E-ARTICLE

Sixty-Five Million Years of Change in Temperature and Topography Explain Evolutionary History in Eastern North American Plethodontid Salamanders

Richard Barnes[1],* and Adam Thomas Clark[2],*,†

1. Energy and Resources Group, University of California, Berkeley, California 94720;   2. Ecology, Evolution, and Behavior, University of Minnesota, Saint Paul, Minnesota 55108

ABSTRACT: For many taxa and systems, species richness peaks at midelevations. One potential explanation for this pattern is that large-scale changes in climate and geography have, over evolutionary time, selected for traits that are favored under conditions found in contemporary midelevation regions. To test this hypothesis, we use records of historical temperature and topographic changes over the past 65 Myr to construct a general simulation model of plethodontid salamander evolution in eastern North America. We then explore possible mechanisms constraining species to midelevation bands by using the model to predict plethodontid evolutionary history and contemporary geographic distributions. Our results show that models that incorporate both temperature and topographic changes are better able to predict these patterns, suggesting that both processes may have played an important role in driving plethodontid evolution in the region. Additionally, our model (whose annotated source code is included as a supplement) represents a proof of concept to encourage future work that takes advantage of recent advances in computing power to combine models of ecology, evolution, and earth history to better explain the abundance and distribution of species over time.

Keywords: paleoclimatology, climate history, agent-based model (ABM), general simulation model (GSM), eco-evo, phylogenetic reconstruction.

Introduction

Peaks in species richness at midelevation bands have been observed in ecosystems and taxa around the globe (Rahbek 1995). A number of ecological processes may contribute to this "hump-shaped" distribution, including varying autotrophic productivity (Rosenzweig 1995), trade-offs between competitive ability and environmental tolerance (Gifford and Kozak 2012), and differences in area and isolation, which often reach their respective maximum and minimum at midelevations (Sanders 2002). Evolutionary processes have also been suggested; however, they are more difficult to support, as it is often unclear how speciation and extinction rates have changed over time (Morlon 2014).

We propose to address this question using plethodontid salamanders as a model system. These salamanders are often used as model organisms for studies combining ecology and evolutionary biology (Bruce et al. 2000). Nearly all of the ~450 extant plethodontid species are found in North or Central America (Min et al. 2005; AmphibiaWeb 2016), and there is substantial evidence that the family originated in the southern Appalachians (Wilder and Dunn 1920; Dowling 1956; Mueller et al. 2004; Herman and Bouzat 2016) 48–79 Mya (Kozak et al. 2009; Martin et al. 2016). In eastern North America, their current range is largely contiguous, stretching across most of the United States and southern Canada east of the Mississippi River (Herman and Bouzat 2016). Plethodontids are therefore a particularly tractable system for studying long-term patterns of speciation, as the clade appears to have originated, diversified, and persisted within a relatively distinct geographic region.

Using phylogenetic and species distribution data for eastern North American plethodontids, Kozak and Wiens (2010) argue that evolutionary stasis could be a mechanism driving worldwide patterns of species richness along elevation gradients. The authors showed that plethodontid richness peaks along midelevation bands in the Appalachians and suggested that this is due to increased "time for speciation" in these regions. They cite several possible mechanisms that could lead to evolutionary stasis within fixed elevation bands, including interspecific competition and daily variation in temperature and moisture. However, the evolutionary history of plethodontids has coincided with dramatic environ-

mental and geographic change, complicating this explanation. Since the early Permian, the Appalachians have likely lost more than half their original elevation (Slingerland and Furlong 1989). Similarly, over the past 65 Myr there have been large and rapid fluctuations in mean global temperature (Zachos et al. 2001). Such events present a dynamic backdrop for evolution, with potentially major impacts on species' evolutionary histories (Ricklefs et al. 1999; Rangel et al. 2007). For example, recent molecular work suggests that both the phylogenetic and the geographic relationships among *Plethodon serratus* subpopulations have been strongly influenced by changes in regional climate throughout the Pleistocene (Thesing et al. 2016).

We hypothesize that variation in temperature and topography, rather than elevation gradients per se, have changed and constrained the ranges of plethodontid salamanders over their evolutionary history. Because salamanders are small, ectothermic, and relatively poor dispersers, they are sensitive to changes in temperature, and species ranges are often restricted to specific microclimates determined by local temperature and moisture (Carfioli et al. 2000; Gifford and Kozak 2012; Thesing et al. 2016). Large-scale changes that have caused the area available at any particular range of temperatures to vary substantially over time have therefore also potentially changed the space and resources available to salamander species with limited temperature tolerances. If conditions currently found in midelevations were historically more common than other conditions, then it would stand to reason that these regions are also disproportionately species rich. This would be particularly evident if processes such as those cited by Kozak and Wiens (2010) prevented species from adapting to other conditions that have been less historically common.

To evaluate the influence of changes in temperature and topography, we focus on 95 Appalachian plethodontid taxon groups found primarily east of the Mississippi River. These include 79 recognized species and 16 undescribed lineages, including all 82 lineages from Kozak and Wiens (2010) and 13 additional species from Kozak et al. (2009; for the full species list, see the appendix, available online). In all, this accounts for all but six of the described species in the region. We used these data, including species' ranges, richness, and phylogenies, to construct a general simulation model (GSM) of plethodontid evolution (Rangel and Diniz-Filho 2005; Rangel et al. 2007; Gotelli et al. 2009). GSMs are useful when processes are too complex to analyze analytically (Colwell et al. 2012) and include a wide class of computational methods to predict species' responses to the joint influences of evolutionary and ecological processes.

We used this model to emulate species trait evolution and community dynamics over 65 Myr of historical climate and topographic change in eastern North America. Next, we parameterized alternate versions of this model for scenarios where temperature and elevation either change through time or remain static. Finally, we compared model fits across scenarios to test whether including the historical information significantly improved predictions.

Methods

The complete, annotated code used to run the simulations described here is available as part of the supplemental material online and is also accessible online at https://github.com/r-barnes/BarnesClark2017-Salamanders.[1] More specific details regarding our simulation, phylogenetic reconstruction, and optimization methods are also available in the appendix.

Our GSM tracks the evolutionary history of a single progenitor species and all of its descendant species from an origin 65 Mya to the present. Since it would be computationally prohibitive to simulate millions of individual salamanders, each species is instead represented by one or more "populations" that span larger spatial and temporal scales than would an individual. Biologically, these populations might be thought of as demes—subgroups whose trait and genetic differences are small relative to the species as a whole. In our model, we assume that individuals within a population are spatially proximal, share the same traits, and are genetically identical. Between populations, these aspects can vary, which leads to within-species variation in traits, relatedness, and geographic range. The "abundance" of a species in our model is therefore determined by the number of populations of that species summed across all spatial bins. If multiple populations of a species share the same spatial bin at the same time, these populations can "breed" (representing interactions between individuals from the two populations), generating a new "child" population, as described in "Breeding." Over time, mutations can accumulate (generally because populations move to new elevation bands and become reproductively isolated from other populations of the same species). If the genetic differences among populations within a single species become large enough as a result of this, the populations diverge into separate species.

The model progresses in discrete, uniform time steps. Within each time step, several events occur: (*a*) populations suffer mortality based on how their optimal temperature differs from their environment as well as the local density of con- and heterospecifics (i.e., the number of populations per unit area in each elevation bin; see details in "Mortality"); (*b*) the surviving populations breed to produce new populations, during which genetic mutations and crossover may result in new species ("Breeding"); and (*c*) both the

---

1. Code that appears in *The American Naturalist* is provided as a convenience to the readers. It has not necessarily been tested as part of the peer review.

parent and the child populations have an opportunity to migrate to higher or lower elevations ("Temperature and Topographic Change"). As topography and temperature vary in simulation time, the total area available at any given temperature fluctuates, thereby subjecting species to changing selective pressures, leading to re-creations of potential evolutionary histories.

Knobs and Scenarios

Because GSMs are complex, it is helpful to design them with "control knobs," which can be toggled to test different scenarios and assumptions (Gotelli et al. 2009). Our GSM contains two primary control knobs: TEMPERATURE and ELE-VATION ("Temperature and Topographic Change"). When these knobs are "off," we hold their conditions constant for the whole simulation; when they are "on," we vary conditions following historical patterns. Our GSM also includes seven parameters that are directly optimized to match observed patterns and five secondary control knobs. We use these optimized parameters and secondary knobs to provide the best possible model fit for testing our primary hypotheses about temperature and topography and to test the effects of various model assumptions

The seven optimized parameters are as follows: μ and σ, which control the mutation rates of species relatedness and temperature optima, respectively (for details, see "Breeding"); p, which controls the threshold for speciation ("Breeding"); w_0 and w_T, which control mortality rates as a function of temperature ("Mortality"); and w_C and w_H, which control the strength of con- and heterospecific competition, respectively ("Mortality"). The secondary control knobs are COMPETITION ("Mortality"), DISPERSAL ("Dispersal"), ANCESTRAL ELEVATION ("Temperature and Topographic Change"), REINVASIONS ("Dispersal"), and TIME STEP (which sets the model's time step to either 0.5 or 0.1 Myr). The effects of these knobs are discussed in the indicated sections.

Temperature and Topographic Change

In our model, temperature at a given time varies depending on local elevation following the adiabatic lapse rate of dry air, $-9.8°C/km$. If the TEMPERATURE knob is on, then temperature also changes over time according to historic data on global temperature change (Zachos et al. 2001; for details, see the appendix). The time series for sea-level temperature is shown in figure 1a.

We discretize space into a series of 50 equally spaced bins representing the total area available at different elevation bands. To characterize the contemporary area available as a function of elevation, we used data from WorldClim corresponding to the contemporary range of plethodontids east of the Mississippi River, collated in Kozak and Wiens
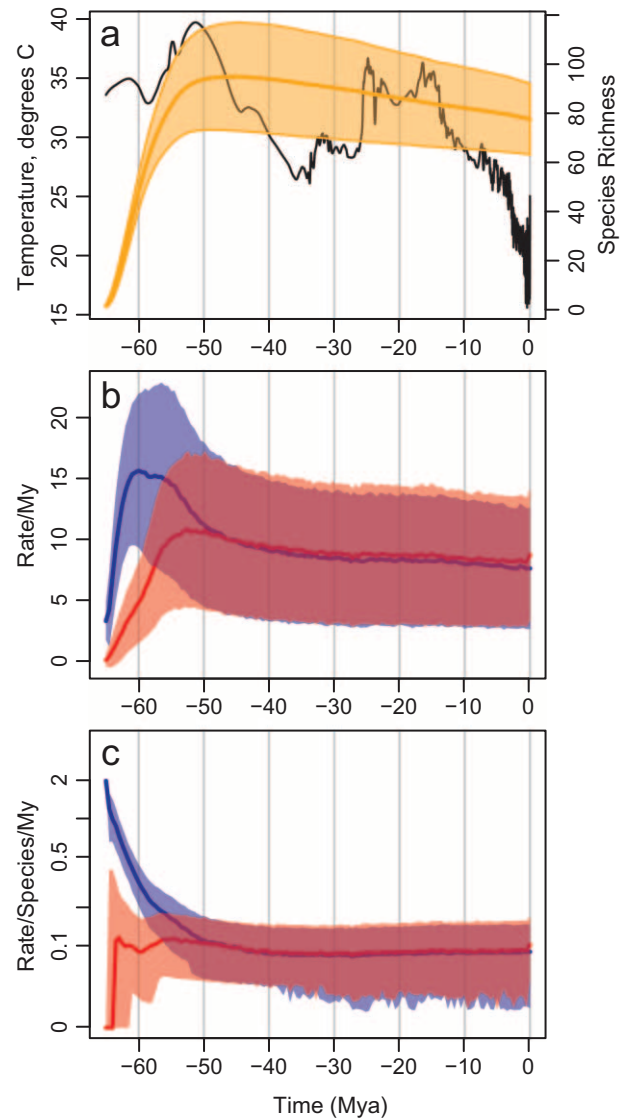

Figure 1: Time trends predicted from the fitted scenario with elevation and temperature change (scenario ET). Lines show means ± 1 SD. *a*, Temperature trend over time (black line) and species richness (orange line). *b, c*, Rates for speciation (blue) and extinction (red). Darker regions show overlap between confidence intervals.

(2010, their table 1). If the ELEVATION knob is off, then these modern elevation data are used to describe topography. If the ELEVATION knob is on, then the area available in each bin varies over time following the region's orogenic history, which is driven largely by the erosion of the Appalachians (for details, see the appendix). As erosion occurs, higher bins are removed. We also tested two settings for the ANCESTRAL ELEVATION knob: the original plethodontid ancestor was placed either (*a*) at sea level or (*b*) at the middle of the initial elevation range.

Breeding

In our model, each salamander population is described by three traits: a species ID, a neutral genome used to test relatedness, and an optimal temperature T_{opt} at which the salamander's temperature-dependent mortality ("Mortality") is at a minimum. Breeding events take place between "parent" populations and give rise to new "child" populations, which can themselves become parent populations in the next time step. Randomly chosen pairs of populations breed with each other, but only if they are conspecific (share a species ID), share at least fraction p of the characters in their genotype, and are located in the same elevation bin, which implies close geographic proximity of the populations. For details, see the appendix.

The neutral genome that determines relatedness is implemented as a 64-bit field that changes independently of the salamander's optimal temperature. Child populations inherit these characters from either of their parents with a 50% probability, plus an additional mutation probability determined by μ. Optimum temperature follows a random walk, whereby child populations inherit the average trait values of their parents plus a random mutation drawn from a normal distribution with mean zero and standard deviation σ.

Mutations may accumulate over time, leading to genotype divergence in reproductively isolated populations. If these differences exceed a fraction p, then the populations can no longer breed, and we consider them to be members of separate species. Speciation is therefore an emergent characteristic of our model, which can arise through any mechanism that imposes reproductive isolation. Most frequently, this occurs through geographic isolation (i.e., allopatric speciation) among populations that have dispersed to different elevation bins. Reproductive isolation could also occur within a single elevation bin through random chance (i.e., sympatric speciation), although geographic proximity makes this unlikely. While technically possible, this second mode of speciation is rare in our model.

Mortality

Each population has a per-time-step mortality probability determined by the density of con- and heterospecifics in its bin as well as the degree to which its optimal temperature is mismatched to its environment. Note that population mortality leads to extinction events only if all populations that make up a particular species die out.

At each time step, a population has a probability of dying out given by the following logistic function:

$$\Lambda = w_T(T_{opt} - T_{bin})^2 + w_H A_H + w_C A_C + w_0, \quad (1)$$

$$\Pr[\text{death}] = 1/(1 + \exp(-\Lambda)). \quad (2)$$

This is effectively a stochastic Lotka-Volterra competition equation, where w_0 describes the base mortality rate, w_T describes changes in mortality attributable to deviations in the population's optimal temperature T_{opt} from the environment's temperature T_{bin}, and w_C and w_H describe, respectively, the influence of competition between con- and heterospecifics, expressed by their respective abundances per unit area, A_C and A_H.

These latter parameters enforce a soft "carrying capacity" for the bin when $w_C > 0$. Thus, bins with smaller areas will generally support fewer populations, and populations are less likely to persist in a bin if the bin's temperature is far from their optimum temperature. When many populations are present, mortality increases. We test three settings for the COMPETITION knob: (*a*) no heterospecific competition ($w_H = 0$), (*b*) equal conspecific/heterospecific competition ($w_H = w_C$), and (*c*) unequal conspecific/heterospecific competition (w_H and w_C are optimized separately as two independent parameters).

Dispersal

We test three settings for the DISPERSAL knob. Depending on this, at each time step a population has a 50% probability of migrating to (*a*) a randomly chosen bin neighboring their current bin (random local dispersal), (*b*) a neighboring bin that more closely matches their optimal temperature (directed local dispersal), or (*c*) a randomly chosen bin anywhere along the elevation gradient (global dispersal). Dispersal rate is therefore controlled by time step length.

Although plethodontids are thought to have originated in the southern Appalachians, there is some evidence that lineages that initially dispersed out of eastern North America later reinvaded the region as many as four times (Martin et al. 2016). The REINVASION knob accounts for this. When it is on, populations have a 0.1% chance per time step of leaving the simulated region. Once they leave, populations persist and mutate but do not compete or experience temperature and elevation fluctuations. They have a 0.1% chance per time step of reinvading the simulated region. In simulations, this typically leads to the first successful reinvasion by an outside lineage around 50 Mya, with additional invasions becoming more frequent as time moves forward, roughly matching the hypothesized historical invasion history of Martin et al. (2016).

Empirical Metrics

We compared our simulation results for extant species (i.e., those present 0 Mya) to empirically observed patterns and molecular phylogenies based on six metrics. Four of these have been suggested previously as important for analyzing

GSMs (Gotelli et al. 2009; Morlon 2014): (*a*) total species richness (i.e., number of species), (*b*) total species with the middle of their elevation range falling in each elevation bin, (*c*) species elevational range-size frequency distributions, and (*d*) lineages-through-time (LTT) plots. We also include two additional metrics: (*e*) Pybus and Harvey's γ statistic (Pybus and Harvey 2000), which better characterizes recent divergence events, and (*f*) Blomberg's *K* statistic (Blomberg et al. 2003), which we use to characterize the relationship between phylogenetic distance among species and differences in temperature at the middle of their elevational range (an empirical proxy for optimum temperature).

For metrics *a–c*, we used the richness and distributional data from Kozak and Wiens (2010) and matched them to the elevation bins from our simulation. For metrics *d–f*, we used the published phylogeny from Kozak et al. (2009). We use this phylogeny because it corresponds directly to the lineages in Kozak and Wiens (2010) and because its topology closely matches that of a more recent study by Martin et al. (2016). To account for uncertainty in divergence times for the molecular phylogeny, we used the aligned sequences and methods from Kozak et al. (2009) to reconstruct 10,000 bootstrapped replicates of the maximum likelihood phylogeny in RAxML (ver. 8.2.9; Stamatakis 2014). We then time-calibrated the resulting trees on the basis of the age constraints from Kozak et al. (2009) using the PATHd8 algorithm (Britton et al. 2007) implemented in phyloGenerator (Pearse and Purvis 2013). We then calculated metrics *d–f* for all 10,000 trees, resulting in an empirical frequency distribution for each metric (for details, see the appendix).

Model Fitting and Scenarios

To optimize the seven parameters ("Knobs and Scenarios") in our model, we used a simulated annealing algorithm to maximize the likelihood of observed data given simulated predictions across all six empirical metrics ("Empirical Metrics"). Because these metrics involve different numbers of comparisons (e.g., there is only a single observed species richness to compare, but there are many observations across time for the LTT plot), likelihoods vary by orders of magnitude across metrics even when models fit well. This leads some metrics to have a disproportionately strong effect on the optimization. To prevent this, we calculated a likelihood ratio for each metric by taking the raw likelihoods for the observed data given the model and dividing them by the likelihood that would be expected for the observed data if it perfectly matched the simulated distribution. We then minimized the value of this metric, which maximized the weighted model likelihood. This ensured that the log-likelihood ratios of each metric were generally within the same order of magnitude, giving the metrics roughly equal weight in our optimization process. For details, see the appendix.

Because there were almost 300 potential configurations of the secondary knobs ("Knobs and Scenarios") in our GSM, we used a forward selection procedure to determine the best position for them (see the appendix). This led us to a configuration with shared conspecific/heterospecific competition, random local dispersal, midelevation ancestral origin, no reinvasions, and a time step of 0.5 Myr.

To test whether including information about temperature and topographic change improved the GSM's fit, we optimized on the above best configuration using the two primary knobs. This gave four different scenarios: TEMPERATURE and ELEVATION knobs off (scenario N), only the ELEVATION knob on (scenario E), only the TEMPERATURE knob on (scenario T), and both knobs on (scenario ET).

To test the effects of some of our model assumptions, we also optimized the ET scenario for all possible single-knob deviations from the best configuration of the five secondary knobs. This gave seven additional scenarios: two for the DISPERSAL knob, two for the COMPETITION knob, one for the ANCESTRAL ELEVATION knob, one for the REINVASION knob, and one for the TIME STEP knob. We then ran 5,000 iterations for each of these four primary and seven secondary scenarios at their maximum likelihood parameter estimates to calculate model likelihoods and distributions for each of our metrics. Finally, to test for significant differences in likelihood between models, within each of the primary and secondary scenarios we bootstrapped by resampling 5,000 iterations with replacement and calculated the likelihood for each iteration.

Results

For the total fit across all six metrics, the ET scenario, which included both elevation and temperature changes through time, provided significantly better fits than the E, T, and N scenarios; the last fit the worst (table 1). For the seven secondary scenarios we tested, all provided significantly worse total fits than the regular ET scenario, with the exception of two (the global and directed DISPERSAL knob scenarios, and the REINVASION knob on) for which the total fit was not significantly different.

For the LTT plot and predictions of extant species richness (which simply reflects the LTT plot at 0 Mya), the ET scenario matched observed trends reasonably well (fig. 2*a*, 2*d*), although the richness predictions followed a bimodal distribution. The best predictions came from the scenario with the DISPERSAL knob set to directed. For both the random and the directed DISPERSAL knob settings, the ET scenario matched the general trend from the molecular phylogeny, with an accumulation of roughly 0.5 lineages per million years for the first 40 Myr of the simulation, increasing to five lineages per million years for the past 15–20 Myr (figs. 2*a*, 3*a*). Setting the ANCESTRAL ELEVATION knob to low technically gave

Table 1: Likelihood for each scenario evaluated at their maximum likelihood parameter estimates

Scenario	Total fit		Richness		Midelevation		Elevation range		LTT plot		γ		K	
	LR	P	LR	P	LR	P	LR	P	LR	P	LR	P	LR	P
Elevation + temperature (ET)	4.71	...	.38	...	.09	...	1.81	...	2.14	...	−.28	...	.57	...
Neither (N)	7.66	<.001***	1.25	<.001***	.11	<.001***	1.89	<.001***	2.01	>.999***	.59	<.001***	1.8	<.001***
Temperature only (T)	6.54	<.001***	1.09	<.001***	.09	.3158	1.74	>.999***	2.11	.9064	.04	<.001***	1.48	<.001***
Elevation only (E)	5.4	<.001***	−.37	>.999***	.03	>.999***	1.88	<.001***	3.02	<.001***	−.41	>.999***	1.25	<.001***
$w_H \neq w_C$	6.73	<.001***	−.38	>.999***	.05	>.999***	1.92	<.001***	3.07	<.001***	2.16	<.001***	−.08	>.999***
$w_H = 0$	4.95	.0388*	−.25	>.999***	.19	<.001***	1.97	<.001***	2.75	<.001***	−.05	<.001***	.35	>.999***
Directed dispersal	4.75	.38	−.43	>.999***	.01	>.999***	2.2	<.001***	1.63	>.999***	1.41	<.001***	−.08	>.999***
Global dispersal	15.09	<.001***	11.22	<.001***	.05	>.999***	2.87	<.001***	1.29	>.999***	−.3	.7426	−.04	>.999***
Low-elevation origin	17.72	<.001***	−.66	>.999***	.22	<.001***	3.8	<.001***	1.55	>.999***	13.12	<.001***	−.31	>.999***
Multiple reinvasions	4.49	.9386	.23	.9032	.07	>.999***	1.92	<.001***	2.89	<.001***	−.35	.9644*	−.27	>.999***
Time step = .1	15.12	<.001***	11.25	<.001***	.12	<.001***	2.94	<.001***	1.31	>.999***	−.27	.3924	−.24	>.999***

Note: Shown is the likelihood ratio (LR) as described in the text. Note that models with a smaller LR have a higher likelihood. P values test LR$_{ET}$ < LR$_i$ for each scenario i. Asterisks indicate significant differences. LTT = lineages-through-time.

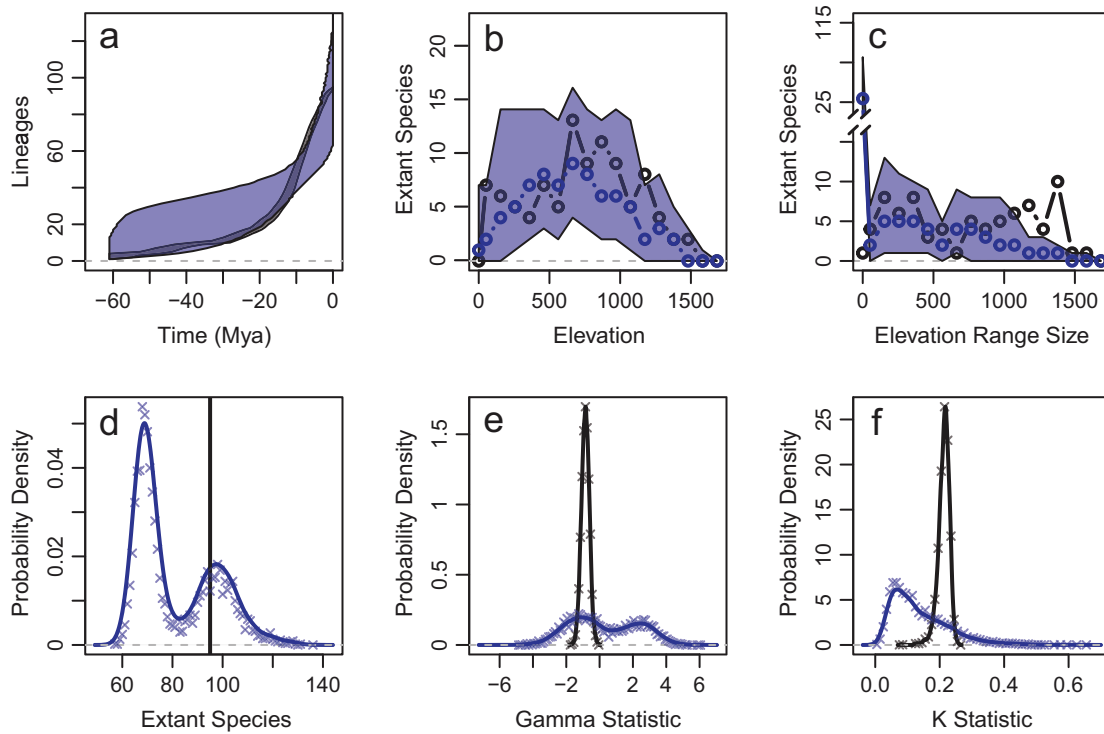* α < .05.

*** α < .001.

Figure 2: Comparison of observed and predicted metrics, based on the fitted scenario with elevation and temperature change (scenario ET). Black lines show observed data, while blue lines show results from simulations. The gray-shaded region in *a* and the blue-shaded regions in *a–c* show 95% confidence intervals for observed and simulated metrics, respectively. Lines in *b* and *c* show values for observations or medians for simulated values. Lines in *d–f* show smoothed kernel density estimates, and points show frequency calculated empirically from binned data.

a lower likelihood ratio, but it predicted an average of 150 extant species and performed well because of a very broad distribution for simulated values.

For species' midelevation distribution, predictions from the ET scenario closely matched the observed hump-shaped trend (fig. 2*b*). Several of the secondary scenarios provided significantly better fits with qualitatively similar results, but differences in likelihood were small. For the species elevation range distribution, the ET scenario significantly outperformed all other primary and secondary scenarios except for the T scenario, although again, these predictions were qualitatively similar to those from the ET scenario. For all models, the fit for elevation range was relatively poor, with significantly too few species with elevation ranges greater than 1,000 m and significantly too many with ranges of 0 m (fig. 2*c*).

As with richness, ET predictions for the γ statistic produced a bimodal distribution including both positive and negative values (fig. 2*e*). For the *K* statistic, the ET model predicted values that were significantly less than 1 but skewed somewhat lower than those from the molecular phylogeny

(fig. 2*f*). Nevertheless, likelihoods for the ET model were significantly greater than those for the other primary scenarios, with the exception of the T scenario, which fit the γ statistic significantly better. Among the secondary scenarios, several outperformed the ET scenario for these metrics, and the multiple-reinvasions scenario provided particularly good predictions for both the γ and the *K* statistic (fig. 3*c*, 3*d*).

The ET scenario predicted smooth time trends for total species richness (fig. 1*a*), speciation and extinction rates (fig. 1*b*), and per capita speciation and extinction rates (fig. 1*c*). In general, speciation exceeded extinction for the first 15 Myr of the simulation, leading to a rapid accumulation of species, followed by a slightly higher extinction than speciation rate, leading to a gradual total decline in richness. This broadly matched trends in the E and T scenarios, although scenarios without elevation change did not experience net loss of species over time. Conversely, speciation and extinction dynamics for many of the secondary scenarios, including the two best-fitting scenarios—directed dispersal (fig. 3*e*) and multiple reinvasions (fig. 3*f*)—appeared to be strongly driven by temperature changes, with peaks in extinction rate corre-
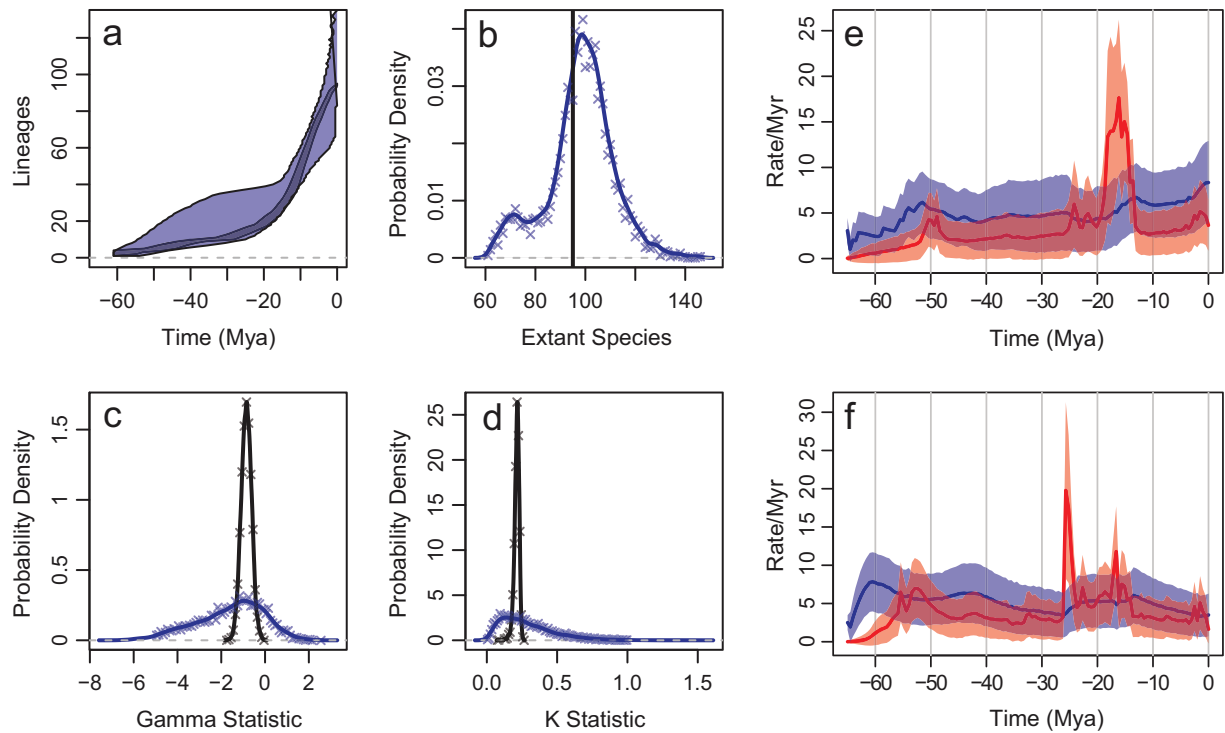
Figure 3: Examples of metric comparisons and time trends for other fitted scenarios. *a*, *b*, and *e* are for the scenario with directed dispersal, and *c*, *d*, and *f* are for the scenario with multiple reinvasions. Symbols and intervals for *a–d* are as described in the figure 2*a*, 2*d*, 2*e*, and 2*f* legends, respectively. Colors and intervals in *e* and *f* are as described in the figure 1*b* legend.

sponding to periods of rapid temperature change around 50, 25, and 15 Mya.

Discussion

Our results suggest that changes in temperature and elevation jointly influenced the evolutionary history of plethodontid salamanders in the populations we consider. This was by no means a foregone conclusion: although they utilize different information about historical environmental dynamics, the N, E, T, and ET scenarios all include the same number of model parameters, and there is no analytical reason to assume that the more "complex" scenarios should yield better predictions. Our findings are consistent with results for several plethodontid species from the North American interior highlands that attribute geographic distributions and phylogenetic relationships to a combination of climate and topographic factors, particularly during the Pleistocene era (Martin et al. 2016; Thesing et al. 2016). This also accords with findings for the *Eurycea bislineata* species complex east of the Mississippi River, which suggests that small changes in topography—and particularly in historical drainage basins—explain phylogenetic relationships (Kozak et al. 2006).

Unlike these existing studies, our results apply to a larger set of lineages and rely on mechanisms that act at substantially larger spatial and temporal scales. Nevertheless, it has been suggested previously that large-scale climate and geographic changes are important drivers of evolutionary change (Ricklefs et al. 1999). For example, dramatic shifts in community composition and species ranges resulting from climate change have been recorded among terrestrial plant and animal communities, particularly over the past 12 kyr in eastern North America (Jackson and Overpeck 2000). Similarly, over the past few hundred thousand years, Milankovitch climate forcing has been posited as a major driver of large-scale shifts in species distributions (Dynesius and Jansson 2000) and may be responsible for promoting rapid speciation among closely related lineages while reducing the formation of deep branches in species phylogenies (Jansson and Dynesius 2002).

Contemporary Diversity and Species Elevation Ranges

As noted above, midelevation bands tend to have greater relative area, which may promote species diversity (Sanders 2002), especially if temperatures are stable. However, the midelevation diversity observed in contemporary popula-

tions in our model cannot be explained by these factors alone: historical temperatures were generally much higher than they are today (Zachos et al. 2001), and temperatures matching those in contemporary midelevation bands were comparatively rare. Despite this, because species midelevation bands were better described by the E scenario and their elevation ranges were better described by the T scenario (table 1), it appears that joint consideration of historical temperature and elevation changes are necessary for explaining species elevational distributions. This may also indicate that elevation changes were important for driving interspecific differences in geographic distribution and temperature tolerance, while temperature fluctuations were important for within-species variability in temperature tolerances.

Several of the secondary scenarios were better able to predict species midelevation bands than was the primary ET model, including scenarios with the COMPETITION knob set with reduced interspecific competition relative to conspecific competition, the DISPERSAL knob set to directed local dispersal or global dispersal, and the scenario with the REINVASION knob turned on. Reduction in the strength of interspecific competition allows more species to coexist within a single elevation band (Gifford and Kozak 2012), which could explain why this scenario was better able to model high diversity at midelevation ranges.

Because the secondary dispersal and reinvasion scenarios allowed species to colonize areas more rapidly than the primary model, these scenarios' superior predictive ability for species midelevation bands may also indicate that rapid dispersal was important for colonizing contemporary midelevation bands, as they came to represent a larger fraction of the total available area. This would match findings suggesting that the large, fragmented geographic ranges of some plethodontids led to rapid divergence during the Pleistocene (Thesing et al. 2016). Interestingly, these scenarios also correspond to predictions of rapid temperature-driven extinction events between 10 and 30 Mya (fig. 3*e*, 3*f*), which suggests that a substantial portion of the contemporary salamander species resulted from relatively recent diversification events. This provides another explanation of why these scenarios were better able to explain high diversity at contemporary midelevation bands, as this diversification would have coincided with the rapid increase in available area at these temperature ranges over the past 10 Myr.

Scenario Results

In accordance with results from other GSMs, our findings suggest that no single model is able to simultaneously match all of the metrics that we test (Gotelli et al. 2009). Although the ET scenario was the best-fitting model when considered across all of the metrics that we tested, it performed poorly for predictions of extant species richness and for the γ and

K statistics; it was outperformed by several other scenarios for both the LTT plot and the extant species midelevation distribution. The same problem persisted even for higher-dimensional interactions among control knobs that we tested. For example, a three-way interaction scenario testing the ET model with directed dispersal and multiple reinvasions performed worse than either of the two-way interactions. Thus, different scenarios of our GSM appear to be best suited for predicting specific kinds of metrics.

The most obvious difference between the primary and secondary scenarios was in the relationship between speciation and extinction rates over time. In the ET model (and most other primary scenarios), w_T was small, and correlations between diversification rates and temperature change were weak (fig. A1, available online). Nevertheless, the K statistic was significantly less than 1, suggesting that selective pressures on temperature optima were sufficiently strong to override niche conservatism (which arises in the models as a low σ value, representing a Brownian process with limited drift; Blomberg et al. 2003). Conversely, in most of the secondary scenarios, effects of temperature on mortality were substantially higher, and periods of rapid temperature change corresponded to spikes in extinction rates. Interestingly, speciation rates changed little through time in these scenarios, which does not match current paradigms of "bursts" of speciation leading to or following from major extinction events (Ricklefs 2014). This suggests that the plethodontid species pool may be determined primarily by a relatively constant background rate of speciation, with periodic bursts of extinction.

Our results for the competition scenarios suggest that competitive effects among con- and heterospecifics are not significantly different. However, assuming no heterospecific competitive interactions led to small but significant reductions in likelihood, suggesting that in addition to environmental changes, competitive interactions among species are important in this system (Rangel et al. 2007). This matches results for global amphibian distributions (Munguía et al. 2012) and for physiological models of the species *Plethodon teyahelee* (Gifford and Kozak 2012), suggesting that a combination of competition, physiological tolerances, and dispersal limitation determines salamander species ranges.

Among the dispersal scenarios, directed dispersal substantially outperformed global dispersal for predictions of extant species richness and outperformed random local dispersal for both richness and LTT plots. This illustrates the importance of metacommunity dynamics with limited dispersal for maintaining diversity in our model (Hanski 1998; Kerr et al. 2002; Holyoak et al. 2005), although it suggests that species' abilities to track favorable climate gradients is important as well (Ackerly 2003; Rangel et al. 2007). The sharp "elbow" in the LTT plots for both the molecular and the simulated phylogenies are consistent with high extinction rates

relative to speciation (figs. 2*a*, 3*a*), match previous results for plethodontids (Nee et al. 1994), and generally match predictions from all scenarios that we tested. Interestingly, given that scenarios that predicted rapid spikes in extinction rates through time generally provided better estimates of the LTT plot than did scenarios that did not, this suggests that periodic extinction events may be important for accurately re-creating the patterns observed in the molecular phylogeny (Morlon 2014).

For the scenario with multiple reinvasions, improved fits for the γ and K statistics suggest that recolonization events from outside eastern North America are important for explaining plethodontid diversification rates and the evolution of their temperature optima; this accords with molecular and biogeographic evidence (Martin et al. 2016). Compared to the ET scenario, the γ statistic for the reinvasion scenario was lower and primarily but not significantly negative, suggesting that the speciation rate has been slowing over time (Pybus and Harvey 2000). Potentially, this is because reinvasions from outside the region allowed for deep branching events in the phylogeny without the need for high speciation and extinction rates.

Modeling Methods

The models that we present here have a number of advantages over existing methods. First, because we separately track relatedness and trait values, traits are generally correlated with phylogenetic relatedness in our model but are not determined by it (Ackerly 2009). Second, our approach allows speciation to emerge as a coupled result of population size, genetic isolation, and drift, without the need to impose an exogenous speciation rate (Hubert et al. 2015). Importantly, our method even allows speciation to take place sympatrically if, by random chance, nearby populations fail to mate for a sufficiently large number of time steps (Losos and Glor 2003), although spatially segregated populations are naturally much more likely to diverge. Third, in our reinvasion scenarios we show a simple but tractable method for incorporating evolutionary dynamics outside the main area of interest, which seems to perform well for our system. Finally, while our optimization methods follow existing methods for GSMs by pooling all prediction errors into a single term (Gotelli et al. 2009), the likelihood ratio approach that we utilize allows us to fit multiple metrics simultaneously without individual components of those metrics dominating the optimization.

There are also a number of important caveats for interpreting our model results. Most importantly, we model temperature, competition, and dispersal at very large scales. Salamanders are strongly influenced by moisture and humidity (Carfioli et al. 2000; Thesing et al. 2016), and their distributions are often strongly associated with microclimates

and small-scale topography, such as river valleys and drainage basins (Bernardo and Reagan-Wallin 2002; Kozak et al. 2006; Johnson et al. 2008). Thus, even before the Pleistocene, when salamander ranges were more contiguous, it is unlikely that a one-dimensional elevation model accurately characterizes dispersal barriers and habitat gradients in the region. While mechanistic models exist that could be used to incorporate these smaller-scale influences on salamander populations (Vasseur and McCann 2005; Kearney and Porter 2009; Buckley et al. 2010; Gifford and Kozak 2012), such methods are typically parameter rich and difficult to fit (Rangel et al. 2007). Moreover, even if we could accurately model species responses, it would be difficult, or perhaps impossible, to hindcast historical changes in small-scale factors, such as microhabitats and drainage basins, across a multi-million-year timeline. Thus, it may not be possible to improve our predictions by incorporating more detailed environmental models.

Additionally, while we do account for within-species variability in optimum temperature, which is potentially important (Bolnick et al. 2003), we do not account for within- or between-species variation in temperature tolerance width. This is partially an attempt to simplify the model and matches methods used in previous GSMs (Rangel et al. 2007; Gotelli et al. 2009). Nevertheless, it also represents a methodological challenge for future studies. Ideally, niche width should be allowed to evolve over time just like other traits. However, unless this is incorporated as part of a trade-off, species will naturally evolve to have the highest temperature tolerance range possible.

A third problem is that for a number of our empirical metrics, even the best-fitting models still provide relatively poor fits. In particular, the elevation range frequency distribution includes too many species with very small ranges and not enough with large ranges. This could be reflective of incomplete sampling for real-world data. Plethodontids are well known for including many cryptic and poorly described species (Camp and Wooten 2016; Thesing et al. 2016), and wide species ranges could result from erroneously lumping multiple species or from incorrectly characterizing recently diverged species (Rosenblum et al. 2012). Nevertheless, it seems likely that our GSM simply does not allow for sufficient migration of isolated populations; one way to address this would be to allow dispersal probability and distance to evolve as a trait, facilitating the evolution of generalist and specialist species.

Similarly, for both γ and K statistics our predictions produce substantially wider and less symmetrical distributions than do the molecular phylogenies. This appears to be a result of the discrete branch lengths that our simulations produce (e.g., the shortest tip is always at least one time step long). In the scenario with shorter time steps, the distributions for both statistics are substantially more Gaussian and

better centered on the observed values. Adding random noise to the branch lengths from simulations with longer time steps gave similarly improved results, which again suggests a discretization effect.

Conclusion

Our results suggest that community-level interactions coupled with geographic and climate changes are important contributors to the observed correspondence between plethodontid salamander phylogenetic and spatial patterns in eastern North America (Kozak and Wiens 2010). Our results also show the practical applicability of linking population dynamics, community composition, and environmental change in a GSM framework to explain evolutionary histories (Gotelli et al. 2009). Recent advances in models that include both species interactions and evolution have provided useful insight into understanding both processes (Rangel and Diniz-Filho 2005; Rangel et al. 2007; Rosindell et al. 2015). We hope that incorporating geological and climate history provides further insight and that our study can help others apply similar tools in the future.

Literature Cited

Ackerly, D. 2009. Conservatism and diversification of plant functional traits: evolutionary rates versus phylogenetic signal. Proceedings of the National Academy of Sciences of the USA 106:19699–19706.

———. 2003. Community assembly, niche conservatism, and adaptive evolution in changing environments. International Journal of Plant Sciences 164:S165–S184.

AmphibiaWeb. 2016. Amphibian species lists. Accessed August 24, 2016.

Bernardo, J., and N. Reagan-Wallin. 2002. Plethodontid salamanders do not conform to "general rules" for ectotherm life histories: insights from allocation models about why simple models do not make accurate predictions. Oikos 97:398–414.

Blomberg, S. P., T. Garland Jr., A. R. Ives, and B. Crespi. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57:717–745.

Bolnick, D. I., R. Svanbäck, J. A. Fordyce, L. H. Yang, J. M. Davis, C. D. Hulsey, and M. L. Forister. 2003. The ecology of individuals: incidence and implications of individual specialization. American Naturalist 161:1–28.

Britton, T., C. L. Anderson, D. Jacquet, S. Lundqvist, and K. Bremer. 2007. Estimating divergence times in large phylogenetic trees. Systematic Biology 56:741–752.

Bruce, R. C., R. G. Jaeger, and L. D. Houck. 2000. Preface. Pages ix–x *in* R. C. Bruce, R. G. Jaeger, and L. D. Houck, eds. The biology of plethodontid salamanders. Kluwer/Plenum, New York.

Buckley, L. B., M. C. Urban, M. J. Angilletta, L. G. Crozier, L. J. Rissler, and M. W. Sears. 2010. Can mechanism inform species' distribution models? Ecology Letters 13:1041–1054.

Camp, C. D., and J. A. Wooten. 2016. Hidden in plain sight: cryptic diversity in the Plethodontidae. Copeia 104:111–117.

Carfioli, M. A., H. M. I. Tiebout, S. A. Pagano, K. M. Heister, and F. C. Lutcher. 2000. Monitoring *Plethodon cinereus* populations. Pages 463–475 *in* R. C. Bruce, R. G. Jaeger, and L. D. Houck, eds. The Biology of plethodontid salamanders. Kluwer/Plenum, New York.

Colwell, R. K., R. R. Dunn, and N. C. Harris. 2012. Coextinction and persistence of dependent species in a changing world. Annual Review of Ecology, Evolution, and Systematics 43:183–203.

Dowling, H. G. 1956. Geographic relations of Ozarkian amphibians and reptiles. Southwestern Naturalist 1:174–189.

Dynesius, M., and R. Jansson. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. Proceedings of the National Academy of Sciences of the USA 97:9115–9120.

Gifford, M. E., and K. H. Kozak. 2012. Islands in the sky or squeezed at the top? ecological causes of elevational range limits in montane salamanders. Ecography 35:193–203.

Gotelli, N. J., M. J. Anderson, H. T. Arita, A. Chao, R. K. Colwell, S. R. Connolly, D. J. Currie, et al. 2009. Patterns and causes of species richness: a general simulation model for macroecology. Ecology Letters 12:873–886.

Hanski, I. 1998. Metapopulation dynamics. Nature 396:41–49.

Herman, T. A., and J. L. Bouzat. 2016. Range-wide phylogeography of the four-toed salamander: out of Appalachia and into the glacial aftermath. Journal of Biogeography 43:666–678.

Holyoak, M., M. A. Leibold, and R. D. Holt. 2005. Metacommunities: spatial dynamics and ecological communities. University of Chicago Press, Chicago.

Hubert, N., V. Calcagno, R. S. Etienne, and N. Mouquet. 2015. Metacommunity speciation models and their implications for diversification theory. Ecology Letters 18:864–881.

Jackson, S. T., and J. T. Overpeck. 2000. Responses of plant populations and communities to environmental changes of the late Quaternary. Paleobiology 26:194–220.

Jansson, R., and M. Dynesius. 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.

Johnson, M. A., M. Leal, L. Rodriguez Schettino, A. C. Lara, L. J. Revell, and J. B. Losos. 2008. A phylogenetic perspective on foraging mode evolution and habitat use in West Indian *Anolis* lizards. Animal Behaviour 75:555–563.

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

Kerr, B., M. A. Riley, M. W. Feldman, and B. J. Bohannan. 2002. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. Nature 418:171–174.

Kozak, K. H., R. A. Blaine, and A. Larson. 2006. Gene lineages and eastern North American palaeodrainage basins: phylogeography and speciation in salamanders of the *Eurycea bislineata* species complex. Molecular Ecology 15:191–207.

Kozak, K. H., R. W. Mendyk, and J. J. Wiens. 2009. Can parallel diversification occur in sympatry? repeated patterns of body-size evolution in coexisting clades of North American salamanders. Evolution 63:1769–1784.

Kozak, K. H., and J. J. Wiens. 2010. Niche conservatism drives elevational diversity patterns in Appalachian salamanders. American Naturalist 176:40–54.

Losos, J. B., and R. E. Glor. 2003. Phylogenetic comparative methods and the geography of speciation. Trends in Ecology and Evolution 18:220–227.

Martin, S. D., D. B. Shepard, M. A. Steffen, J. G. Phillips, and R. M. Bonett. 2016. Biogeography and colonization history of plethodontid salamanders from the interior highlands of eastern North America. Journal of Biogeography 43:410–422.

Min, M. S., S.-Y. Yang, R. Bonett, D. Vieites, R. Brandon, and D. Wake. 2005. Discovery of the first Asian plethodontid salamander. Nature 435:87–90.

Morlon, H. 2014. Phylogenetic approaches for studying diversification. Ecology Letters 17:508–525.

Mueller, R. L., J. R. Macey, M. Jaekel, D. B. Wake, and J. L. Boore. 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. Proceedings of the National Academy of Sciences of the USA 101:13820–13825.

Munguía, M., C. Rahbek, T. F. Rangel, J. A. F. Diniz-Filho, and M. B. Araújo. 2012. Equilibrium of global amphibian species distributions with climate. PLoS ONE 7:e34420.

Nee, S., E. C. Holmes, R. M. May, and P. H. Harvey. 1994. Extinction rates can be estimated from molecular phylogenies. Philosophical Transactions of the Royal Society B 344:77–82.

Pearse, W. D., and A. Purvis. 2013. phyloGenerator: an automated phylogeny generation tool for ecologists. Methods in Ecology and Evolution 4:692–698.

Pybus, O. G., and P. H. Harvey. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. Proceedings of the Royal Society B 267:2267–2272.

Rahbek, C. 1995. The elevational gradient of species richness: a uniform pattern? Ecography 18:200–205.

Rangel, T. F. L., and J. A. F. Diniz-Filho. 2005. An evolutionary tolerance model explaining spatial patterns in species richness under environmental gradients and geometric constraints. Ecography 28:253–263.

Rangel, T. F. L., J. A. F. Diniz-Filho, and R. K. Colwell. 2007. Species richness and evolutionary niche dynamics: a spatial pattern–oriented simulation experiment. American Naturalist 170:602–616.

Ricklefs, R. E. 2014. Reconciling diversification: random pulse models of speciation and extinction. American Naturalist 184:268–276.

Ricklefs, R. E., R. E. Latham, and H. Qian. 1999. Global patterns of tree species richness in moist forests: distinguishing ecological influences and historical contingency. Oikos 86:369–373.

Rosenblum, E. B., B. A. Sarver, J. W. Brown, S. Des Roches, K. M. Hardwick, T. D. Hether, J. M. Eastman, M. W. Pennell, and L. J. Harmon. 2012. Goldilocks meets Santa Rosalia: an ephemeral speciation model explains patterns of diversification across time scales. Evolutionary Biology 39:255–261.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge.

Rosindell, J., L. J. Harmon, and R. S. Etienne. 2015. Unifying ecology and macroevolution with individual-based theory. Ecology Letters 18:472–482.

Sanders, N. J. 2002. Elevational gradients in ant species richness: area, geometry, and Rapoport's rule. Ecography 25:25–32.

Slingerland, R., and K. P. Furlong. 1989. Geodynamic and geomorphic evolution of the Permo-Triassic Appalachian mountains. Geomorphology 2:23–37.

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

Thesing, B. D., R. D. Noyes, D. E. Starkey, and D. B. Shepard. 2016. Pleistocene climatic fluctuations explain the disjunct distribution and complex phylogeographic structure of the southern red-backed salamander, *Plethodon serratus*. Evolutionary Ecology 30:89–104.

Vasseur, D. A., and K. S. McCann. 2005. A mechanistic approach for modeling temperature-dependent consumer-resource dynamics. American Naturalist 166:184–198.

Wilder, I., and E. Dunn. 1920. The correlation of lunglessness in salamanders with a mountain brook habitat. Copeia 84:63–68.

Zachos, J., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292:686–693.

References Cited Only in the Online Appendixes

Matmon, A., P. Bierman, J. Larsen, S. Southworth, M. Pavich, and M. Caffee. 2003. Temporally and spatially uniform rates of erosion in the southern Appalachian Great Smoky Mountains. Geology 31:155–158.

NWS (National Weather Service). 2011. Temperature records for Greensboro, North Carolina. Accessed May 1, 2011.

USGS (United States Geological Survey). 2004. Geologic provinces of the United States: Appalachian highlands province. Accessed May 1, 2011.

Wiens, J. 2007. Global patterns of species richness and diversification in amphibians. American Naturalist 170:S86–S106.