The Past as a Window to the Future: Phylogeography, Historical Demography, and Conservation Genetics of *Plethodon jordani*, a Terrestrial Salamander Endemic to the Great Smoky Mountains

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Amy Marie Luxbacher

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advised by Dr. Kenneth H. Kozak

December 2014

ACKNOWLEDGMENTS

I am grateful to many people that helped make this research come to fruition. First and foremost, thank you to my advisor, Dr. Ken Kozak, who took a chance by accepting me as his first student in his first semester at UMN. He has been a great advocate and provided terrific support, always encouraging me when I needed it, offering insightful feedback, and helping me keep sight of the "big picture." Thanks also to current and former Kozak lab members: Dr. Matt Gifford provided inspiration to get my project going and ran the mechanistic models discussed in Chapter 1; Dr. Don Shepard shared his experience and wisdom on seemingly everything and helped me get started with analyses covered in Chapter 2; and Matt Dufort, Ben Lowe, Marta Lyons, Tricia Markle, Chris Smith, and Ben Weinstein provided lab and computational assistance and have been all-around great friends, lab mates, and colleagues.

Thank you to my committee members, Dr. Sharon Jansa, Dr. Peter Tiffin, and Dr. Bob Zink, who have encouraged me and provided helpful feedback to make me a better writer and scientist. I am indebted to Dr. Keith Barker who patiently took me under his proverbial wing, shared his resources and extensive knowledge, and taught me how to "do" lab work. Dr. Clarence Lehman brightened many a long winter afternoon by distributing chocolates on the third floor of the Ecology building. Dr. Sarah Boyer has been a generous, kind, and inspiring mentor who helped me solidify my career aspirations.

I am also grateful for the support of the Ecology, Evolution and Behavior Graduate Program (especially Dr. Sarah Hobbie, Dr. Jacques Finlay, and Lisa Wiggins) and the College of Biological Sciences. Fellow and former graduate students, including Dr. Mike Dixon, Lorissa Fujishin, Dr. Tom Giarla, Dr. Dominick Halas, Dr. Chih-Ming Hung, Dr. Alexei Powell, Dr. Hernán Vásquez Miranda, Dr. Liz Wallace, and Mike Wells, helped me troubleshoot in the lab and engaged in all types of worldly discussions. Dr. Dawn Tanner was a tremendous help in the lab and provided refreshing conversation and perspective.

Funding for my research has been generously provided by the American Society of Ichthyologists and Herpetologists Gaige Fund, the Bell Museum of Natural History Dayton Fund, the Minnesota Herpetological Society, the Society for Systematic Biology, the University of Minnesota Louise T. Dosdall Fund, National Science Foundation DEB-0949590 (awarded to Ken Kozak), and a National Parks Service George Melendez Wright Climate Change Fellowship. Thanks to Paul Super for his assistance in securing the latter fellowship, as well as collecting permits for Great Smoky Mountains National Park.

Paul Super and Susan Sachs both helped with field logistics, and Emily Darling and Emily Guss coordinated local field assistants in Great Smoky Mountains National Park. Joy Absher, Emily Darling, Matt Gifford, Ken Kozak, Kimberlee Lisicki, Monica Looze, Ben Lowe, Jimmy Luxbacher, Scott McCloskey, Tricia Markle, Korey Schneider, Jordan Shealy, and many Great Smoky Mountains National Park summer interns and ranger-teachers graciously offered hours, days, and weeks of their time to help me find and sample salamanders in the field.

I have been honored with two opportunities to share my research broadly. First, I thank Susan Sachs and Keith Ellington for producing the "Climate Change and the Red-

cheeked Salamander" podcast (http://www.nps.gov/grsm/photosmultimedia/climate-video1.htm). Second, I am grateful to Mary Kay Carson and Tom Uhlman for the opportunity to have my research featured in chapter 5 of their book, *The Park Scientists: Gila Monsters, Geysers, and Grizzly Bears in America's Own Backyard* (2014, Houghton Mifflin Harcourt). Working with them was one of the most meaningful projects in which I have ever been involved because the end product has been so far-reaching.

Finally, I am deeply appreciative of my family and friends, who have shared personally and vicariously in my adventures in graduate school, fieldwork, and everything else. My mother, especially, fostered my love of natural history and encouraged me to get my hands dirty from a young age. My father instilled in me an appreciation for the value of higher education and supported me even though I aspired to be "the other type of doctor." My amazing husband Jimmy believed in me always, and my daughter Audrey has brought perspective, motivation, and fun to each day since her birth. I extend many thanks to Marilu Thomas whose mentorship and friendship has been a saving grace, and to Lindsay Hickey, Amy McWhite, Sarah Sacido, and Melissa Wallace whose camaraderie has kept me laughing and sane for more than half my life.

DEDICATION

To Jimmy, whose love and support has been unwavering

and

Audrey, who is my mirror and changed my life for the better

ABSTRACT

My dissertation research explores historic and climatic factors underlying geographic patterns of genetic structure in *Plethodon jordani*, a salamander that is endemic to the Great Smoky Mountains of eastern North America. Chapter 1 is collaborative work with Ben Weinstein, Dr. Matt Gifford, and Dr. Ken Kozak. We use mechanistic and correlative distribution models to predict the current and past (Pleistocene) distributions of *P. jordani* and evaluate these predictions alongside a mitochondrial phylogeny. Our analyses reveal three reciprocally monophyletic lineages that correspond to three geographic regions in the Great Smoky Mountains. The mechanistic model is consistent with genetic data, and together these results suggest that *P. jordani* maintained populations within its current distribution across multiple glacial cycles by tracking suitable microclimates via small changes in elevational range.

I expand upon these findings in Chapter 2 by investigating geographic and demographic changes over the course of *P. jordani*'s diversification. I find that divergence within lineages corresponds to mid- to late-Pleistocene glaciations when populations shifted to lower elevations. Effective population sizes were relatively constant over time, indicating long-term persistence and stability of populations even as distributions were shifting in response to glacial cycles. Phylogeographic reconstructions show that the range of ancestral *P. jordani* was in the east-central Great Smoky Mountains and populations subsequently expanded, predominantly to the west along the main ridgeline. These results suggest that topography and niche conservatism have been the main constraints on the geographic distribution of *P. jordani* through time.

In Chapter 3, I synthesize results of the two previous chapters, along with new results based on microsatellite data, to establish conservation priorities for *P. jordani*. Microsatellite analyses reveal three genetic clusters that are partially congruent with phylogeographic patterns based on other markers, but show an additional genetic break in the western distribution. I propose that *P. jordani* comprises three evolutionarily significant units that reflect the patterns inferred from different genetic markers, ecological niche models, and historical distributions. My results highlight the importance of integrating information from multiple data sources when devising conservation and monitoring strategies for species that may be threatened by climate change.

Each chapter of this dissertation will be submitted to a different scientific journal, as indicated in the footnotes at the beginning of each chapter; any variations in formatting within chapters reflect the specifications of the different journals.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Integrating mechanistic modeling and phylogeography to infer Pleistocene refugia of a montane endemic salamander, *Plethodon jordani*[1]

SUMMARY

Aim We investigate the extent and locations of glacial refugia in the southern Appalachian Mountains and consider their influence on genetic structure and diversity in a narrowly endemic terrestrial salamander, *Plethodon jordani*. The goals of this study are to: (1) infer potential refugia by modeling the species' distribution since the Last Glacial Maximum (LGM), (2) identify lineages and characterize genetic diversity within populations across the range, and (3) assess whether higher diversity occurs in predicted refugia compared to non-refugia.

Location Southern Appalachian Mountains of eastern North America (North Carolina and Tennessee, USA)

Methods We used mechanistic and correlative distribution models to predict the current, LGM, and refugial distributions of *P. jordani*. We sampled 362 individuals from 70 localities, inferred the phylogeny from 1116 bp of the mitochondrial genome, and tested

[1] A version of this chapter is being prepared for publication:
Luxbacher AM, Weinstein BG, Gifford ME, Kozak KH. Integrating mechanistic modeling and phylogeography to infer Pleistocene refugia of a montane endemic salamander, *Plethodon jordani*. *Journal of Biogeography*

whether refugia predicted by models exhibited higher genetic diversity than recently colonized areas.

Results Both models inferred similar current distributions, but the mechanistic model predicted a broader distribution under LGM conditions than the correlative model. Both models predicted climatically-stable areas (i.e., refugia) at mid-elevations. Phylogenetic analyses revealed three reciprocally monophyletic lineages that correspond to three geographic regions. The distributions of these lineages are not spatially congruent with refugia inferred from the models; however, the mechanistic model is consistent with genetic data given that phylogeographic structure was established before the LGM. High diversity was not associated with the inferred locations of refugia.

Main Conclusions The mechanistic model output and phylogeographic structure reflect the distribution of *P. jordani* at different times in its evolutionary history. Our results suggest that the species maintained populations within its current distribution across multiple glacial cycles by tracking suitable microclimates via small changes in elevational range. These findings support the idea that long-term climatic stability in the southern Appalachians has promoted high regional species richness and endemism.

INTRODUCTION

Pleistocene glacial cycles caused profound changes in species' distributions and influenced the genetic diversity and biodiversity we see today (Avise, 2000; Taberlet &

Cheddadi, 2002; Hewitt, 2004). During glaciations, northern temperate species were generally displaced to southern refugia where environmental conditions were more suitable (Hewitt, 1996; Taberlet *et al.*, 1998; Schmitt, 2007; Stewart *et al.*, 2010). Many species subsequently shifted or broadened their distributions northward as they tracked suitable habitat and microclimate during interglacial periods, leaving a signature of expansion in DNA (Hewitt, 1996, 1999). Increasingly, studies have also reported genetic structure that is not explained by expansion from southern refugia (Pfenninger & Posada, 2002; Rowe *et al.*, 2004; Stewart *et al.*, 2010), and noted the disparity between dispersal rates and the distances species would have had to travel from southern refugia to their current northern distributions (Reid's paradox; Clark *et al.*, 1998). Glacial microrefugia, small areas of suitable habitat at northern latitudes in which isolated populations could survive unfavorable regional climate conditions (Rull, 2009), have been proposed as an explanation for these observations, and many studies have found support for their existence in temperate Europe (Stewart & Lister, 2001; Provan & Bennett, 2008; Schmitt, 2009; Edwards *et al.*, 2012; Schmitt & Varga, 2012) and northwestern North America (Soltis *et al.*, 1997; Brunsfeld *et al.*, 2001; Shafer *et al.*, 2010).

The southeastern United States is another recognized southern glacial refugium for temperate species (Walker & Avise, 1998; Avise, 2000; Waltari *et al.*, 2007). Yet, many regional species exhibit deep phylogeographic structure (Austin *et al.*, 2004; Soltis *et al.*, 2006; Lee-Yaw *et al.*, 2008), suggesting that they may have also maintained disjunct populations outside of main refugia, such as in the southern Appalachian Mountains (Church *et al.*, 2003; Walker *et al.*, 2009). These mountains were not glaciated during the Pleistocene, but high elevations were likely covered by tundra and boreal forest

(Whittaker, 1956; Delcourt & Delcourt, 1980), raising the possibility that mid-elevation areas provided suitable microhabitat for temperate species (Nalepa *et al.*, 2002; Taberlet & Cheddadi, 2002; Dobrowski, 2010; Garrick, 2011). Today the southern Appalachians are characterized by high species richness and endemism for a variety of plant and animal taxa (Sharkey, 2001; Soltis *et al.*, 2006). This pattern of richness, coupled with the paleo-geographic history of the region (Linzey, 2008), suggests that mid-elevation areas may have harbored populations in refugia or microrefugia during recent glacial cycles.

The southern Appalachian Mountains are a global hotspot for salamander diversity, with many endemic species occurring on mountain isolates (Duellman & Sweet, 1999). The geographic distribution of *Plethodon jordani*, for example, is generally restricted to elevations above 1000 m within the boundaries of Great Smoky Mountains National Park (GRSM) in North Carolina and Tennessee (Highton & Peabody, 2000). Previous work based on eight populations identified two cryptic lineages in *P. jordani* and suggested that diversification occurred during the Pleistocene (Weisrock & Larson, 2006). Given that most *Plethodon* species are thought to be highly philopatric and have limited dispersal abilities due to narrow climatic tolerances (Feder, 1983; Petranka, 1998; Kozak & Wiens, 2006, 2010), we expect that this species has not experienced much recent change in its distribution (Wake, 2009). Yet, we know that areas currently suitable for *P. jordani* were considerably cooler than current conditions at the Last Glacial Maximum (LGM; Delcourt & Delcourt, 1980). These observations suggest a history of range shifting that, coupled with the topographic complexity of the area, may have promoted allopatric isolation and divergence among populations. Further investigation into climatic suitability and more comprehensive sampling is needed to characterize geographic

patterns of genetic variation within *P. jordani* and uncover where refugia could have existed in the Great Smoky Mountains or in the southern Appalachian Mounatins at large.

Paleoclimate models are a powerful tool for generating hypotheses about past refugia (Hugall *et al.*, 2002; Knowles *et al.*, 2007; Kozak *et al.*, 2008). These models are typically built following a correlative approach in which climate data are extracted from known species occurrence records and projected to Pleistocene climate scenarios (Carstens & Richards, 2007). However, they can lead to incorrect conclusions if theoretical assumptions are violated, such as when a species does not occur in all climatically-suitable areas, when biotic interactions influence a species' range limit, or if the models are projected onto non-analogous past climate conditions (Nogués-Bravo, 2009). Moreover, they are unable to account for processes that may be limiting species ranges, such as physiological constraints related to temperature (Kearney & Porter, 2009). Mechanistic models, which infer a species' distribution from physiological and microclimate data (Buckley *et al.*, 2010), may overcome these challenges. One mechanistic approach is to quantify the relationship between physiological traits and species-specific energy budget equations to infer a species' distribution, thus capturing interactions between species biology and habitat conditions that are important in determining where the species can occur (reviewed in Kearney & Porter, 2009). In this way, mechanistic models offer a promising means of generating alternative phylogeographic hypotheses that account for relevant constraints on species distributions through time.

Here, we apply a mechanistic bioenergetic model developed for *P. jordani* (Gifford & Kozak, 2012) to identify areas that fit within the physiological tolerances of this species

since the LGM. We consider these results alongside those obtained from a correlative model and use mitochondrial DNA (mtDNA) to investigate phylogeographic patterns across the range of *P. jordani*. The goals of this study are to (1) infer potential refugia by modeling the species' distribution since the LGM, (2) identify lineages and characterize genetic diversity within populations, and (3) assess whether higher diversity occurs in predicted refugia compared to non-refugia. We test the hypothesis that phylogeographic structure is spatially correlated with refugia inferred from the models. Our results improve our understanding of how a narrow-ranging temperate species responded to Pleistocene climate changes and reveal potential locations of refugia in the southern Appalachian Mountains that can be investigated in other dispersal-limited species in the region.

MATERIALS AND METHODS

Mechanistic distribution modeling

We predicted the spatial distribution of suitable microclimate for *P. jordani* using a mechanistic bioenergetic model developed by Gifford and Kozak (2012) that estimates where individuals can obtain sufficient energy to meet their metabolic and reproductive needs. This model solves species-specific energy balance equations by calculating the difference between energy gain (measured as thermal sensitivity of food consumption and potential foraging time) and energy loss (measured as thermal sensitivity of basal metabolic rate and the energy content of an average clutch of eggs). The model output

takes the form of a raster in which grid cells represent discretionary energy in kilojoules; a negative value in a cell indicates that energetic costs exceed energetic inputs for a salamander at that location, whereas a positive value indicates that costs are less than inputs (i.e., an individual obtains more energy through foraging than it loses through metabolic activity). Environmental variables in the model are parameterized using the output from a microclimate model developed for GRSM that approximates the near-ground air temperature variation in 90 m grid cells (Fridley, 2009). These values are used to calculate nighttime variation in the operative body temperature of *P. jordani*, which determines the amount of time that an individual salamander can actively forage at a given location and is used to derive the thermally dependent parameters in the model (Gifford & Kozak, 2012).

Modeling the current distribution of suitable microclimate for *P. jordani* was performed as described in Gifford and Kozak (2012). To model the distribution at the LGM, we first adapted the microclimate model for GRSM to downscale LGM synoptic temperature data to near-ground level. The microclimate model provides estimates of daily minimum and maximum temperatures ca. 1 m above ground at sites across GRSM after accounting for topographic variables, including solar radiation, slope, and soil moisture (Fridley, 2009). Modifications made to the model and details about LGM data are described in Appendix S1. Operative body temperature estimates derived from the microclimate model were input to calculate the thermal sensitivity of food intake and metabolic rate, and resulting values were used in the calculations of energetic input and costs, respectively. All microclimate and bioenergetic modeling was conducted in R v 2.10.1 (R Development Core Team 2009).

Correlative distribution modeling

We also modeled the current and past distributions of *P. jordani* using a correlative approach implemented in MAXENT v 3.3.3a (Phillips *et al.*, 2006). We obtained current climate data at 30 arc-second resolution from the WorldClim database (http://www.worldclim.org; Hijmans *et al.*, 2005). We reduced the data set to nine bioclimatic variables that we consider relevant to the ecological and physiological requirements of *P. jordani* (Appendix S1; Nogues-Bravo, 2009; Elith *et al.*, 2010, 2011). We clipped current climate layers to an area extending from 35.0° to 36.3° latitude and -84.5° to -82.5° longitude, which includes the current distribution of *P. jordani* within GRSM as well as surrounding mountains occupied by closely related species (Weisrock & Larson, 2006; Kozak *et al.*, 2008). We ran the model using 319 occurrence records (Appendix S2) and the current climate variables, then projected it onto LGM climate layers clipped to a larger area that extends from the southern Appalachians surrounding GRSM to northern Florida and includes areas that have been proposed as glacial refugia for other species (Davis, 1983; Avise, 2000; Soltis *et al.*, 2006, Lee-Yaw *et al.*, 2008). Details about occurrence records and LGM data layers can be found in Appendix S1.

We assessed predictive performance of the model with 10-fold (replicate) cross-validation (Elith *et al.*, 2010), which generates different sets of localities for training and testing for each replicate and evaluates them based on the area under the receiver operating characteristic curve (AUC). In MAXENT, this statistic quantifies how well a model identifies known localities more accurately than random background occurrences

across the study region (Phillips *et al.*, 2006). For each replicate, we visually inspected the clamping output to assess the extent to which projected variables reached the limit of their training range due to novel climate conditions at the LGM. We also assessed the extent of model extrapolation using multivariate environmental similarity surfaces (MESS) that indicate areas where at least one variable has a value outside of its training range (Elith *et al.*, 2010).

Assessment of mechanistic and correlative model outputs

We applied decision thresholds to each model output to make binary predictions about current and LGM distributions. For the mechanistic models, we used a discretionary energy value of zero as the threshold (Gifford & Kozak, 2012). When discretionary energy values are less than zero, individuals experience an energy deficit (i.e., energetic demands exceed energetic costs of foraging) and grid cells were considered to be unsuitable; when values are greater than zero, individuals may experience an energy surplus and grid cells were considered to be suitable. For the correlative models, we considered three logistic thresholds (see Liu *et al*., 2005 for a comparison of the performance of different threshold criteria): 10^{th} percentile training presence, equal training sensitivity and specificity, and maximum training sensitivity plus specificity. Grid cells with logistic probability values greater than the threshold value were considered to be suitable.

For each type of model, we determined the area of predicted current and LGM distributions, and calculated the area ratio (current range size / LGM range size; Ruegg

et al., 2006) to quantify amount of change in predicted distribution under each model since the LGM. We also identified stable climatic areas that may have allowed *P. jordani* to persist throughout the Pleistocene, thus functioning as refugia (Hugall *et al.*, 2002; Hewitt, 2004; Carnaval *et al.*, 2009). Following Devitt *et al.* (2013), we reclassified the suitability maps by assigning a score of 0 to unsuitable areas and 1 to suitable areas. For each type of model, we summed these maps to estimate stability over time; areas with a value of 2 were considered stable and to represent potential refugia. These procedures were conducted in ARCGIS 10.1 (ESRI).

Taxon sampling and DNA amplification

We collected tail tip tissue from 324 *P. jordani* individuals from 66 localities across its range between 2009 and 2011 (Appendix S3). Tail tips were stored in 95% ethanol initially, and later frozen at -80°C for long-term storage in the Bell Museum of Natural History. We extracted DNA from these samples using a Blood and Tissue Genomic DNA Miniprep System (Viogene) following the manufacturer's instructions. We also extracted DNA from 31 liver tissue samples that had previously been accessioned to the Bell Museum. These samples were collected in 2008 from four additional localities in GRSM (Appendix S3).

We amplified approximately 1116 nucleotide bases from the NADH dehydrogenase subunit 2 (*ND2*) mitochondrial gene and adjacent tRNAs. Sequencing protocols are described in Appendix S1. We edited sequences in GENEIOUS v 6.1.4 (Biomatters Ltd.). We added 37 *P. jordani* sequences, and one *P. metcalfi* sequence that clustered with *P.*

jordani, from the Weisrock and Larson (2006) study to the data set, for a total of 362 individuals from 70 localities (Figure 1). We aligned all sequence data using the ClustalW algorithm in GENEIOUS, and translated protein-coding regions into amino acids to check alignment. Sequences generated for this study have been deposited in GenBank (Appendix S4).

Phylogenetic analysis

We used JMODELTEST v 2.1.3 to determine the most appropriate model of nucleotide substitution (Guindon & Gascuel, 2003; Darriba *et al.*, 2012), and inferred phylogenetic relationships among haplotypes using Maximum-likelihood and Bayesian-inference methods. We included haplotypes from *P. metcalfi*, *P. shermani*, and *P. teyahalee* in the phylogenetic analyses, and *Plethodon ouachitae* was selected as an outgroup. We constructed the Maximum-likelihood phylogeny using RAXML-HPC2 (Stamatakis, 2014) on the CIPRES Science Gateway (Miller *et al.*, 2010). We partitioned the data by codon position, applied the general time-reversible model with gamma-distributed rate heterogeneity (GTRGAMMA) to each partition, and ran 1000 bootstrap replicates to evaluate nodal support.

We estimated the phylogeny in a Bayesian framework using MRBAYES v 3.2.2 (Ronquist *et al.*, 2012) on the CIPRES Science Gateway. We conducted two independent runs of four Markov chains, set the branch length prior as an unconstrained exponential with a mean of 0.01 to reduce bias in branch-length estimates (Brown *et al.*, 2010), and lowered the value of the temperature parameter to 0.10 to facilitate better mixing over

the joint posterior probability density. Markov chain Monte Carlo (MCMC) analyses were run for 2.5 million generations, sampling every 250 generations.

We assessed MCMC convergence in three ways. First, we checked that the average standard deviation of split frequency values between runs approached 0.005 and that the potential scale reduction factor (PSRF) for each parameter was greater than or equal to 1.0. Second, we plotted the log-likelihood and sampled parameter values for each run against generation time using TRACER v 1.5 (Rambaut & Drummond, 2007) to ensure that posterior probabilities had been sampled sufficiently for each parameter (ESS > 200). Third, we used the online program AWTY (Wilgenbusch *et al.*, 2004; Nylander *et al.*, 2008) to (1) assess the stability of the cumulative split frequency across 10 increments for each run, (2) verify a strong correlation between the posterior probabilities of all splits in each run, and (3) confirm similar symmetric tree-difference scores (Penny & Hendy, 1985) within and between runs. The first 25% of sampled trees were discarded as burn-in, and post-burn-in trees from both runs were combined to estimate a 50% majority-rule consensus tree.

We calculated haplotype diversity (h) and nucleotide diversity (π), and estimated corrected nucleotide divergence among lineages (*Dxy* and *Da*; Nei, 1987) in DNASP v 5 (Librado & Rozas, 2009). We tested for significant associations (Spearman's rho) between stability and genetic diversity (h and π) over all *P. jordani* populations and among populations in the two largest lineages.

RESULTS

Distribution modeling outputs

The mechanistic model predicted the current distribution of *P. jordani* to be approximately half the size of the LGM distribution, with an overlap of 477 km^2 (Table 1), suggesting that the range has shrunk and shifted to higher elevations since the LGM (Figure 2). Despite having a broad distribution of suitable microclimate at the LGM, the maximum energetic yield per unit foraging was less than under current climatic conditions (LGM range: -23.9 – 2.3 kJ, current range: -24 – 10 kJ). Stable areas (i.e., those that were predicted suitable at the LGM and today) comprised 68% of the current distribution and were predicted to occur at mid-elevations (677.87 – 1597.94 m; Table 1).

We present results of the correlative model based on the 10th percentile training presence logistic threshold (0.294), which is the lowest threshold value that did not predict the bottom 10% of localities as suitable (Kozak & Wiens, 2006); predicted distributions in GRSM obtained by applying the other thresholds were similar (results not shown). The average test AUC for the ten replicate runs was 0.968 (± 0.005). Under current climate conditions, the correlative model generally exhibited spatial concordance with the mechanistic model (Figure 2), but predicted a larger suitable area (Table 1). Under LGM conditions, the distribution was predicted to shift to southern latitudes (Appendix S5), with small patches of suitable climate occurring at low- and mid-elevations in western GRSM at the LGM (Figure 2; Table 1). The stability distribution was minimal (ca. 5% of the current distribution; Table 1) and fragmented (Figure 2; Appendix S5). The elevational range of these areas in GRSM was similar to, but more restricted than, the stable area predicted by the mechanistic model (Table 1).

All ten of the correlative model runs showed high clamping in the area corresponding to present-day GRSM, meaning that variables outside of the model training range were truncated to the limit of the training range (Appendix S5; Phillips *et al.*, 2006). High clamping can be an artifact of building models using a relatively small extent surrounding species localities (Anderson & Raza, 2010); however, model runs projected onto larger (and smaller) study regions also exhibited high clamping (results not shown), indicating that variables reach the limits of their training range regardless of the extent of the climate data used to build the models. The MESS maps revealed negative values across the projected area at the LGM because at least one variable was outside the range of the training data (Appendix S5). The most dissimilar variable (i.e., the variable with greatest deviation from the training range) was temperature annual range (the difference between the maximum temperature of the warmest month and the minimum temperature of the coldest month). Taken together, these results show that transferring the correlative model across time entails considerable extrapolation into non-analogous climate conditions driven by differences in temperature extremes from the LGM to present. Predictions of the past distribution of *P. jordani* based on this model should therefore be treated with caution (Elith *et al.*, 2011; Anderson, 2013). We keep this caveat in mind in our discussion.

mtDNA phylogeny and diversity

Model selection based on Akaike information criteria favored the TIM2 model with a gamma distribution ($\alpha = 0.299$). Phylogenetic trees from Maximum-likelihood and

Bayesian inference analyses exhibited identical topologies at all major nodes and revealed three reciprocally monophyletic lineages of *P. jordani* (Figure 3). All major nodes were strongly supported by Bayesian posterior probabilities > 0.99.

The widespread western and central clade (Clade 1; Figure 3) includes 34 haplotypes. It exhibits the highest nucleotide and haplotype diversity of the three lineages, but the latter may be a function of large sample size (Table 2). Clade 2, which is sister to Clade 1, contains two haplotypes reported by Weisrock and Larson (2006) and eight novel haplotypes. Clade 3 comprises four novel haplotypes in addition to the metcalfi-1 haplotype reported by Weisrock and Larson (2006), and exhibits the lowest nucleotide diversity of the three lineages (Table 2). In addition, 18 individuals fall outside of the three main *P. jordani* lineages and its sister species *P. metcalfi* (Weisrock & Larson, 2006; but see Wiens *et al.*, 2006), clustering more closely with *P. metcalfi* and *P. shermani* haplotypes. These individuals are likely a result of past introgression, which is well documented in the *P. jordani* complex (Weisrock *et al.*, 2005; Weisrock & Larson, 2006; Chatfield *et al.*, 2010).

Net corrected sequence divergence between Clades 1 and 2 was 0.61% and between Clades 1 and 2 and Clade 3 was 2.24% (Table 3). Assuming a molecular clock estimate for mitochondria in salamanders of 1.28% per million years (Weisrock *et al.*, 2001), divergences between these clades occurred during the Pleistocene, but much deeper than the LGM. Divergence times estimated using a Bayesian coalescent approach (Luxbacher & Kozak, Chapter 2) also support this conclusion. Sequence divergence between *P. jordani* and *P. metcalfi* (10.8%) agrees with estimates reported by Weisrock and Larson (2006). We found no significant association between stability (i.e., inferred refugia)

and haplotype or nucleotide diversity in *P. jordani* (h: $r = 0.174$, $P = 0.184$; π: $r = 0.168$, $P = 0.200$) or within the two largest clades (Clade 1, h: $r = 0.233$, $P = 0.146$; Clade 1, π: $r = 0.268$, $P = 0.090$; Clade 2, h: $r = 0.389$, $P = 0.123$; Clade 2, π: $r = 0.386$, $P = 0.126$).

DISCUSSION

Previous studies identified large glacial refugia for disparate taxa in the southeastern United States (e.g., Avise, 2000; Waltari *et al.*, 2007), a pattern that predicts regional species expanded from these refugia as climate conditions warmed after the LGM to occupy their current ranges. However, many regionally-distributed taxa exhibit phylogeographic structure that suggests populations persisted in multiple disjunct areas (e.g., Austin *et al.*, 2004; Soltis *et al.*, 2006; Fontanella *et al.*, 2008). The occurrence of refugia within the southern Appalachians has been proposed (Church *et al.*, 2003; Lee-Yaw *et al.*, 2008; Walker *et al.*, 2009), but the extent and location of such refugia had not previously been considered. Our modeling output, phylogeographic structure, and genetic diversity results suggest that *P. jordani* has maintained populations near its current distribution since its origins and through Pleistocene glacial cycles, and support the idea that the southern Appalachian Mountains harbored glacial refugia or microrefugia (Rull, 2009).

Distribution modeling outputs

Outputs from correlative and mechanistic models suggest that suitable microclimate

for *P. jordani* could be found within the Great Smoky Mountains since the LGM, but the extent is considerably less for the correlative model. The distribution of *P. jordani* at the LGM inferred from the mechanistic model coincides with pollen data, which suggest the highest elevations in the Smokies were covered by tundra and boreal forest (Whittaker, 1956; Delcourt & Delcourt, 1980), habitats that were likely too cold for *P. jordani*. The finding that predicted refugia occurred at mid-elevations fits with our understanding of how montane species shift their elevational range during periods of climate change (e.g., Weisrock & Larson, 2006; Galbreath *et al.*, 2009). The broader distribution predicted by the mechanistic model at the LGM compared to current conditions may seem surprising, but plethodontids undergo dramatic seasonal variation in body temperature, and temperate species can be found at near-freezing temperatures in winter (Feder, 1983). Because they spend much of their time in moist covers that can differ from atmospheric temperature and ground temperature (Feder, 1983), they can remain sheltered from broad-scale changes in climate (Wake, 2009).

Other studies that have compared mechanistic and correlative models found that predicted distributions between them varied widely when forecasted onto future climate conditions (Buckley *et al.*, 2010; Elith *et al.*, 2010). Our mechanistic and correlative model outputs are not directly comparable because they were built using different sets of climatic variables; however, we can draw general conclusions by considering the extent to which independent data support each model. We hold greater confidence in the mechanistic model output because it matches predictions based on pollen data, is consistent with phylogeographic data (see below), and does not require extrapolation onto novel climate conditions. The failure of the correlative model to predict a past

distribution for *P. jordani* that is consistent with independent data suggests that correlative modeling may be insufficient at forecasting the future distribution of this species (Milanovich *et al.*, 2010).

One potential caveat with mechanistic modeling is the assumption that a species' climatic tolerances have not changed significantly over the time period being modeled. *Plethodon jordani* meets this assumption because allopatric sister species of plethodontids that diverged during the Pliocene continue to occupy similar climatic niches that do not span lower-elevation habitats occurring between their distributions (Kozak & Wiens, 2006). This pattern suggests that climatic tolerances in *Plethodon* species are evolutionarily conserved over the time scale for which we have modeled changes in the distribution of *P. jordani*.

Phylogeographic patterns

Although *P. jordani* occupies a narrow range across a continuous chain of mountains, our phylogeographic analysis showed three geographically structured clades, including one not revealed in previous phylogenetic studies of the species (Highton & Peabody, 2000; Weisrock & Larson, 2006). Clade 1 is widespread, including individuals from 46 sampled localities in the western and central parts of the range. Clade 2 includes individuals from 24 localities in the northeast part of the range, and is sister to Clade 1. The steep topography and a stream (Bradley Fork) appear to define a southern boundary between these two lineages, but no apparent geographic or environmental boundary occurs to the north. Clade 3 is sister to Clades 1 and 2 and comprises individuals from

nine localities in the southeastern corner of the range. Descending topography separating two ridgelines (Balsam Mountain and Hyatt Ridge) may form a western barrier to gene flow for individuals in this clade. Weisrock and Larson's (2006) metcalfi-1 haplotype is nested within this lineage, suggesting introgression (e.g., *P. metcalfi* nuclear DNA and *P. jordani* mtDNA). Balsam Mountain, where the individual was sampled, is a known hybrid zone between these two species as well as between these and *P. teyahalee* (Chatfield *et al.*, 2010).

We interpret the pattern of genetic structure in a continuous montane landscape lacking many obvious barriers to gene flow to be the result of past isolation and subsequent range expansion (Avise, 2000; Hewitt, 2004; but see Irwin, 2002 for discussion of the possible formation of phylogeographic breaks in the absence of geographic isolation in dispersal-limited species). As climatic tolerances are a key factor underlying the dispersal ability and distribution of *P. jordani*, individuals likely tracked their climatic niche over time (Feder, 1983; Kozak & Wiens, 2006). Cyclical climate fluctuations during the Pleistocene would have caused expansions and contractions of suitable habitat that influenced diversification by isolating populations (Webb & Bartlein, 1992; Soltis *et al.*, 2006; Parra-Olea *et al.*, 2012; Devitt *et al.*, 2013). Given the different ages of the three clades, it is likely that the current genetic pattern was established during two different glacial advances.

During expansion, populations that became fixed for different haplotypes may have re-established contact and interbred, resulting in introgression between differentiated clades (Jansson & Dynesius, 2002). Our results are suggestive of such secondary contact in two areas that correspond with the phylogeographic breaks. At the geographic

junction of Clades 1 and 2, individuals exhibit haplotypes from both clades at three localities (40, 48, and 54; Figure 1). Because incomplete lineage sorting is not expected to leave a predictable biogeographic pattern of this nature (Funk & Omland, 2003), we interpret this pattern as secondary contact. Likewise, individuals sampled from five localities in the southeastern region of GRSM share haplotypes from Clades 2 and 3, suggesting secondary contact between these divergent lineages (Figure 1). Secondary contact is also supported by the fact that populations in this zone all exhibit higher than average measures of nucleotide diversity (Appendix S3; Petit *et al.*, 2003, Canestrelli *et al.*, 2007). Both of these junctions between clades occur in areas that were not climatically stable from the LGM to today, suggesting they were areas of contact between lineages as they expanded from refugia (Moritz *et al.*, 2009).

In many cases, mtDNA alone is sufficient to recover phylogeographic patterns (reviewed in Zink & Barrowclough, 2008; Barrowclough & Zink, 2009), though nuclear loci can improve descriptions of phylogeographic history and elucidate underlying evolutionary processes (Hare, 2001; Edwards & Bensch, 2009). We acknowledge the potential limitations of drawing conclusions from a single locus in this study, but do so to outline hypotheses that can be further evaluated using independent loci (Luxbacher & Kozak, Chapter 2).

Possible explanations for differing results between distribution models & mtDNA patterns

The three geographically structured clades suggest past isolation in three refugia, but

neither stability map depicts three discrete refugia. These results do not support our hypothesis that phylogeographic structure is spatially correlated with refugia inferred from the models, but we do not consider them to be incompatible. Rather, we conclude that the mechanistic model output and phylogeographic structure are both capturing aspects of the history of *P. jordani* at different points in time.

The LGM was a major climatic change that affected the distribution of contemporary temperate species (Webb & Bartlein, 1992; Stewart & Lister, 2001) and is often used to approximate conditions that caused populations to shift repeatedly across glacial cycles (e.g., Lawson, 2013). Paleoclimate model outputs closely correspond to phylogeographic signatures in many species (e.g., Graham *et al.*, 2006; Cordellier & Pfenniger, 2009; Knowles & Alvarado-Serrano, 2010), but the time frames for which paleoclimate data are available are too recent to match older divergences in other species (e.g., Thomé *et al.*, 2010; Werneck *et al.*, 2012). The three *P. jordani* lineages represent diversification that occurred earlier than the LGM, whereas the mechanistic model output reflects the possible distribution of these lineages 21 ka. The deep phylogeographic structure we observe conflicts with the correlative model output, which predicts that *P. jordani* colonized its range recently (i.e., since the LGM). Other range-restricted plethodontids have also exhibited incongruence between phylogeographic structure and refugial predictions derived from correlative models (Waltari *et al.*, 2007), a shared finding that highlights the importance of cross-validating distribution model output with independent data.

Climatic stability is a strong predictor of genetic and species diversity in some dispersal-limited taxa (Graham *et al.*, 2006; Carnaval *et al.*, 2009), but stability was not

correlated with diversity within *P. jordani*, nor was mtDNA diversity localized to any specific area. Although a general trend of northern expansion has been observed many wide-ranging species in eastern North America (e.g., Lee-Yaw *et al.*, 2008), the mechanistic model suggests climate-driven range shifts within the Great Smoky Mountains that may have maintained diversity (Hewitt, 1996). In addition, the broader distribution predicted for *P. jordani* at the LGM may have sustained large population sizes that would not leave a signature of reduced diversity in contemporary populations (Mayr, 1954; Hewitt, 1999). Further investigation into the divergence times within *P. jordani* and demographic signatures of expansion based on multiple genetic markers will provide insight into population dynamics during the Pleistocene (Luxbacher & Kozak, Chapter 2).

CONCLUSIONS

The Great Smoky Mountains encompass the most species-diverse area in temperate North America (Sharkey, 2001), and are positioned within a concentration of contact zones and phylogeographic breaks among salamander and invertebrate taxa (Rissler & Smith, 2010; Garrick, 2011). The potential existence of climatically stable areas in the southern Appalachian Mountains had not been investigated formally or considered at a fine enough spatial scale to determine whether they could have harbored refugia for narrowly endemic species. We show through mechanistic modeling, phylogeographic structure, and genetic diversity that *P. jordani* maintained populations within its current distribution across multiple Pleistocene glacial cycles by tracking suitable

microclimates via small changes in elevational range. Our results provide support for the idea that long-term climatic stability has promoted high species richness and endemism in the region.

Other regional species that are dispersal-limited and restricted to high elevations (e.g., gastropods, insects and entognathous hexapods) may have experienced similar patterns of past isolation during Pleistocene climate fluctuations (Nalepa *et al.*, 2002), and may show similar patterns of structure (Thomas & Hedin, 2008). Montane species are considered most at risk due to climate change (Williams *et al.*, 2003; Thomas *et al.*, 2004; La Sorte & Jetz, 2010), and the southern Appalachians are predicted to experience significant shifts in species composition under future climate conditions (Delcourt & Delcourt, 1998). Investigating the generality of phylogeographic patterns in other regional montane species is warranted to elucidate processes driving diversification and maintaining biodiversity in this topographically complex area.

Table 1 Elevational range, area and area ratio of predicted distributions of *P. jordani* inferred from mechanistic and correlative models. Area ratio is calculated as current suitable area / LGM suitable area

Model	Elevational range (m a.s.l.)	Suitable area (km^2)	Area ratio
Mechanistic			
Current	677.87 – 2022.53	699.77	0.55
LGM	266.15 – 1597.95	1262.21	
Stability	677.87 – 1597.94	476.92	
Correlative			
Current	694.86 – 2022.53	968.67	2.8
LGM	266.05 – 1576.24	344.88	
Stability	855.39 – 1576.24	47.98	

Table 2 Sample sizes and molecular diversity indices of three *P. jordani* clades inferred from *ND2* haplotypes

Grouping	N	H	h	π
Clade 1	228	34	0.844	0.00232
Clade 2	92	10	0.720	0.00112
Clade 3	22	4	0.398	0.00039
All *P. jordani*	361	62	0.918	0.01753

N: number of individuals sequenced; H: number of haplotypes; PH: number of private haplotypes; h: haplotype diversity; π: nucleotide diversity

Table 3 Corrected nucleotide divergence between lineages. *Dxy* is the average number of pairwise nucleotide substitutions per site between lineages calculated using Nei's equation 10.20; *Da* is the net number of nucleotide substitutions between lineages calculated using Nei's equation 10.21 (Nei, 1987)

Comparison	*Dxy* ± SD	*Da* ± SD
Clade 1 – Clade 2	0.00782 ± 0.00043	0.00607 ± 0.00043
Clades 1 & 2 – Clade 3	0.02491 ± 0.00318	0.02237 ± 0.00318
Clades 1, 2, & 3 – *P. metcalfi*	0.11153 ± 0.05876	0.10831 ± 0.05876

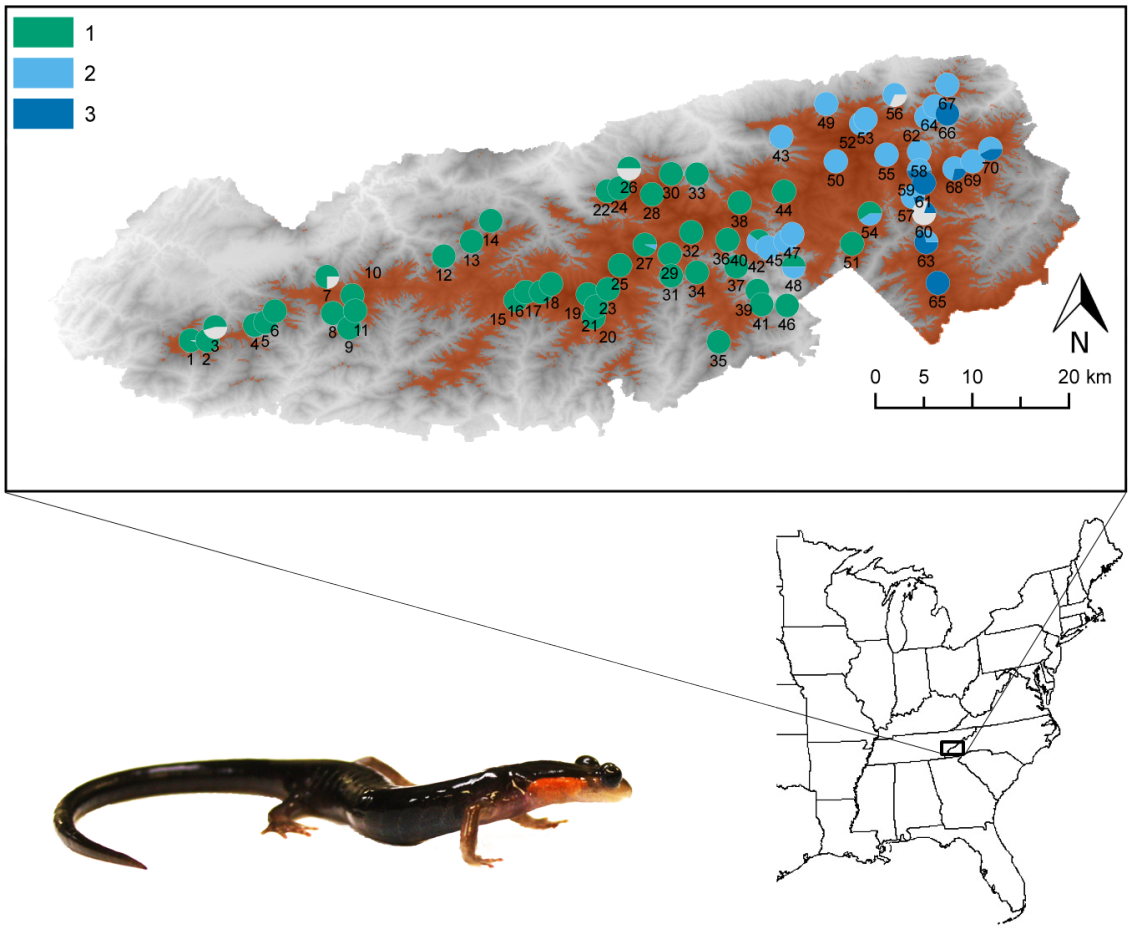


Figure 1 Map of geographic distribution of *P. jordani* and 70 sampling localities in Great Smoky Mountains National Park (North Carolina and Tennessee). Distribution (based on a mechanistic bioenergetic model; Gifford & Kozak, 2012) is depicted in vermilion and overlaid on a 30 m digital elevation model. Localities are numbered as in Appendix S3 and presented as pie diagrams with slice size proportional to the frequency of the major clades identified by phylogenetic analyses. Clade colors are the same as in Figure 3; proportion of individuals not included in three major clades is shown in pale gray

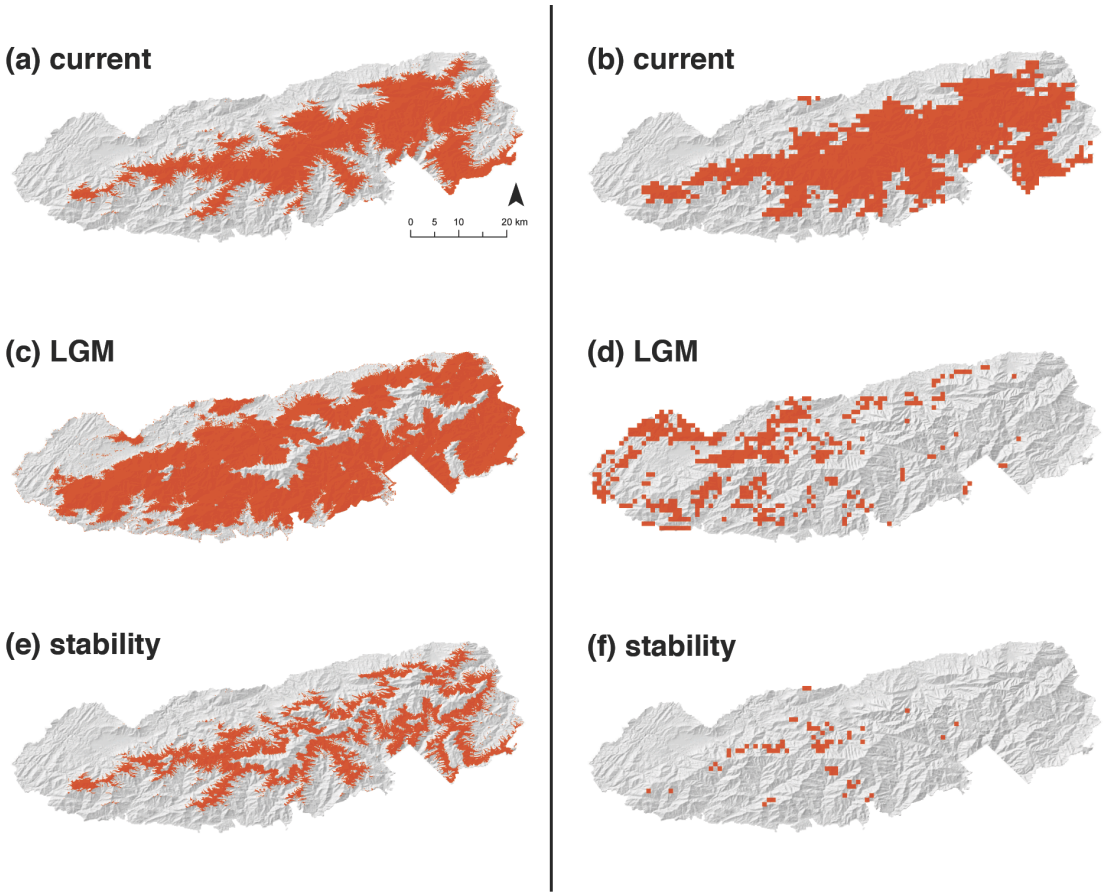


Figure 2 Predicted distributions of *P. jordani* based on mechanistic (left panel) and correlative (right panel) models. (a) and (b), current conditions; (c) and (d), Last Glacial Maximum (LGM; ca. 21,000 ka) conditions (see Appendix S1 for details); (e) and (f), stability (pixels predicted to be suitable under LGM and current conditions). MAXENT model outputs were built on larger extents, but clipped to Great Smoky Mountains National Park for illustration (see Appendix S5 for full extent of predicted distributions). Predicted suitable areas (shown in vermillion) were based on positive energy budget (discretionary energy > 0 kJ) and the 10th percentile training presence logistic threshold (0.294) for mechanistic and correlative model outputs, respectively. Distributions are overlaid on a 30 m digital elevation model

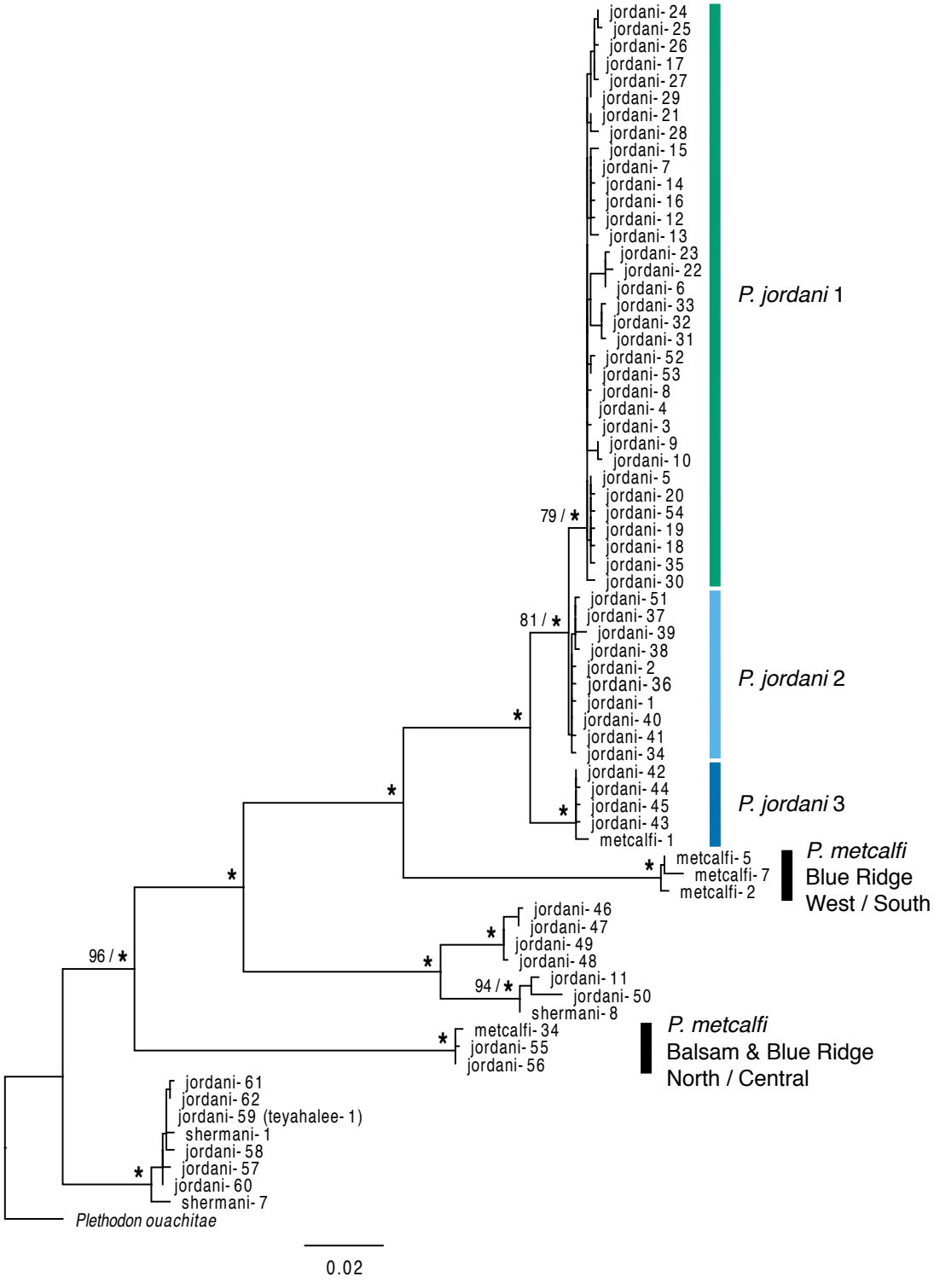

jordani- 24
jordani- 25
jordani- 26
jordani- 17
jordani- 27
jordani- 29
jordani- 21
jordani- 28
jordani- 15
jordani- 7
jordani- 14
jordani- 16
jordani- 12
jordani- 13
jordani- 23
jordani- 22
jordani- 6
jordani- 33
jordani- 32
jordani- 31
jordani- 52
jordani- 53
jordani- 8
jordani- 4
jordani- 3
jordani- 9
jordani- 10
jordani- 5
jordani- 20
jordani- 54
jordani- 19
jordani- 18
jordani- 35
jordani- 30
79 / *
jordani- 51
jordani- 37
81 / *
jordani- 39
jordani- 38
jordani- 2
jordani- 36
jordani- 1
jordani- 40
jordani- 41
jordani- 34
jordani- 42
jordani- 44
jordani- 45
jordani- 43
metcalfi- 1
metcalfi- 5
metcalfi- 7
metcalfi- 2
jordani- 46
jordani- 47
jordani- 49
jordani- 48
96 / *
jordani- 11
94 / *
jordani- 50
shermani- 8
metcalfi- 34
jordani- 55
jordani- 56
jordani- 61
jordani- 62
jordani- 59 (teyahalee- 1)
shermani- 1
jordani- 58
jordani- 57
jordani- 60
shermani- 7
Plethodon ouachitae
0.02
P. jordani 1
P. jordani 2
P. jordani 3
P. metcalfi
Blue Ridge
West / South
P. metcalfi
Balsam & Blue Ridge
North / Central

Figure 3 Maximum-likelihood and Bayesian phylogeny for *P. jordani* based on *ND2* haplotypes with *P. ouachitae* used as an outgroup. Numbers above branches indicate bootstrap support and posterior probabilities, respectively; strongly supported nodes (bootstrap support ≥ 99 and posterior probabilities ≥ 0.99) are denoted by asterisks. *Plethodon jordani* haplotypes 1-11, and *P. metcalfi* and *P. shermani* haplotypes are the same as in Weisrock and Larson (2006); jordani-59 is identical to Weisrock and Larson's teyahalee-1 haplotype

CHAPTER 2

Rangewide multilocus phylogeography and historical demography of a temperate montane endemic salamander, *Plethodon jordani*[2]

SUMMARY

We investigated the consequences of Pleistocene climatic changes on narrow-ranging temperate mountain species by focusing on *Plethodon jordani*, a salamander that is endemic to the Great Smoky Mountains in eastern North America. We inferred genealogies from three anonymous nuclear loci, estimated divergence times, modeled past demographic processes, and reconstructed the phylogeographic history and ecological niches of cryptic lineages within *P. jordani*. A multilocus nuclear network supports differentiation into two phylogeographic groups in the east and west of the distribution. Divergence within lineages corresponds to mid- to late-Pleistocene glaciations when populations presumably shifted to lower elevations. Effective population sizes were relatively constant, indicating long-term persistence and stability of populations in spite of shifting distributions in response to glacial cycles. Phylogeographic reconstruction shows that the range of ancestral *P. jordani* was in the east-central Great Smoky Mountains and subsequently expanded, mainly to the west. Topography and niche conservatism appear to have constrained the distribution of *P. jordani* to along the main ridgeline. These findings highlight the importance of complex topography in creating opportunities for divergence in montane regions that do not

[2] A version of this chapter is being prepared for publication:
Luxbacher AM, Kozak KH. Rangewide multilocus phylogeography and historical demography of a temperate montane endemic salamander, *Plethodon jordani*. *Molecular Ecology*

exhibit strong environmental gradients. Our results provide insight into factors influencing lineage diversification in a North American biodiversity hotspot, and support the idea that the southern Appalachian Mountains harbored Pleistocene refugia for montane taxa.

INTRODUCTION

During periods of climatic change, montane endemic species typically shift their elevational ranges to track suitable climatic and habitat conditions (e.g., Davis & Shaw, 2001; Guralnick, 2007; Tingley *et al.*, 2009), and these distributional shifts have genetic consequences for populations (Hewitt, 1996, 2004). In many montane regions, periods of cool climate cause lineages that were once allopatric to shift to lower elevations, resulting in a mixing of gene pools that can obscure signatures of local differentiation (e.g., Weisrock & Larson, 2006; Galbreath *et al.*, 2009; Velo-Anton *et al.*, 2013); periods of warm climate result in shifts to higher elevations that can promote allopatric divergence (e.g., Knowles, 2001; Ribas *et al.*, 2007; Qiu *et al.*, 2009; Schoville *et al.*, 2012). Distributional shifts are often accompanied by population expansions or contractions that produce demographic changes over time (Davis & Shaw, 2001; Lessa *et al.*, 2003; Pauls *et al.*, 2006; Schoville *et al.*, 2011). These patterns have been recognized in areas with steep environmental gradients, including sky islands (Knowles, 2000, 2001; Carstens & Knowles, 2007, Shepard & Burbrink, 2008, 2009; Lawson, 2010; Qu *et al.*, 2011; Schultheis *et al.*, 2012) and tropical mountains (Streicher *et al.*, 2009; Bell *et al.*, 2010; Barker *et al.*, 2011; Parra-Olea *et al.*, 2012). However, mountain ranges worldwide

differ in their topography, connectivity, and climatic zonation (Graham *et al.*, 2014). Investigating the extent to which species' distributions, diversification, and demography are affected by climatic changes in other montane systems will improve our understanding of factors influencing population persistence and connectivity in topographically complex regions. This information is necessary to assess species' responses to projected future climate change in these particularly diverse and vulnerable regions (Ohlemüller *et al.*, 2008; La Sorte & Jetz, 2010, McCain & Colwell, 2011).

The southern Appalachian Mountains in eastern North America include a number of smaller chains of mountains that are characterized by a complex topography of gorges and valleys separated by ridges, but lack the climatic zonation of tropical mountains (Janzen, 1967; Cadena *et al.*, 2011). The Great Smoky Mountains are one chain that includes some of the highest peaks and species richness in the region (Sharkey, 2001; Soltis *et al.*, 2006), with many narrow endemic species that continue to be described (Nichols & Langdon, 2007). During Pleistocene glacial periods, high elevations in these mountains were considerably cooler than today (Whittaker, 1956) and vegetation communities shifted to lower elevations (Delcourt & Delcourt, 1980). As the climate warmed since the Last Glacial Maximum (LGM), spruce-fir, hemlock, and deciduous forests replaced tundra and boreal forest ecosystems at high elevations (Whittaker, 1956). There is also some evidence to suggest the existence of pockets of stable habitat or microclimate in these mountains that may have sheltered populations from broad-scale climate fluctuations that occurred since the LGM (Chapter 1; Delcourt & Delcourt, 1998), raising the possibility that some species may not have undergone extensive range shifts. These mountains thus present an interesting system to investigate the extent that

shifting distributions may have influenced population stability and genetic connectivity during periods of past climatic change.

Climatic niche conservatism played an important role in the speciation and formation of elevational patterns of species richness and endemism among plethodontid salamanders in the southern Appalachians (Kozak & Wiens, 2006, 2010). However, the genetic consequences of niche conservatism on intraspecific diversification and population dynamics on recent (i.e., Pleistocene and Holocene) time scales are uncertain (Crespi *et al.*, 2003). *Plethodon jordani* is a montane endemic salamander species whose distribution lies entirely within the boundaries of Great Smoky Mountains National Park (GRSM). In Luxbacher *et al.* (Chapter 1), we identified three geographically-circumscribed monophyletic mitochondrial lineages within *P. jordani* that suggest the existence of three Pleistocene refugia. However, refugia inferred from two different models of the LGM distribution did not coincide with this phylogeographic pattern, suggesting deeper divergences. The correlative model output indicated that much of *P. jordani*'s distribution was displaced to southern latitudes during the LGM. In contrast, phylogeographic data and the mechanistic model output showed that *P. jordani* maintained populations within its current distribution across multiple glacial cycles by tracking suitable microclimates via small changes in elevational range. In this study, we aim to clarify these conflicting ideas and gain insight into historical processes underlying this phylogeographic pattern using multilocus data.

We address four objectives. First, because multiple markers can provide greater resolution of phylogeographic patterns than a single locus (Edwards & Bensch, 2009), we infer genealogies from anonymous nuclear loci and assess their concordance with

patterns of structure observed in mitochondrial DNA (mtDNA). Second, we estimate divergence times to evaluate the temporal context of diversification within *P. jordani*. Third, we model past demographic processes to test for changes in effective population size over time. For this objective, we evaluate support for two alternative hypotheses based on results presented in Chapter 1: (1) Ancestral populations of *P. jordani* were located in the east before expanding and diverging westward, as inferred from phylogeographic structure and supported by the mechanistic model; and (2) *P. jordani* recently colonized GRSM from the west, as suggested by the correlative model output (Figure 1). The first hypothesis predicts a signature of expansion in the western part of the range, but relative stability since at least the LGM; the second hypothesis predicts expansion across GRSM, particularly since the LGM and in the eastern part of the range. Our fourth objective is to reconstruct the phylogeographic history and ecological niches of *P. jordani*. For this objective, we test the hypothesis that dispersal rates varied across lineages against a null hypothesis of uniform dispersal, and consider whether lineages occupy different niches. Overall, our results offer insight into the demographic and phylogeographic history of *P. jordani* that help us better understand how narrow-ranging montane species responded to Pleistocene climate fluctuations in topographically complex regions, and have implications for predicting species responses to future climate change.

MATERIALS AND METHODS

Sampling and DNA extraction

We collected tail tip tissue from 324 *P. jordani* individuals from 66 localities in GRSM between 2009 and 2011. Localities were chosen to cover the known geographic distribution of the species (Figure 2). Tail tips were stored in 95% ethanol initially, and later frozen at -80°C for long-term storage in the Bell Museum of Natural History. We extracted DNA from these samples using a Blood and Tissue Genomic DNA Miniprep System (Viogene) following the manufacturer's instructions. We also extracted DNA from 31 *P. jordani* liver samples that were housed at the Bell Museum; these samples were collected in 2008 from four additional localities in GRSM.

DNA amplification and analysis

We amplified approximately 1116 nucleotide bases from the *ND2* mitochondrial gene and adjacent tRNAs using primers and conditions described in Chapter 1 (Table 1). We also sequenced three anonymous nuclear loci from individuals that were selected to cover the geographic distribution and represent the three mitochondrial clades (Table 1; Shepard *et al.*, in review). For these loci, we performed PCR in 10 μL reaction volumes consisting of 1 μL DNA template, 5 μL 2X GoTaq® Green Master Mix (Promega), 0.5 μL each of locus-specific forward and reverse primers, and 3 μL ddH_2O. Thermocycler conditions for amplification were as follows: 2 min initial denaturation at 94°C; 35 cycles of denaturation at 94°C for 20 s, annealing at 54°C (56°C for locus 16) for 30 s, and extension at 72°C for 90 s; final extension at 72°C for 5 min. We purified PCR products with ExoSAP-IT® (Affymetrix). Products were sequenced at the University of

Minnesota BioMedical Genomics Center on an ABI 3730xl DNA Analyzer.

We visualized chromatograms and edited sequences in GENEIOUS v 6.1.4 (Biomatters Ltd.). For the nuclear loci, ambiguity codes were used to represent polymorphisms in heterozygous individuals. Haplotypes for individuals with heterozygous indels were resolved in the program INDELLIGENT v 1.2 (Dmitriev & Rakitov, 2008). All sequences for each locus were then aligned using the ClustalW algorithm in GENEIOUS.

We inferred nuclear haplotypes computationally from genotypic data using PHASE v 2.1.1 (Stephens *et al.*, 2001; Stephens & Donnelly, 2003) with input and output files prepared using SEQPHASE (Flot, 2010). Each locus was analyzed with 1000 burn-in steps, a thinning interval of 1, and 1000 final steps. We specified a parent-independent mutation mechanism, and included a file containing known phases of heterozygotes that were resolved with INDELLIGENT. Following results of a simulation study (Garrick *et al.*, 2010), we reduced the phase call threshold to 0.6; any heterozygous sites below this threshold were coded with IUPAC ambiguity codes in both alleles. We performed five separate runs for each locus using different random starting seeds to check for consistency across runs, and inferred haplotypes from the run that received the highest average value for goodness of fit.

We calculated nucleotide diversity (π) in DNASP v 5 (Librado & Rozas, 2009). Ambiguous sites were recoded as "N" and not used in calculations (Zasada *et al.*, 2014). We used TOPALI v 2.5 (Milne *et al.*, 2004) to test for recombination in nuclear loci using the difference of sums of squares (DSS) method with a 100 bp sliding window and 10 bp step size. The statistical significance of DSS peaks was assessed with 100 parametric bootstrap replicates. We tested whether selection was acting on *ND2* with

McDonald–Kreitman tests (McDonald & Kreitman, 1991) in DNASP. We excluded non-coding sequences from the analyses, and ran comparisons of the three mitochondrial clades to each other as well as all three clades combined to the sister species, *P. metcalfi*. We assessed statistical significance with Fisher's exact tests. A Hudson–Kreitman–Aguadé test (HKA test; Hudson *et al.*, 1987) was used to evaluate evidence for selection acting on the nuclear loci. Sequences from *P. ouachitae* (the closest species for which we could obtain data from all three loci) were used for the outgroup.

Phylogenetic analyses

We selected the best-fit model of nucleotide substitution for each locus based on Akaike Information Criterion using JMODELTEST v 2.1.3 (Guindon & Gascuel, 2003; Darriba *et al.*, 2012). We inferred phylogenies for each nuclear locus independently using Maximum Likelihood and Bayesian methods. Maximum-likelihood analyses were performed using RAXML-HPC2 (Stamatakis, 2014) on the CIPRES Science Gateway (Miller *et al.*, 2010). For each locus, we applied a GTR model with gamma-distributed rate heterogeneity (GTRGAMMA) and ran 1000 bootstrap replicates to evaluate nodal support.

Bayesian analyses were conducted using MRBAYES v 3.2.2 (Ronquist *et al.*, 2012) on the CIPRES Science Gateway. For each locus, we ran two independent runs of four Markov chains. We set the branch length prior as an unconstrained exponential with a mean of 0.01 to reduce bias in branch-length estimates (Brown *et al.*, 2010), and lowered the value of the temperature parameter to 0.10 to facilitate better mixing over the joint

posterior probability density. MCMC analyses were run for 20 million generations, sampling every 1000 generations.

We assessed MCMC convergence in three ways. First, we checked that the average standard deviation of split frequency values between runs approached 0.002 and that the potential reduction scale factor (PSRF) for each parameter was equal to 1.0. Second, we plotted the log-likelihood and sampled parameter values for each run against generation time using TRACER v 1.5 (Rambaut & Drummond, 2007) to ensure that posterior probabilities had been sampled sufficiently for each parameter (ESS > 200). Third, we used the online program AWTY (Wilgenbusch *et al.*, 2004; Nylander *et al.*, 2008) to (1) assess the stability of the cumulative split frequency across 10 increments for each run, (2) verify a strong correlation between the posterior probabilities of all splits in each run, and (3) confirm similar symmetric tree-difference scores (Penny & Hendy, 1985) within and between runs. The first 5000 samples of each run were discarded as burn-in. For each locus, post-burn-in trees from both runs were combined to estimate a 50% majority-rule consensus tree.

Because networks can represent intraspecific genetic variation more appropriately than phylogenetic trees (Posada & Crandall, 2001), we also investigated genealogical relationships by constructing separate statistical parsimony networks for *ND2* and each nuclear locus using a 95% connection limit in TCS v 1.21 (Clement *et al.*, 2000). To build a multilocus nuclear network, we first generated individual allele distance matrices for each nuclear locus under the Tamura-Nei model (Tamura & Nei, 1993) in MEGA v 5.2 (Tamura *et al.*, 2011). We then combined the individual nuclear matrices into a multilocus distance matrix that was equally weighted across loci using POFAD v 1.03

(Joly & Bruneau, 2006). We only included individuals for which we had sequence data for all three loci in this analysis (N = 118). Results were visualized using the NeighborNet algorithm (Bryant & Moulton, 2004) in SPLITSTREE v 4.13.1 (Huson & Bryant, 2006).

Assessment of concordance between mtDNA and anonymous loci

We evaluated how well nuclear data supported patterns of structure inferred from mtDNA using two approaches. First, we tested for associations between nuclear and mitochondrial data with Mantel tests, following Barlow *et al.* (2013). We constructed a matrix in which individuals from the same mitochondrial clade were coded as 1 and individuals from different mitochondrial clades were coded as 2. We tested for associations between the genetic (multilocus nuclear) distance matrix and this mitochondrial clade matrix using Mantel tests, and accounted for the influence of geographic distance with a partial Mantel test. The geographic distance matrix was comprised of Euclidean distances between individual sampling localities that we calculated using the 'Near' tool in ARCGIS v 9.3 (ESRI). We tested for isolation by distance by conducting Mantel tests between genetic and Euclidean distances among pairs of individuals, and controlled for the effect of mitochondrial clade with a partial Mantel test. These analyses were run using the Isolation By Distance Web Service v 3.23 (Jensen *et al.*, 2005), and significance was assessed from 10,000 randomizations.

Second, we assessed the relative degree of divergence in nuclear loci using the genealogical sorting index (*gsi*), a statistic that quantifies exclusive ancestry of specified

groups on a genealogy (Cummings *et al.*, 2008). We used individual ML trees as input and grouped individuals according to mitochondrial clades. The degree of exclusive ancestry is calculated on a scale from 0 (mixed ancestry of groups) to 1 (reciprocal monophyly), and has been shown to detect exclusive ancestry even when groups are in early stages of divergence (i.e., not monophyletic; Cummings *et al.*, 2008). We calculated the *gsi* statistic for each nuclear locus separately and for an ensemble of all three loci (gsi_T). Statistical significance was assessed from 10,000 permutations using the Genealogical Sorting Index web server (http://www.genealogicalsorting.org).

Divergence time and historical demographic analyses

We estimated divergence times within *P. jordani* in BEAST v 1.7.5 (Drummond *et al.*, 2012) using two approaches based on *ND2* sequence data (Schoville *et al.*, 2013). First, we derived an estimate of the split between *P. jordani* and its sister species, *P. metcalfi*, from divergence times estimated by Kozak *et al.* (2009). We set a normal prior on the root (mean = 10.3 Ma and standard deviation = 1.44 Ma to encompass the 95% confidence interval). Second, we placed a normal prior on the substitution rate (mean = 6.4×10^{-3}/site/Myr; Weisrock *et al.*, 2001) and applied a standard deviation of 1.6×10^{-3}/site/million years to encompass the range of *ND2* rates reported by other studies of plethodontids and salamandrids (Mueller, 2006; Wu *et al.*, 2013). For both analyses, we used a GTR model with gamma-distributed rate variation, a strict molecular clock, and a constant coalescent tree model. We used a strict clock because preliminary runs using an uncorrelated lognormal relaxed clock yielded a marginal distribution of the standard

deviation that extended to 0. We used a constant coalescent model because the marginal posterior of the exponential growth rate parameter included 0 in preliminary runs applying an exponential model (Marino *et al.*, 2011). For each analysis, we conducted two MCMC runs of 50 million generations each, sampling every 5000 generations. We combined the log files with a 10% burn-in for each file. We ensured that ESS values > 200 and checked convergence of topologies between the runs in AWTY as described above. We generated a maximum clade credibility (MCC) tree using TREEANNOTATOR v 1.7.5 (Drummond *et al.*, 2012) and visualized it in FIGTREE v 1.2.3 (http://tree.bio.ed.ac.uk/software/figtree).

We examined evidence of historical demographic changes using multiple approaches. First, we calculated Tajima's (1989) *D*, Fu's (1997) *Fs*, and Ramos-Onsins and Rozas' (2002) R_2 statistics for each locus and each mtDNA lineage in DNASP. We tested for significant deviations from constant population size using 10,000 coalescent simulations. Second, we reconstructed demographic histories for each mtDNA lineage using GMRF Bayesian skyride plots implemented in BEAST (Minin *et al.*, 2008). These plots improve upon the summary statistics because they use the coalescent properties of a gene tree to infer fluctuations in effective population size over time. We constructed the plots using time-aware smoothing for each mitochondrial clade. We estimated the rate and specified a *ND2* mutational rate of 1.28% sequence divergence per million years (Weisrock *et al.*, 2001). For clades 1 and 2, all other parameters were identical to the divergence dating analysis described above; for clade 3, we used a HKY model. We conducted two independent MCMC runs of 50 million generations, sampling every 5000 generations, and excluded the first 10% of samples in each run as burn-in. We used TRACER to

assess MCMC convergence by comparing parameter trends between runs and verifying that adequate effective sample sizes (ESS > 200) were reached for each parameter, and to visualize the skyride plots. Following Eytan and Hellberg (2010), we considered changes in effective population size to be significant if the 95% HPD intervals at the root and tip of the plots were non-overlapping.

Third, we conducted an extended Bayesian skyline plot (EBSP) analysis that integrated data from all four loci. Substitution models, clock models, and trees were unlinked for each locus. For *ND2*, we applied the same settings used in the GMRF Bayesian skyride plot analyses; for each nuclear locus, we applied the HKY model. We used strict clock models for all loci, assigned nuclear clock rates a uniform [0,1] prior distribution, and estimated nuclear rates relative to *ND2*, which was fixed at a rate of 0.0064. We performed three independent runs to check for consistency between runs. We ran each analysis for 300 million generations each, sampling every 30,000 generations, and discarded the first 10% as burn-in. We assessed MCMC convergence using TRACER as described above.

Phylogeographic diffusion modeling

We reconstructed the phylogeographic history of *P. jordani* with continuous diffusion models (Lemey *et al.*, 2010) implemented in BEAST. These models use the geographic coordinates associated with sampled individuals to reconstruct localities of internal nodes in a phylogeny and estimate spatial diffusion (i.e., phylogeographic dispersal) across a continuous landscape over time. We compared the fit of a standard Brownian diffusion

(BD) model and a relaxed random walk (RRW) model with gamma distributed rate variation. The BD model assumes a uniform diffusion rate across all branches of the phylogeny; the RRW model allows the diffusion rate to vary along individual branches. We limited these analyses to the mtDNA data set because it includes more extensive individual and geographic sampling than the nuclear loci. We included only individuals from the three mtDNA clades and only one individual representing each haplotype from a given locality (N = 132). Because RRW models cannot accommodate duplicate geographic coordinates among samples, we added random noise to samples sharing the same locality; new coordinates were assigned from a window of 0.001 x 0.001 decimal degrees centered over the original locality (Barlow *et al.*, 2013). We applied the same settings used in the GMRF Bayesian skyride plot analyses. For each model, we conducted two independent MCMC runs of 50 million generations, sampling every 5000 generations. We used TRACER to ensure adequate effective sample sizes for each parameter (ESS > 200) and determine burn-in.

We compared the fit of BD and RRW models with marginal-likelihood estimation. Specifically, we used path and stepping-stone sampling estimators, which have been shown to provide more accurate and consistent results than harmonic-mean estimators (Baele *et al.*, 2012). Analyses consisted of a chain of 1-million steps, sampled every 1000 steps and 100 path steps. We calculated log Bayes factors by subtracting the log marginal likelihood estimates for each estimator (Camargo *et al*., 2013), and selected the model that best fit the data following the guidelines proposed by Kass & Raftery (1998). We used the selected model to reconstruct the phylogeographic diffusion pattern for *P. jordani*.

After discarding the first 10% of trees from the run of the selected model that yielded the highest (i.e., least negative) log likelihood as burn-in, we estimated a MCC tree from the remaining trees using TREEANNOTATOR. We projected this tree onto geographic space and generated 80% HPD polygons representing uncertainty in the coordinates of ancestral locations using SPREAD v 1.0.6 (Bielejec *et al.*, 2011). We visualized the resulting phylogeographic diffusion pattern in Google Earth (http://earth.google.com).

Ecological niche modeling

We modeled the ecological niches of the three mtDNA lineages using MAXENT v 3.3.3a (Phillips *et al.*, 2006). We obtained current climate data at 30 arc-second resolution from the WorldClim database (http://www.worldclim.org; Hijmans *et al.*, 2005). We reduced the data set to nine bioclimatic variables that we consider relevant to the ecological and physiological requirements of *P. jordani*. These variables included: annual mean diurnal range, annual temperature range, mean temperature of wettest quarter, mean temperature of driest quarter, mean temperature of warmest quarter, precipitation of wettest month, precipitation seasonality, precipitation of driest quarter, and precipitation of warmest quarter. We clipped climate layers to an area extending from 35.0° to 36.3° latitude and -84.5° to -82.5° longitude, which includes the current distribution of *P. jordani* within GRSM and surrounding mountains. Details about occurrence records and data layer preparation are described in Luxbacher *et al.* (Chapter 1).

We used ENMTOOLS v 1.4.3 to quantify niche overlap from Schoener's *D* (Schoener, 1968), Warren's *I* (Warren *et al.*, 2008), and relative rank (Warren & Seifert, 2011)

statistics, and to perform niche identity and background tests. Niche identity tests evaluate whether the ecological niche models (ENMs) generated for the three lineages are more different than expected if they had been generated from the same underlying distribution (Warren *et al.*, 2010). Observed niche overlap values calculated for each pair of models were compared to null distributions generated from 100 pseudoreplicates to assess statistical significance (Warren *et al.*, 2010). Background tests determine whether the ENMs for each lineage are more different than expected considering the underlying environmental differences between the geographic regions where they each lineage occurs (Warren *et al.*, 2008). For this test, the three overlap statistics were calculated between ENMs based on one clade and localities randomly placed within the range of another clade; one hundred pseudoreplicates were conducted to generate a null distribution to which we compared the observed overlap values to assess statistical significance. We generated the random background points for each analysis from ASCII raster files that masked the range of each lineage.

RESULTS

Phylogenetic analyses

The number of nuclear haplotypes per locus ranged from 27 to 67 (Table 1). Nucleotide diversity estimates ranged from 0.010 to 0.019, which were similar to mtDNA (0.018; Table 1). No significant recombination was detected in any of the nuclear loci, and we found no significant departures from neutrality in mtDNA based on the McDonald–

Kreitman tests ($P > 0.05$ for all comparisons; results not shown) or in the nuclear loci based on the HKA test ($\chi^2 = 0.006$, $P = 0.94$).

The nuclear genealogies inferred from Maximum-likelihood and Bayesian methods were concurrent, but are characterized by poor resolution for each locus and do not recover the three clades inferred from mtDNA (Figures S2-S4). The mtDNA parsimony network revealed strong geographic structuring of haplotypes (Figure 3). In contrast, networks for each nuclear locus showed considerable sharing of haplotypes among geographic regions (Figure 3) and mtDNA clades (Figure S5). Locus 9 grouped into three networks. The main network is composed of 36 haplotypes, including one dominant ancestral haplotype represented in 35 localities, and all three mitochondrial clades (Figure S5a). A second network includes five haplotypes that all share two large insertions (27 and 14 bps) and two small deletions (2 and 3 bps). The four individuals with these haplotypes were sampled from three localities in the central part of the distribution and are all part of mtDNA Clade 1. A third network includes two haplotypes sampled from two localities that exhibit a unique 15-bp insertion. The two individuals sharing these haplotypes were sampled from Clades 2 and 3 in the eastern part of the distribution. Loci 16 and 41 both exhibit several dominant haplotypes across the range (Figure 3b,c).

Although individual nuclear networks show that alleles are not clustered according to geography or mitochondrial clades, the multilocus nuclear network shows an east-west break, with considerable mixing of alleles in the center of the distribution (Figure 3d). No east-central individuals occur west of locality 12 and no west-central individuals occur east of locality 48.

Concordance between mtDNA and anonymous loci

Results of Mantel tests revealed a significant positive association between nuclear genetic distance and mitochondrial clade that remained significant after controlling for the effect of geographic (Euclidean) distance ($r = 0.144$, $P = 0.007$; partial Mantel test: $r = 0.098$, $P = 0.041$). These results indicate that nuclear distances between individuals in the same mitochondrial clade are significantly smaller than distances between individuals in different mitochondrial clades, and support a correlation between nuclear and mitochondrial data. Results of Mantel tests also revealed a significant positive association between genetic and log geographic distance supporting a pattern of isolation by distance in *P. jordani* ($r = 0.096$, $P = 0.0002$). These results remained significant after accounting for the influence of mitochondrial clade ($r = 0.077$, $P = 0.003$).

Measures of exclusive ancestry for each mtDNA clade calculated from the genealogical sorting index were generally low (*gsi*: 0.137–0.22; gsi_T: 0.141–0.168), which suggests that the nuclear loci are capturing early stages of divergence in *P. jordani* (Cummings *et al.*, 2008). We reject the null hypothesis of no divergence for two of the three nuclear loci and the ensemble of all three loci grouped according to Clade 1 (Table 2). Two loci also support exclusive ancestry of Clades 2 and 3 (Table 2). Because simulations have shown that small sample sizes decrease statistical power to reject the null hypothesis (Cummings *et al.*, 2008) and each of these groupings included fewer individuals than Clade 1, we expect the third locus would support the hypothesis of exclusive ancestry if sample sizes were increased.

Divergence and demographic history

Both approaches used to estimate divergence times yielded similar estimates, but the estimates based on substitution rate exhibited broader confidence intervals (Table 3). Divergence within *P. jordani* dates to the late Pliocene when Clade 3 split from Clades 1 and 2 (Table 3). The latter two clades diverged in the middle Pleistocene, with the mean date falling just after the Middle Pleistocene Transition (MPT; Clark *et al.*, 2006). The common ancestor of each lineage dates to the late Pleistocene (Table 3).

Tests of historical demographic changes show an overall signature of demographic expansion for *P. jordani* based on mtDNA. All three metrics indicate a significant signature of expansion for Clade 1 (Table 1). Tajima's *D* and Fu's *Fs* were significantly negative for Clade 3, but Ramos-Onsins & Rozas' R_2, which is more robust for small samples (Ramos-Onsins & Rozas, 2002), showed no significant evidence of demographic expansion (Table 1). All nuclear loci had negative Fu's *Fs* values, but only locus 16 was statistically significant.

The GMRF Bayesian skyride plots showed increasing population trends for all three clades (Figures 4-6), but only Clade 1 exhibited marginally significant expansion (Figure 4). The median estimate of expansion was similar in clades 1 and 3 (Clade 1: 0.068 Ma, 95% HPD 0.031-0.12; Clade 3: 0.064 Ma, 95% HPD 0.012-0.159), and more recent for Clade 2 (0.03 Ma, 95% HPD 0.009-0.06). The effective population size inferred from the EBSP analysis of all four loci remains constant over the course of its demographic history (Figure 7). The population size change parameter has a median of 0, indicating that the direction of population size has not changed during its coalescent history (95% HPD 0-

1).

Phylogeographic diffusion and ecological niche modeling

Bayes factors indicate that the variable diffusion rate (RRW) model fits the data better than the constant rate (BD) model (Table 4). The RRW model suggests an east-central ancestor for *P. jordani*, with dispersal mainly occurring to the west (Figure 8). The mean diffusion rate was 69.09 km/Myr (95% HPD 31.99 – 109.99), with faster rates occurring during initial expansion along the ridgeline.

Niche identity tests showed that Clades 1 and 2 exhibit statistically significant ecological differences with respect to the nine climatic variables used to build the models for all three overlap metrics (Table S1; Figs. S4-S6). Considering the geographic distribution of these lineages, these results suggest that the western clade occupies a different ecological niche than its sister lineage in the northeastern part of the range; however, this result may simply reflect different underlying environmental conditions available to each clade (Warren *et al.*, 2008). Background tests address this possibility and indicated that samples from Clade 1 are more similar to background from the range of Clade 2 than expected by chance, and samples from Clade 2 are no more or less similar to the range of Clade 1 than expected (Table S2). This asymmetric result is not uncommon when the availability of suitable microclimate differs between the geographic regions of the two taxa under comparison (e.g., Nakazato *et al.*, 2010). In this case, it appears to be due to cooler and wetter conditions in the northeastern distribution that coincides with the distribution of Clade 2 compared to the western distribution that

coincides to Clade 1. Taken together, the niche identity and background tests indicate that the Clades 1 and 2 do not exhibit significant ecological differentiation, nor do comparisons of other lineages (Tables S1-S2).

DISCUSSION

In this study, we investigated the historical processes shaping spatial patterns of genetic structure and demography in *Plethodon jordani*, a salamander that is endemic to the Great Smoky Mountains. These mountains are a continuous temperate chain that encompass a different suite of environmental conditions than sky island or tropical mountain systems that have been the focus of previous research (e.g., Knowles, 2000; Shepard & Burbrink, 2008; Bell *et al.*, 2010; Parra-Olea *et al.*, 2012), and are predicted to have exhibited areas of stable habitat or microclimate through time (Luxbacher *et al.*, Chapter 1; Delcourt & Delcourt, 1998). We demonstrate that divergences in *P. jordani* occurred during Pleistocene glacial periods, and results of our separate analyses based on nuclear and mitochondrial data indicate that ancestral *P. jordani* occurred in the east-central region of the Smokies and expanded its range along the ridgeline, predominantly to the west. Glacial cycles fragmented ancestral populations of *P. jordani*, resulting in its current phylogeographic patterns, but there has been little change in population sizes since the formation of lineages.

Phylogeographic history

The multilocus nuclear network showed a pattern of genetic distinction between eastern and western regions, with considerable mixing in the central region. This pattern coincides with the diffusion model, which suggests that lineages expanded their ranges slightly to the east and extensively to the west from the east-central Great Smoky Mountains. The direction and size of range expansion appears to be due to topography and the limited availability of suitable microclimate. Even in periods of cooler climate, *P. jordani* would have had few opportunities to colonize adjacent mountains because they were separated by lowlands that remained climatically unsuitable (Kozak & Wiens, 2006; Ohlemüller *et al.*, 2008). The phylogeographic dispersal model supports the hypothesis that dispersal rates varied across the landscape and within and among lineages. The inferred diffusion rate is low, but reasonable given that *P. jordani* is thought to be highly philopatric and have limited dispersal abilities (Petranka, 1998), with average home range sizes of less than five square meters (Nishikawa, 1990). The pattern of isolation-by-distance observed within mitochondrial clades also attests to the low dispersal ability of individuals.

The split between Clades 1 and 2 and divergence within clades follow the occurrence of the MPT when glaciations shifted from low-amplitude cycles occurring ca. every 41,000 years to high-amplitude 100,000-year cycles (Clark *et al.*, 2006; Bintanja & van de Wal, 2008). Shepard and Burbrink (2008, 2009) also found divergence after the MPT in *Plethodon* species endemic to the Ouachita Mountains. They suggested that the longer cycles and greater climatic extremes after the MPT provided opportunities for expansion during interglacial periods and accumulation of genetic differences among lineages that were isolated during glacial periods. Our results appear to show a similar pattern for *P.*

jordani in the Great Smoky Mountains.

Mean estimated divergence times within the three main clades date to different glacial periods in the late Pleistocene (Gibbard & van Kolfschoten, 2004): the common ancestor of Clades 1, 2, and 3 are estimated prior to glacial terminations 5 (0.424 Ma), 3 (0.243 Ma), and 2 (0.130 Ma), respectively (Raymo, 1997; Lisiecki & Raymo, 2005). These results support the idea that the current genetic structure in *P. jordani* was established during different glacial advances (Luxbacher *et al.*, Chapter 1). If populations shifted to mid- and low elevations during past glacial advances as predicted at the LGM (Luxbacher *et al.*, Chapter 1), they would have experienced limited connectivity during these glaciations, providing opportunities for divergence. Subsequent expansion up the mountain during brief interglacials in the late Pleistocene (Lisiecki & Raymo, 2005) may not have afforded enough time for extensive gene flow, considering *P. jordani*'s slow rate of dispersal. These relatively recent coalescent times may mean that earlier signatures of isolation and expansion were erased by more recent glacial cycles (Bonatelli *et al.*, 2014).

The Great Smokies are a continuous chain of mountains that do not appear to exhibit areas of disparate environmental conditions that act as barriers to gene flow to *P. jordani*, yet they still exhibited conditions causing isolation of populations during Pleistocene glacial cycles. Our findings highlight the importance of complex topography in the diversification of this species and may help explain patterns of divergence in other species in the southern Appalachian Mountains. Comparative phylogeographic studies have not been conducted at the scale of our research, but one study of harvestmen (*Fumontana deprehendor*) endemic to mid-elevations in the region showed divergence

between one western and several eastern population in the Smokies (Thomas & Hedin, 2008). As studies consider this region in finer detail, more examples of this east-west break may emerge among montane endemic species.

Demographic history

The hypothesis that ancestral populations of *P. jordani* were located in the east before expanding westward predicts a signature of demographic expansion in the western part of the range. This hypothesis is supported by summary statistics that show significant expansion and BSPs that show marginally significant expansion for mtDNA Clade 1. Considered alongside the lack of significant demographic expansion detected in Clades 2 and 3, these results suggest relative stability of population sizes in the eastern part of the range and expansion to the west. Demographic results based on three nuclear loci do not show evidence of expansion, with the exception of Fu's *Fs* for locus 16, and the EBSP based on combined data from all four loci also shows constant population size. We note that mtDNA may have detected westward demographic expansion due to its smaller effective population size (hence, shorter coalescence time; see below), whereas the nuclear loci did not due to their longer coalescence time; however, we also acknowledge that a single locus reflects only one of a multitude of histories (Karl *et al.*, 2012) and that multiple loci are more robust in detecting population size dynamics (Heled & Drummond, 2008), and discuss the implications of demographic stability further.

The relatively constant population size inferred from the combined data implies a lack of significant demographic response to Pleistocene climatic changes (Barker *et al.*,

2011; Ding *et al.*, 2011). This finding corroborates our predictions that regions within the Great Smoky Mountains have remained climatically stable since at least the LGM (Luxbacher *et al.*, Chapter 1). Other studies have also reported constant population sizes through the LGM and during cooler climatic conditions for a range of taxa in different regions of the world (e.g., Shepard & Burbrink, 2008; Ding *et al.*, 2011; Qu *et al.*, 2011; Schoville *et al.*, 2013; Chabarria & Pezold, 2013). The finding that population sizes were relatively constant throughout recent glaciations is somewhat surprising for *P. jordani* because the output of our best-supported distribution model (Luxbacher *et al.*, Chapter 1) predicted the area of suitable microclimate for this species at the LGM to be larger than today, suggesting that the size of the range may have decreased since then. The pattern of having a relatively constant effective population size, but occupying a smaller area may suggest that *P. jordani* is currently living at higher densities than at the LGM (McKay *et al.*, 2010). The output of the mechanistic bioenergetic model offers some evidence in favor of this possibility. We found that the range of energetic yield per unit foraging was narrower at the LGM than under current climatic conditions (LGM: -23.9 – 2.3 kJ, current: -24 – 10 kJ; Luxbacher *et al.*, Chapter 1), which could indicate that individuals were less surface active and may have sustained smaller populations during Pleistocene glaciations. As temperatures warmed since the LGM, more individuals living at an energy surplus could have resulted in greater surface activity and higher population densities.

In any case, our results do not support the alternative hypothesis that *P. jordani* recently colonized GRSM from the west, as suggested by the correlative model output (Figure 1). We questioned the output of this model in Luxbacher *et al.* (Chapter 1)

because it required extrapolating onto novel climate conditions and lacked consistency with independent phylogeographic and palynological data. The demographic results presented here further support that correlative modeling may be insufficient for generating predictions about the past or future distribution of this species.

Concordance among markers

On one hand, the phylogeographic patterns inferred from mtDNA suggest that populations were more isolated in the past than they are now (Jockusch & Wake, 2002). On the other hand, because the likelihood of detecting phylogeographic breaks increases as dispersal distance decreases (Irwin, 2002) and *P. jordani* is a dispersal-limited species (Petranka, 1998), mitochondrial breaks may have formed by chance and not reflect past isolation. However, considering the responses of other montane plethodontids to past climate change (Shepard & Burbrink, 2008, 2009; Rovito, 2010; Devitt *et al.*, 2013; Velo-Anton *et al.*, 2013) and our LGM distribution modeling results (Luxbacher *et al.*, Chapter 1), it appears more likely that *P. jordani* tracked suitable climate conditions up and down mountains over time. Consequently, populations experienced periods of isolation resulting in divergent lineages.

In contrast to the mitochondrial phylogeny, genealogies for each nuclear locus differ from one another and none exhibit reciprocally monophyletic groups that correspond to mitochondrial clades or geography. However, the Mantel test and *gsi* results indicate that nuclear data do not conflict with mitochondrial data. The lack of resolution in the nuclear trees is concurrent with the idea that not enough time has elapsed for distinct lineages

to sort, which is not unexpected given that the coalescent size of nuclear markers is four times longer mitochondrial markers due to differences in inheritance and effective population sizes (Avise, 2000; Zink & Barrowclough, 2008). In addition, individual nuclear networks revealed extensive sharing of haplotypes among mitochondrial clades and geographic regions, suggesting that nuclear loci may reflect incomplete lineage sorting (Chan *et al.*, 2012). These lineages appear to be in early stages of divergence, as supported by the low *gsi* values calculated when the nuclear alleles were grouped according to mtDNA clades (Cummings *et al.*, 2008).

Male-biased dispersal can also result in a pattern of strong structure in maternally-inherited markers and a lack of structure in nuclear markers, as has been observed in other plethodontids (Jockusch & Wake, 2002; Liebgold *et al.*, 2011). We cannot rule out the possibility that male-biased dispersal is driving the pattern we observe in *P. jordani*, but incomplete lineage sorting is more plausible given the general biogeographic agreement between mitochondrial and nuclear loci (Toews & Brelsford, 2012). Individual data on movement patterns will be integral to address the mechanisms underlying patterns of dispersal and gene flow in this species (Lowe *et al.*, 2008).

Conservation implications

Montane species are particularly vulnerable to climate change (Ohlemüller *et al.*, 2008; La Sorte & Jetz, 2010); therefore, understanding factors influencing persistence and connectivity among populations in topographically complex regions is critical for assessing species' responses to projected future climate change. In general, our results

highlight the importance of topography in maintaining suitable microclimate for *P. jordani*. Niche conservatism clearly played a role in preventing populations from expanding outside of the Great Smoky Mountains in the north and east where inhospitable low elevations separate these mountains from adjacent chains, even during glacial periods when populations shifted to mid- and lower elevations. *Plethodon jordani* expanded its range to the west by following suitable climatic conditions along the ridgeline during past interglacials. Niche conservatism may spell dire consequences for this and other montane plethodontid species, as they track suitable conditions to higher elevations in the future (Milanovich *et al.*, 2010). Indeed, the extent of microclimate that is physiologically suitable for *P. jordani* in the far western part of its range (from Parson Bald to Clingmans Dome) is predicted to decrease under climate change (K. H. Kozak, unpublished data). In addition, low elevation populations across the distribution are already living near the limit of their physiological tolerances (Gifford & Kozak, 2012), and may be particularly susceptible to declines. However, we note that because individuals spend much of their time in moist covers that can differ from atmospheric temperature and ground temperature (Feder, 1983), they can remain sheltered from broad-scale changes in climate (Wake, 2009) and may have some resiliency to endure future climatic changes. Our results showing relatively stable demographic trends over past periods of climatic change also attest to potential resiliency of this species. Still, monitoring populations will be critical to gauge their response. The core of suitable microclimate, as well as genetic diversity, for *P. jordani* remains in the east-central region of GRSM and this area should be the focus of future monitoring and conservation efforts.

Table 1 Locus information, samples sizes, and summary statistics of sequences used in this study

Locus	Primer source	bp	Model of sequence evolution	N	N_A	H	S	k	π	Tests of demographic expansion		
										D	F_S	*R2*
ND2												
All[#]				361	-	62	254	18.859	0.018	-1.691*	-1.166	0.037*
Clade 1	Luxbacher *et al.*,	1116	TIM2+G	228	-	34			0.002	-1.789*	-20.800*	0.031*
Clade 2	Chapter 1			92	-	10			0.001	-0.787	-3.129	0.067
Clade 3[#]				22	-	4			0.0004	-1.240*	-1.827*	0.104
9	Shepard *et al.*, in review	601	GTR+G	121	168	43	98	10.020	0.019	-1.359	-5.112	0.051
16	Shepard *et al.*, in review	562	TPM1uf+I+G	132	236	67	64	6.659	0.013	-1.258	-33.822*	0.051
41	Shepard *et al.*, in review	534	TPM1uf+I	136	218	27	27	5.246	0.010	0.280	-2.015	0.097

bp: number of base pairs; N: number of individuals sequenced; N_A: number of unambiguous sequences; H, number of haplotypes; S, number of segregating sites; k, average number of differences between sequences; π, nucleotide diversity *D*, Tajima's *D*; *Fs*, Fu's *Fs*; *R2*, Ramos-Onsins & Rozas' *R2*

* *P* value < 0.05

metcalfi-1 haplotype was excluded due to a large portion of missing data

Table 2 *Gsi* values for individual nuclear loci grouped according to mtDNA clades and gsi_T calculated for the ensemble of all three loci. Statistical significance (denoted by asterisks) was assessed from 10,000 permutations

Grouping	*gsi*			gsi_T
	Locus 09	Locus 16	Locus 41	
Clade 1	0.168*	0.137	0.192**	0.168**
Clade 2	0.151	0.182*	0.220**	0.151
Clade 3	0.141	0.166*	0.207**	0.141

*P value < 0.05
**P value < 0.01

Table 3 Estimated divergence dates based on two different methods using 1116 bp of the *P. jordani* mitochondrial genome. Nodes follow Figure S1. Median divergence times based on combined runs are given in millions of years followed by 95% HPD intervals

Node	Inference method	
	Root height	Substitution rate
P. jordani - *P. metcalfi*	9.86 (6.98 - 12.76)	8.93 (4.77 - 16.35)
P. jordani	2.22 (1.24 - 3.40)	2.06 (1.02 - 3.81)
Clades 1 & 2	0.84 (0.44 - 1.38)	0.78 (0.36 - 1.50)
Clade 1	0.46 (0.23 - 0.75)	0.43 (0.19 - 0.83)
Clade 2	0.26 (0.01 - 0.53)	0.24 (0.01 - 0.55)
Clade 3	0.15 (0 - 0.43)	0.14 (0 - 0.43)

Table 4 Results of phylogeographic diffusion modeling implemented in BEAST. Marginal likelihood estimates (MLE) from Brownian diffusion (BD) and relaxed random walk (RRW) models were compared based on path sampling and stepping-stone sampling. Analyses were run two times. In all runs, Bayes factors indicate strong support for the variable diffusion rate (RRW) model (Kass & Raftery, 1995)

	MLE Path sampling		MLE Stepping-stone sampling	
Model	Run 1	Run 2	Run1	Run 2
BD	-2614.42	-2613.49	-2615.26	-2613.93
RRW	-2253.02	-2250.79	-2254.82	-2251.99
Bayes Factors	361.4	362.7	360.44	361.94

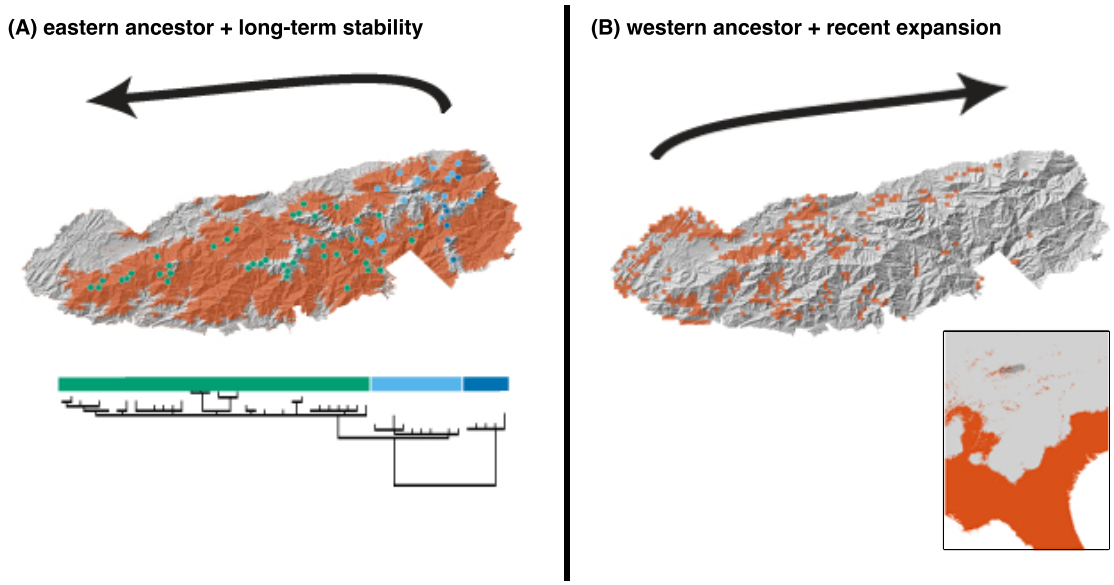


Figure 1 Two hypotheses we considered to investigate historical demographic changes in *P. jordani*. In both panels, arrows indicate the direction of predicted range expansion for each hypothesis. (A) is based on phylogenetic data from mtDNA, which suggest that ancestral *P. jordani* was located in the eastern Great Smoky Mountains, and the distribution at the Last Glacial Maximum (LGM) inferred from a mechanistic model, which suggests *P. jordani* was widespread at mid-elevations during the most recent glacial advance (Luxbacher *et al.*, Chapter 1). This hypothesis predicts a signature of demographic expansion in the western part of the range, but relative stability since at least the LGM. Localities on the map are colored according to mtDNA lineage, and the LGM distribution from the mechanistic model is depicted in vermilion. (B) is based on the LGM distribution inferred from correlative model (shown in vermilion; Luxbacher *et al.*, Chapter 1), and predicts expansion across the Great Smoky Mountains, particularly since the LGM and in the eastern part of the range. The full extent of the LGM distribution, which was mainly shifted far south of the Great Smoky Mountains, is shown in the inset

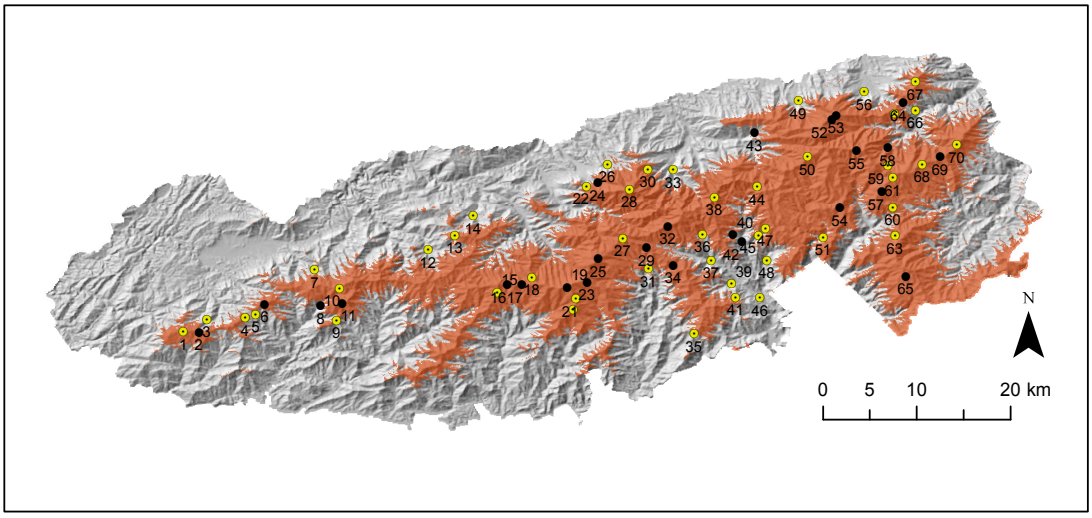


Figure 2 Map of sampling localities plotted on a 30 m digital elevation model of Great Smoky Mountains National Park. The known distribution of *P. jordani* (based on a mechanistic bioenergetic model; Gifford & Kozak 2012) is depicted in vermilion. Black dots indicate localities sampled for *ND2*; yellow dots indicate localities sampled for *ND2* and anonymous nuclear loci. Localities are numbered as in Appendix S3 of Chapter 1

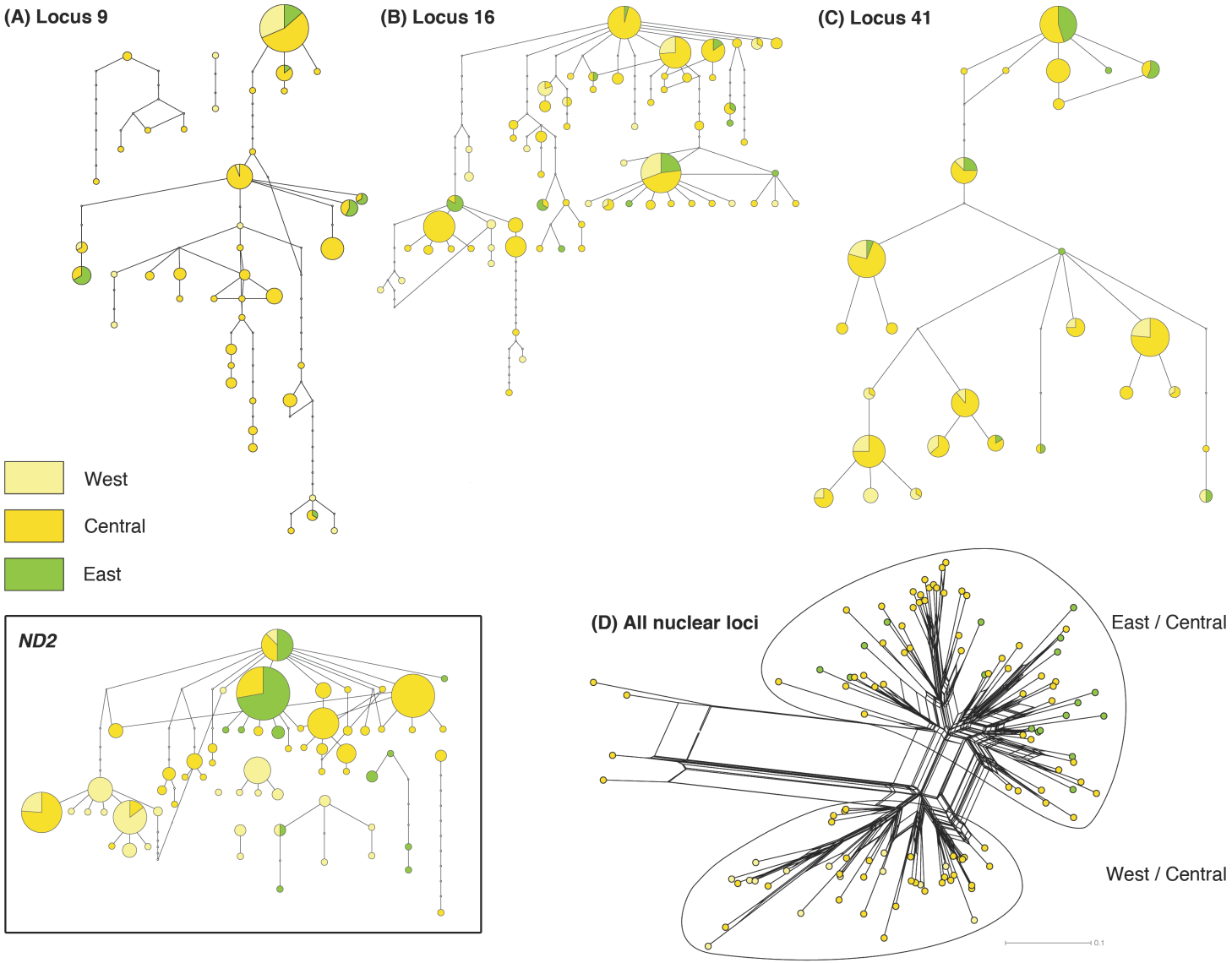

(A) Locus 9
(B) Locus 16
(C) Locus 41
West
Central
East
ND2
(D) All nuclear loci
East / Central
West / Central
0.1

Figure 3 Individual haplotype networks for each nuclear locus used in this study and a multilocus network constructed from all three nuclear loci. Haplotypes are colored according to geographic region: west (localities 1-11), central (Localities 12-48), east (localities 49-70); see Chapter 2, Figure 2. (A) through (C), Individual networks were estimated using statistical parsimony with a 95% connection significance. The size of each circle is proportional its haplotype frequency and open circles represent inferred unsampled haplotypes. Haplotype circles are depicted as pie diagrams with slice size colored according to geographic region. (D), Multilocus network based on Tamura-Nei distances from all three nuclear loci. The four individuals offset to the left are highly divergent at locus 9. Inset, Haplotype network for *ND2*

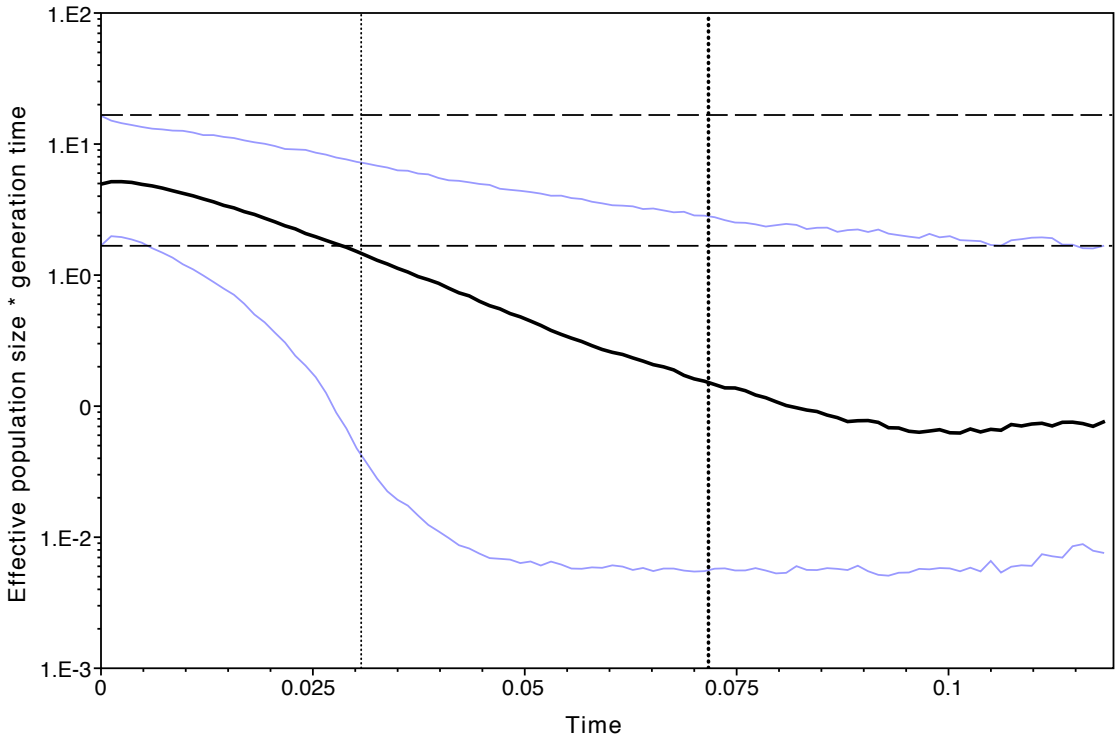


Figure 4 GMRF Bayesian skyride plots showing the demographic history of mtDNA Clade 1 (Figure S1). Time, in millions of years, is shown on the *x*-axis; effective population size in log number of individuals and generation time is shown on the *y*-axis. Estimated median population size (central horizontal line) and 95% confidence intervals (upper and lower horizontal lines) are based on a substitution rate of 1.28% per million years. The vertical dotted line represents the median age estimate. The upper 95% HPD on age estimate aligns with the right edge of the plot; the lower 95% HPD is the vertical dotted line to the left of the median. Horizontal dashed lines represent the cutoffs used to assess significance of population size change

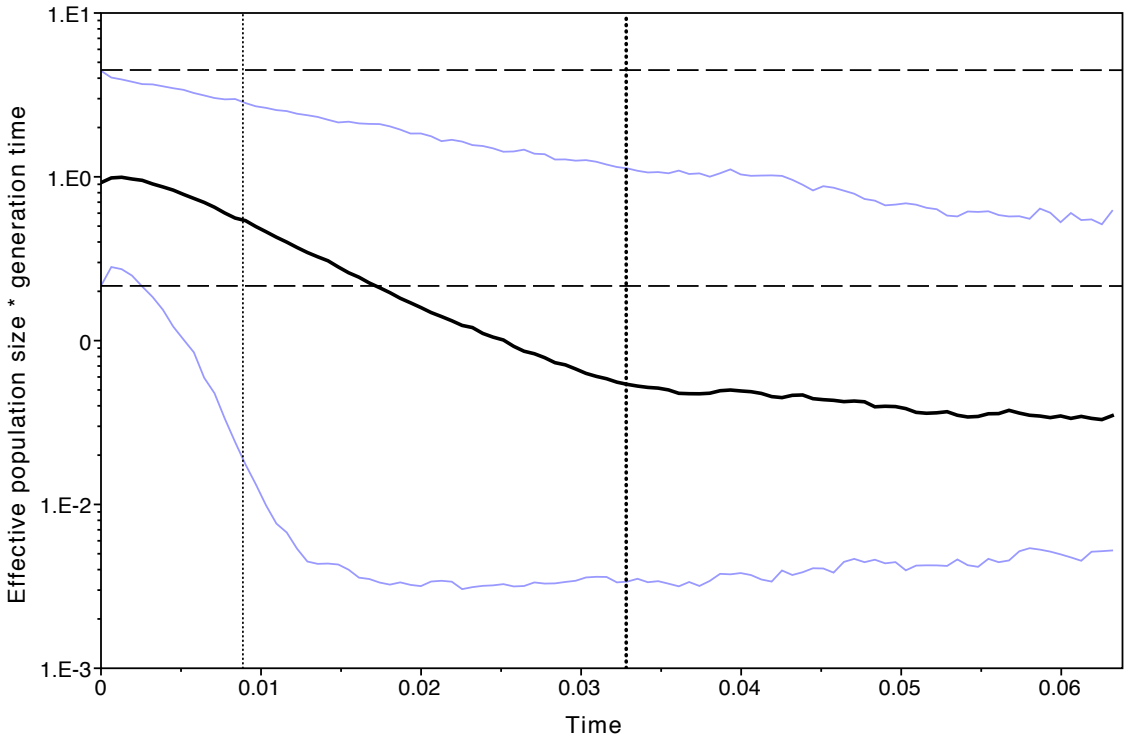


Figure 5 GMRF Bayesian skyride plots showing the demographic history of mtDNA Clade 2 (Figure S1). Time, in millions of years, is shown on the *x*-axis; effective population size in log number of individuals and generation time is shown on the *y*-axis. Estimated median population size (central horizontal line) and 95% confidence intervals (upper and lower horizontal lines) are based on a substitution rate of 1.28% per million years. The vertical dotted line represents the median age estimate. The upper 95% HPD on age estimate aligns with the right edge of the plot; the lower 95% HPD is the vertical dotted line to the left of the median. Horizontal dashed lines represent the cutoffs used to assess significance of population size change

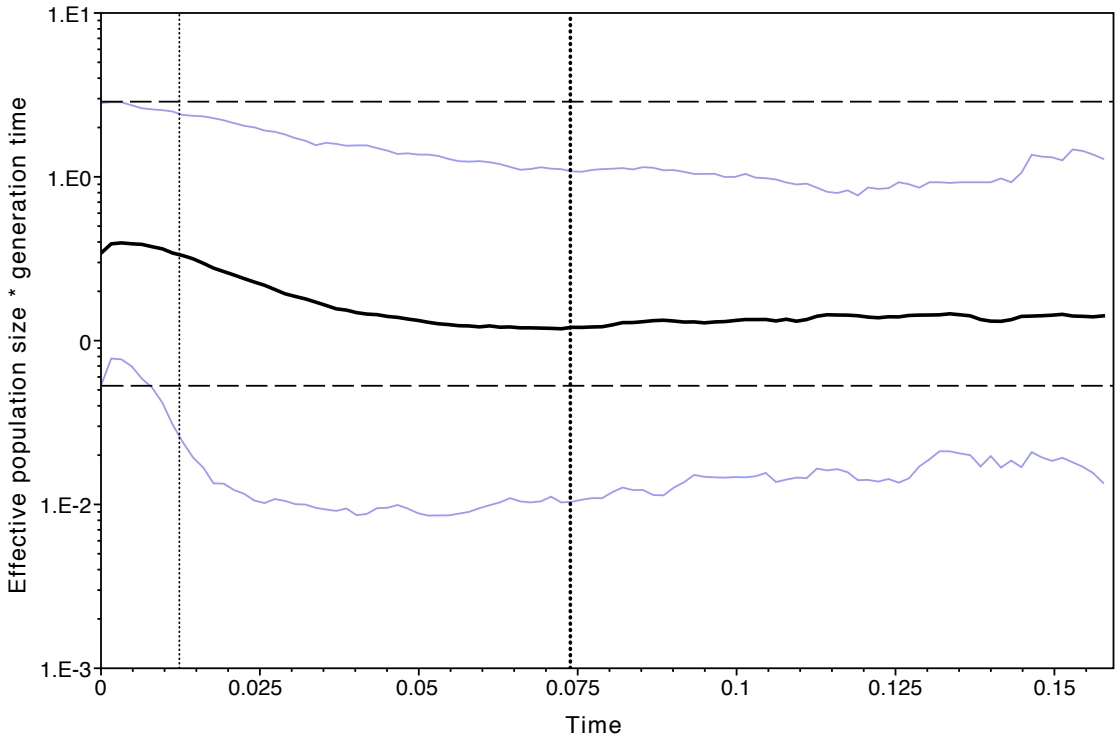


Figure 6 GMRF Bayesian skyride plots showing the demographic history of mtDNA Clade 3 (Figure S1). Time, in millions of years, is shown on the *x*-axis; effective population size in log number of individuals and generation time is shown on the *y*-axis. Estimated median population size (central horizontal line) and 95% confidence intervals (upper and lower horizontal lines) are based on a substitution rate of 1.28% per million years. The vertical dotted line represents the median age estimate. The upper 95% HPD on age estimate aligns with the right edge of the plot; the lower 95% HPD is the vertical dotted line to the left of the median. Horizontal dashed lines represent the cutoffs used to assess significance of population size change

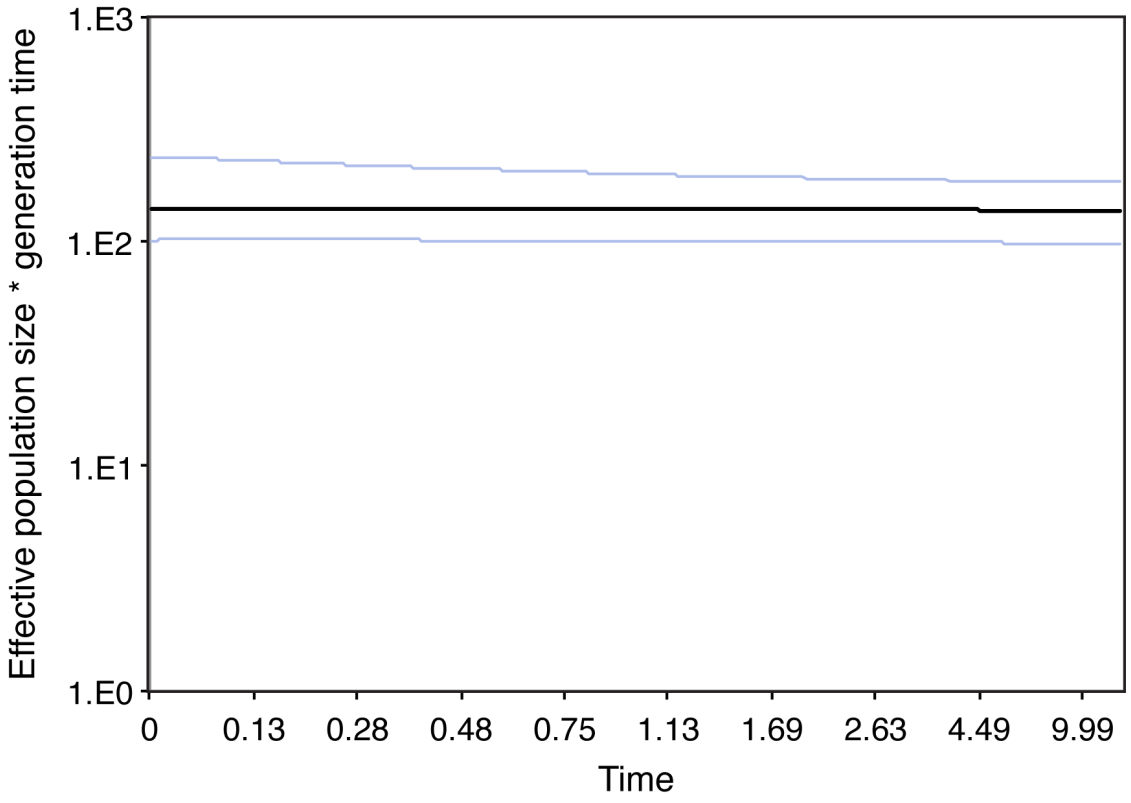


Figure 7 Extended Bayesian skyline plot (EBSP) based on mtDNA and three anonymous nuclear loci showing the demographic history of *P. jordani*. Time, in millions of years, is shown on the *x*-axis; effective population size in log number of individuals and generation time is shown on the *y*-axis. The central horizontal line represents the estimated median population size and upper and lower horizontal lines show 95% confidence intervals

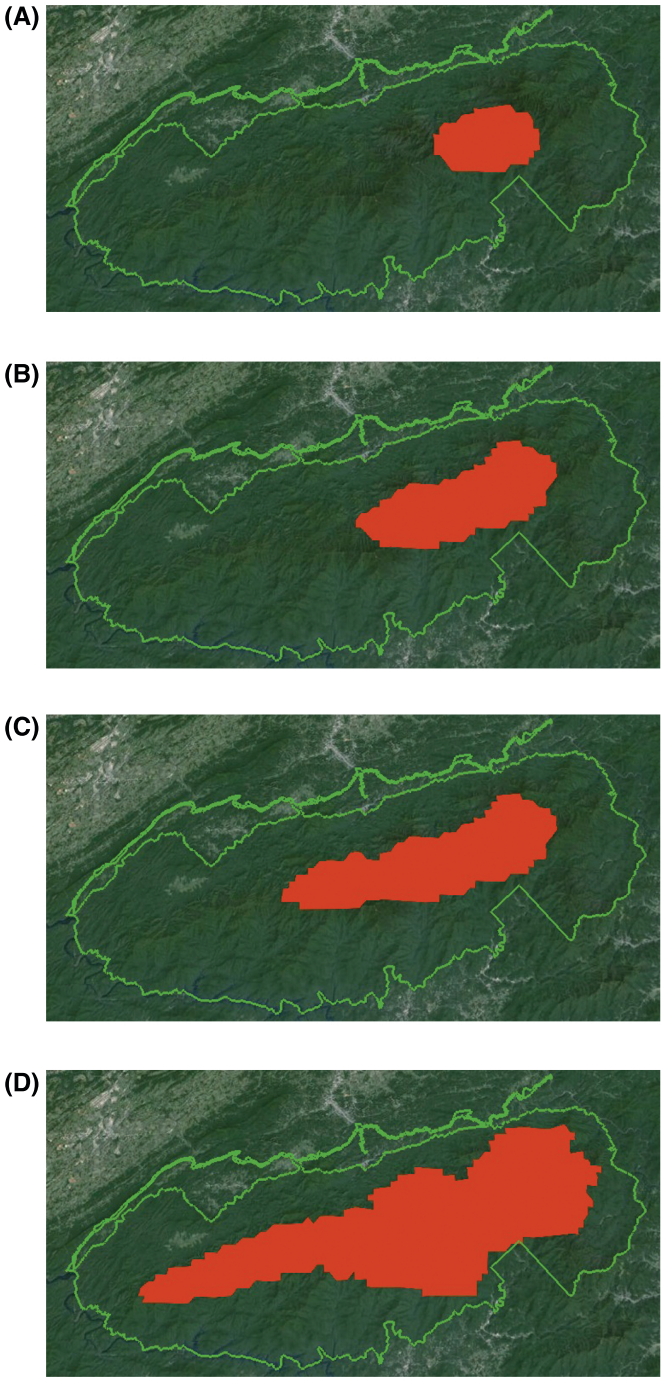


Figure 8 Relaxed random walk continuous phylogeographic model for *P. jordani*. (A) through (D), Inferred distribution at four time periods beginning with the ancestral distribution of mtDNA lineages (A). Polygons represent the 80% highest posterior density intervals for the inferred ancestral range at each time period

CHAPTER 3

Geographic genetic structure and conservation units of a narrow endemic salamander, *Plethodon jordani,* in a continuous montane environment[3]

SUMMARY

Range-restricted montane species are considered particularly vulnerable to climate change. Consequently, documenting genetic diversity and structure within and among populations, respectively, is an essential step in developing conservation plans to maintain their evolutionary potential in the face of climate change. We report geographic patterns of genetic variation in *Plethodon jordani*, a terrestrial salamander that is endemic to high elevations in Great Smoky Mountains National Park. We use data from nine species-specific microsatellite loci collected from 237 individuals distributed among 38 populations to (1) characterize genetic diversity within populations and estimate genetic structure among populations, (2) evaluate ecological exchangeability of inferred genetic clusters, and (3) compare these results to patterns recovered from mtDNA. We find high genetic diversity in the south-central and western parts of the distribution. Bayesian clustering analyses indicate that three genetic clusters best explain the data. These clusters are partially congruent with phylogeographic patterns based on previous studies, but we detect an additional genetic break in the western part of the distribution not inferred by earlier methods. We also find mito-nuclear discordance in the eastern region

[3] A version of this chapter is being prepared for publication:
Luxbacher AM, Kozak KH. Geographic genetic structure and conservation units of a narrow endemic salamander, *Plethodon jordani*, in a continuous montane environment. *Conservation Genetics*

of the distribution. Despite the occurrence of these genetic breaks, ecological niche models (ENMs) for each genetic cluster exhibit high overlap, suggesting ecological exchangeability. We propose that *P. jordani* comprises three evolutionarily significant units that reflect the patterns inferred from different genetic markers, ENMs, and historical distributions. Our results highlight the importance of integrating information from multiple sources when considering conservation strategies for species that may be threatened by climate change.

INTRODUCTION

Mountains are centers for diversity for many taxonomic groups (Myers, 2000; Lomolino, 2001; McCain, 2005; Guo *et al.*, 2013), and frequently harbor cryptic diversity that scientists are only beginning to uncover (e.g., Ribas *et al.*, 2007; Barata *et al.*, 2012; Kieswetter & Schneider, 2013; Demos *et al.*, 2014). This lag in our understanding of current levels of genetic variation is troubling given that montane species are particularly susceptible to declines due to environmental changes (La Sorte & Jetz, 2010), and have been among the first to show local extirpation and extinction that can be attributed to climate change (Parmesan, 2006; Wake, 2009). When considered alongside projections of future widespread range contractions and extinctions of montane species (Williams *et al.*, 2003; Thomas *et al.*, 2004), these declines underscore the need to document genetic diversity and structure so that conservation strategies can focus resources on unique evolutionary units and maximize the potential of species to respond to climate change (Petit *et al.*, 1998; Moritz, 2002; Bálint *et al.*, 2011).

Many montane species are expected to respond to future climatic changes by shifting their distributions to track suitable habitat and microclimate, as they have responded to past periods of climatic fluctuation (Hewitt, 2000; Brunsfeld *et al.*, 2001; Knowles, 2001; Swenson & Howard, 2005; Kozak & Wiens, 2006; Waltari *et al.*, 2007). In the southern Appalachian Mountains of eastern North America, these distributional shifts are predicted to cause significant changes in species composition and extinction of rare species (Delcourt & Delcourt, 1998). The southern Appalachians are topographically complex and include the largest tracts of forest in the eastern United States, much of which is protected in Great Smoky Mountains National Park (GRSM). GRSM is considered one of the most biologically-diverse temperate areas, with many narrow endemic and cryptic species that continue to be described (Sharkey, 2001; Nichols & Langdon, 2007).

Plethodon jordani is one of the endemic species to GRSM and occurs in a small, continuous range spanning high elevations (Highton & Peabody, 2000; Figure 1). This species is not currently considered threatened, but range-restricted species are vulnerable to stochastic events (Purvis *et al.*, 2000; Charrette *et al.*, 2006; Sodhi *et al.*, 2008), and climate change is likely to affect *P. jordani* by decreasing its available habitat in the coming decades (Gifford & Kozak, 2012). A recent study suggested complete extinction as soon as 2080 due to habitat loss associated with climate change (Milanovich *et al.*, 2010), and some evidence indicates that *P. jordani* may already be declining in parts of its range. For example, the mean number of individuals detected at one high-elevation site surveyed in 1994 declined compared to the mean number detected during nine surveys conducted from 1961 to 1975 (Highton, 2005). Another study found evidence

of decline in five out of 18 of Highton's (2005) sites that were re-surveyed in 2009 (Caruso & Lips, 2012). Given that one mission of National Parks is to conserve wildlife for the enjoyment of future generations, documenting diversity in this species and identifying evolutionarily significant units (ESUs; Ryder, 1986) is a critical step in developing plans to maintain its evolutionary potential in the face of climate change (Moritz, 2002).

Management efforts should target intraspecific units for conservation that are genetically and ecologically distinct (May *et al.*, 2011), but the boundaries of such units in continuously distributed species, like *P. jordani*, are not always clear (Diniz-Filho & Telles, 2002). Previous research outlined several different approaches that could be used to determine appropriate conservation units for *P. jordani*. One strategy may be to consider the three geographically circumscribed lineages (Luxbacher *et al.*, Chapter 1) as separate conservation units (Zink *et al.*, 2010). Alternatively, populations at mid-elevations could receive priority for conservation because they are predicted to have remained stable during past periods of climatic change (Luxbacher *et al.*, Chapter 1) and generally harbor greater species diversity and phylogenetic diversity per species (Kozak & Wiens, 2010). Conservation strategies could also focus on areas in the east-central part of the distribution that have been inhabited the longest (Luxbacher & Kozak, Chapter 2). This area is also high in mitochondrial DNA (mtDNA) diversity, another criterion that has been proposed to prioritize populations for conservation (e.g., Petit *et al.*, 1998; Frankham *et al.*, 2002; Moritz, 2002). Finally, efforts could focus on maintaining any populations that are ecologically differentiated (Crandall *et al.*, 2000), as they may exhibit local adaptations (Rader *et al.*, 2005). Considering that the western region of

GRSM is characterized by lower elevations, hence warmer and less humid conditions than high elevations to the east, it is possible that western populations occupy different ecological niches from other populations, and may warrant preservation.

In this study, we aim to clarify a strategy. We compare data from newly-developed microsatellites with results from previous phylogeographic studies based on mtDNA and anonymous nuclear loci (Luxbacher *et al.*, Chapter 1; Luxbacher & Kozak, Chapter 2). Microsatellites have proven to be highly informative for addressing issues of conservation, including assigning individuals to genetic groups, documenting patterns of genetic diversity and differentiation, and detecting subtle reductions in gene flow in the recent past (Jehle & Arntzen, 2002; Beebee, 2005; Selkoe & Toonen, 2006; Emel & Storfer, 2012; Toews & Brelsford, 2012). The goals of this study are to (1) characterize genetic diversity within populations of *P. jordani* and estimate levels of geographic genetic structure among populations from microsatellite loci, (2) evaluate ecological exchangeability of inferred genetic clusters, and (3) compare these results to patterns recovered from mtDNA. We synthesize our findings with results from previous studies to make conservation recommendations for *P. jordani* and discuss implications for other montane endemics in the southern Appalachian Mountains.

METHODS

Microsatellite loci isolation and optimization

We followed a protocol for constructing a genomic DNA library enriched for

microsatellite loci that was developed by Glenn and Schable (2005). Briefly, we extracted genomic DNA from liver tissue from two individuals collected from the type locality for *P. jordani* (Indian Gap, GRSM; Dunn, 1926) using a phenol-chloroform technique (F. K. Barker, pers. comm.). We digested the genomic DNA with RsaI and XmnI restriction enzymes (New England Biolabs), and ligated SuperSNX linkers (SuperSNX24 Forward: 5'-GTTTAAGGCCTAGCTAGCAGAATC-3'; SuperSNX24+4P Reverse: 5'-GATTCTGCTAGCTAGGCCTTAAACAAAA-3') to the digested fragments. We hybridized the linker-ligated DNA fragments with tri- and tetranucleotide biotinylated probes, and used streptavidin-coupled Dynabeads® (Invitrogen) to capture fragments containing microsatellite sequences. Hybridized sequences were recovered through PCR with the SuperSNX-24 Forward primer. After performing a second round of enrichment on the amplified enriched DNA, we cloned the double-enriched DNA with a pGEM®-T vector system (Promega). We plated the transformed bacteria (i.e., cells containing an enriched DNA fragment and a cloning vector) on LB/ampicillin/IPTG/X-Gal agarose plates and incubated the plates at 37°C for 16 hours. We screened the resulting bacterial colonies and picked 288 colonies that contained inserts to amplify by PCR. Products containing the amplified inserts were run on 1% agarose gels in Tris-Borate EDTA buffer. We identified the PCR products with fragments greater than 500 base pairs, and purified these products using Exonuclease I and Shrimp Alkaline Phosphatase enzymes (Affymetrix). We sequenced 115 purified PCR products at the University of Minnesota BioMedical Genomics Center on an ABI 3730xl DNA Analyzer (Applied Biosystems) using BigDye Terminator v 3.1 chemistry.

We screened sequences for microsatellite repeats and designed primer pairs from

24 sequences using the program PRIMER3 v 0.4.0 (Rozen & Skaletsky, 2000). For each primer pair, we attached a M13F tail to the 5' end of one primer (Boutin-Ganache *et al.*, 2001) and a "pigtail" to the non-tailed primer (Brownstein *et al.*, 1996). Fluorescent labels were added to each fragment during PCR following the three-primer protocol of Schuelke (2000), in which a third fluorescent dye-labeled M13 primer is added to each reaction. We conducted preliminary analyses in which we amplified the same samples separately with M13 primers labeled with each of the four fluorescent dyes (6-FAM™, NED™, PET™, VIC®). Results of these analyses (not shown) revealed length variation of up to three nucleotides, depending on the dye used to label each locus; therefore, we used the same dye-labeled primer in all reactions for a given locus (Table 1).

We screened the 24 loci for length heterogeneity on 20 individuals from two localities. Preliminary screening resulted in nine polymorphic loci that we amplified for 237 individuals from 38 localities (Tables 1, S1). We performed PCR in 12.5 μL reaction volumes. For four loci (PLJOb, PLJOe, PLJOk, and PLJOr), reactions consisted of 1 μL DNA template, 10X CoralLoad PCR Buffer (Qiagen), 0.4 mM dNTP Mix (Invitrogen), locus-specific forward and reverse primers (Table 1), 0.4 μM fluorescent-labeled M13F primer (Table 1), 0.3125 U *Taq* DNA Polymerase (Qiagen), and ddH_2O to attain the final reaction volume. For the remaining loci, reactions consisted of 1 μL DNA template, 5X Green GoTaq® Flexi Buffer (Promega), 2 mM $MgCl_2$ (Promega), 0.4 mM dNTP Mix (Invitrogen), locus-specific forward and reverse primers (Table 1), 0.4 μM fluorescent-labeled M13F primer (Table 1), 0.3125 U GoTaq® Hot Start Polymerase (Promega), and ddH_2O to achieve the final reaction volume. Thermocycler conditions for amplification were as follows: 2 min initial denaturation at 95°C (3 min for reactions with GoTaq®);

25 cycles of denaturation at 95°C for 15 s, annealing at primer-specific temperature (Table 1) for 15 s, and extension at 72°C for 20 s; 10 cycles of 95°C for 15 s, 53°C (M13 primer annealing temperature) for 15 s, and 72°C for 20 s; final extension at 72°C for 7 min. Amplified fragments were analyzed on an ABI 3730xl DNA Analyzer alongside the LIZ®-500 size standard (Applied Biosystems). Approximately five percent of samples were PCR-amplified and genotyped a second time to verify repeatability in scoring.

We genotyped individuals using GENEMARKER software v 1.85 (SoftGenetics); all scoring was confirmed by visual inspection. We binned alleles using the program TANDEM (Matschiner & Salzburger, 2009). For each locus, we sequenced one homozygous individual to verify the length of its microsatellite. We then specified these known lengths for each locus as fixpoints in TANDEM, allowing the transformation of allele sizes to be adjusted so that the observed allele size was rounded to the actual allele size.

Genetic diversity

We calculated observed and expected heterozygosity in FSTAT v 2.9.3 (Goudet, 2001). We tested for deviation from Hardy-Weinberg equilibrium using an exact test in the software GENEPOP v 4.1.4 (Rousset, 2008). We specified 10,000 dememorization steps, and ran 1,000 batches with 10,000 iterations per batch. We used a sequential Bonferroni correction for multiple tests to evaluate *P* values for significant departures from Hardy-Weinberg equilibrium at a table wide significance value of 0.05 (Rice, 1989). GENEPOP was also used to test for linkage disequilibrium for each pair of loci across all

populations (Markov chain parameters: 10,000 dememorization steps, 1,000 batches, 10,000 iterations/batch). We estimated genotyping errors with MICRO-CHECKER v 2.2.3 (van Oosterhout *et al.*, 2004), and calculated the rarefied allelic richness (A) and the number of private alleles (P) in each population, assuming a sample of 18 alleles, with HP-RARE (Kalinowski, 2005). These statistics account for the potential bias resulting from different sample sizes among populations.

Population structure and genetic differentiation

We conducted Bayesian clustering analyses of the microsatellite data using GENELAND v 3.2.2 (Guillot *et al.*, 2005a, b; Guillot, 2008) and STRUCTURE v 2.3.3 (Pritchard *et al.*, 2000). In GENELAND, we first ran the uncorrelated allele-frequency model, per the authors' recommendations. We specified 2,000,000 MCMC sample iterations, a thinning interval of 100, and burn-in of 10,000 iterations, and ran ten independent runs for $K = 1$ to 20. We followed this initial run with the correlated allele-frequency model using the same settings.

In STRUCTURE, we assumed an admixture model with uncorrelated allele frequencies, and applied the LOCPRIOR model, which uses the sampling locations of each individual as prior information to improve clustering (Hubisz *et al.*, 2009). We specified 1,000,000 MCMC sample iterations after a burn-in of 1,000,000 iterations, and ran ten independent runs for $K = 1$ to 20. The most likely number of distinct genetic clusters was assessed from the mean log likelihood of the data [Ln(X|K)] and the ΔK statistic as described in Evanno *et al.* (2005), as implemented in STRUCTURE

HARVESTER (Earl & vonHoldt, 2012). We summarized results of the ten runs for the best-supported *K* value with CLUMPP v 1.1.2 (Jakobsson & Rosenberg, 2007) and visualized the genetic groups with DISTRUCT v 1.1 (Rosenberg, 2004). We followed the initial STRUCTURE run with a run using the correlated allele-frequency model and the same settings.

We tested each of the inferred clusters for deviations from Hardy-Weinberg and linkage equilibrium using exact tests in GENEPOP v 4.1.4 (Rousset, 2008; Guillot *et al.*, 2009). We tested for differentiation between the clusters using exact G tests based on allele frequencies (Goudet *et al.*, 1996). To evaluate concordance between mitochondrial and nuclear markers, we also grouped the microsatellite data according to mtDNA lineages (Luxbacher *et al.*, Chapter 1) and tested for differentiation using exact G tests. For all of these analyses, we assessed statistical significance with MCMC simulations (10,000 dememorization steps, 1,000 batches, 10,000 iterations/batch).

F_{ST} values were calculated for all population pairs in ARLEQUIN v 3.5.1.2 (Excoffier *et al.*, 2005). Due to missing data in three loci (PLJOe, PLJOk, PLJOq), analyses were based on six loci. Statistical significance was assessed from 1,000 permutations and adjusted by sequential Bonferroni correction for multiple comparisons (Rice, 1989). Because high heterozygosity can depress F_{ST} values (Hedrick, 1999; Karl *et al.*, 2012), we calculated pairwise and global values of Jost's D as an alternative method for quantifying genetic differentiation (Jost, 2008; Meirmans & Hedrick, 2011). These calculations were based on the same six loci used to calculate F_{ST}, and were made using SMOGD v 1.2.5 (Crawford, 2010).

We tested for isolation-by-distance (IBD) based on microsatellite measures of

differentiation (Slatkin's linearized F_{ST} and Jost's D) with the Isolation By Distance Web Service v 3.23 (Jensen *et al.*, 2005). We first calculated Euclidean and topographic distances for all population pairs using ARCGIS v 9.3 (ESRI). Euclidean distances were determined utilizing the 'Near' tool, and topographic distances were calculated from a 30 m digital elevation model using the 'Surface Length' tool, available in the Spatial Analyst extension. We conducted Mantel tests between genetic and Euclidean distances among all population pairs. If a significant correlation was found, then we conducted partial Mantel tests between genetic and topographic distances, controlling for the effect of Euclidean distance. We also tested for IBD by running separate Mantel tests on populations grouped according to clusters inferred from GENELAND and STRUCTURE (see below). In all tests, significance was assessed from 10,000 randomizations.

Ecological niche modeling

If discrete genetic units also exhibit ecological differentiation, they may harbor adaptive variation that would be important to maintain from a conservation standpoint (Crandall *et al.*, 2000; Rader *et al.*, 2005). To determine whether the three inferred genetic clusters are ecologically differentiated or ecologically exchangeable (Templeton, 1989; Rader *et al.*, 2005), we modeled their ecological niches using MAXENT v 3.3.3a (Phillips *et al.*, 2006). We obtained current climate data at 30 arc-second resolution from the WorldClim database (http://www.worldclim.org; Hijmans *et al.*, 2005). We reduced the data set to the following nine bioclimatic variables that we consider relevant to the ecological and physiological requirements of *P. jordani*: annual mean diurnal range, annual

temperature range, mean temperature of wettest quarter, mean temperature of driest quarter, mean temperature of warmest quarter, precipitation of wettest month, precipitation seasonality, precipitation of driest quarter, and precipitation of warmest quarter. We clipped climate layers to an area extending from 35.0° to 36.3° latitude and -84.5° to -82.5° longitude, which includes the current distribution of *P. jordani* within GRSM and surrounding mountains. Details about occurrence records and data layer preparation are described in Luxbacher *et al.* (Chapter 1).

We used ENMTOOLS v 1.4.3 to quantify niche overlap using Schoener's *D* (Schoener, 1968), Warren's *I* (Warren *et al.*, 2008), and relative rank (Warren & Seifert, 2011) statistics, and to perform a niche identity test, which evaluates whether the ecological niche models (ENMs) generated for the three clusters are more different than expected if they had been generated from the same underlying distribution (Warren *et al.*, 2010). Observed niche overlap values for each pair of models were compared to null distributions generated from 100 pseudoreplicates to assess statistical significance (Warren *et al.*, 2010).

Comparison with mtDNA

We amplified approximately 1116 nucleotide bases from the *ND2* mitochondrial gene and adjacent tRNAs from 324 individuals representing 70 populations (Figure 1). Primers and PCR conditions are described in Luxbacher *et al.* (Chapter 1). We constructed a statistical parsimony network for *ND2* using a 95% connection limit in TCS v 1.21 (Clement *et al.*, 2000), and expanded upon results presented in previous chapters by

considering geographic locations of haplogroups within each lineage. We also tested for IBD within each mtDNA clade using the Isolation By Distance Web Service. We estimated F_{ST} for all populations with two or more individuals in DNASP v 5 (Librado & Rozas, 2009), and conducted Mantel tests between F_{ST} and log Euclidean distance among all population pairs in each lineage. Negative genetic distances were set to zero (Meirmans, 2006). If a significant correlation was found, then we conducted partial Mantel tests between genetic and topographic distances, controlling for the effect of Euclidean distance. In all tests, significance was assessed from 10,000 randomizations.

RESULTS

Genetic diversity

The number of alleles per locus ranged from five (PLJOr) to 47 (PLJOg) with an average of 27.4 per locus (Table S1). Observed heterozygosities were variable, ranging from 0.21 to 0.854 with an average of 0.682 per locus (Table S1). Randomization tests indicated significant deviation from Hardy–Weinberg equilibrium for one locus (PLJOe) at one locality (56). There were no significant deviations from linkage equilibrium after sequential Bonferroni correction. Analysis with MICRO-CHECKER indicated no evidence of possible scoring errors due to stuttering or large allele dropout, but indicated seven loci (PLJOb, PLJOe, PLJOk, PLJOl, PLJOo, PLJOq, PLJOw) with possible null alleles. However, because no individual samples failed to amplify for these loci, we do not view null alleles as a pervasive problem and we retained all loci in subsequent

analyses.

Allelic richness ranged from an average of 1.99 to 7.6 alleles per locus. Averaged over all nine loci, richness ranged from 3.56 to 14.67 (mean = 7.31) per locality (Table S1). Average private allelic richness ranged from 0 to 0.21 alleles per locus. Averaged over all loci, private allelic richness ranged from 0 to 0.33 (mean = 0.065) per locality (Table S1). Four localities (12, 25, 40, 43) exhibited private alleles at multiple loci.

Bayesian clustering analyses

The uncorrelated allele-frequency model implemented in GENELAND inferred three genetic clusters that follow a west-east orientation (Figure 2). We followed this analysis with one using the correlated allele-frequency model. This analysis inferred sixteen clusters, but some clusters included or were entirely composed of "ghost populations" that had no individuals assigned to them (Guillot, 2008). Because the correlated allele frequency model can yield spurious results when data exhibit IBD as our data do (see below; Guillot *et al.*, 2009), we do not hold confidence in these clusters and only discuss the results of the uncorrelated allele frequency model.

The model implemented in STRUCTURE based on uncorrelated allele frequencies also suggests that three clusters provide the best fit to the data, as indicated by the mean probability and the ad hoc ΔK statistic (Mean LnP(K) = -9983.9 and ΔK = 58.06 for K = 3). These three clusters coincide geographically with the clusters inferred by the GENELAND analysis (Figure 2), and show admixture occurring at interior margins (Figure 3). The model based on correlated allele frequencies inferred two clusters

(results not shown). The split between them aligns with the western split inferred by the uncorrelated model (Figure 3), but this model failed to differentiate the eastern cluster in the uncorrelated model. Because the uncorrelated model is appropriate to use when one expects populations to differ in their allele frequencies as we do here given the expectation of low dispersal abilities among *P. jordani* (Petranka, 1998; Pritchard *et al.*, 2000), we focus our discussion on results of the uncorrelated allele frequency model.

Each of the three inferred clusters exhibited significant deviations from HWE at four to seven loci (Table S2). There were no significant deviations from linkage equilibrium in any clusters after sequential Bonferroni correction (results not shown). The three clusters are significantly differentiated based on exact G tests using allele frequencies ($p < 0.001$ for all pairwise comparisons; Table S3). Microsatellite allele frequencies remained significantly differentiated when grouped according to the three mtDNA lineages ($p < 0.001$ for all pairwise comparisons).

Genetic differentiation and ecological niche modeling

Pairwise F_{ST} from microsatellite loci ranged from near zero to 0.255 (Table S4). Tests of pairwise differentiation among sampling localities based on F_{ST} were significant for 13 of 780 comparisons after Bonferroni correction. Significant pairwise comparisons support a pattern of differentiation between western and eastern localities. Jost's D was generally high, ranging from 0.5 (PLJOb) to 0.76 (PLJOl) per locus, with the exception of PLJOr, which was near zero due to low number of alleles. Global genetic differentiation among 35 populations was $D_{est} = 0.197$ (95% CI 0.222–0.259). Pairwise comparisons of the

harmonic mean of D across all loci ranged from near zero to 0.59 (Table S4).

Results of Mantel tests revealed a significant positive association between genetic and geographic distance regardless of which measure of genetic differentiation was used (linearized F_{ST}: $r = 0.262$; $P < 0.001$; D_{est}: $r = 0.197$; $P = 0.002$). These results support a general pattern of IBD across the range of *P. jordani*. However, we found no significant pattern of isolation-by-topographic-distance after controlling for effect of Euclidean distance (linearized F_{ST}: $P = 0.276$; D_{est}: $P = 0.151$), and separate Mantel tests in which populations were grouped according to the three inferred clusters showed no significant association between linearized F_{ST} or Jost's D and Euclidean distance within each genetic cluster (results not shown).

ENMs for each cluster were generally similar, but Cluster 1 exhibited higher suitability in the western part of the distribution (Figure S1). Niche identity tests showed that none of the clusters exhibit statistically significant ecological differences with respect to the nine climatic variables used to build the models for any of the three overlap metrics (Figures S2-S4; Table S5).

mtDNA analysis

The mtDNA network revealed distinct haplogroups that correspond to geographic regions within GRSM (Figure 1). Genetic distances were strongly correlated with log Euclidean distance for Clade 1 ($r = 0.353$, $P < 0.001$) and Clade 2 ($r = 0.372$, $P < 0.001$). After accounting for the effect of Euclidean distance, genetic distance remained significantly correlated with topography in Clade 1 ($r = 0.275$, $P < 0.001$), but not Clade 2 ($r =$

0.069, $P = 0.266$). We found no significant correlation between genetic and geographic distance in Clade 3.

DISCUSSION

Cryptic diversity and structure recovered from microsatellites

Microsatellites revealed different areas of diversity across the range of *P. jordani* than presented in previous studies based on mtDNA (Luxbacher *et al.*, Chapter 1), but this finding is not surprising because mitochondrial and nuclear DNA can differ in their responses to evolutionary events due to their different modes of inheritance (Brito, 2007). In general, high microsatellite diversity was localized in two areas of GRSM. In the south-central part of the range, Localities 42 and 47 exhibited the highest allelic richness, and localities 45 and 47 had high heterozygosity and showed admixture with Cluster 3. In the western part of the range, localities with the highest private alleles (17, 23, 25) occur surrounding the genetic break detected by GENELAND and STRUCTURE (Figures 2, 3).

The Bayesian clustering analyses inferred three genetic clusters that correspond to the western, central, and eastern regions of the distribution. Cluster 1 includes 12 localities from the western part of the range, Cluster 2 includes 16 localities in the central part of the range, and Cluster 3 includes 10 localities in the eastern part of the range. The break between Clusters 1 and 2 was not detected in previous studies based on mtDNA or anonymous nuclear loci and may reflect recent restrictions in gene flow in this part of the

range (Selkoe & Toonen, 2006). Clingmans Dome, the highest elevation in the Great Smoky Mountains occurs within this break, and may have presented a barrier to gene flow during the Pleistocene when high elevations were too cold to sustain *P. jordani* populations (Luxbacher *et al.*, Chapter 1).

Microsatellite metrics of differentiation revealed a pattern of IBD across the range of *P. jordani*, but this trend is driven by differentiation within each cluster, as separate Mantel tests for each cluster showed no association between genetic and geographic distance. These findings suggest that gene flow is not restricted within genetic clusters and contrasts with the pattern of IBD detected in mtDNA Clades 1 and 2. Although slope and topography play an important role in structuring some montane amphibian populations (e.g., Funk *et al.*, 2005; Giordano *et al.*, 2007), we found no effect of isolation-by-topographic-distance for *P. jordani* based on microsatellite data, but did find an isolation-by-topographic-distance effect in mtDNA Clade 1. These contrasting patterns could be due to differences in dispersal between sexes (see below), but we also note that because sample sizes were smaller in microsatellite analyses, the difference could also be due to reduced statistical power in tests conducted on microsatellite clusters relative to mtDNA clades (Meirmans, 2012).

Comparison between mtDNA and microsatellites

We explicitly draw comparisons between mtDNA and microsatellites here, but note that three anonymous nuclear loci generally affirm mtDNA patterns (Luxbacher & Kozak, Chapter 2). The Bayesian clustering analyses inferred three genetic clusters that were

partially congruent with phylogeographic patterns based on mtDNA (Figures 2, 3; Luxbacher *et al.*, Chapter 1; Luxbacher & Kozak, Chapter 2), specifically at the boundary of microsatellite Clusters 2 and 3 that coincides with the break between mtDNA Clades 1 and 2. Although inferring evolutionary relationships from microsatellites alone can be complicated, the agreement between both marker types here provides strong support for long-term isolation and continued restricted gene flow among *P. jordani* in this region (Zhang & Hewitt, 2003).

Another pattern emerges in the western part of the range where both clustering algorithms detected structure that is not represented as a deep split in mtDNA (Luxbacher *et al.*, Chapter 1). Because the uniparental inheritance of mtDNA results in a fourfold smaller effective population size relative to nuclear DNA (Avise, 2000), we expect mtDNA to coalesce faster than nuclear loci, thus reinforcing any genetic break inferred from microsatellites. One possible explanation for the observed pattern is that a coinciding genetic break in mtDNA may have been obscured due to a selective sweep that homogenized mitochondrial variation (Ballard & Whitlock, 2004), but we previously found no evidence for selection acting on *ND2* (Luxbacher and Kozak, Chapter 2). Instead, a closer look at mtDNA haplogroups presented in this study shows that most of the localities sampled west of the break fall into a single haplogroup (Figure 1). Thus, both marker types show differentiation in this region and we do not view this as a case of mito-nuclear discordance.

In contrast, we do find an instance of mito-nuclear discordance in the eastern region of GRSM where mtDNA detects a deep split along Balsam Mountain and Sterling Ridge that is not inferred from microsatellites. One explanation for this pattern is that the

lineages are diverging, but not enough time has elapsed for nuclear markers to show the signature of divergence. Nuclear markers can be considered lagging indicators of demographic change due to their fourfold larger effective population size compared to mtDNA; therefore, they are more likely to retain ancestral polymorphism than mtDNA (Zink & Barrowclough, 2008).

Alternatively, this discordant pattern between mitochondrial and nuclear markers may be due to sex-biased dispersal (Jockusch & Wake, 2002; Johansson *et al.*, 2008). Sex-specific patterns of dispersal have not been studied in *P. jordani*, but if dispersal in this species is male-biased, as in other *Plethodon* species (Liebgold *et al.*, 2011), then we would expect the rate of nuclear gene flow to be higher than the rate of mitochondrial gene flow. This difference would generate the observed nuclear homogeneity, and is consistent with a hypothesis of historic genetic drift that occurred in isolated refugia combined with recent gene flow between the formerly isolated populations (Luxbacher *et al.*, Chapter 1; Monsen & Blouin, 2003; Yang & Kenagy, 2009). Other recent studies have shown that mtDNA can retain a signature of historical separation following secondary contact and hybridization longer than a limited number of nuclear loci (e.g., Ng & Glor, 2011).

An important caveat to any potential explanations for this mito-nuclear discordance is the relatively limited sampling in the southeastern region of GRSM. Four of the nine localities with clade 3 haplotypes have limited data (1-2 individuals sampled) and we did not genotype individuals from these sites at microsatellite loci. Future work, including more extensive sampling of individuals in this region will be necessary to localize any possible secondary contact.

Synthesis of data to delimit ESUs

Many approaches have been proposed to establish conservation strategies at the intraspecific level (e.g., Petit *et al.*, 1998; Pearse & Crandall, 2004; Bonin *et al.*, 2007; Diniz-Filho *et al.*, 2012). Here we aim to synthesize data generated from different molecular markers, ENMs, and historical distributions (Luxbacher *et al.*, Chapter 1; Luxbacher & Kozak, Chapter 2) to diagnose ESUs for *P. jordani* and propose monitoring and conservation recommendations (May *et al.*, 2011). ESUs can be defined as groups that exhibit reciprocal monophyly for mitochondrial alleles and significant divergence of allele frequencies at nuclear loci (Moritz, 1994). These criteria distinguish mtDNA Clades 1 and 2 as distinct ESUs. We contend that Clade 3 also meets the criteria for an ESU even though analyses of microsatellite data do not cluster individuals from localities corresponding to Clade 3 separately from those corresponding to Clade 2. The most plausible explanations for this pattern suggest that it is not a case of conflict between the two marker types (see above). Moreover, exact G tests confirm that microsatellite allele frequencies are significantly differentiated when individuals are grouped according to mtDNA Clades 2 and 3. We thus find sufficient support that Clade 3 meets Moritz's second criterion of significant divergence of allele frequencies at nuclear loci.

We further suggest that the ESU corresponding to Clade 1 comprises two separate management units (MUs). MUs include populations with significant divergence of allele frequencies at mitochondrial or nuclear loci, and are a useful unit for monitoring because they characterize current population structure (Moritz, 1994). Mitochondrial DNA

does not detect a deep break in the western distribution of *P. jordani*, but microsatellite loci meet the significant divergence criterion through the GENELAND and STRUCTURE analyses that recover two geographically clustered groups within Clade 1 and the exact G tests that confirm significant differentiation between these clusters.

The ESUs and MUs have highly-localized geographic distributions. The first ESU extends from the western edge of the distribution of *P. jordani* near Parson Bald to Hughes Ridge in the eastern part of GRSM. Within in this area, the western MU extends from Parson Bald to near Clingmans Dome, and the eastern MU runs from Clingmans Dome to Hughes Ridge. Because suitable microclimate for *P. jordani* in this region is predicted to decrease with climate change (K. H. Kozak, unpublished data), we suggest that populations in the western MU are at high risk of future decline or local extirpation. Populations at low-elevation range margins in this MU should thus be the focus of monitoring, as they may be the first to show responses to climate change (e.g., shifting distributions, changing population densities). The western MU was colonized relatively recently in the evolutionary history of *P. jordani* (Luxbacher & Kozak, Chapter 2). Given that Clusters 1 and 2 are ecologically exchangeable based on our ENM analyses, we suggest that conservation resources should focus instead on maintaining populations in the eastern MU, as well as the other two ESUs that have been occupied for a longer period of time and exhibit higher microclimate suitability for *P. jordani* (Luxbacher *et al.*, Chapters 1; Luxbacher & Kozak, Chapter 2).

The second ESU occurs in the northeastern part of GRSM along the main ridgeline and the northern ridge of Balsam Mountain extending to Mt. Sterling. The third ESU extends up Balsam Mountain and along Sterling Ridge. Divergence estimates suggests

that populations in this area have a long history of isolation from other *P. jordani* (Luxbacher & Kozak, Chapter 2), making them of high interest for conservation. This area lies in a relatively remote part of GRSM (minimal road and trail access), and warrants future sampling to determine the geographic extent of *P. jordani* and the contact zone between Clades 2 and 3. Particular attention should be paid to individuals in the regions of apparent secondary contact here, and at the boundary between Clades 1 and 2, because interaction between divergent lineages may facilitate novel evolutionary outcomes that can benefit the future of species (Moritz *et al.*, 2009).

Beyond these ESUs and MUs, focusing monitoring efforts generally on populations at their elevational range limits will be a critical aspect of future conservation plans for *P. jordani*. Gifford and Kozak (2012) investigated the physiological traits of *P. jordani* in relation to climate using a mechanistic biophysical model that predicts the spatial distribution of climatically-suitable habitat for *P. jordani* based on where individuals can be surface active for long enough to meet their energetic needs. They found that low elevation populations are already living near the limit of their physiological tolerances (Gifford & Kozak, 2012). Therefore, as we suggest for the western MU, monitoring populations at their low elevational margins may provide the first indications of how *P. jordani*, as well as other montane endemic species, will respond to climatic changes.

Over 30 species of salamanders occur in GRSM, but *P. jordani* is the Park's only endemic salamander (Dodd, 2004). Because its range lies within the boundaries of a National Park, *P. jordani* is protected from many of the factors causing declines in other species (e.g., habitat loss due to land use change; Collins & Storfer, 2003; Beebee & Griffiths, 2005), but the threat of climate change transcends reserve boundaries, and

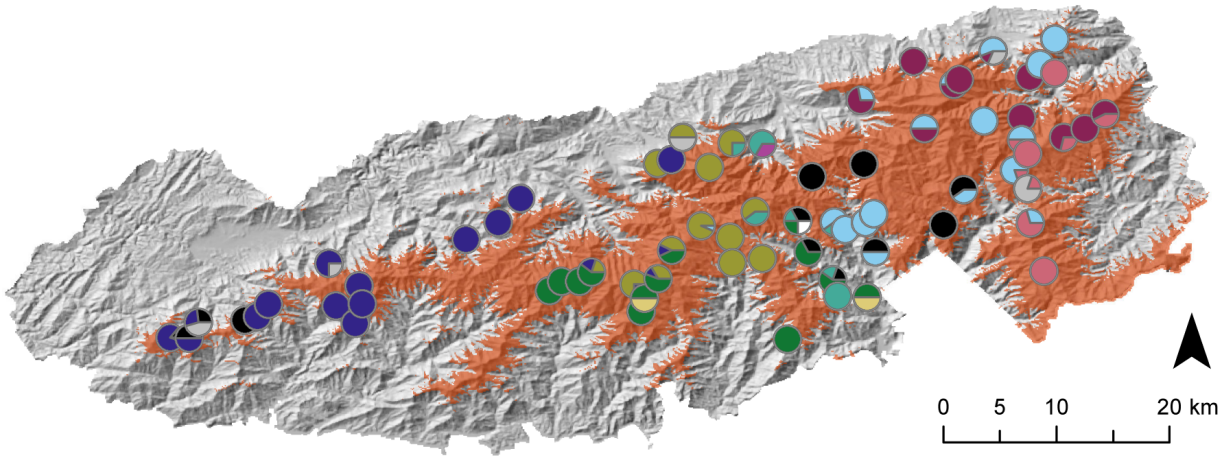
the effects of climate change on montane salamanders are predicted to be profound (Parra-Olea *et al.*, 2005; Wake, 2009; Milanovich *et al.*, 2010). Plethodontid salamanders are generally considered indicators of biodiversity and forest health (Noss, 1999; Welsh & Droege, 2001; Davic & Welsh, 2004); therefore, tracking how species like *P. jordani* respond to climate change may provide some indication about how other species may respond, as well as broader changes in temperate forest communities. The Great Smoky Mountains includes many other regional species that are dispersal-limited and restricted to high elevations, and may show similar patterns of genetic structure to *P. jordani* (Thomas & Hedin, 2008). As patterns of genetic variation and structure in other regional species are studied, the ESUs we recommend can be incorporated into a broader complementarity approach for managing species' responses to the effects of climate change (Margules & Pressey, 2000; Groves *et al.*, 2012).

Table 1 Characteristics of nine microsatellite loci developed for *Plethodon jordani*

Locus	Repeat motif	Primer sequences (5'—3')	T_A	Primer concentration	M13 label	Allele size range	N	N_A	H_o	H_e
PLJOb	GACA[i]	F: GTTTCTTGCAGTGACAGCTGTGGTGAT R: TGTAAAACGACGGCCAGTTGCAGGAAGAACCAAAGAGG	58	0.2 μM	PET	48-144	238	18	0.595	0.776
PLJOe	TAGA	F: TGTAAAACGACGGCCAGTAATTCCACTCTGCACCCTTG R: GTTTCTTAGTGTTGGCTGCCTGTGTCT	62	0.2 μM	NED	28-200	207	36	0.822	0.889
PLJOg	TAGA[i]	F: GTTTCTTACCACATGATTTTGGTCGTG R: TGTAAAACGACGGCCAGTGACCTTGGTATGATGTGCTCAA	52	0.24 μM	PET	52-208	238	30	0.854	0.914
PLJOk	TCTA[i]	F: GTTTCTTTCCCCTTAATTCACATTTACCA R: TGTAAAACGACGGCCAGTTAAAAGCAATCAGCCTTTGG	50	0.16 μM	NED	24-116	219	22	0.765	0.785
PLJOl	GAA	F: TGTAAAACGACGGCCAGTTGGGGATGGGAAGAATAAGA R: GTTTCTTCACTGGTTATGGGGCACAC	54	0.16 μM	VIC	42-183	232	43	0.777	0.940
PLJOo	ATCT[i]	F: TGTAAAACGACGGCCAGTTCAAATCATGGTTAGTCCACA R: GTTTCTTGCCTATCTGCTTTGGTGTTT	52	0.24 μM	NED	24-296	236	47	0.778	0.837
PLJOq	ATC	F: TGTAAAACGACGGCCAGTGAATCACCACTCTGTGGGATAA R: GTTTCTTGTGATGGCTGTGGCTAAAGT	54	0.16 μM	PET	9-261	180	17	0.573	0.777
PLJOr	ATG	F: GTTTCTTAGCTCTGATGGCCTCTTGTG R: TGTAAAACGACGGCCAGTACCAAACACTCAGAGCATCG	57	0.24 μM	6-FAM	18-30	239	5	0.210	0.251
PLJOw	ATCT	F: TGTAAAACGACGGCCAGTCCAGATCAGGAGGTAAAGGTG R: GTTTCTTAGCCGACAAAAGGTAAAGGA	52	0.16 μM	6-FAM	12-128	230	29	0.766	0.908

T_A: annealing temperature (°C); N: number of individuals genotyped; N_A: number of alleles; H_o: observed heterozygosity; H_e: expected heterozygosity

[i]: imperfect repeat motif based on sequence


0 5 10 20 km

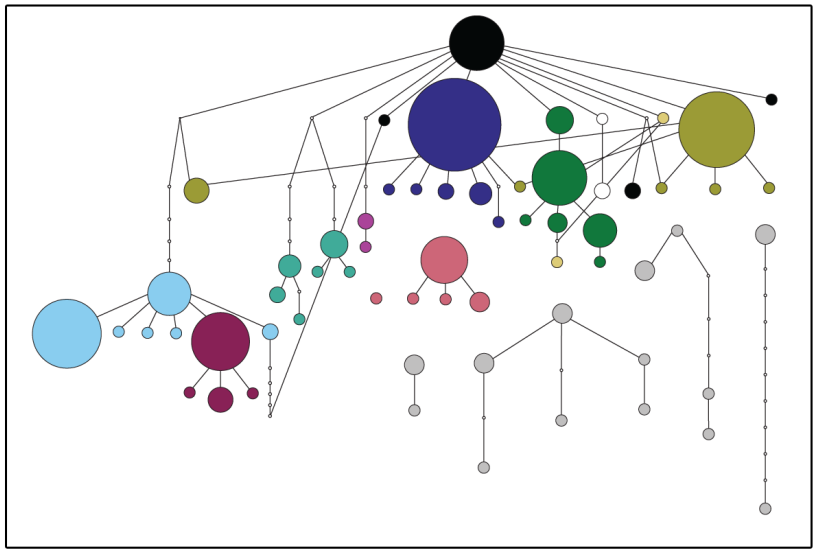


Figure 1 Map of geographic distribution of *P. jordani* and 70 sampling localities for mtDNA in Great Smoky Mountains National Park (North Carolina and Tennessee). Distribution (based on a mechanistic bioenergetic model; Gifford & Kozak, 2012) is depicted in vermilion and overlaid on a 30 m digital elevation model. Localities are presented as pie diagrams with slices colored and sized according to the frequency of individuals representing different haplogroups (inset). The area of the enlarged region in the eastern United States is shown on map in lower left. **Inset** Haplotype network for *ND2* estimated using statistical parsimony with a 95% connection significance. The size of each circle is proportional its haplotype frequency and open circles represent inferred unsampled haplotypes. Haplotypes are colored according to haplogroups identified by phylogenetic analyses (Luxbacher *et al.*, Chapter 1). Individuals not represented in the three major clades are shown in pale gray

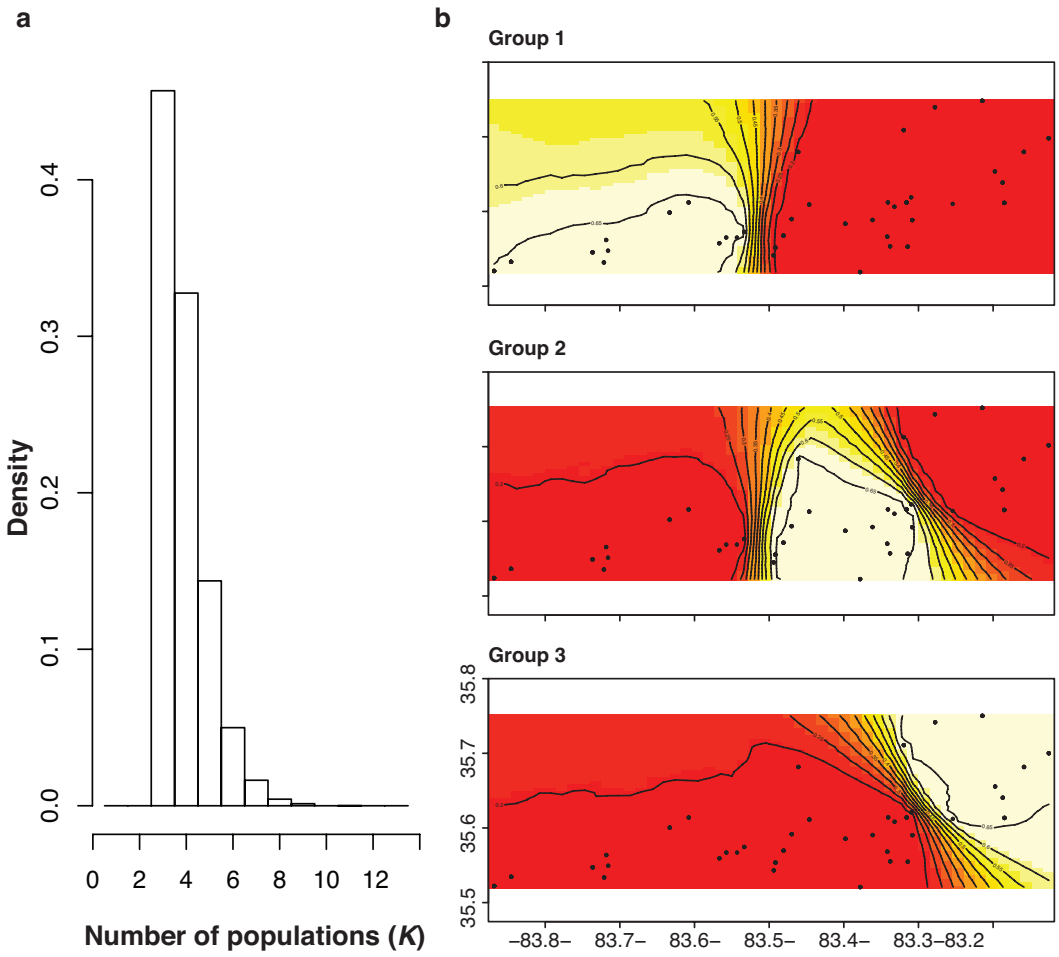


Figure 2 GENELAND results for 237 *P. jordani* from 38 localities based on 9 microsatellite loci. **a** posterior probability (after burn-in) of the number of genetic groups showing support for K = 3. **b** posterior probability of membership to the three inferred groups (K). High values are shown in the yellow area surrounded by 0.65 isocline. Sampling localities are plotted by latitude and longitude

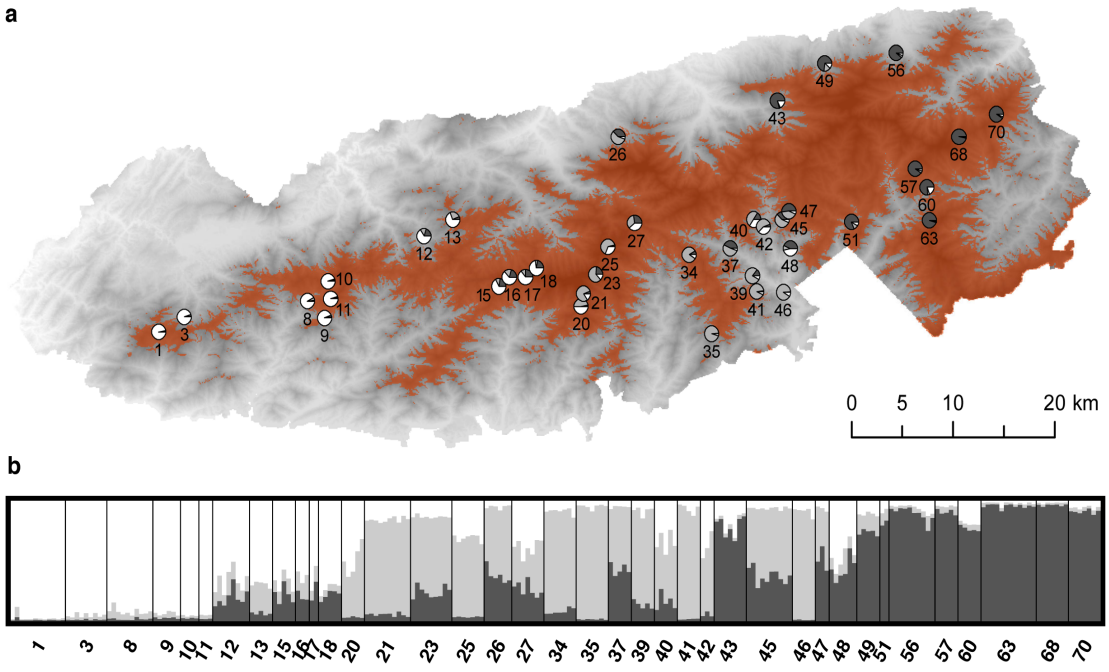


Figure 3 STRUCTURE population clustering results for 237 *P. jordani* based on 9 microsatellite loci. **a** Map of 38 localities from which individuals were genotyped. Localities are numbered as in Table S1 and presented as pie diagrams with slice size proportional to the estimated membership coefficients in each of the three inferred clusters. **b** Each bar represents the proportion of membership in each inferred population cluster for each individual. Individuals are grouped by sampling localities ordered by latitude

BIBLIOGRAPHY

Anderson RP (2013) A framework for using niche models to estimate impacts of climate change on species distributions. *Annals of the New York Academy of Sciences*, **1297**, 8–28.

Anderson RP, Raza A (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*, **37**, 1378–1393.

Austin, JD, Lougheed SC, Boag PT (2004) Discordant temporal and geographic patterns in maternal lineages of eastern north American frogs, *Rana catesbeiana* (Ranidae) and *Pseudacris crucifer* (Hylidae). *Molecular Phylogenetics and Evolution*, **32**, 799–816.

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

Baele G, Lemey P, Bedford T *et al.* (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, **29**, 2157–2167.

Bálint M, Domisch S, Engelhardt CHM *et al.* (2011) Cryptic biodiversity loss linked to global climate change. *Nature Climate Change*, **1**, 313–318.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Barata M, Carranza S, Harris DJ (2012) Extreme genetic diversity in the lizard *Atlantolacerta andreanskyi* (Werner, 1929): a montane cryptic species. *BMC Evolutionary Biology*, **12**, 167.

Barker BS, Waide RB, Cook JA (2011) Deep intra-island divergence of a montane forest endemic: phylogeography of the Puerto Rican frog *Eleutherodactylus portoricensis* (Anura: Eleutherodactylidae). *Journal of Biogeography*, **38**, 2311–2325.

Barlow A, Baker K, Hendry CR *et al.* (2013) Phylogeography of the widespread African puff adder (*Bitis arietans*) reveal multiple Pleistocene refugia in southern Africa. *Molecular Ecology*, **22**, 1134–1157.

Barrowclough GF, Zink RM (2009) Funds enough and time: mtDNA, nuDNA, and the discovery of divergence. *Molecular Ecology*, **18**, 2934–2936.

Beebee TJC (2005) Conservation genetics of amphibians. *Heredity*, **95**, 423–427.

Beebee TJC, Griffiths RA (2005) The amphibian decline crisis: a watershed for

conservation biology? *Biological Conservation*, **125**, 271–285.

Bell RC, Parra JL, Tonione M *et al.* (2010) Patterns of persistence and isolation indicate resilience to climate change in montane rainforest lizards. *Molecular Ecology*, **19**, 2531–2544.

Bielejec F, Rambaut A, Suchard MA, Lemey P (2011) SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics*, **27**, 2910–2912.

Bintanja R, van de Wal RSW (2008) North American ice-sheet dynamics and the onset of 100,000-year glacial cycles. *Nature*, **454**, 869–872.

Bonatelli IAS, Perez MF, Peterson AT *et al.* (2014) Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology*, **23**, 3044–3063.

Bonin A, Nicole F, Pompanon F, Miaud C, Taberlet P (2007) Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conservation Biology*, **21**, 697–708.

Boutin-Ganache I, Raposo M, Raymond M, Deschepper CF (2001) M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *BioTechniques*, **31**, 25–28.

Brito PH (2007) Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (*Strix aluco*). *Molecular Ecology*, **16**, 3423–3437.

Brown JM, Hedtke SM, Lemmon AR, Lemmon EM (2010) When trees grow too long: investigating the causes of highly inaccurate Bayesian branch-length estimates. *Systematic Biology*, **59**, 145–161.

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. *BioTechniques*, **20**, 1004–1010.

Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS (2001) Comparative phylogeography of northwestern North America: a synthesis. *Integrating ecology and evolution in a spatial context* (ed. by J. Silvertown and J. Antonovics), pp 319–339. Blackwell Publishing, Williston.

Bryant D, Moulton V (2004) Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, **21**, 255–265.

Buckley LB, Urban MC, Angilletta MJ *et al.* (2010) Can mechanism inform species distribution models? *Ecology Letters*, **13**, 1041–1054.

Cadena CD, Kozak KH, Gómez JP *et al.* (2011) Latitude, elevational climatic zonation, and speciation in New World vertebrates. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 194–201.

Camargo A, Werneck FP, Morando M, Sites JW Jr, Avila LJ (2013) Quaternary range and demographic expansion of *Liolaemus darwinii* (Squamata: Liolaemidae) in the Monte Desert of Central Argentina using Bayesian phylogeography and ecological niche modeling. *Molecular Ecology*, **22**, 4038–4054.

Canestrelli D, Cimmurata R, Nascetti G (2007) Phylogeography and historical demography of the Italian treefrog, *Hyla intermedia*, reveals multiple refugia, population expansions and secondary contacts within peninsular Italy. *Molecular Ecology*, **16**, 4808–4821.

Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C (2009) Stability predicts genetic diversity in the Brazilian Atlantic Forest Hotspot. *Science*, **323**, 785–789.

Carstens BC, Knowles LL (2007) Shifting distributions and speciation: divergence during rapid climate change. *Molecular Ecology*, **16**, 619–627.

Carstens BC, Richards CL (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution*, **61**, 1439–1454.

Caruso NM, Lips KR (2012) Truly enigmatic declines in terrestrial salamander populations in Great Smoky Mountains National Park. *Diversity and Distributions*, **19**, 38–48.

Chabarria RE, Pezold F (2013) Phylogeography and historical demography of *Sicydium salvini* in the eastern Pacific. *Ichthyology Research*, **60**, 353–362.

Chan LM, Choi D, Raselimanana AP *et al.* (2012) Defining spatial and temporal patterns of phylogeographic structure in Madagascar's iguanid lizards (genus *Oplurus*). *Molecular Ecology*, **21**, 3839–3851.

Charrette NA, Cleary DFR, Mooers AØ (2006) Range-restricted, specialist Bornean butterflies are less likely to recover from ENSO-induced disturbance. *Ecology*, **87**, 2330–2337.

Chatfield MWH, Kozak KH, Fitzpatrick BM, Tucker PK (2010) Patterns of differential introgression in a salamander hybrid zone: inferences from genetic data and ecological niche modeling. *Molecular Ecology*, **19**, 4265–4282.

Church SA, Kraus JM, Mitchell JC, Church DR, Taylor DR (2003) Evidence for multiple Pleistocene refugia in the postglacial expansion of the eastern tiger salamander, *Ambystoma tigrinum tigrinum*. *Evolution*, **57**, 372–383.

Clark JS, Fastie C, Hurtt, G *et al.* (1998) Reid's paradox of rapid plant migration. *Bioscience*, **48**, 13–24.

Clark PU, Archer D, Pollard D *et al*. (2006) The Middle Pleistocene transition: characteristics, mechanisms, and implications for long-term changes in atmospheric pCO2. *Quaternary Science Reviews*, **25**, 3150–3184.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Collins JP, Storfer A (2003) Global amphibian declines: sorting the hypotheses. *Diversity and Distributions*, **9**, 89–98.

Cordellier M, Pfenninger M (2009) Inferring the past to predict the future: climate modeling predictions and phylogeography for the freshwater gastropod *Radix balthica* (Pulmonata, Basommatophora). *Molecular Ecology*, **18**, 534–544.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution*, **15**, 290–295.

Crawford NG (2010) SMOGD: software for the measurement of genetic diversity. *Molecular Ecology Resources*, **10**, 556–557.

Crespi EJ, Rissler LJ, Brown RA (2003) Testing Pleistocene refugia theory: phylogeographical analysis of *Desmognathus wrighti*, a high-elevation salamander in the southern Appalachians. *Molecular Ecology*, **12**, 969–984.

Cummings MP, Neel MC, Shaw KL (2008) A genealogical approach to quantifying lineage divergence. *Evolution*, **62**, 2411–2422.

Darriba D, Taboada GL, Doallo R, Posada D (2012) JMODELTEST 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

Davic RD, Welsh HH Jr (2004) On the ecological roles of salamanders. *Annual Review of Ecology and Systematics*, **35**, 405–434.

Davis MB (1983) Quaternary history of deciduous forests of eastern North America and Europe. *Annals of the Missouri Botanical Garden*, **70**, 550–563.

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673–679.

Delcourt PA, Delcourt HR (1980) Vegetation maps for eastern North America: 40,000 yr B.P. to the present. In: *Geobotany II* (ed Romans RC), pp. 123–165. Plenum Press,

New York.

Delcourt PA, Delcourt HR (1998) Paleoecological insights on conservation of biodiversity: a focus on species, ecosystems, and landscapes. *Ecological Applications*, **8**, 921–934.

Demos TC, Peterhans JCK, Agwanda B, Hickerson MJ (2014) Uncovering cryptic diversity and refugial persistence among small mammal lineages across the Eastern Afromontane biodiversity hotspot. *Molecular Phylogenetics and Evolution*, **71**, 41–54.

Devitt T, Devitt SEC, Hollingsworth B, McGuire J, Moritz C (2013) Montane refugia predict population genetic structure in the Large-blotched *Ensatina* Salamander. *Molecular Ecology*, **22**, 1650–1665.

Ding L, Gan X, He S, Zhao E (2011) A phylogeographic, demographic and historical analysis of the short-tailed pit viper (*Gloydius brevicaudus*): evidence for early divergence and late expansion during the Pleistocene. *Molecular Ecology*, **20**, 1905–1922.

Diniz-Filho JAF, Telles MPC (2002) Spatial autocorrelation analysis and the identification of operational units for conservation in continuous populations. *Conservation Biology*, **16**, 924–935.

Diniz-Filho JAF, Melo DB, de Oliviera G *et al.* (2012) Planning for optimal conservation of geographical genetic variability within species. *Conservation Genetics*, **13**, 1085–1093.

Dmitriev DA, Rakitov RA (2008) Decoding of superimposed traces produced by direct sequencing of heterozygous indels. *PLoS Computational Biology*, **4**, e1000113.

Dobrowski SZ (2010) A climatic basis for microrefugia: the influence of terrain on climate. *Global Change Biology*, **17**, 1022–1035.

Dodd CK Jr (2004) *The Amphibians of Great Smoky Mountain National Park*. University of Tennessee Press, Knoxville.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Duellman WE, Sweet SS (1999) Distribution patterns of amphibians in the Nearctic region of North America. In: *Patterns of Distribution of Amphibians: A Global Perspective* (ed Duellman WE), pp 31–109. Johns Hopkins University Press, Baltimore.

Dunn ER (1926) *The Salamanders of the Family Plethodontidae*. Smith College Publications, Northampton.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Edwards CJ, Soulsbury CD, Statham MJ *et al.* (2012) Temporal genetic variation in the red fox, *Vulpes vulpes*, across western Europe and the British isles. *Quaternary Science Reviews*, **57**, 95–104.

Edwards S, Bensch S (2009) Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Molecular Ecology*, **18**, 2930–2933.

Elith J, Kearney M, Phillips S (2010) The art of modeling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

Elith J, Phillips SJ, Hastie T *et al.* (2011) A statistical explanation of MAXENT for ecologists. *Diversity and Distributions*, **17**, 43–57.

Emel SL, Storfer A (2012) A decade of amphibian population genetic studies: synthesis and recommendations. *Conservation Genetics*, **13**, 1685–1689.

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

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Eytan RI, Hellberg ME (2010) Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. *Evolution*, **64**, 3380–3397.

Feder M (1983) Integrating the ecology and physiology of plethodontid salamanders. *Herpetologica*, **39**, 291–310.

Flot JF (2010) SEQPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. *Molecular Ecology Resources*, **10**, 162–166.

Fontanella FM, Feldman CR, Siddall ME, Burbrink FT (2008) Phylogeography of *Diadophis punctatus*: extensive lineage diversity and repeated patterns of historical demography in a trans-continental snake. *Molecular Phylogenetics and Evolution*, **46**, 1049–1070.

Frankham R, Ballou JD, Briscoe DA (2002) *An Introduction to Conservation Genetics*. Cambridge University Press, New York.

Fridley JD (2009) Downscaling climate over complex terrain: high fine-scale spatial variation of near-ground temperatures in a montane forested landscape (Great Smoky Mountains, USA). *Journal of Applied Meteorology and Climatology*, **48**, 1033–1049.

Fu Y (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 397–423.

Funk WC, Blouin MS, Corn PS *et al.* (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology*, **14**, 483–496.

Galbreath KE, Hafner DJ, Zamudio KR (2009) When cold is better: climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American pika, *Ochotona princeps*). *Evolution*, **63**, 2848–2863.

Garrick RC (2011) Montane refuges and topographic complexity generate and maintain invertebrate biodiversity: recurring themes across space and time. *Journal of Insect Conservation*, **15**, 469–478.

Garrick RC, Sunnucks P, Dyer RJ (2010) Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. *BMC Evolutionary Biology*, **10**, 118.

Gibbard P, Van Kolfschoten T (2004) The Pleistocene and Holocene epochs. In: *A Geologic Time Scale (*eds Gradstein FM, Ogg JG & Smith AG), pp. 441–452. Cambridge University Press, London.

Gifford ME, Kozak KH (2012) Islands in the sky or squeezed at the top? Ecological causes of elevational range limits in montane salamanders. *Ecography*, **35**, 193–203.

Giordano AR, Ridenhour BJ, Storfer A (2007) The influence of altitude and topography on genetic structure in the long-toed salamander *(Ambystoma macrodactulym*). *Molecular Ecology*, **16**, 1625–1637.

Glenn TC, Schable NA (2005) Isolating microsatellite DNA loci. In*: Molecular Evolution: Producing the Biochemical Data, Part B* (eds Zimmer EA & Roalson E), pp 202–222. Academic Press, San Diego.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www2.unil.ch/popgen/softwares/fstat.htm

Goudet J, Raymond M, DeMeeüs T, Rousset F (1996) Testing genetic differentiation in diploid populations. *Genetics*, **144**, 1933–1940.

Graham CH, Moritz C, Williams SE (2006) Habitat history improves prediction of biodiversity in rainforest fauna. *Proceedings of the National Academy of Sciences, USA*, **103**, 632–636.

Graham CH, Carnaval AC, Cadena CD *et al.* (2014) The origin and maintenance of montane diversity: integrating evolutionary and ecological processes. *Ecography*, **37**, 711–719.

Groves CR, Game ET, Anderson MG *et al.* (2012) Incorporating climate change into systematic conservation planning. *Biodiversity and Conservation*, **21**, 1651–1671.

Guillot G (2008) Inference of structure in subdivided populations at low levels of genetic differentiation—the correlated allele frequencies model revisited. *Bioinformatics*, **24**, 2222–2228.

Guillot G, Mortier F, Estoup, A (2005a) GENELAND: A computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 708–711.

Guillot G, Estoup A, Mortier F, Cosson JF (2005b) A spatial statistical model for landscape genetics. *Genetics*, **170**, 1261–1280.

Guillot G, Leblois R, Coulon A, Frantz AC (2009) Statistical methods in spatial genetics. *Molecular Ecology*, **18**, 4734–4756.

Guindon S, Gascuel O (2003) A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology*, **52**, 696–704.

Guo Q, Kelt DA, Sun Z *et al.* (2013) Global variation in elevational diversity patterns. *Science Reports*, **3**, 3007.

Guralnick R (2007) Differential effects of past climate warming on mountain and flatland species distributions: a multispecies North American mammal assessment. *Global Ecology and Biogeography*, **16**, 14–23.

Hare M (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology and Evolution*, **16**, 700–706.

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.

Heled J, Drummond AJ (2008) Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, **8**, 289.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt GM (1999) Post-glacial recolonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 183–195.

Highton R (2005) Declines of eastern North American woodland salamanders (*Plethodon*). In: *Status and Conservation of U.S. Amphibians* (ed Lannoo MJ), pp 34–46. University of California Press, Berkeley.

Highton R, Peabody RB (2000) Geographic protein variation and speciation in salamanders of the *Plethodon jordani* and *Plethodon glutinosus* complexes in the southern Appalachian Mountains with the descriptions of four new species. In: *The Biology of Plethodontid Salamanders* (eds Bruce RC, Jaeger RG & Houck LD), pp. 31–93. Kluwer Academic/Plenum Publishers, New York.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Hudson RR, Kreitman M, Aguadé M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153–159.

Hugall A, Moritz C, Moussalli A, Stanisic J (2002) Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). *Proceedings of the National Academy of Sciences, USA*, **99**, 6112–6117.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.

Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution*, **56**, 2383–2394.

Jakobsson M, Rosenberg N (2007) CLUMPP: a cluster matching and permutation program for dealing with multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jansson R, Dynesius M (2002) The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. *Annual Review of Ecology and Systematics*, **33**, 741–777.

Janzen DH (1967) Why mountain passes are higher in the tropics. *American Naturalist*, **101**, 233–249.

Jehle R, Arntzen JW (2002) Microsatellite markers in amphibian conservation genetics. *Herpetological Journal*, **12**, 1–9.

Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. *BMC Genetics*, **6**, 13.

Jockusch EL, Wake DB (2002) Falling apart and merging: diversification of slender salamanders (Plethodontidae: *Batrachoseps*) in the American West. *Biological Journal of the Linnean Society*, **76**, 361–391.

Johansson H, Surget-Groba Y, Thorpe RS (2008) The roles of allopatric divergence and natural selection in quantitative trait variation across a secondary contact zone in the lizard *Anolis roquet*. *Molecular Ecology*, **17**, 5146–5156.

Joly S, Bruneau A (2006) Incorporating allelic variation for reconstructing the evolutionary history of organisms from multiple genes: an example from *Rosa* in North America. *Systematic Biology*, **55**, 623–636.

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.

Karl SA, Toonen RJ, Grant WS, Bowen BW (2012) Common misconceptions in molecular ecology: echoes of the modern synthesis. *Molecular Ecology*, **21**, 4171–4189.

Kass RE, Raftery AE (1995) Bayes Factors. *Journal of the American Statistical Association*, **90**, 773–795.

Kearney M, Porter WP (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334–350.

Kieswetter CM, Schneider CJ (2013) Phylogeography in the northern Andes: Complex history and cryptic diversity in a cloud forest frog, *Pristimantis w-nigrum* (Craugastoridae). *Molecular Phylogenetics and Evolution*, **69**, 417–429.

Knowles LL (2000) Tests of Pleistocene speciation in montane grasshoppers (Genus

Melanoplus) from the sky islands of western North America. *Evolution*, **54**, 1337–1348.

Knowles LL (2001) Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, **10**, 691–701.

Knowles LL, Alvarado-Serrano DF (2010) Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: insights from coupled ecological, demographic and genetic models in montane grasshoppers. *Molecular Ecology*, **19**, 3727–3745.

Knowles LL, Carstens BC, Keat ML (2007) Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Current Biology*, **17**, 940–946.

Kozak KH, Wiens JJ (2006) Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution*, **60**, 2604–2621.

Kozak KH, Wiens JJ (2010) Niche conservatism drives elevational diversity patterns in Appalachian salamanders. *American Naturalist*, **176**, 40–54.

Kozak KH, Graham CH, Wiens JJ (2008) Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology and Evolution*, **23**, 141–148.

Kozak KH, Mendyk RW, Wiens JJ (2009) Can parallel diversification occur in sympatry? Repeated patterns of body-size evolution in co-existing clades of North American salamanders. *Evolution*, **63**, 1769–1784.

La Sorte FA, Jetz W (2010) Projected range contractions of montane biodiversity under global warming. *Proceedings of the Royal Society B: Biological Sciences,* **277**, 3401–3410.

Lawson LP (2010) The discordance of diversification: evolution in the tropical-montane frogs of the Eastern Arc Mountains of Tanzania. *Molecular Ecology*, **19**, 4046–4060.

Lawson LP (2013) Diversification in a biodiversity hot spot: landscape correlates of phylogeographic patterns in the African spotted reed frog. *Molecular Ecology*, **22**, 1947–1960.

Lee-Yaw JA, Irwin JT, Green DM (2008) Postglacial range expansion from northern refugia by the wood frog, *Rana sylvatica*. *Molecular Ecology*, **17**, 867–884.

Lemey P, Rambaut A, Welch JJ, Suchard MA (2010) Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution*, **27**, 1877–1885.

Lessa EP, Cook JA, Patton JL (2003) Genetic footprints of demographic expansion in

North America, but not Amazonia, during the Late Quaternary. *Proceedings of the National Academy of Sciences*, **100**, 10331–10334.

Librado P, Rozas J (2009) DNASP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Liebgold EB, Brodie ED III, Cabe PR (2011) Female philopatry and male-biased dispersal in a direct-developing salamander, *Plethodon cinereus*. *Molecular Ecology*, **20**, 249–257.

Linzey DW (2008) *A Natural History Guide to Great Smoky Mountains National Park*. University of Tennessee Press, Knoxville, TN.

Lisiecki LE, Raymo ME (2005) A Pliocene–Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleooceanography*, **20**, PA1003.

Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385–393.

Lomolino MV (2001) Elevation gradients of species-density: historical and prospective views. *Global Ecology and Biogeography*, **10**, 3–13.

Lowe WH, McPeek MA, Likens GE, Cosentinos BJ (2008) Linking movement behaviour to dispersal and divergence in plethodontid salamanders. *Molecular Ecology*, **17**, 4459–4469.

Margules CR, Pressey RL (2000) Systematic conservation planning. *Nature*, **405**, 243–253.

Marino IAM, Pujolar JM, Zane L (2011) Reconciling deep calibration and demographic history: Bayesian inference of post glacial colonization patterns in *Carcinus aestuarii* (Nardo, 1847) and *C. maenas* (Linnaeus, 1758). *PLoS One*, **6**, e28567.

Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics work-flows. *Bioinformatics*, **25**, 1982–1983.

May SE, Medley KA, Johnson SA, Hoffman EA (2011) Combining genetic structure and ecological niche modeling to establish units of conservation: a case study of an imperiled salamander. *Biological Conservation*, **144**, 1441–1450.

Mayr E (1954) Change of genetic environment and evolution. In: *Evolution as a Process* (eds Huxley J, Hardy AC & Ford EB), pp. 157–180. E.B. Allen & Unwin, London.

McCain CM (2005) Elevational gradients in diversity of small mammals. *Ecology*, **86**, 366–372.

McCain CM, Colwell RK (2011) Assessing the threat to montane biodiversity from discordant shifts in temperature and precipitation in a changing climate. *Ecology Letters*, **14**, 1236–1245.

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, **351**, 652–654.

McKay BD, Mays HL Jr, Peng Y *et al.* (2010) Recent range-wide demographic expansion in a Taiwan endemic montane bird, Steere's Liocichla (*Liocichla steerii*). *BMC Evolutionary Biology*, **10**, 71.

Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*, **60**, 2399–2402.

Meirmans PG (2012) The trouble with isolation by distance. *Molecular Ecology*, **21**, 2839–2846.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Milanovich JR, Peterman WE, Nibbelink NP, Maerz JC (2010) Projected loss of a salamander diversity hotspot as a consequence of projected global climate change. *PLoS ONE*, **5**, e12189.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA, pp. 1–8.

Milne I, Wright F, Rowe G *et al.* (2004) TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics*, **20**, 1806–1807.

Minin VN, Bloomquist EW, Suchard MA (2008) Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Molecular Biology and Evolution*, **25**, 1459–1471.

Monsen KJ, Blouin MS (2003) Genetic structure in a montane ranid frog: restricted gene flow and nuclear-mitochondrial discordance. *Molecular Ecology*, **12**, 3275–3286.

Moritz C (1994) Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology and Evolution*, **9**, 373–375.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238–254.

Moritz C, Hoskin CJ, MacKenzie JB *et al.* (2009) Identification and dynamics of a

cryptic suture zone in tropical rainforest. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1235–1244.

Mueller RL (2006) Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Systematic Biology*, **55**, 289–300.

Myers N, Mittermeier R, Mittermeier C, Fonseca G, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Nakazato T, Warren DL, Moyle LC (2010) Ecological and geographic modes of species divergence in wild tomatoes. *American Journal of Botany*, **97**, 680–693.

Nalepa CA, Luykx P, Klass K, Deitz LL (2002) Distribution of karyotypes of the *Cryptocercus punctulatus* species complex (Dictyoptera: Cryptocercidae) in the Southern Appalachians: relation to habitat and history. *Annals of the Entomological Society of America*, **95**, 276–287.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Ng J, Glor RE (2011) Genetic differentiation among populations of a Hispaniolan trunk anole that exhibit geographical variation in dewlap colour. *Molecular Ecology*, **20**, 4302–4317.

Nichols BJ, Langdon KR (2007) The Smokies all taxa biodiversity inventory: history and progress. *Southeastern Naturalist*, 6(special issue 1), 27–34.

Nishikawa KC (1990) Intraspecific spatial relationships of two species of terrestrial salamanders. *Copeia*, **1990**, 418–426.

Nogués-Bravo D (2009) Predicting the past distribution of species climatic niche. *Global Ecology and Biogeography*, **18**, 521–531.

Noss RF (1999) Assessing and monitoring forest biodiversity: a suggested framework and indicators. *Forest Ecology and Management*, **115**, 135–146.

Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*, **24**, 581–583.

Ohlemüller R, Anderson BJ, Araújo MB *et al.* (2008). The coincidence of climatic and species rarity: high risk to small-range species from climate change. *Biology Letters*, **4**, 568–572.

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 637–669.

Parra-Olea G, Martínez-Meyer E, Pérez Ponce de León G (2005) Forecasting climate change effects on salamander distribution in the highlands of Central Mexico. *Biotropica*, **37**, 202–208.

Parra-Olea G, Windfield JC, Velo-Antón G, Zamudio KR (2012) Isolation in habitat refugia promotes rapid diversification in a montane tropical salamander. *Journal of Biogeography*, **39**, 353–370.

Pauls SU, Lumbsch HT, Haase P (2006) Phylogeography of the montane caddisfly *Drusus discolor*: evidence for multiple refugia and periglacial survival. *Molecular Ecology*, **15**, 2153–2169.

Pearse DE, Crandall KA (2004) Beyond F_{ST}: analysis of population genetic data for conservation. *Conservation Genetics*, **5**, 585–602.

Penny D, Hendy MD (1985) The use of tree comparison metrics. *Systematic Zoology*, **34**, 75–82.

Petit RJ, Aguinagalde I, deBeaulieu J (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Petranka JW (1998) *Salamanders of the Unites States and Canada*. Smithsonian Institution Press, Washington D. C.

Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution*, **56**, 1776–1788.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution*, **16**, 37–45.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, **23**, 564–571.

Purvis A, Gittleman JL, Cowlishaw G, Mace GM (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society B: Biological Sciences*, **267**, 1947–1952.

Qiu Y, Guan B, Fu C, Comes HP (2009) Did glacials and/or interglacials promote allopatric incipient speciation in East Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in Dysosma versipellis. *Molecular Phylogenetics and Evolution*, **51**, 281–293.

Qu Y, Luo X, Zhang R, Song G, Zou F, Lei F (2011) Lineage diversification and historical demography of a montane bird *Garrulax elliotii* – implications for the Pleistocene evolutionary history of the eastern Himalayas. *BMC Evolutionary Biology*, **11**, 174.

R Development Core Team (2009) The R Project for Statistical Computing. Available at: http://www.r-project.org/.

Rader RB, Belk MC, Shiozawa DK, Crandall KA (2005) Empirical tests for ecological exchangeability. *Animal Conservation*, **8**, 239–247.

Rambaut A, Drummond A (2007) TRACER v 1.5. Computer program available from http://beast.bio.ed.ac.uk/tracer.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Raymo M (1997) The timing of major climate terminations. *Paleoceanography*, **12**, 577–585.

Ribas CC, Moyle RG, Miyaki CY, Cracraft J (2007) The assembly of montane biotas: linking Andean tectonics and climatic oscillations to independent regimes of diversification in *Pionus* parrots. *Proceedings of the Royal Society B: Biological Sciences,* **274**, 2399–2408.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Rissler LJ, Smith WH (2010) Mapping amphibian contact zones and phylogeographical break hotspots across the United States. *Molecular Ecology*, **19**, 5404–5416.

Ronquist F, Teslenko M, Van der Mark P *et al.* (2012) MRBAYES 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

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

Rovito SM (2010) Lineage divergence and speciation in the Web-toed Salamanders (Plethodontidae: *Hydromantes*) of the Sierra Nevada, California. *Molecular Ecology*, **19**, 4554–4571.

Rowe KC, Heske EJ, Brown PW, Paige KN (2004) Surviving the ice: northern refugia and postglacial colonization. *Proceedings of the National Academy of Sciences, USA*, **101**, 10355–10359.

Rozen S, Skaletsky H (2000) PRIMER3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, **132**, 365–386.

Ruegg KC, Hijmans RJ, Moritz C (2006) Climate change and the origin of migratory pathways in the Swainson's thrush, *Catharus ustulatus*. *Journal of Biogeography*, **33**, 1172–1182.

Rull V (2009) Microrefugia. *Journal of Biogeography*, **36,** 481–484.

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, 1, 9–10.

Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, **4**, 11.

Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. *Frontiers in Zoology*, **6**, 9.

Schmitt T, Varga Z (2012) Extra-Mediterranean refugia: the rule and not the exception? *Frontiers in Zoology*, **9**, 22.

Schoener TW (1968) *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology*, **49**, 704–726.

Schoville SD, Roderick GK, Kavanaugh DH (2012) Testing the 'Pleistocene species pump' in alpine habitats: lineage diversification of flightless ground beetles (Coleoptera: Carabidae: *Nebria*) in relation to altitudinal zonation. *Biological Journal of the Linnean Society*, **107**, 95–111.

Schoville SD, Stuckey M, Roderick GK (2011) Pleistocene origin and population history of a neoendemic alpine butterfly. *Molecular Ecology*, **20**, 1233–1247.

Schoville SD, Uchifune T, Machida R (2013) Colliding fragment islands transport independent lineages of endemic rock-crawlers (Grylloblattodea: Grylloblattidae) in the Japanese archipelago. *Molecular Phylogenetics and Evolution*, **66**, 915–927.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, **18**, 233–234.

Schultheis AS, Booth JY, Perlmutter LR, Bond JE, Sheldon AL (2012) Phylogeography and species biogeography of montane Great Basin stoneflies. *Molecular Ecology*, **21**, 3325–3340.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, **9**, 615–629.

Shafer AA, Cullingham CI, Côté AD, Coltman DW (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**, 4589–4621.

Sharkey MJ (2001) The All Taxa Biological Inventory of the Great Smoky Mountains National Park. *Florida Entomologist*, **84**, 556–564.

Shepard DB, Burbrink FT (2008) Lineage diversification and historical demography of a sky island salamander, *Plethodon ouachitae*, from the Interior Highlands. *Molecular Ecology*, **17**, 5315–5335.

Shepard DB, Burbrink FT (2009) Phylogeographic and demographic effects of Pleistocene climate fluctuations in a montane salamander, *Plethodon fourchensis*. *Molecular Ecology*, **18**, 2243–2262.

Shepard, DB, Chong RA, Sun C, Mueller RL, Kozak KH. Nuclear markers for shallow-scale phylogenetics in taxa with giant genomes: resolving relationships within the plethodontid salamander genus *Desmognathus*. Submitted to: *Molecular Phylogenetics and Evolution.*

Sodhi NS, Bickford D, Diesmos AC *et al.* (2008) Measuring the meltdown: drivers of global amphibian extinction and decline. *PLoS ONE*, **3**, e1636.

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353–373.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261–4293.

Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30,** 1312–1313.

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–1169.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.

Stewart JR, Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology and Evolution*, **16**, 608–613.

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 661–671.

Streicher JW, Crawford AJ, Edwards CW (2009) Multilocus molecular phylogenetic analysis of the montane *Craugastor podiciferus* species complex (Anura: Craugastoridae) in Isthmian Central America. *Molecular Phylogenetics and Evolution*, **53**, 620–630.

Taberlet P, Cheddadi R (2002) Quaternary refugia and persistence of biodiversity. *Science*, **297**, 2009–2010.

Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *American Naturalist*, **166**, 581–591.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial recolonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Tamura K, Peterson D, Peterson N *et al.* (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

Templeton AR (1989) The meaning of species and speciation: a genetic perspective. In: *Speciation and Its Consequences* (eds Otte D & Endler JA), pp. 3–27. Sinauer Press, Sunderland.

Thomas CD, Cameron A, Green RE *et al.* (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.

Thomas SM, Hedin M (2008) Multigenic phylogeographic divergence in the paleoendemic southern Appalachian opilionid *Fumontana deprehendor* Shear (Opiliones, Laniatores, Triaenonychidae). *Molecular Phylogenetics and Evolution*, **46**, 645–658.

Thomé MTC, Zamudio KR, Giovanelli JGR *et al.* (2010) Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. *Molecular Phylogenetics and Evolution*, **55**, 1018–1031.

Tingley MW, Monahan WB, Beissinger SR, Moritz C (2009) Birds track their Grinnellian niche through a century of climate change. *Proceedings of the National Academy of Sciences*, **106**, 19637–19643.

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.

van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4, 535–538.

Velo-Antón G, Parra JL, Parra-Olea G, Zamudio KR (2013) Tracking climate change in a dispersal-limited species: reduced spatial and genetic connectivity in a montane salamander. *Molecular Ecology*, **22**, 3261–3278.

Wake DB (2009) What salamanders have taught us about evolution. *Annual Review of Ecology, Evolution and Systematics*, **40**, 33–352.

Walker D, Avise JC (1998) Principles of phylogeography as illustrated by freshwater and terrestrial turtles in the southeastern United States. *Annual Review of Ecology and Systematics*, **29**, 23–58.

Walker MJ, Stockman AK, Marek PE, Bond JE (2009) Pleistocene glacial refugia across the Appalachian Mountains and coastal plain in the millipede genus *Narceus*: evidence from population genetic, phylogeographic, and paleoclimatic data. *BMC Evolutionary Biology*, **9**, 25.

Waltari E, Hijmans RJ, Peterson AT *et al.* (2007) Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE*, **7**, e563.

Warren DL, Seifert SN (2011) Ecological niche modeling in MAXENT: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**, 335–342.

Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868–2883.

Warren DL, Glor RE, Turelli M (2010) ENMTOOLS: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607–611.

Webb T, Bartlein PJ (1992) Global changes during the last 3 million years: climatic controls and biotic response. *Annual Review of Ecology and Systematics*, **23**, 141–

173.

Weisrock DW, Larson A (2006) Testing hypotheses of speciation in the *Plethodon jordani* species complex with allozymes and mitochondrial DNA sequences. *Biological Journal of the Linnean Society*, **89**, 25–51.

Weisrock DW, Kozak KH, Larson A (2005) Phylogeographic analysis of mitochondrial gene flow and introgression in the salamander, *Plethodon shermani*. *Molecular Ecology*, **14**, 1457–1472.

Weisrock DW, Macey JR, Ugurtas IH, Larson A, Papenfuss TJ (2001) Molecular phylogenetics and historical biogeography among salamandrids of the 'true' salamander clade: rapid branching of numerous highly divergent lineages in *Mertensiella luschani* associated with the rise of Anatolia. *Molecular Phylogenetics and Evolution*, **18,** 434–448.

Welsh HH, Droege S (2001) A case for using plethodontid salamanders for monitoring biodiversity and ecosystem integrity of North American forests. *Conservation Biology*, **15**, 558–569.

Werneck FP, Gamble T, Colli GR, Rodrigues MT, Sites JW Jr (2012) Deep diversification and long-term persistence in the South American 'dry diagonal': integrating continent-wide phylogeography and distribution modeling of geckos. *Evolution*, **66**, 3014–3034.

Whittaker RH (1956) Vegetation of the Great Smoky Mountains. *Ecological Monographs*, **26**, 2–80.

Wiens JJ, Engstrom TN, Chippindale PT (2006) Rapid diversification, incomplete isolation, and the "speciation clock" in North American salamanders (genus *Plethodon*): testing the hybrid swarm hypothesis of rapid radiation. *Evolution*, **60**, 2585–2603.

Wilgenbusch JC, Warren DL, Swofford DL (2004) AWTY: A System for Graphical Exploration of MCMC Convergence in Bayesian Phylogenetic Inference. Available from http://ceb.csit.fsu.edu/awty

Williams SE, Bolitho EE, Fox S (2003) Climate change in Australian tropical rainforests: an impending environmental catastrophe. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 1887–1892.

Wu Y, Wang Y, Jiang K, Hanken J (2013) Significance of pre-Quaternary climate change for montane species diversity: insights from Asian salamanders (Salamandridae: *Pachytriton*). *Molecular Phylogenetics and Evolution*, **66**, 380–390.

Yang D, Kenagy GJ (2009) Nuclear and mitochondrial DNA reveal contrasting evolutionary processes in populations of deer mice (*Peromyscus maniculatus*).

Molecular Ecology, **18**, 5115–5125.

Zasada IA, Peetz A, Howe DK *et al.* (2014) Using mitogenomic and nuclear ribosomal sequence data to investigate the phylogeny of the *Xiphinema americanum* species complex. *PLoS One*, **9**, e90035.

Zhang D, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563–584.

Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, **17**, 2107–2121.

Zink RM, Jones AW, Farquhar CC, Westberg MC, Rojas JIG (2010) Comparison of molecular markers in the endangered black-capped vireo (*Vireo atricapilla*) and their interpretation in conservation. *Auk*, **127**, 797−806

Appendix S1 Chapter 1 Supplemental modeling and DNA amplification methods.

Modification of microclimate model for mechanistic distribution modeling

We adapted the microclimate model for GRSM (Fridley, 2009) to downscale synoptic temperature data from the Last Glacial Maximum (LGM) to near-ground level, and used these data to parameterize environmental variables in the bioenergetic model. We obtained LGM temperature data at 2.5 arc-minute resolution from the WorldClim database (http://www.worldclim.org; Hijmans *et al.*, 2005). These data are based on general circulation model (GCM) simulations from the Community Climate System Model (CCSM; Collins *et al.*, 2004) conducted by the Paleoclimate Modelling Intercomparison Project Phase II (PMIP2). The microclimate model provides estimates of daily minimum and maximum temperatures ca. 1 m above ground at sites across GRSM after accounting for topographic variables, including solar radiation, slope, and soil moisture (Fridley, 2009). To facilitate computational efficiency, we modified the microclimate model to estimate quarterly minimum and maximum temperatures. These quarterly estimates were chosen because they capture the range of variation in annual temperature conditions in GRSM and encompass the seasonal sensitivity of lapse rates (Fig. 4 in Fridley, 2009).

The microclimate model requires lapse rate and temperature parameters as input for each grid cell based on elevation. The original model derived these parameters by plotting the elevation of ten weather stations in and adjacent to GRSM (Table A1 in Fridley, 2009) against the minimum and maximum temperatures measured at those stations, then fitting regression lines to them. We adjusted the model to LGM conditions by extracting minimum and maximum temperature data from the CCSM model for those ten localities and regressing them against elevation (assumed to be constant over time; Linzey, 2008) to generate lapse rate (slope) and temperature (intercept) for each quarter. We used the fixed-effect coefficients determined to best fit the original model (Table 2 in Fridley, 2009). We also maintained the annual shortwave radiation, stream distance, and topographic convergence index (a measure of soil moisture based on slope; Beven & Kirkby, 1979) rasters used in the original model. We calculated new values for daily shortwave radiation for one day representing each quarter [Julian (ordinal) days 15, 105, 196, 288] from a digital elevation model using the r.sun function in GRASS GIS v 5.4.

Data preparation for correlative distribution modeling

We reduced the full data set of 19 bioclimatic variables to variables that we consider relevant to *P. jordani* (Nogues-Bravo, 2009; Elith *et al.*, 2010, 2011). First, we first generated a correlation matrix for the 19 variables using the 'Band Collection Statistics' tool available in the Spatial Analyst extension of ARCGIS v 10.1 (ESRI). Then, we removed highly correlated variables ($r > 0.85$; Elith *et al.*, 2006) and retained variables that describe long-term seasonal patterns and moisture availability based on the observations that terrestrial plethodontids spend much of their time in moist microhabitats and evaporative water loss is a major constraint on their activity (Spight, 1968; Feder, 1983). These variables included: annual mean diurnal range, annual

temperature range, mean temperature of wettest quarter, mean temperature of driest quarter, mean temperature of warmest quarter, precipitation of wettest month, precipitation seasonality, precipitation of driest quarter, and precipitation of warmest quarter.

We used occurrence records from our genetic sampling (Table 1 in the text), from a previous study (Weisrock & Larson, 2006), and from the collection of R. Highton (deposited in the Smithsonian National Museum of Natural History) for the model (Appendix S1). To reduce sampling bias (Phillips *et al.*, 2009; Kramer-Schadt *et al.*, 2013; Boria *et al.*, 2014), we filtered our occurrence records and retained only one locality per grid cell, for a total of 319 unique localities. We ran the model using these records and the current climate variables, then projected it onto a climate scenario from the LGM. The LGM climate data for the nine selected bioclimatic variables were based on GCM simulations from the CCSM models (described above), and obtained at 2.5 arc-minute resolution from the WorldClim database. We downscaled the LGM layers to the resolution of the current data layers using the 'delta method' (Ramirez-Villegas & Jarvis, 2010; Bonatelli *et al.*, 2014). We clipped the LGM layers to an area extending from 29° to 37° latitude and -86° to -80° longitude that encompasses the southern Appalachians surrounding GRSM south to northern Florida. We calibrated the model on the extent of the current climate layers (described in the text), then projected it on this larger area. This approach has been shown to result in more realistic predictions of distributions and to reduce over-fitting that can confound model transferability onto different time periods (Anderson & Raza, 2010).

Protocols for DNA amplification

We used the METf.6 forward primer (5'-AAGCTTTCGGGCCCATACC-3'; Macey *et al.*, 1997) and ALAr.9 (5'-AAAGTGTTTGAGTTGCATTCA-3'; Kozak *et al.*, 2006) or ND2.R1 (5'-GTTGCATTCATGAGATGTAGGA-3'; this study) reverse primers in the polymerase chain reactions (PCR). We performed PCR in 10 μL reaction volumes consisting of 1 μL DNA template, 2X GoTaq® Green Master Mix (Promega), 0.5 μM forward and reverse primers, and ddH_2O to attain the final reaction volume. Thermocycler conditions for amplification were as follows: 3 min initial denaturation at 96°C; 30 cycles of denaturation at 96°C for 30 s, annealing at 50°C for 30 s, and extension at 70°C for 90 s; final extension at 72°C for 4 min. We purified PCR products with ExoSAP-IT® (Affymetrix). Products were sequenced at the University of Minnesota BioMedical Genomics Center on an ABI 3730xl DNA Analyzer that uses BigDye Terminator v 3.1 chemistry.

REFERENCES

Anderson, R.P. & Raza, A. (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*, **37**, 1378–1393.

Beven, K.J. & Kirkby, M.J. (1979) A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, **24**, 43–69.

Bonatelli, I.A.S., Perez, M.F., Peterson, A.T., Taylor, N.P., Zappi, D.C., Machado, M.C., Koch, I., Pires, A.H. & Moraes, E.M. (2014). Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology*, **23**, 3044–3063.

Boria, R.A., Olsen, L.E., Goodman, S.M. & Anderson, R.P. (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, **275**, 73–77.

Collins, W.D., Bitz, C.M., Blackmon, M.L., Bonan, G.B., Bretherton, C.S., Carton, J.A., Chang, P., Doney, S.C., Hack, J.J., Henderson, T.B., Kiehl, J.T., Large, W.G., McKenna, D.S., Santer, B.D. & Smith, R.D. (2004) The community climate system model version 3 (CCSM3). *Journal of Climate*, **19**, 2122–2143.

Elith, J., Graham, C.H., Anderson, R.P., *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modeling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Feder, M. (1983) Integrating the ecology and physiology of plethodontid salamanders. *Herpetologica*, **39**, 291–310.

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

Kozak, K.H., Weisrock, D.W. & Larson, A. (2006) Rapid lineage accumulation in a non-adaptive radiation: phylogenetic analysis of diversification rates in eastern North American woodland salamanders (Plethodontidae: *Plethodon*). *Proceedings of the Royal Society B: Biological Sciences*, **273**, 539–546.

Kramer-Schadt, S., Niedballa, J. & Pilgrim, J.D. (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*, **19**, 1366–1379.

Macey, J.R., Larson, A., Ananjeva, N.B., Fang, Z. & Papenfuss, T.J. (1997) Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. *Molecular Biology and Evolution*, **14**, 91–104.

Nogués-Bravo, D. (2009) Predicting the past distribution of species climatic niche. *Global Ecology and Biogeography*, **18**, 521–531.

Phillips, S.J., Dudík, M. & Elith, J. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, **19**, 181–197.

Ramirez-Villegas, J. & Jarvis, A. (2010) Downscaling global circulation model outputs: the delta method. Decision and Policy Analysis Working Paper No. 1. Cali, Columbia: Centro Internacional de Agricultura Tropical (CIAT).

Spight, T.M. (1968) The water economy of salamanders: evaporative water loss. *Physiological Zoology*, **41**, 195–203.

Weisrock, D.W. & Larson, A. (2006) Testing hypotheses of speciation in the *Plethodon jordani* species complex with allozymes and mitochondrial DNA sequences. *Biological Journal of the Linnean Society*, **89**, 25–51.

Appendix S3 Chapter 1 Geographic information and measures of diversity for *P. jordani*.

Table S1 Geographic information and measures of mtDNA diversity for each sampled population of *P. jordani*. Numbers in parentheses after the locality number correspond to localities used by Weisrock and Larson (2006). * indicates that the number of individuals includes sequences used in Weisrock & Larson's study. Haplotypes from introgressed individuals were excluded from calculations of haplotype and nucleotide diversity

Locality Number	Latitude	Longitude	Elevation (m)	N	H	PH	h	π
1 (94)	35.5203	-83.8685	1474	22*	3	1	0.338	0.0003
2	35.5201	-83.8532	1405	2	2	0	1	0.0009
3	35.5326	-83.8458	1166	11	2	0	0.600	0.0005
4	35.5342	-83.8090	1269	5	1	0	0	0
5	35.5372	-83.7991	1296	5	1	0	0	0
6	35.5470	-83.7905	1438	4	1	1	0	0
7	35.5794	-83.7421	1004	4	2	1	0.833	0.0552
8	35.5454	-83.7366	1008	10	1	0	0	0
9	35.5317	-83.7213	1348	6	1	0	0	0
10	35.5618	-83.7184	1495	4	1	0	0	0
11	35.5474	-83.7159	1510	2	1	0	0	0
12	35.5985	-83.6332	908	7	1	0	0	0
13	35.6120	-83.6078	1377	5	2	1	0.600	0.0005
14	35.6310	-83.5899	927	1	1	1	-	-
15	35.5575	-83.5669	1654	5	2	1	0.400	0.0004
16	35.5651	-83.5575	1639	3	2	0	0.667	0.0006
17	35.5651	-83.5432	1702	2	1	0	0	0
18	35.5726	-83.5336	1585	5	3	0	0.700	0.0023
19 (93)	35.5628	-83.4997	2019	4*	2	0	0.500	0.0009
20	35.5412	-83.4940	1789	5	1	0	0	0
21	35.5515	-83.4918	1781	2	2	1	1	0.0009
22	35.6588	-83.4812	1324	6	2	1	0.333	0.0003
23	35.5674	-83.4808	1806	9	4	1	0.694	0.0018
24 (92)	35.6618	-83.4703	1213	5*	1	0	0	0
25	35.5899	-83.4700	1797	7	3	1	0.714	0.0023
26	35.6797	-83.4611	998	4	1	0	0	0
27 (91)	35.6092	-83.4465	1703	15*	2	0	0.133	0.0010
28	35.6559	-83.4399	1927	1	1	0	-	-
29	35.6001	-83.4237	1443	2	1	0	0	0
30	35.6750	-83.4220	1152	4	2	1	0.500	0.0040
31	35.5800	-83.4220	1097	2	2	1	1	0.0009
32	35.6205	-83.4034	1773	5	4	2	0.900	0.0034
33 (88)	35.6744	-83.3981	816	9*	4	4	0.778	0.0042
34 (90)	35.5834	-83.3980	1405	12*	1	0	0	0
35	35.5189	-83.3781	1278	7	2	1	0.286	0.0003
36	35.6134	-83.3696	1171	12	6	2	0.879	0.0024
37	35.5886	-83.3615	892	3	2	0	0.667	0.0006
38	35.6482	-83.3585	1676	1	1	0	-	-

39	35.5662	-83.3419	778	5	3	1	0.700	0.0036
40	35.6122	-83.3409	1100	5	3	0	0.700	0.0056
41	35.5533	-83.3381	1203	2	1	0	0	0
42	35.6063	-83.3322	889	3	1	0	0	0
43	35.7091	-83.3198	992	7	3	1	0.667	0.0011
44	35.6583	-83.3172	1629	1	1	0	-	-
45 (89)	35.6121	-83.3158	1186	15*	1	0	0.133	0.0136
46	35.5527	-83.3145	726	2	2	1	1	0.0027
47	35.6189	-83.3097	1239	3	1	0	0	0
48	35.5885	-83.3083	868	6	2	0	0.600	0.0038
49	35.7396	-83.2779	938	5	2	1	0.400	0.0004
50	35.6866	-83.2691	1898	2	2	0	1	0.0018
51	35.6101	-83.2541	1160	2	1	0	0	0
52	35.7217	-83.2458	1863	2	2	0	1	0.0009
53	35.7256	-83.2415	1808	2	1	0	0	0
54 (87)	35.6381	-83.2378	1544	5*	4	3	0.900	0.0041
55	35.6924	-83.2221	1803	2	2	2	1	0.0018
56	35.7482	-83.2146	816	16	3	0	0.691	0.0008
57	35.6535	-83.1976	1672	5	2	0	0.400	0.0083
58	35.6951	-83.1919	1283	2	1	0	0	0
59	35.6778	-83.1918	1691	2	2	0	1	0.0207
60	35.6386	-83.1872	1455	5	1	0	0	0
61	35.6665	-83.1872	1680	1	1	1	-	-
62	35.7277	-83.1853	1525	5	1	1	0	0
63	35.6114	-83.1850	1440	17	5	3	0.728	0.0097
64	35.7376	-83.1774	1255	1	1	0	-	-
65 (85)	35.57333	-83.1744	1609	1*	1	1	-	-
66	35.7306	-83.1655	1001	1	1	0	-	-
67	35.7576	-83.1654	1518	3	1	0	0	0
68	35.6797	-83.1588	1702	7	2	0	0.476	0.0103
69	35.6864	-83.1418	1581	2	1	0	0	0
70	35.6981	-83.1255	1739	7	2	0	0.571	0.0123

N: number of individuals sequenced; H: number of haplotypes; PH: number of private haplotypes; h: haplotype diversity; π: nucleotide diversity

Appendix S5 Chapter 1 Full extent of predicted distributions of *P. jordani* since the Last Glacial Maximum (LGM), and clamping, MESS, and most dissimilar variable (MoD) maps generated from MAXENT output.

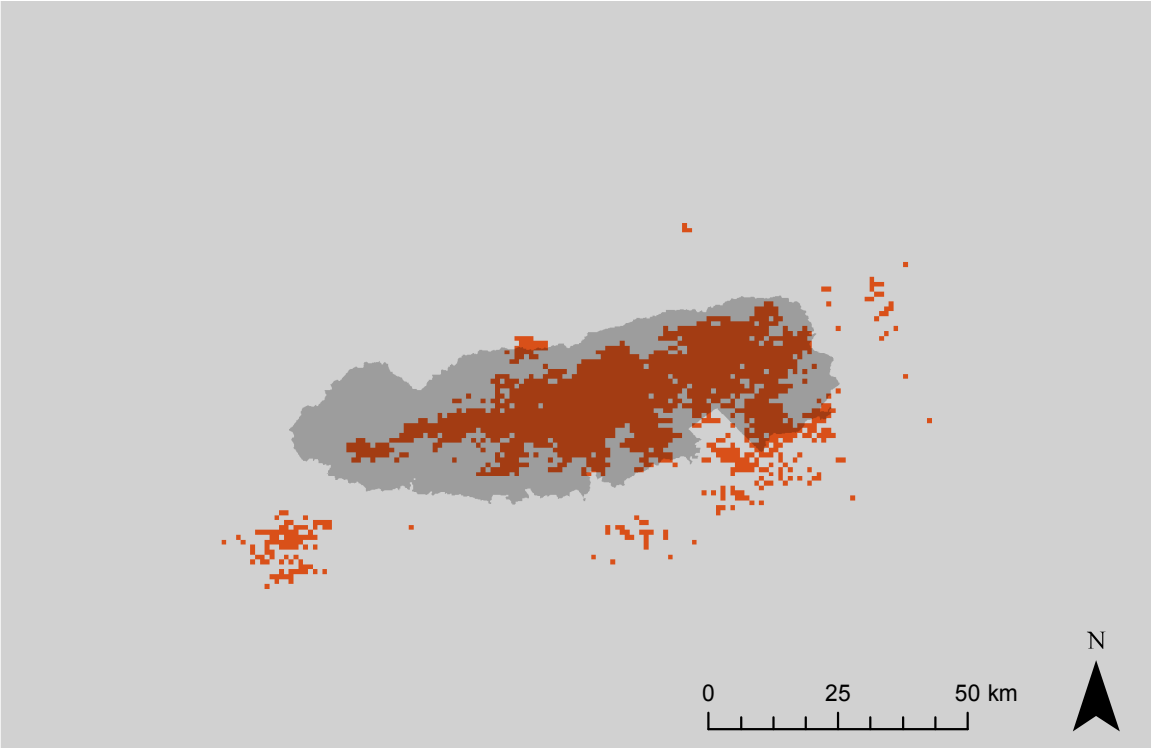


Figure S1 Full extent of predicted current distribution of *P. jordani* from the MAXENT model. Predicted suitable areas are shown in vermillion; the area within Great Smoky Mountains National Park (GRSM; depicted in Fig. 2b Chapter 1) is shown in dark gray. Climatically suitable areas outside of the current distribution in GRSM are presumably not occupied because they are separated by unsuitable climate at low elevations or occupied by a closely-related species (Kozak *et al.*, 2008)

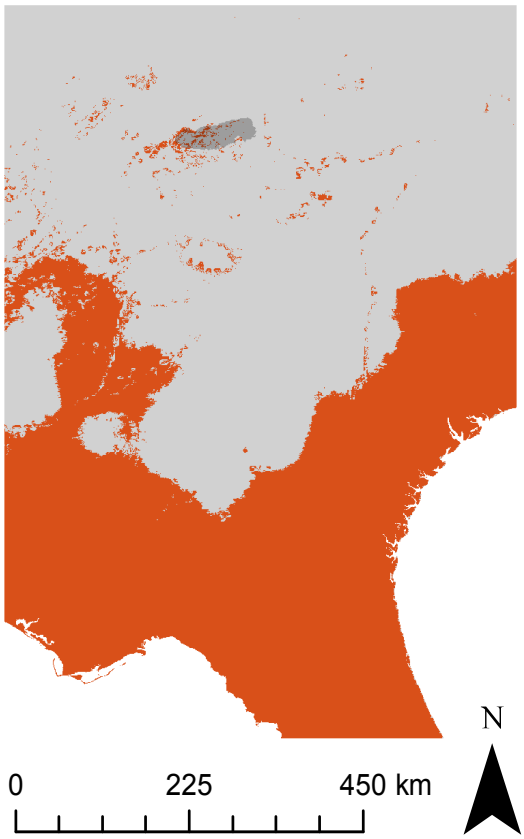


Figure S2 Full extent of predicted distribution of *P. jordani* at the LGM from the MAXENT model (see Appendix S1 for details on LGM data). Predicted suitable areas (shown in vermillion) were based on the 10th percentile training presence logistic threshold. The area within Great Smoky Mountains National Park (depicted in Fig. 2d of Chapter 1) is shown in dark gray

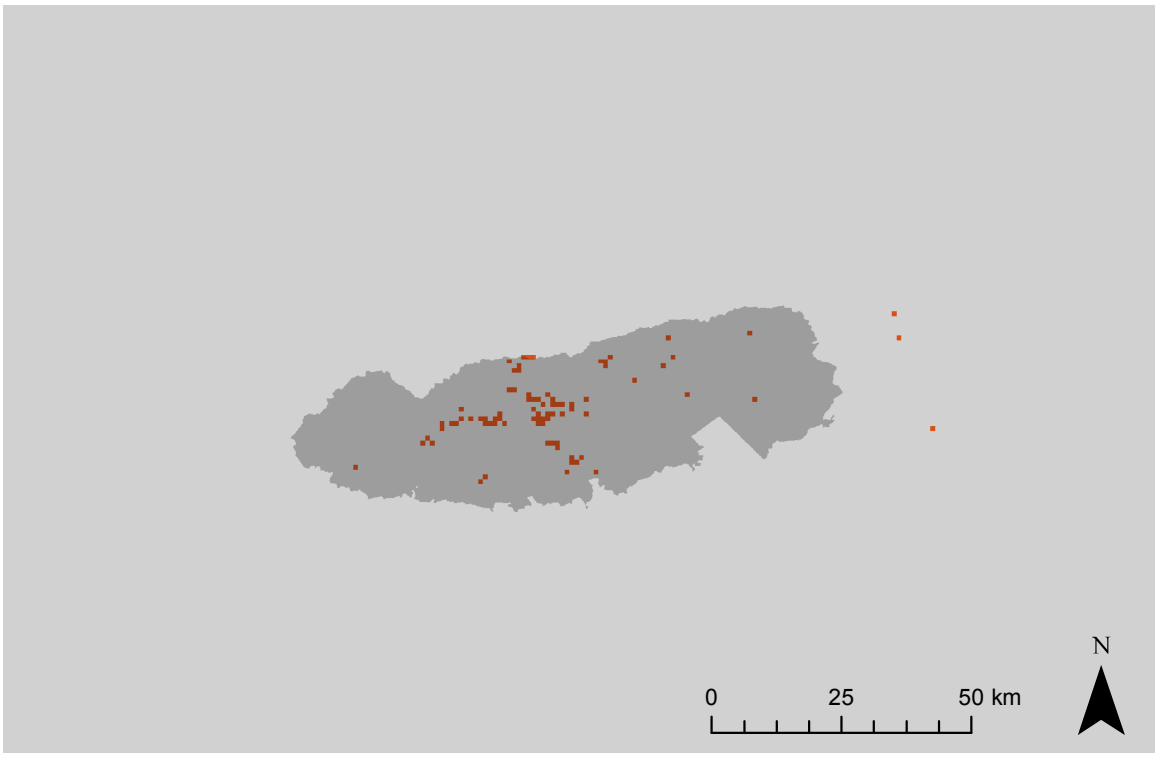


Figure S3 Full extent of predicted stability distribution of *P. jordani* from the LGM to present from the MAXENT model. Predicted suitable areas are shown in vermillion; the area within Great Smoky Mountains National Park (depicted in Fig. 2f of Chapter 1) is shown in dark gray

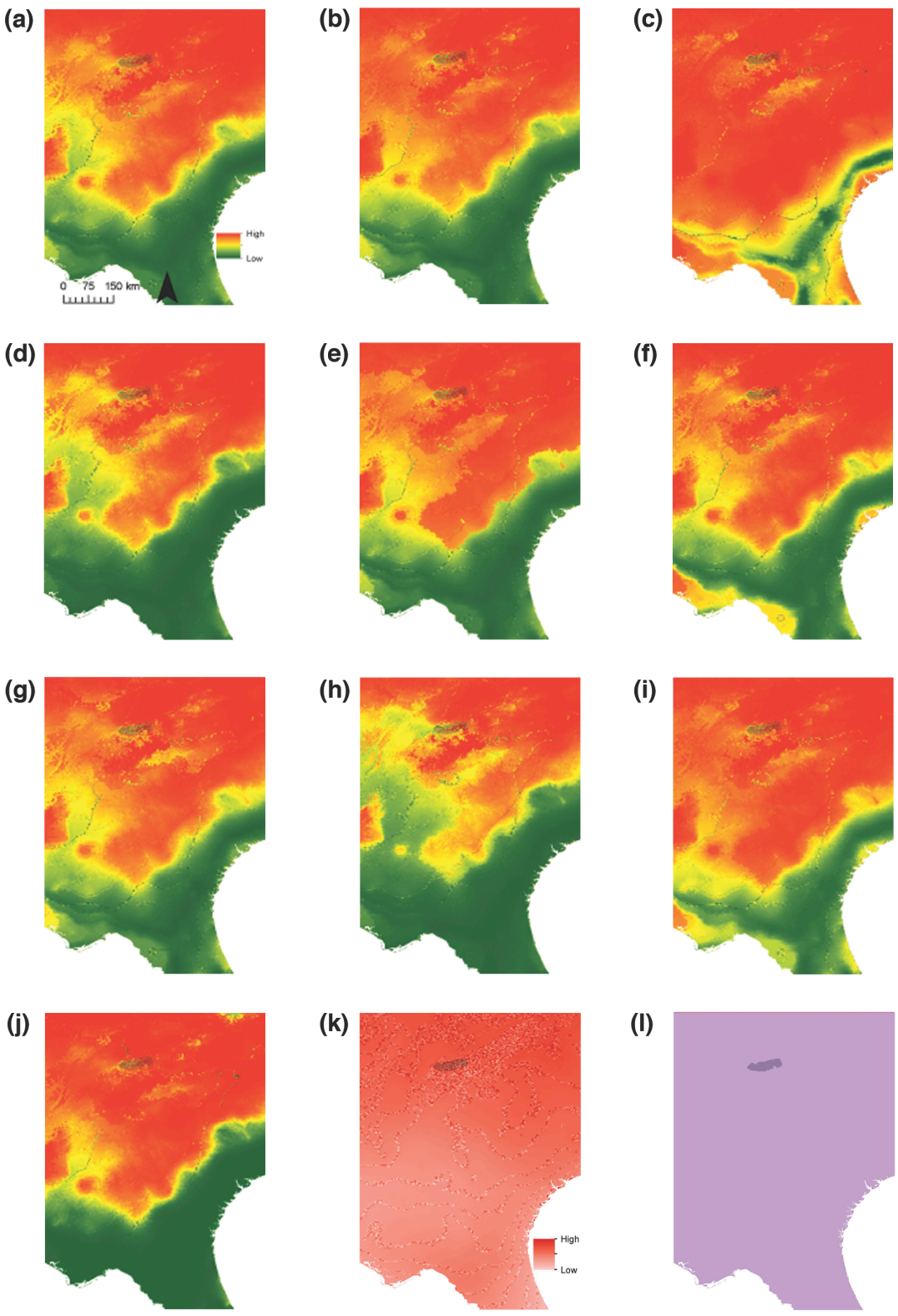

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
High
Low
0 75 150 km
High
Low

Figure S4 Clamping, multivariate environmental similarity surfaces (MESS), and most dissimilar variables (MoD) from MAXENT model. **(a)** through **(i)**, The difference between clamped and unclamped predictions for each of the ten replicate runs. Legend is shown in the bottom right of panel of (a); clamping values range from 0 (all variables within their training range; green) to 1 (all variables outside their training range; red). All runs show high clamping in Great Smoky Mountains National Park. **(k)**, MESS maps indicating areas where at least one variable has a value outside of its training range. All MESS values are negative across the projected region, meaning that they show dissimilarity in environmental space compared to the current model, hence transferability of the model across time should be treated with caution. Results were identical for all ten replicates; therefore, only one representative map is shown. **(l)**, MoD maps indicating the variable furthest outside of its training range when projected on the LGM (i.e., the variable that has shifted the most across time points): temperature annual range. Results were identical for all ten replicates; therefore, only one representative map is shown.

REFERENCE

Kozak, K.H., Graham, C.H. & Wiens, J.J. (2008) Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology and Evolution*, **23**, 141–148.

Appendix S6 Chapter 2 Supplemental Tables S1–S2.

Table S1 Niche overlap values based on three statistics calculated for each pair of lineages using ENMTOOLS. Values in bold are statistically significant based on comparisons between the computed values and null distributions generated from 100 pseudoreplicates (Figures S7-S9)

Comparison	Niche overlap		
	Schoener's *D*	Warren's *I*	Relative Rank
Clade 1 - Clade 2	**0.643**	**0.881**	**0.768**
Clade 1 - Clade 3	0.545	0.801	0.678
Clade 2 - Clade 3	0.623	0.877	0.840

Table S2 Results from background tests based on three statistics calculated for each pair of lineages using ENMTOOLS. Asterisks indicate comparisons that are more similar than expected by chance

Focal taxon	Background	Overlap statistic		
		Schoener's *D*	Warren's *I*	Relative Rank
Clade 1	Clade 2	0.643*	0.881*	0.787
Clade 2	Clade 1	0.643	0.881	0.865
Clade 1	Clade 3	0.545	0.801	0.711
Clade 3	Clade 1	0.545	0.801	0.710
Clade 2	Clade 3	0.623	0.877	0.826
Clade 3	Clade 2	0.623*	0.877*	0.787

Appendix S7 Chapter 2 Supplemental Figures S1–S9.

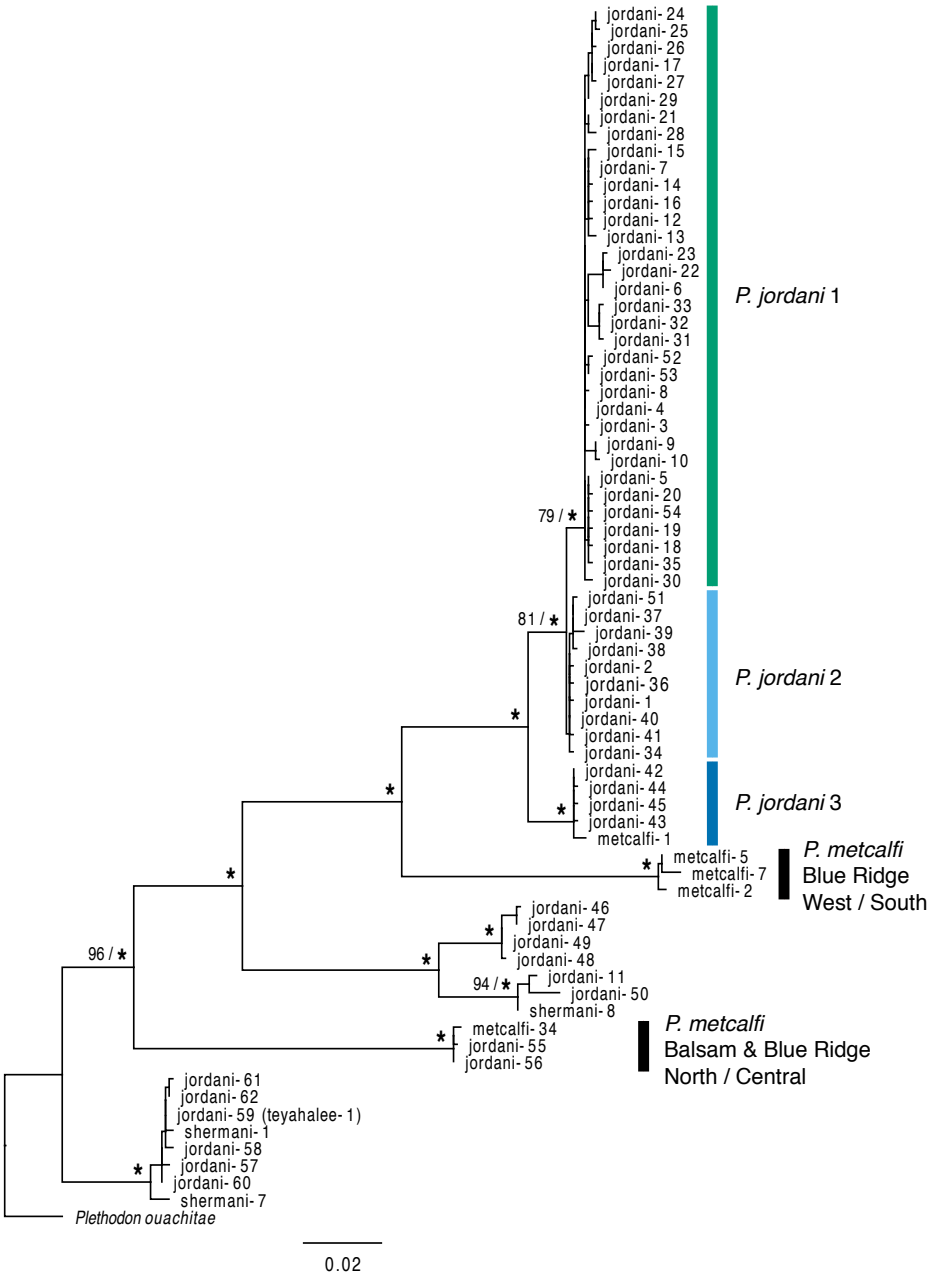


Figure S1 *ND2* phylogeny based on haplotypes for *P. jordani* inferred from Maximum-likelihood and Bayesian inference, with *P. ouachitae* used as an outgroup (reproduced from Luxbacher *et al.*, Chapter 1). Numbers above branches indicate bootstrap support and posterior probabilities, respectively; strongly supported nodes (bootstrap support ≥ 99 and posterior probabilities ≥ 0.99) are denoted by asterisks. *Plethodon jordani* haplotypes 1-11, and *P. metcalfi* and *P. shermani* haplotypes are the same as in Weisrock and Larson (2006); jordani-59 is identical to Weisrock and Larson's teyahalee-1 haplotype

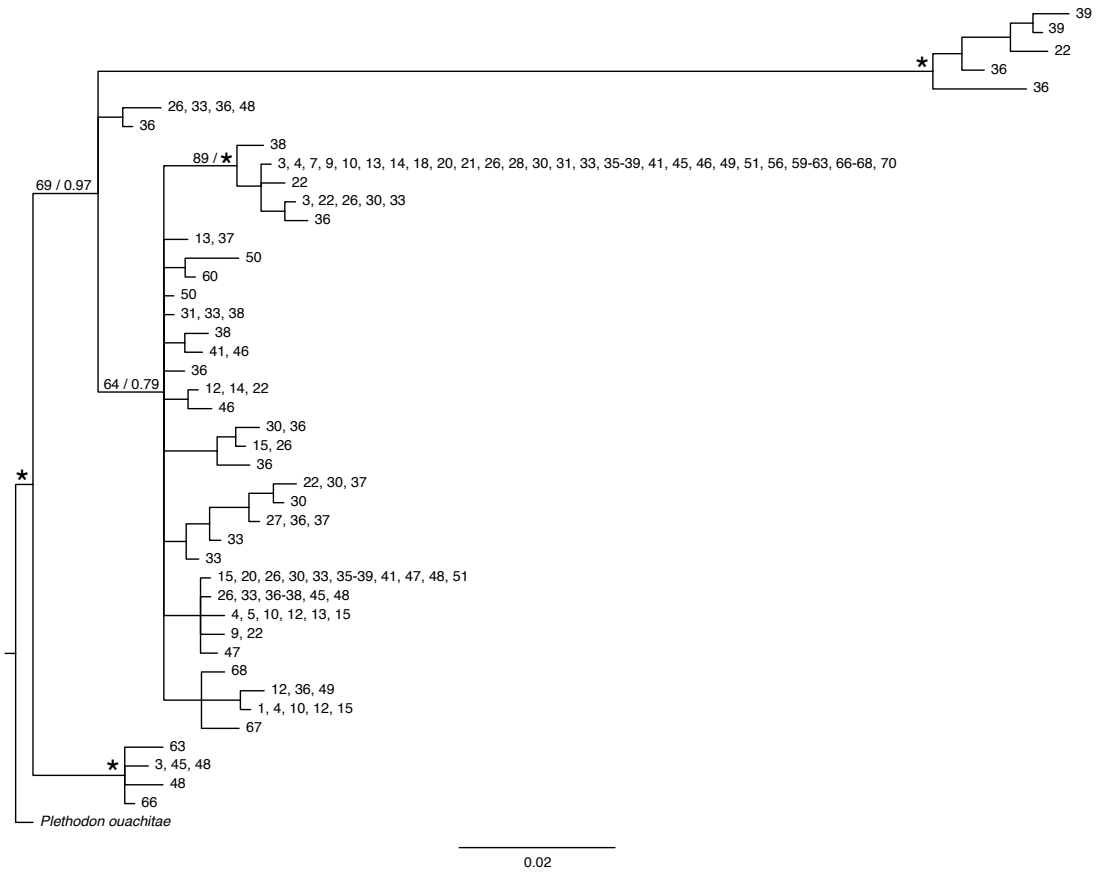


Figure S2 Nuclear genealogy for *P. jordani* inferred from Maximum-likelihood and Bayesian inference based on locus 9 alleles, with *P. ouachitae* used as an outgroup. Numbers above branches indicate bootstrap support and posterior probabilities, respectively; strongly supported nodes (bootstrap support ≥ 99 and posterior probabilities ≥ 0.99) are denoted by asterisks. Each tip is labeled with locality numbers (Chapter 2, Figure 2) for individuals possessing that allele

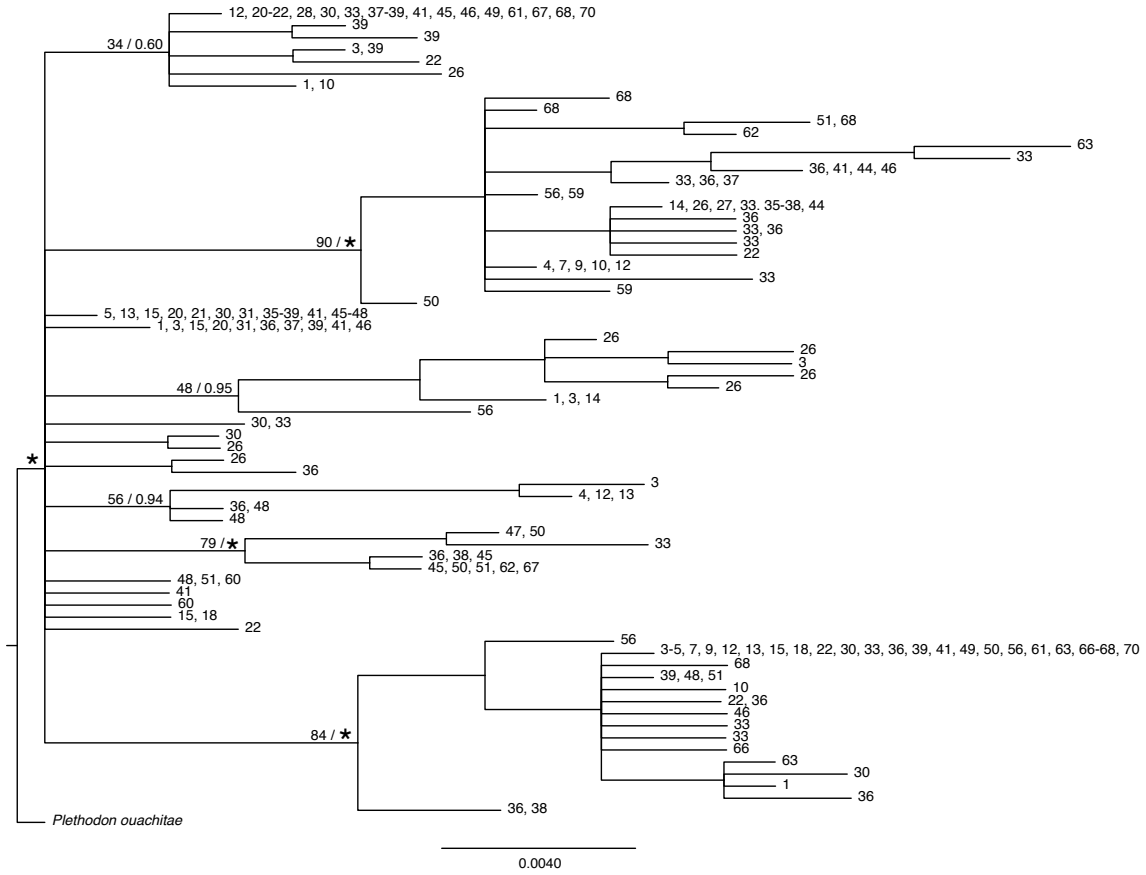


Figure S3 Nuclear genealogy for *P. jordani* inferred from Maximum-likelihood and Bayesian inference based on locus 16 alleles, with *P. ouachitae* used as an outgroup. Numbers above branches indicate bootstrap support and posterior probabilities, respectively; strongly supported nodes (bootstrap support ≥ 99 and posterior probabilities ≥ 0.99) are denoted by asterisks. Each tip is labeled with locality numbers (Chapter 2, Figure 2) for individuals possessing that allele

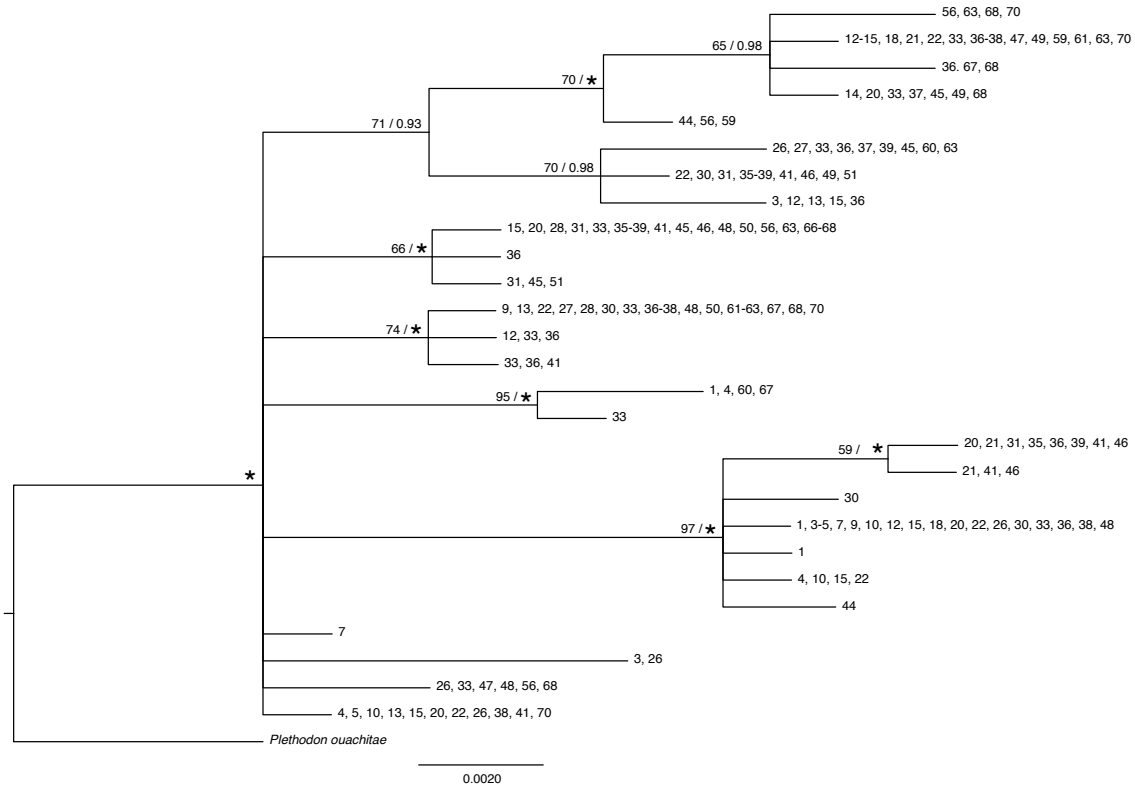


Figure S4 Nuclear genealogy for *P. jordani* inferred from Maximum-likelihood and Bayesian inference based on locus 41 alleles, with *P. ouachitae* used as an outgroup. Numbers above branches indicate bootstrap support and posterior probabilities, respectively; strongly supported nodes (bootstrap support ≥ 99 and posterior probabilities ≥ 0.99) are denoted by asterisks. Each tip is labeled with locality numbers (Chapter 2, Figure 2) for individuals possessing that allele

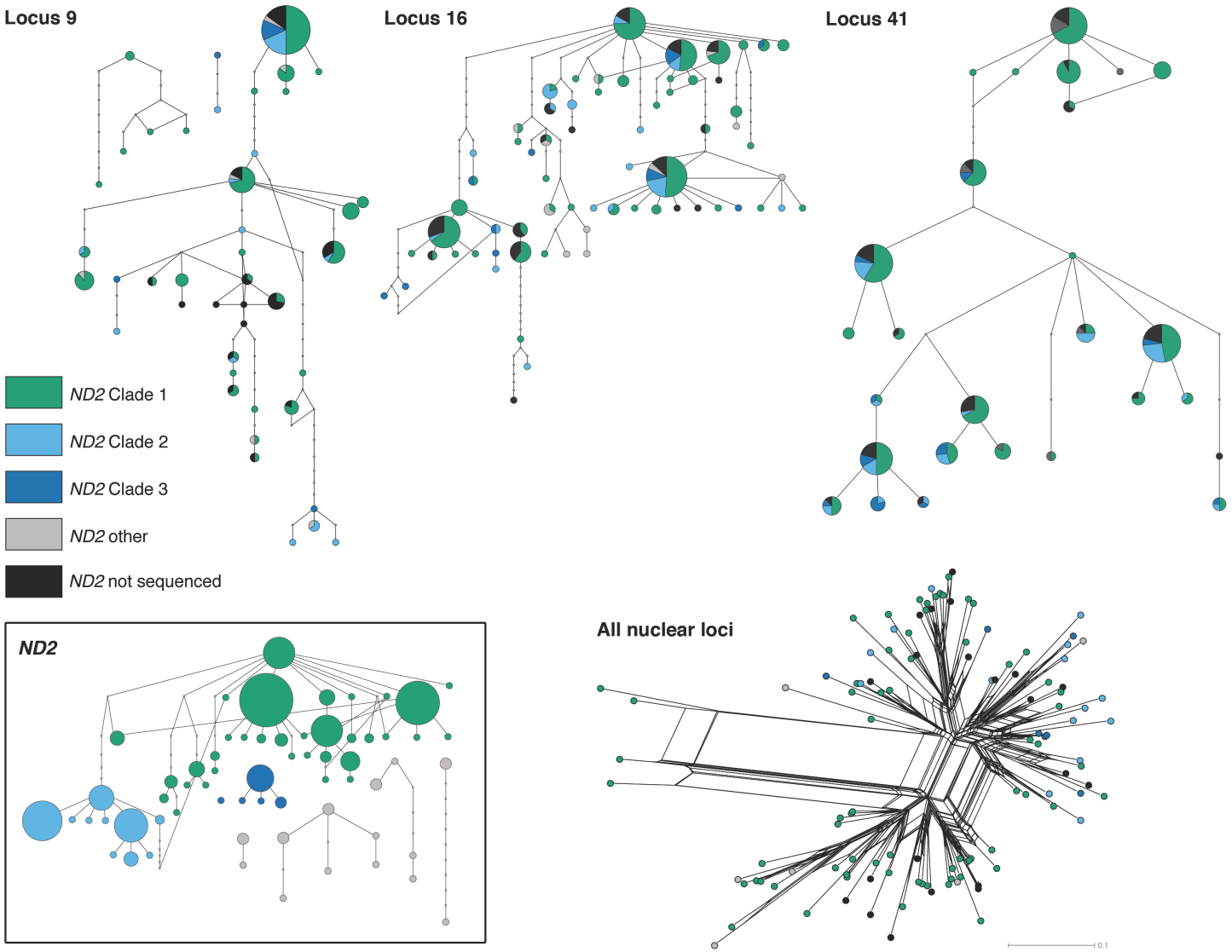

Locus 9
Locus 16
Locus 41
ND2 Clade 1
ND2 Clade 2
ND2 Clade 3
ND2 other
ND2 not sequenced
ND2
All nuclear loci
0.1

Figure S5 Individual haplotype networks for each nuclear locus used in this study and multilocus network constructed from all three nuclear loci. Haplotypes are colored according to mitochondrial clade (Figure S1). (A) through (C), Individual networks were estimated using statistical parsimony with a 95% connection significance. The size of each circle is proportional its haplotype frequency and open circles represent inferred unsampled haplotypes. Haplotype circles are depicted as pie diagrams with slices colored according to geographic region. (D), Multilocus network based on Tamura-Nei distances from all three nuclear loci. The four individuals offset to the left are highly divergent at locus 9. Inset, Haplotype network for *ND2*

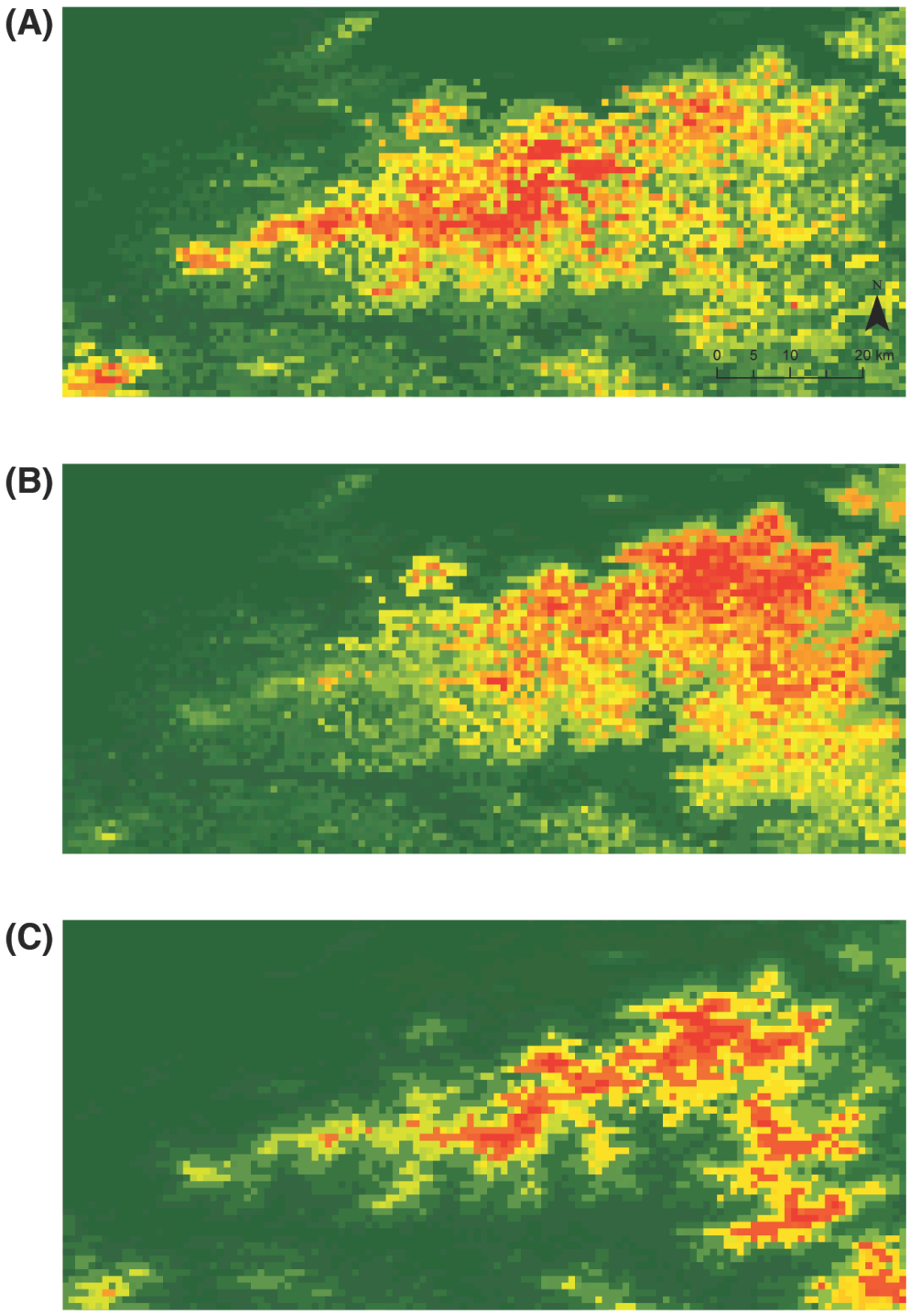


Figure S6 Ecological niche models for each mtDNA clade generated in MAXENT. (A), Clade 1. (B), Clade 2. (C), Clade 3

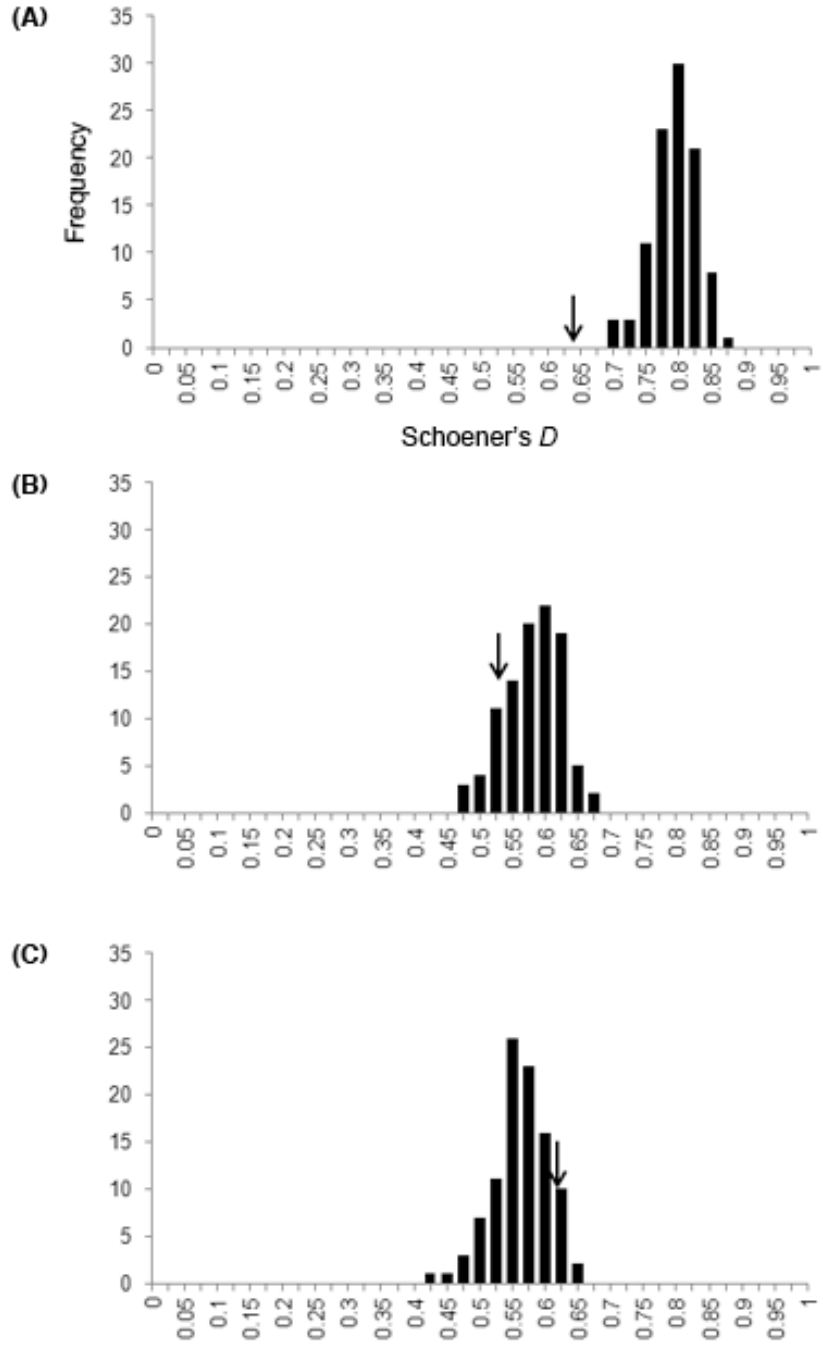


Figure S7 Niche identity test based on Schoener's *D* (Schoener, 1968) statistic for comparisons between each mtDNA clade: (A) Clade 1 – Clade 2, (B) Clade 1 – Clade 3, (C) Clade 2 – Clade 3. Axes for each plot are identical to (A). In each plot, observed identity values (Table S1) are indicated with an arrow. Suitability scores generated by ecological niche models for each lineage exhibit statistically significant ecological differences if the observed overlap value falls outside of the distribution of values obtained from 100 pseudoreplicates

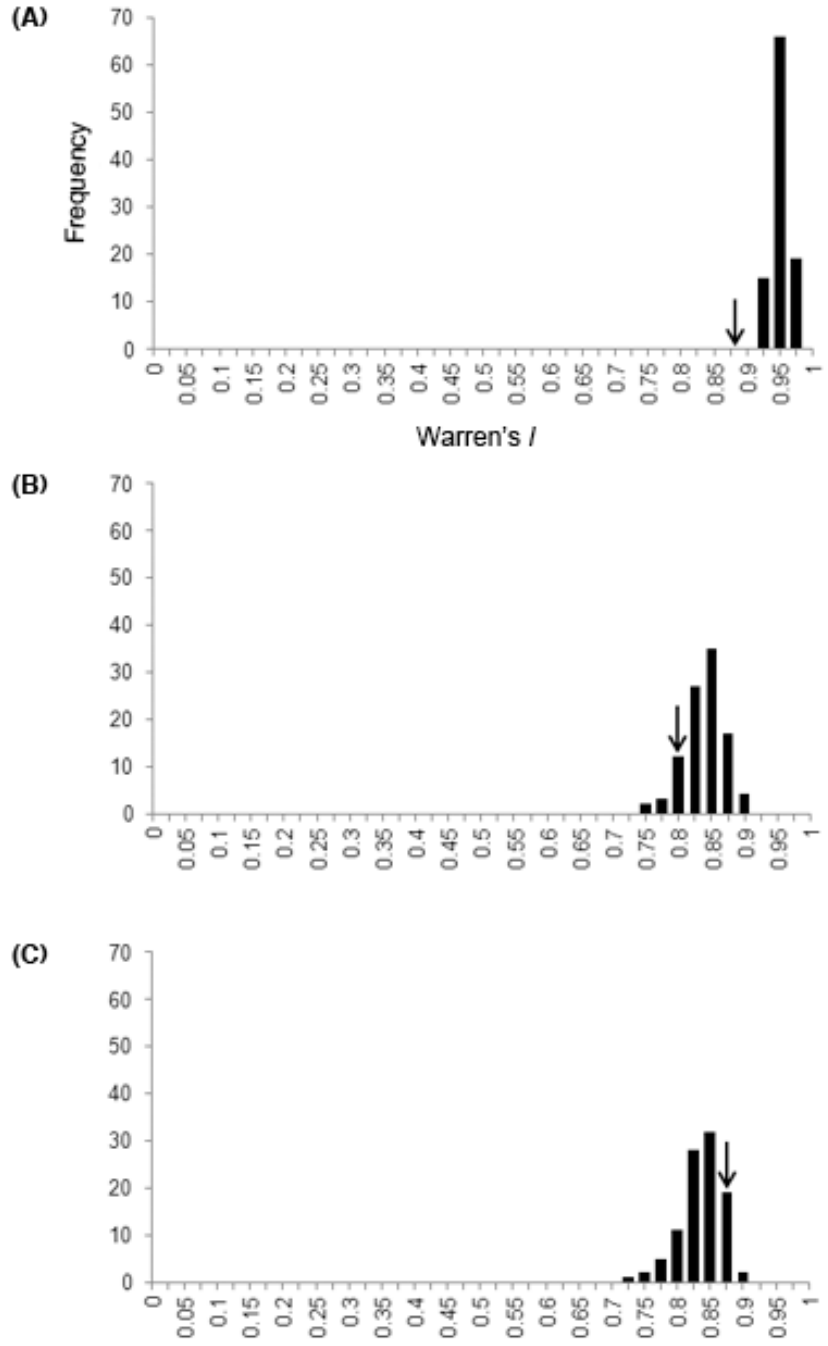


Figure S8 Niche identity test based on Warren's *I* (Warren *et al.*, 2008) statistic for comparisons between each mtDNA clade: (A) Clade 1 – Clade 2, (B) Clade 1 – Clade 3, (C) Clade 2 – Clade 3. Axes for each plot are identical to (A). In each plot, observed identity values (Table S1) are indicated with an arrow. Suitability scores generated by ecological niche models for each lineage exhibit statistically significant ecological differences if the observed overlap value falls outside of the distribution of values obtained from 100 pseudoreplicates

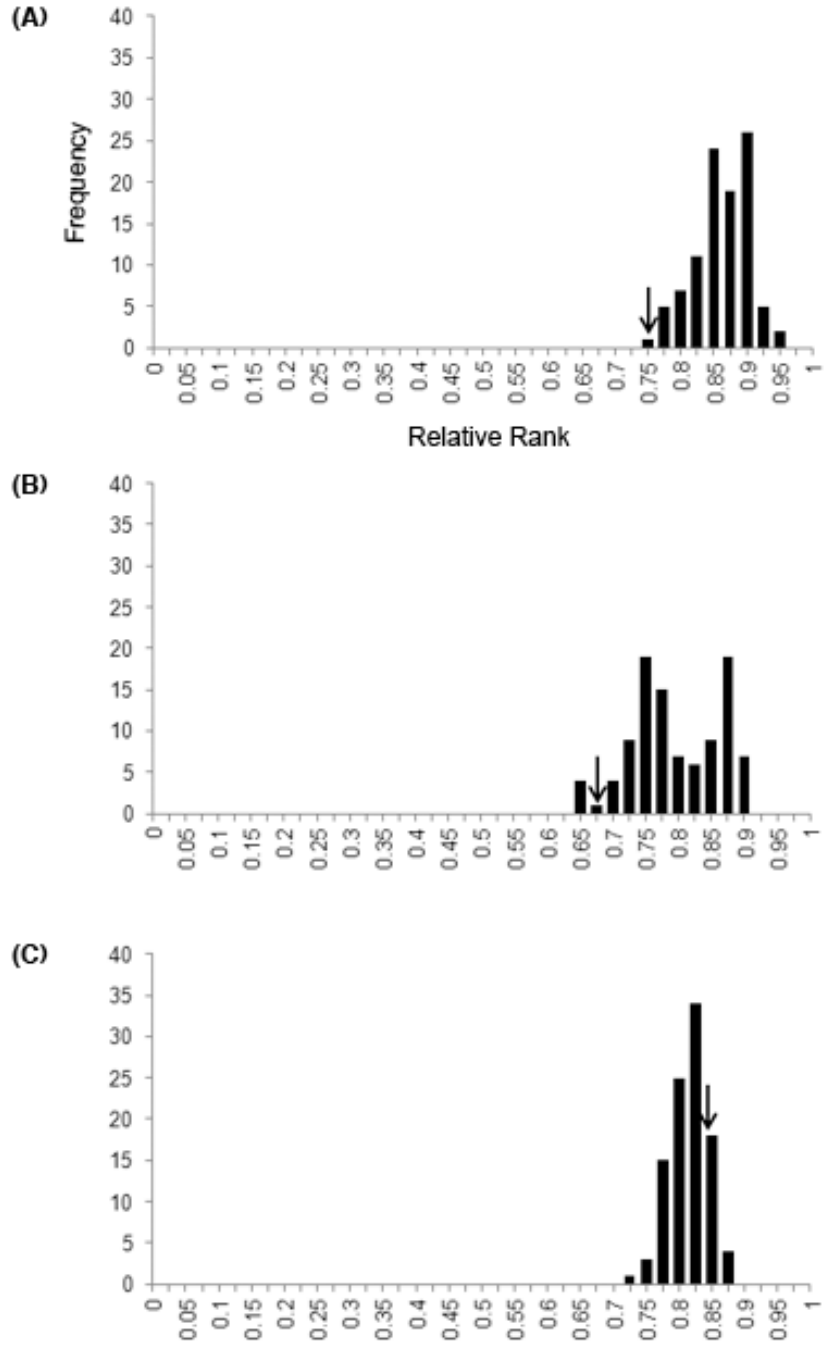


Figure S9 Niche identity test based on relative rank (Warren & Seifert, 2011) statistic for comparisons between each mtDNA clade: (A) Clade 1 – Clade 2, (B) Clade 1 – Clade 3, (C) Clade 2 – Clade 3. Axes for each plot are identical to (A). In each plot, observed identity values (Table S1) are indicated with an arrow. Suitability scores generated by ecological niche models for each lineage exhibit statistically significant ecological differences if the observed overlap value falls outside of the distribution of values obtained from 100 pseudoreplicates

Appendix S8 Chapter 3 Supplemental Tables S1–S3, S5.

Table S1 Geographic information and measures of genetic diversity for *P. jordani* populations based on nine microsatellite loci. Numbers in parentheses after the locality number correspond to localities used in Weisrock and Larson's (2006) study. Diversity measures from mtDNA for these and additional localities are reported in Luxbacher *et al.* (Chapter 1; Table S1)

Locality Number	Latitude	Longitude	Elevation (m)	Microsatellite				
				N	A	PA	Ho	He
1 (WL94)	35.5203	-83.8685	1474	12	7.27	0.08	0.683	0.735
3	35.5326	-83.8458	1166	9	7.67	0	0.713	0.768
6	35.5470	-83.7905	1438	1	9.44	0	1	1
8	35.5454	-83.7366	1008	10	7.63	0.1	0.756	0.864
9	35.5317	-83.7213	1348	6	5	0	0.688	0.737
10	35.5618	-83.7184	1495	4	5	0	0.719	0.875
11	35.5474	-83.7159	1510	3	3.56	0.11	0.625	0.800
12	35.5985	-83.6332	908	8	7.89	0.14	0.708	0.808
13	35.6120	-83.6078	1377	5	8.78	0.11	0.7	0.767
14	35.6310	-83.5899	927	1	5.44	0	1	1
15	35.5575	-83.5669	1654	5	4.67	0.03	0.756	0.803
16	35.5651	-83.5575	1639	3	7.67	0.11	0.857	0.886
17	35.5651	-83.5432	1702	2	5.78	0.22	0.800	0.800
18	35.5726	-83.5336	1585	5	9.44	0.11	0.625	0.825
20	35.5412	-83.4940	1789	5	5.56	0	0.725	0.867
21	35.5515	-83.4918	1781	10	7.85	0.1	0.695	0.798
23	35.5674	-83.4808	1806	9	9.11	0.33	0.704	0.819
25	35.5899	-83.4700	1797	7	7.22	0.22	0.706	0.764
26	35.6797	-83.4611	998	6	7.11	0.11	0.793	0.878
27 (WL91)	35.6092	-83.4465	1703	7	10.22	0	0.613	0.732
34 (WL90)	35.5834	-83.3980	1405	7	6.89	0.01	0.619	0.752
35	35.5189	-83.3781	1278	7	5.56	0	0.696	0.727
37	35.5886	-83.3615	892	5	11.33	0	0.700	0.799
39	35.5662	-83.3419	778	5	6.11	0.11	0.778	0.830
40	35.6122	-83.3409	1100	5	9.22	0.12	0.694	0.797
41	35.5533	-83.3381	1203	5	5.89	0.11	0.733	0.800
42	35.6063	-83.3322	889	3	14.67	0	0.667	0.752
43	35.7091	-83.3198	992	7	7.11	0.03	0.706	0.815
45 (WL89)	35.6121	-83.3158	1186	10	7.03	0	0.863	0.854
46	35.5527	-83.3145	726	5	3.78	0	0.533	0.610
47	35.6189	-83.3097	1239	3	14.56	0.11	0.810	0.862
48	35.5885	-83.3083	868	6	5.78	0.11	0.630	0.803
49	35.7396	-83.2779	938	5	6.56	0	0.817	0.837
51	35.6101	-83.2541	1160	2	6.44	0.11	0.688	0.792
56	35.7482	-83.2146	816	10	8.01	0.01	0.722	0.792

57	35.6535	-83.1976	1672	5	5.89	0.11	0.746	0.871
60	35.6386	-83.1872	1455	5	4.67	0	0.739	0.789
63	35.6114	-83.1850	1440	12	7.69	0	0.635	0.814
68	35.6797	-83.1588	1702	7	6.78	0	0.714	0.816
70	35.6981	-83.1255	1739	7	6.11	0	0.704	0.748

N: number of individuals sequenced or genotyped; H: number of haplotypes; PH: number of private haplotypes; h: haplotype diversity; π: nucleotide diversity; A: rarefied allelic richness; PA: rarefied private allelic richness; H_o: observed heterozygosity; H_e: expected heterozygosity

Table S2 Hardy-Weinberg tests for three clusters inferred by GENELAND and STRUCTURE. Significance was assessed from 1,000 Markov chain Monte Carlo simulations with 10,000 iterations per batch and 10,000 dememorization steps using GENEPOP

Cluster	N	Locus p-value								
		PLJOb	PLJOe	PLJOg	PLJOk	PLJOl	PLJOo	PLJOq	PLJOr	PLJOw
1	72	**<0.0001**	**0.0082**	0.5527	0.1022	0.1484	**0.0006**	0.0636	0.2353	**<0.0001**
2	99	**0.0001**	0.0659	**<0.0001**	0.5294	**<0.0001**	**0.0095**	**<0.0001**	**0.0104**	0.0895
3	66	**0.0204**	**<0.0001**	0.4632	0.0650	**<0.0001**	**0.0016**	**<0.0001**	**0.0003**	**0.0047**

Table S3 Statistical significance of exact G tests of population differentiation for the three clusters inferred by GENELAND and STRUCTURE. Significance was assessed from 1,000 Markov chain Monte Carlo simulations with 10,000 iterations per batch and 10,000 dememorization steps using GENEPOP. Values in bold indicate that the distribution of alleles between each pair of clusters at a given locus was significantly different after Bonferroni correction

Comparison	Locus p-value									
	PLJOb	PLJOe	PLJOg	PLJOk	PLJOl	PLJOo	PLJOq	PLJOr	PLJOw	Overall
Cluster 1 - Cluster 2	**<0.0001**	**<0.0001**	0.0523	**0.0013**	**0.0290**	**<0.0001**	**<0.0001**	0.0306	**<0.0001**	**<0.0001**
Cluster 1 - Cluster 3	**<0.0001**	0.0661	**0.0026**	**<0.0001**	**0.0002**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**
Cluster 2 - Cluster 3	**<0.0001**	**0.0140**	0.1012	**<0.0001**	**0.0002**	**<0.0001**	**<0.0001**	**0.0002**	**<0.0001**	**<0.0001**

Table S4 Pairwise F_{ST} values (below diagonal) and harmonic D (above diagonal) for *P. jordani* localities

(See separate Excel file)

Table S5 Niche overlap values based on three statistics calculated for each pair of clusters using ENMTOOLS. None are statistically significant

Comparison	Niche overlap		
	Schoener's *D*	Warren's *I*	Relative Rank
Cluster 1 - Cluster 2	0.700	0.933	0.818
Cluster 1 - Cluster 3	0.643	0.899	0.734
Cluster 2 - Cluster 3	0.767	0.933	0.800

Appendix S9 Chapter 3 Supplemental Figures S1–S4.

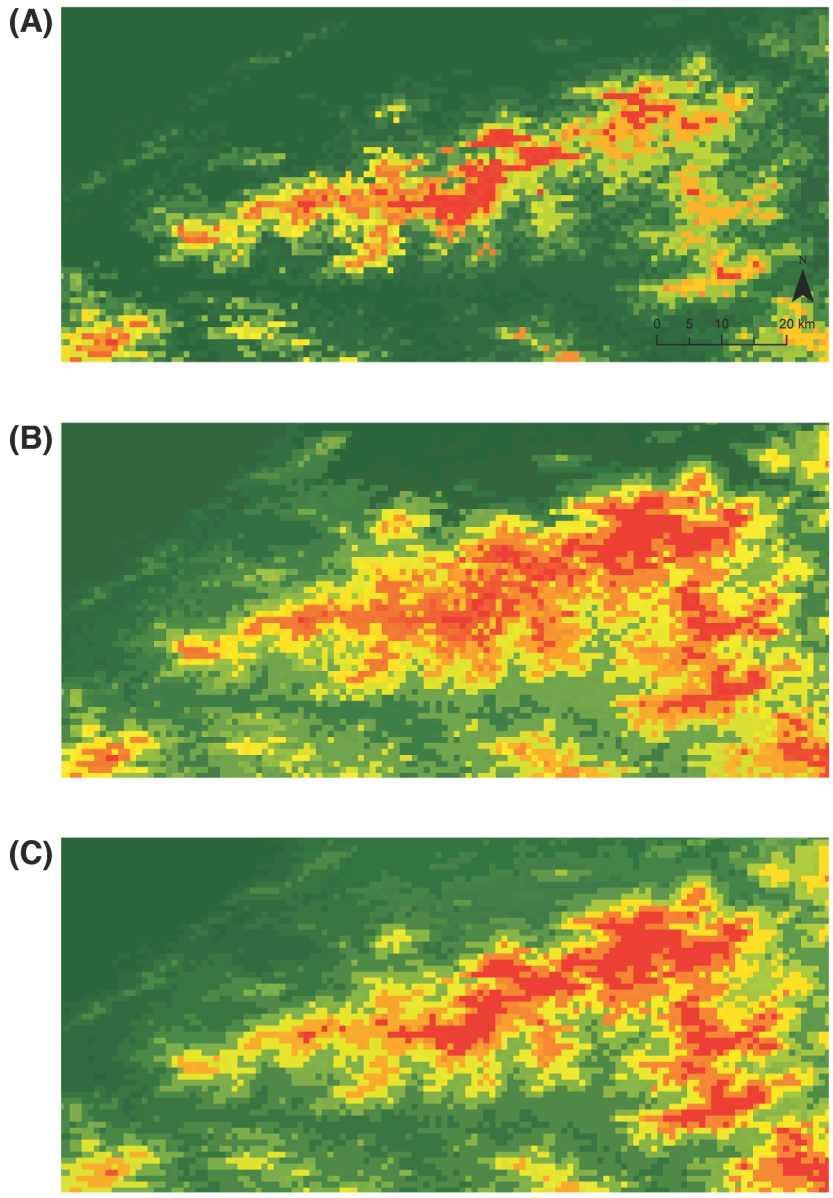


Figure S1 Ecological niche models for each microsatellite cluster generated in MAXENT. (A), Cluster 1. (B), Cluster 2. (C), Cluster 3

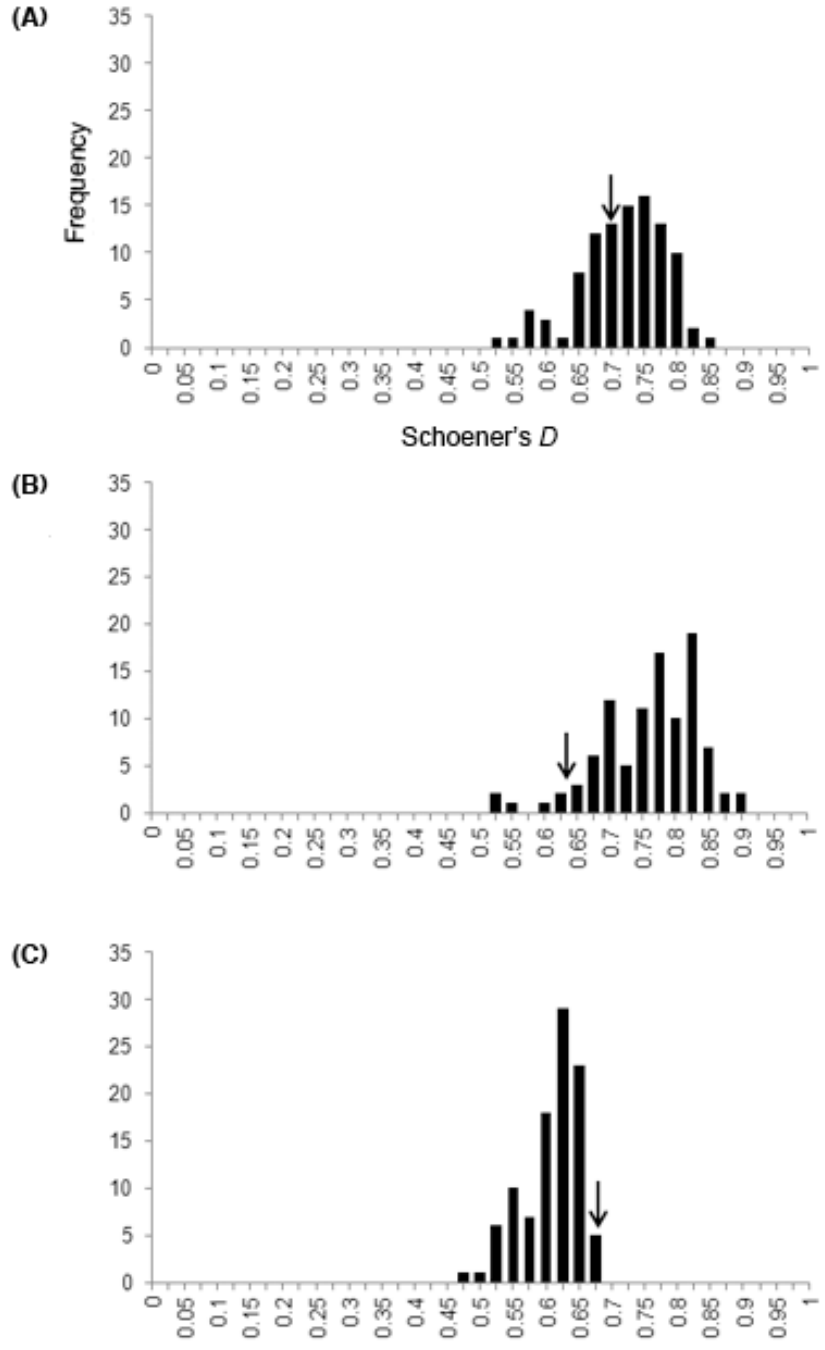


Figure S2 Niche identity test based on Schoener's *D* (Schoener, 1968) statistic for comparisons between each cluster: (A) Cluster 1 – Cluster 2, (B) Cluster 1 – Cluster 3, (C) Cluster 2 – Cluster 3. Axes for each plot are identical to (A). In each plot, observed identity values (Chapter 3, Table S5) are indicated with an arrow. Suitability scores generated by ecological niche models for each lineage exhibit statistically significant ecological differences if the observed overlap value falls outside of the distribution of values obtained from 100 pseudoreplicates

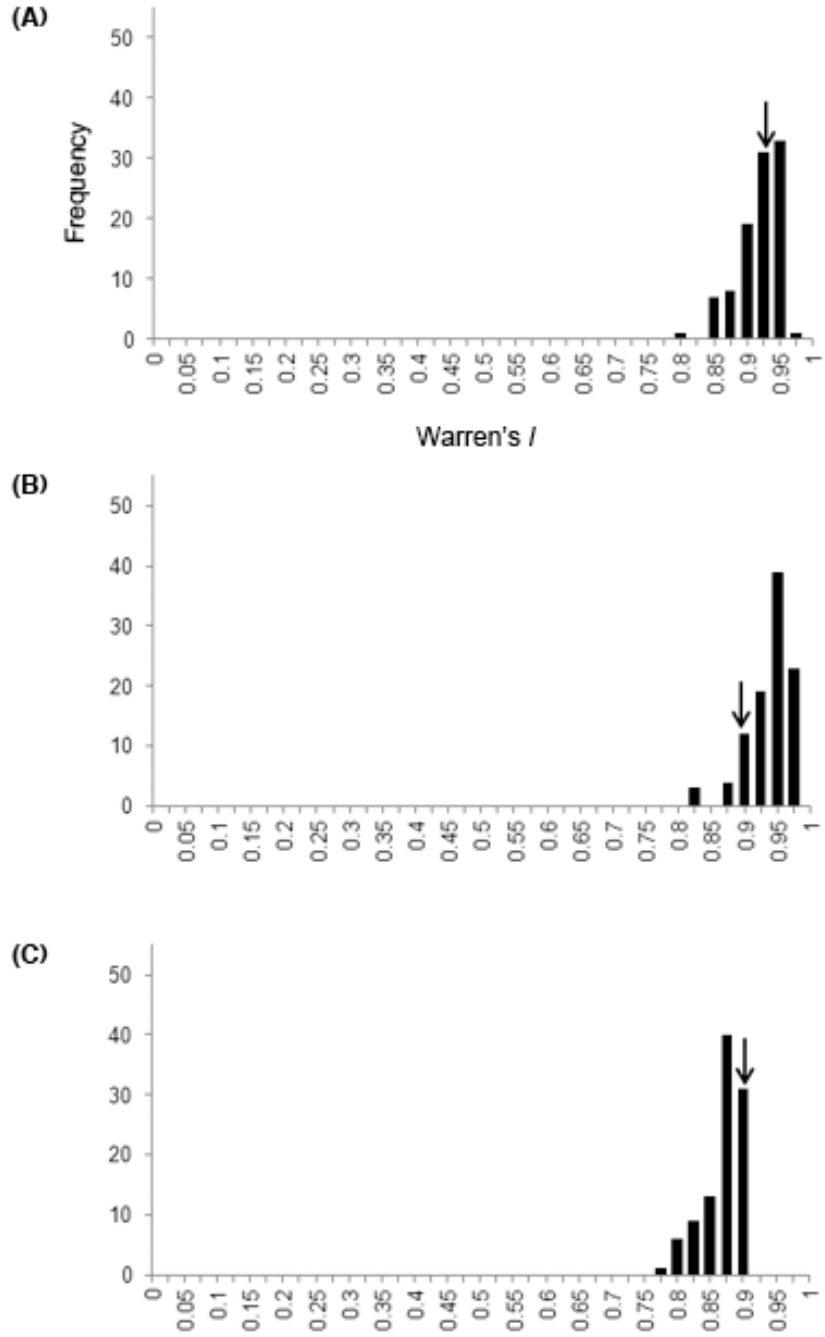


Figure S3 Niche identity test based on Warren's *I* (Warren *et al.*, 2008) statistic for comparisons between each cluster: (A) Cluster 1 – Cluster 2, (B) Cluster 1 – Cluster 3, (C) Cluster 2 – Cluster 3. Axes for each plot are identical to (A). In each plot, observed identity values (Chapter 3, Table S5) are indicated with an arrow. Suitability scores generated by ecological niche models for each lineage exhibit statistically significant ecological differences if the observed overlap value falls outside of the distribution of values obtained from 100 pseudoreplicates

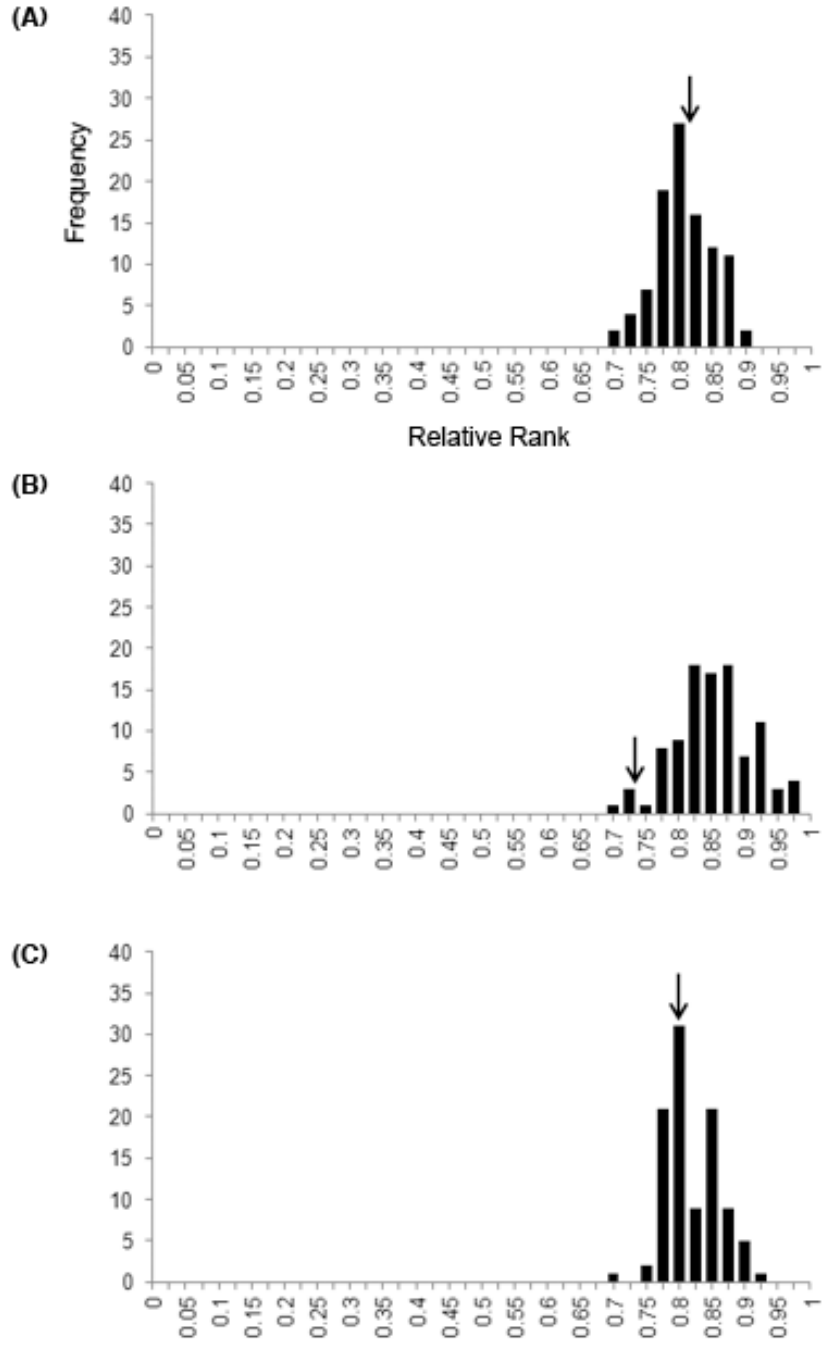


Figure S4 Niche identity test based on relative rank (Warren & Seifert, 2011) statistic for comparisons between each cluster: (A) Cluster 1 – Cluster 2, (B) Cluster 1 – Cluster 3, (C) Cluster 2 – Cluster 3. Axes for each plot are identical to (A). In each plot, observed identity values (Chapter 3, Table S5) are indicated with an arrow. Suitability scores generated by ecological niche models for each lineage exhibit statistically significant ecological differences if the observed overlap value falls outside of the distribution of values obtained from 100 pseudoreplicates