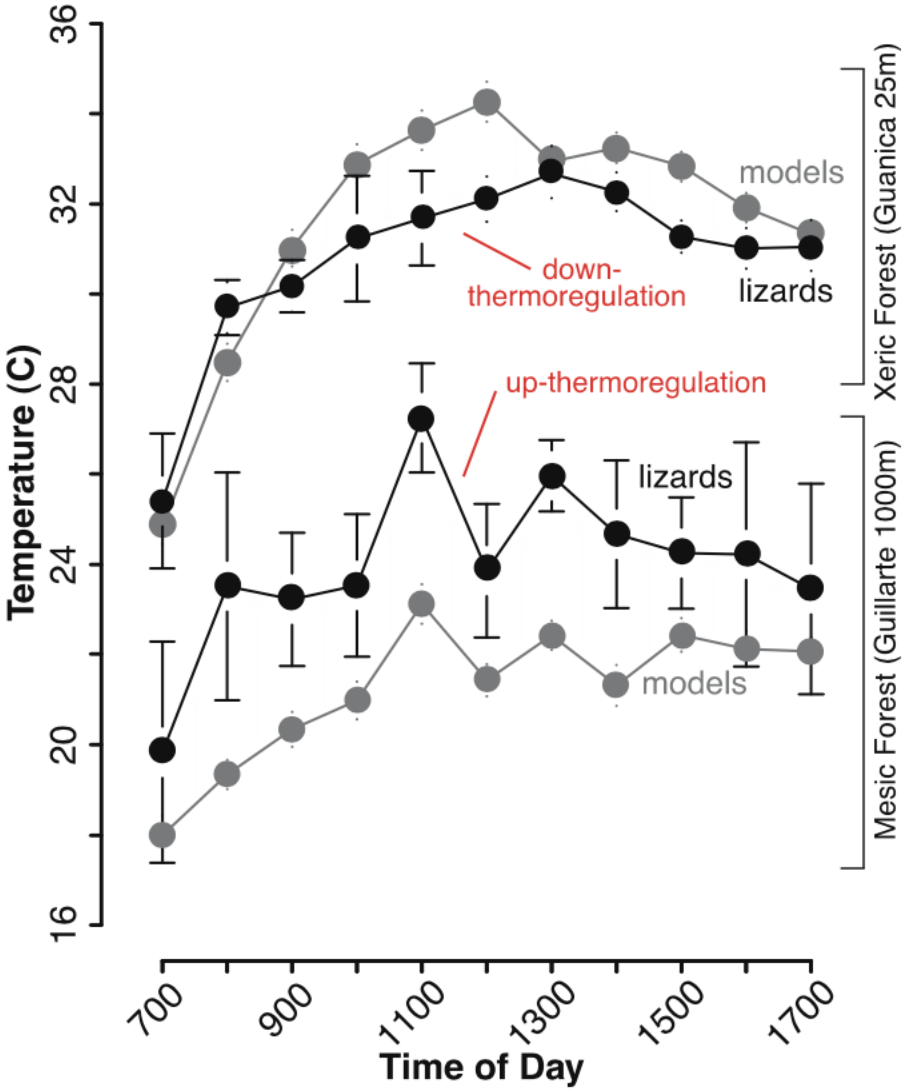


Figure 1. Field-active body temperatures ('lizards') and operative temperatures ('models') throughout the day for *A. cristatellus* at high- and low-elevation sites.

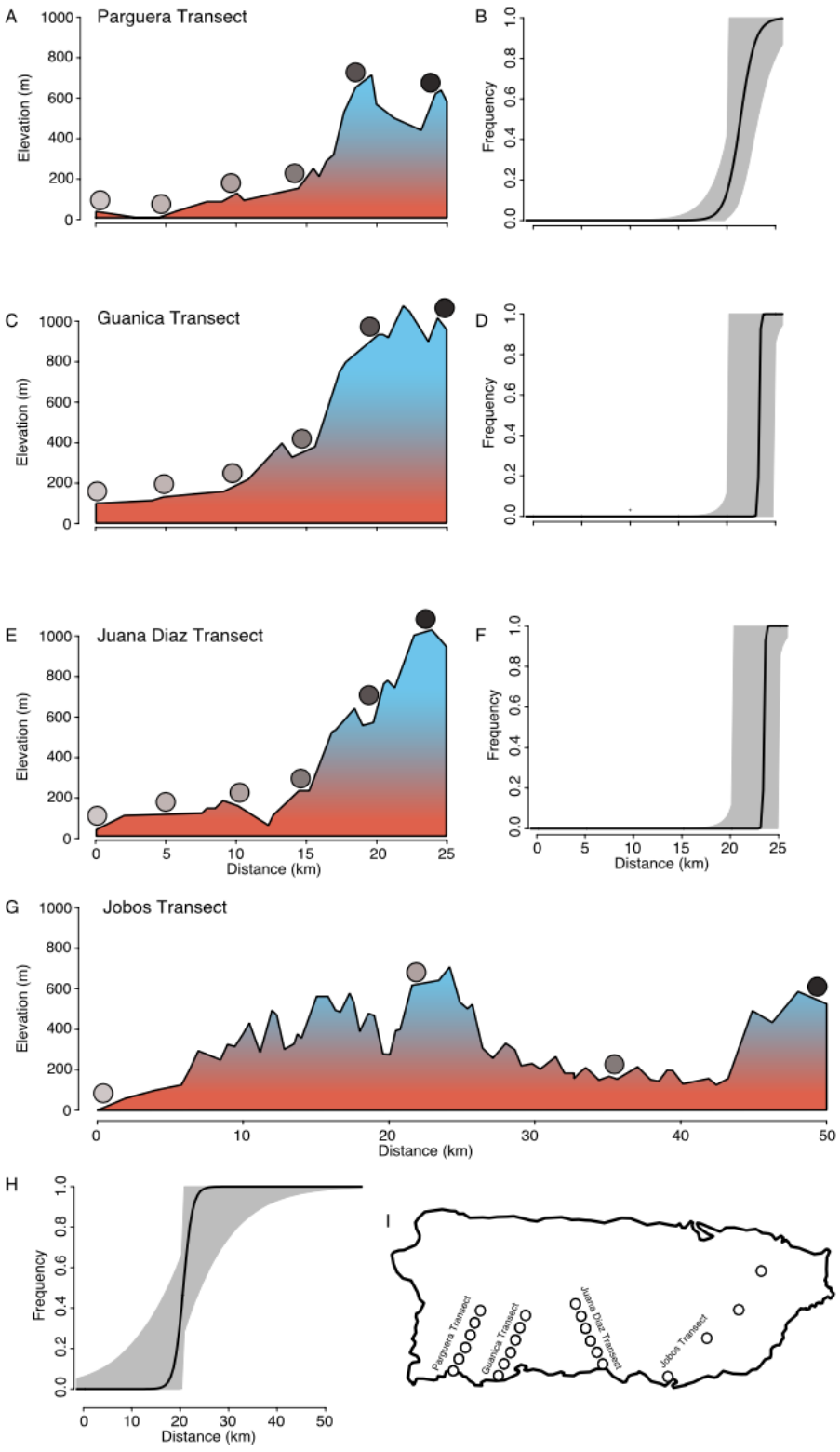


Figure 2. Elevational profiles (ACEG), transect sampling localities (I), and maximum likelihood SNP clines (BDFH).