

RESEARCH  
PAPERS



# Niche breadth and geographic range size as determinants of species survival on geological time scales

Erin E. Saupe<sup>1,\*</sup>, Huijie Qiao<sup>2</sup>, Jonathan R. Hendricks<sup>3,4</sup>, Roger W. Portell<sup>5</sup>, Stephen J. Hunter<sup>6</sup>, Jorge Soberón<sup>7</sup> and Bruce S. Lieberman<sup>7</sup>

<sup>1</sup>Department of Geology & Geophysics, Yale University, New Haven, CT 06511, USA, <sup>2</sup>Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China, <sup>3</sup>Department of Geology, San José State University, San José, CA 95192, USA, <sup>4</sup>Paleontological Research Institution, Ithaca, NY 14850, USA, <sup>5</sup>Division of Invertebrate Paleontology, Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA, <sup>6</sup>School of Earth and Environment, University of Leeds, West Yorkshire LS2 9JT, UK, <sup>7</sup>Biodiversity Institute and Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, KS 66045, USA

\*Correspondence: Erin E. Saupe, Department of Geology & Geophysics, Yale University, 210 Whitney Ave., New Haven, CT 06511, USA. E-mail: erin.saupe@yale.edu

## ABSTRACT

**Aim** Determining which species are more prone to extinction is vital for conserving Earth's biodiversity and for providing insight into macroevolutionary processes. This paper utilizes the Pliocene to Recent fossil record of mollusks to identify determinants of species' extinction over the past three million years of Earth history.

**Location** Western Atlantic.

**Methods** We focus on 92 bivalve and gastropod species that lived during the mid-Pliocene Warm Period (mPWP; ~3.264–3.025 Ma) and have either since gone extinct or are still extant. We used ecological niche modeling (ENM) to assess the vulnerability of these species to extinction as a function of both fundamental (FN) and realized (RN) niche breadth proxies, geographic range size, and amount of suitable area available to them during the Last Glacial Maximum (LGM; ~21 Ka).

**Results** Geographic range size emerged as a key predictor of extinction for the studied mollusk species, with RN breadth and amount of suitable area available during the LGM as secondary predictors. By contrast, FN breadth was not a significant predictor of extinction risk.

**Main conclusions** The failure to recover FN breadth as a predictor of extinction may suggest that extinction resistance is achieved when species are more successful in filling the geographic extent of their fundamental tolerances. That is, when it comes to species' survival, being a generalist or specialist *sensu stricto* may be secondary to the unique historical, dispersal, and biotic constraints that dictate a species' occupation of suitable environments, and consequently of geographic space, at a particular time. Identifying the factors that promote extinction is important because of the time-intensive nature of estimating extinction risk for individual species and populations, and because of the rising concerns about the future of marine ecosystems and biodiversity.

## Keywords

ecological niche modeling, extinction selectivity, fossils, fundamental niche, last glacial maximum, macroecology, macroevolution, niche volume, realized niche.

## INTRODUCTION

Determining which species are more prone to extinction is vital for conserving Earth's biodiversity (McKinney, 1997; Schwartz *et al.*, 2006; Lee & Jetz, 2011) and for providing insight into macroevolutionary processes over geological time scales (Kiessling & Aberhan, 2007; Payne & Finnegan, 2007; Meseguer

*et al.*, 2014). Although several traits have been identified as correlating with extinction risk (e.g. McKinney, 1997; Mace *et al.*, 2008), one of the more robust is geographic range size. Both neontological (e.g. Thomas *et al.*, 2004; Schwartz *et al.*, 2006; Harris & Pimm, 2008) and paleontological (e.g. Kiessling & Aberhan, 2007; Payne & Finnegan, 2007; Harnik *et al.*, 2012) studies have found that large geographic range sizes enhance

taxon survivorship. Different causal mechanisms, however, have been invoked to explain why large geographic ranges provide extinction resistance. Often, geographic range and environmental tolerance are assumed to co-vary, even though species with large geographic ranges can have small niche breadths if the underlying environment is homogeneous and geographically common, whereas species with broad niche breadths can have small geographic ranges if they are limited by dispersal, biotic, or other geographic barriers (Gaston, 2003; Peterson *et al.*, 2011).

An important question, then, is whether geographic range is merely a proxy for species' environmental tolerances, or whether it serves as a buffer against extinction *in and of itself*. If the latter, species with large geographic ranges would be unlikely to experience environmental perturbations across their entire distributional area. As a result, these species would survive such disturbances irrespective of their tolerances. Alternatively, environmental tolerances may reign supreme and dictate species' geographic ranges, supporting the long-held notion that specialist species (stenotopes) are more prone to extinction than generalist species (eurytopes) (e.g. McKinney, 1997; Fernandez & Vrba, 2005). Over the last several decades, numerous efforts have been made to understand the relative effects of geographic range size and niche breadths on macroevolutionary dynamics (e.g. Kammer *et al.*, 1997; Thuiller *et al.*, 2005; Harnik *et al.*, 2012; Nürnberg & Aberhan, 2013). Often, these studies have found opposing patterns with respect to the relative roles that geographic range size and habitat breadth have in determining extinction risk, with geographic range size emerging as both primary (e.g. Harnik *et al.*, 2012) and as secondary (e.g. Nürnberg & Aberhan, 2013) to habitat breadth. Missing from these studies is an explicit analysis of the different manifestations of species' tolerances, such as the environment occupied by a species *versus* the full suite of environments that a species can tolerate.

We add to this body of work by utilizing the exceptional record of Pliocene to Recent mollusks from the Western Atlantic to test the relative effects of geographic range size and niche breadths on survivorship of species from the Pliocene to the present day. These records provide a well-characterized system for analyzing the dynamics of niche breadths and geographic range size over time. Molluscan remains are abundant and particularly well studied with respect to patterns of extinction, species turnover, and ecological change across this interval (Allmon, 2001; Todd *et al.*, 2002), and both bivalves and gastropods have proven excellent study systems for analyzing diversity dynamics through time (Jablonski, 2008; Nürnberg & Aberhan, 2013). We focus on 92 bivalve and gastropod species that lived during the mid-Pliocene Warm Period (mPWP; ~3.264–3.025 Ma); 47 of these are now extinct, while 45 are extant. We assess the vulnerability of these species to extinction as a function of both fundamental (FN) and realized (RN) niche breadth proxies, geographic range size, and the amount of suitable area available to them during the Last Glacial Maximum (LGM; ~21 Ka).

The FN is considered here in Grinnellian terms (Grinnell, 1917), defined as the set of all combinations of abiotic environ-

mental conditions in which a species can survive and maintain viable populations (Peterson *et al.*, 2011). The FN has been suggested to be a species-level trait that is stable both within and across lineages (Strubbe *et al.*, 2013; Saupe *et al.*, 2014a), whereas the realized niche (RN) is a subset of the FN that can change through time for reasons relating to resource use, biotic factors such as competition, availability of suitable environments, and/or dispersal barriers (Peterson *et al.*, 2011).

Ultimately, identification of the traits that promote or inhibit extinction provides insight into the causal mechanisms generating patterns of diversity over geological time scales (McKinney, 1997; Jablonski, 2008). The fossil record provides a ledger of such evolutionary winners and losers, which can be used to generate a list of threat factors leading to species' extinctions. These rule sets are important because of the time-intensive nature of estimating extinction risk for individual species and populations (Keith *et al.*, 2008), and because of rising concerns about the future of marine ecosystems and biodiversity (Jones *et al.*, 2013; Saupe *et al.*, 2014b).

## METHODS

### Taxa

We used the *Neogene Atlas of Ancient Life* (Hendricks *et al.*, 2014) to select taxonomically vetted Pliocene species from eight gastropod families (Bursidae, Conidae, Fasciolaridae, Muricidae, Personidae, Ranellidae, Strombidae, and Tonnidae) and six bivalve families (Arcidae, Cardiidae, Carditidae, Lucinidae, Tellinidae, and Veneridae). We utilized all species in the *Neogene Atlas* with Pliocene representatives to avoid potential biases based on taxon selection. The only species that were excluded from our analyses were those with poor sampling, extreme micro-endemism, or poor model quality, detailed below. We eliminated species with very restricted distributions to account for potential artifacts that might arise from sampling bias using a two-step process: we retained species with  $\geq$  four spatially unique occurrences at 2.5 arc-minute (~4.5 km) resolution, and we excluded species with only one spatially unique occurrence at 1.25° (~140 km) resolution (i.e. the scale of the environmental data). This procedure ensured that poorly sampled species were removed from niche estimations. A total of 47 now-extinct (of 80) and 45 still-extant (of 65) species remained after removing poorly sampled species and those with non-significant niche models (see niche modeling section for details). The remaining taxa belonged to 16 gastropod and 50 bivalve genera (Tables S1 & S2).

We studied species living during the mPWP that have either since gone extinct or are still extant. It is not surprising that the extant species have fossil records that extend into the Pliocene, given that many marine invertebrates have durations greater than three million years (Eldredge *et al.*, 2005). Moreover, species that disappeared from the fossil record sometime between the Pliocene and the present almost surely went extinct during this interval, since the Western Atlantic is one of the best-sampled marine regions in the world.

## Distributional data

We gathered fossil occurrence data from Pliocene strata of the Atlantic and Gulf Coastal Plains, USA, and surrounding regions. Distributional data were incorporated only from those formations that straddled the time interval of our environmental data (~3.264–3.025 Ma), which included the Duplin, Goose Creek, Guaiguaza Clay, Hopegate, Intracoastal, Jackson Bluff, Mare, Raysor, Tamiami (Pinecrest Beds), and Yorktown formations (Tables S3 & S4). Most of these data derived from holdings at the *Florida Museum of Natural History*, with supplements from the *Academy of Natural Sciences of Drexel University*, *Paleontological Research Institution*, *Yale Peabody Museum*, and *Virginia Museum of Natural History*. For a comprehensive study on the Pliocene stratigraphy of the Atlantic Coastal Plain, see Saupe *et al.* (2014a). For each species, we subsampled occurrence data to leave one occurrence per environmental grid cell (Tables S1 & S2). This procedure ensured that localities with multiple occurrences were not unduly weighted in niche modeling analyses (Yackulic *et al.*, 2013), but it does not, however, eliminate aggregation of occurrences due to heterogeneity in survey efforts or differences in fossilization probability (Varela *et al.*, 2011).

The spatial resolution of the environmental data ( $1.25 \times 1.25^\circ$  or  $\sim 140 \times 140$  km at the equator, described below) limited the number of spatially unique occurrences available for niche modeling. We performed analyses on two datasets to test the sensitivity of observed patterns: (1) utilizing species that had  $\geq$  two spatially unique occurrences at  $1.25^\circ$  resolution, and (2) utilizing species with  $\geq$  three spatially unique occurrences at  $1.25^\circ$  resolution (Tables S1 & S2). We did not eliminate species with three or fewer spatially unique occurrences at  $1.25^\circ$  resolution, as this would remove species with small geographic range sizes and/or small niche breadths, potentially biasing our results. That is to say, species with truly small geographic range sizes were almost surely excluded from our analyses, given the large range sizes obtained for the studied species (average of 165,065 km<sup>2</sup>; maximum of 691,023 km<sup>2</sup>; minimum of 644 km<sup>2</sup>). Importantly, the number of spatially unique occurrences at  $1.25^\circ$  resolution only impacted the niche modeling analyses: all occurrence data were used to calculate geographic ranges and most RN parameters (see Tables S1 & S2). The frequency distribution of occurrences per species was right-skewed on a normal plot (Fig. S1), which is a pattern found for many other extant terrestrial and marine taxa (Gaston & He, 2002), and suggests our data may be broadly commensurate with neontological data. Additional corrections for potential sampling biases are described in the ecological niche modeling protocols.

## Environmental data

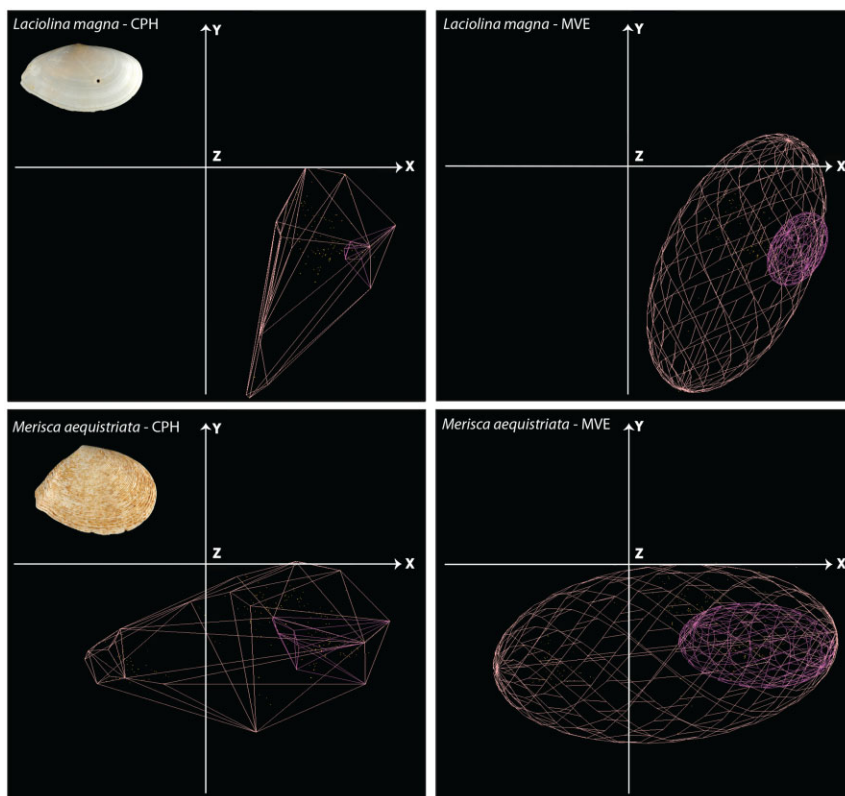
Environmental data were derived from the coupled atmosphere-ocean HadCM3 global climate model (GCM) (Gordon *et al.*, 2000) using the alternative PRISM3D PlioMIP dataset for the mPWP (~3.264–3.025 Ma) (Haywood *et al.*, 2011). The GCM was also run with LGM (~21 Ka) boundary conditions from Singarayer *et al.* (2011). Sea level changed significantly between

the mPWP and LGM, with GCM data reflecting these large-scale differences (e.g. a submerged southern portion of the Florida peninsula during the mPWP). We converted the monthly salinity and temperature outputs from the two GCM runs into maximum, minimum, and average yearly fields for both surface- and bottom-water conditions using ArcGIS 10.1 (Table S5). We centered and standardized data in each of these 12 coverages, performing a principal component analysis (PCA) on the correlation matrix using the PCARaster function in the ENMGadgets package in R (Barve & Barve, 2014). We retained the first four principal components for model calibration, which explained cumulatively  $\geq 97\%$  of the variance in the dataset. Models were calibrated using Pliocene data and projected to LGM conditions, since distributional data for species during the LGM are limited. Accordingly, the same PCA structure for the Pliocene was enforced for the LGM using the PCAProjections function in the ENMGadgets package in R (Barve & Barve, 2014). Model output was available to us only from HadCM3, and thus we were unable to utilize an ensemble-modeling approach (Diniz-Filho *et al.*, 2009).

## Ecological niche modeling

Pliocene niche models for each species were generated using Maxent v. 3.3.3, a maximum entropy algorithm that finds suitable environmental combinations for species under a null expectation that suitability is proportional to availability. Thus, Maxent minimizes the relative entropy of observed environments relative to those in the background (Elith *et al.*, 2011). Models were calibrated within a region bounded by the Americas and  $34^\circ$  W longitude, and  $48^\circ$  N and  $44^\circ$  S latitude. Following preferred methodological protocol (Phillips *et al.*, 2009; VanDerWal *et al.*, 2009; Barve *et al.*, 2011), we sought the union of the area sampled by researchers and that which was most likely accessible to the species during the Pliocene.

We enabled only the linear and quadratic features in Maxent to produce realistic response curves that match those known from physiological experiments of plants and animals (Hooper *et al.*, 2008; Angilletta, 2009). To correct for biases in fossil distributional data, we implemented what is termed a 'bias file' within Maxent (Dudík *et al.*, 2005). The bias file describes the probability that an area was sampled; thus, regions with rock outcrop (i.e. areas where species may actually be detected or sampled) were weighted twice as heavily as regions without rock outcrop. Maxent will then factor out this bias during the modeling process to account for incomplete knowledge of a species' distribution (Dudík *et al.*, 2005). Pliocene models were projected to LGM conditions to determine if species retained suitable area during glacial times. We could not directly assess distributional dynamics during the LGM because of low sea levels, which prevented occurrence records from being preserved on what is now dry land. To avoid inaccurate projections to the LGM as a result of novel environmental conditions, we deactivated clamping but allowed for extrapolation following Owens *et al.* (2013). Previous research on these and other species has suggested that large-scale parameters of species'



**Figure 1** Niche volumes in three-dimensional principal component space for representative species *Laciolina magna* and *Merisca aequistriata*. The left column depicts the minimum volume complex polyhedrons (CPH) for both thresholded mean niche models (fundamental niche proxy; light pink) and Least Training Presence niche models (realized niche proxy; dark pink). The right column depicts the minimum volume ellipsoids (MVE) for thresholded mean niche models (fundamental niche proxy; light pink) and LTP niche models (realized niche proxy; dark pink).

niches are stable on both long and short time scales (Strubble *et al.*, 2013; Saupe *et al.*, 2014a). As such, projections can – at minimum – provide estimates of the amount of suitable area available to species during the LGM. An important caveat, however, is that models built in one time period may not fare well when projected across significant periods of climatic change (Blois *et al.*, 2013). Consequently, we used these projections solely as an estimate of suitability dynamics during the LGM.

For the majority of species, we assessed model quality using a jackknife procedure from Pearson *et al.* (2007) (Tables S1 & S2). For those species with only two spatially unique occurrences, models that correctly predicted both occurrences under a mean suitability threshold (Liu *et al.*, 2013) were deemed significant. Note that both occurrences will always be correctly predicted under a Least Training Presence (LTP) threshold (Pearson *et al.*, 2007). Species were eliminated if models were not significant using either of these approaches. Although all non-significant models were removed from analyses, models built with fewer than five occurrences may also have the potential to exhibit poor performance (Pearson *et al.*, 2007).

### Niche breadth calculations

We consider the term niche in the classic Grinnellian (Grinnell, 1917) sense of abiotic environmental variables, defined by Hutchinson (1957) as a hyper-volume in multi-dimensional space, within which a species can maintain viable populations. Niche breadths were quantified in three dimensions by calculat-

ing volumes from the thresholded niche models, such that the dimensions for the three niche axes were considered simultaneously. We calculated volumes using NicheA (available at: <http://nichea.sourceforge.net>), which is an open-source, cross-platform application that combines several toolkits to analyze ecological niches in both environmental and geographic space. We quantified niche volumes from the first three principal component axes within environmental space, which explained > 89% of the variance in the dataset (Fig. 1). To evaluate the potential sensitivity of our results, we used two different methods to calculate niche volumes: minimum volume complex polyhedrons (CPH) and minimum volume ellipsoids (MVE) around suitable area in three-dimensional environmental space.

*Realized niche characterizations.* To examine whether the breadth of environment occupied by a species (i.e. a species' RN) impacts survival, we calculated RN breadth estimates in three ways. First, we calculated CPH and MVE niche breadth volumes for the area occupied by a species in environmental space. Second, we tested the sensitivity of these calculations to potential sampling biases by degrading the available fossil record to the lowest number of occurrences for all species (i.e. two) using 100 bootstrap replicates. For each bootstrap replicate, we calculated niche breadth as the environmental distance between the two occurrences on the first principal component layer, which explained ~39% of the variance in the dataset. We used the median value from the 100 replicates as the niche breadth for a given species. Niche breadths for species with only two occurrences were calculated in the same fashion, but without bootstrap replication. This process is similar to rarefaction, which is

commonly applied to paleontological data and attempts to correct for differing abundance values (Kiessling & Aberhan, 2007). Finally, estimates from continuous model suitability surfaces were converted to binary output (i.e. 1 = suitable; 0 = unsuitable) using the least training presence (LTP) method, which constricts estimates of suitability to the lowest value associated with an occurrence (Pearson *et al.*, 2007). These binary models were then used to calculate CPH and MVE niche breadth volumes.

**Fundamental niche proxies.** To obtain proxies of the FN, we calculated CPH and MVE niche breadth volumes using the mean suitability score from the Maxent model as a threshold (Liu *et al.*, 2013). On average, this method produces broader niche estimates when prevalence is low to moderate compared to the LTP method (Liu *et al.*, 2013). Such was the case for our study: the mean number of pixels predicted to be suitable across all species increased over 350% using a mean threshold (for proxies of the FN) compared to a LTP threshold (for proxies of the RN) (see Fig. S2).

We acknowledge that accurately estimating FNs is difficult even when utilizing mechanistic approaches (Kearney & Porter, 2009), and our method probably produced niche estimates closer to the existing niche or which were simply larger than the RN (Peterson *et al.*, 2011). Still, recent biophysical approaches have suggested that FNs can be represented by limited parameters such as temperature (Kearney *et al.*, 2013), and we used model parameters that match known physiological response curves of species (Hooper *et al.*, 2008; Angilletta, 2009). Consequently, estimates may at least be congruent with a species' potential suitable area.

### Geographic area calculations

The geographic area occupied by a species was calculated in two different ways: by summing the number of suitable pixels from projected niche models using both model threshold approaches, and by creating minimum convex polygons around the unfiltered (all) occurrence data for each species. To estimate the latter, we used the minimum bounding geometry function in ArcGIS 10.1 and calculated polygon areas using the USA Contiguous Albers Equal Area Conic projection. We also estimated the amount of area projected to remain suitable for these species during the LGM (~21 Ka) by tabulating the number of suitable pixels in model projections using both the LTP and mean threshold approach. Results were equivalent across these two approaches, and thus we present results only for the mean suitability method.

### Statistical tests

**Mann-Whitney *U* analyses.** Variances were similar between groups (now-extinct and still-extant species) for each parameter as determined from both a Levene and a Brown-Forsythe test. Area and volume calculations, however, were not normally distributed, and thus we used one-way nonparametric Mann-Whitney *U*-tests for analyses (Table 1). We tested four hypoth-

eses using Pliocene data for still-extant and now-extinct species: whether still-extant species (1) had greater FN breadth proxies than now-extinct species; (2) had larger geographic ranges than now-extinct species; (3) had greater RN breadth proxies than now-extinct species, such that they occupied broader realized environmental space; and (4) had more suitable area remaining during the LGM than now-extinct species. Ten Mann-Whitney *U*-tests were performed in order to test the sensitivity of our results to the different estimation methods, given that FN breadth was calculated in two ways, RN breadth in five ways, geographic area in two ways, and LGM projections in one way (see Table 1). We measured the effect size of our analyses using the methodology of Wendt (1972), following the equation  $r = 1 - (2U)/(n1 \times n2)$ .

**Binary logistic regressions.** In addition to the Mann-Whitney *U* analyses, we tested whether these parameters predict which species survive using both multivariate and univariate binary logistic regressions. Multivariate results were informed by backward stepwise regressions using the Akaike Information Criterion (AIC) in the R package MASS (see Table S6). Most variables were log-transformed (see Table 2) in order to conform to the assumption of linearity of the independent variables and the log odds.

**Hierarchical partitioning.** Multivariate approaches may be affected by multicollinearity among explanatory variables (Fig. S3). In these instances, hierarchical partitioning is useful in identifying the most likely explanatory factors in a model, while minimizing the influence of multicollinearity. The approach provides a measure of the strength of effect of each factor that is largely independent from the effects of other factors (Chevan & Sutherland, 1991; Mac Nally, 2002). Hierarchical partitioning was conducted for all estimation combinations of the four variables using logistic regression and log-likelihood as the goodness-of-fit measure. Following Mac Nally (2002), statistical significance of the independent contribution of variables was tested using a randomization routine. All analyses were implemented in the hier.part package in R (Mac Nally & Walsh, 2004).

## RESULTS

### Mann-Whitney *U* analyses

Contrary to our original hypothesis, extant species did not have significantly larger FN breadth proxies than extinct species, although values are close to alpha ( $\alpha$ ) level of 0.05 ( $P = 0.052$  and  $P = 0.066$  for CPH and MVE calculations, respectively) (Table 1). Extant species, however, did have significantly larger geographic ranges than extinct species, both when considering suitable area predicted by the model ( $P = 0.021$ ) and the area calculated from polygons ( $P = 0.001$ ). Consequently, there does not seem to be a one-to-one correspondence between FN breadth proxies predicted by the models and the amount of suitable area available geographically (Table 1).

The RN, however, does seem to be a significant predictor of extinction (Table 1). Extant species occupied significantly larger volumes of environmental space than extinct species, using esti-

**Table 1** Results from Mann–Whitney *U*-tests. Tests assessed whether extant species (1) had greater fundamental niche (FN) breadth proxies than extinct species, (2) had greater realized niche (RN) breadth proxies (that is, occupied broader realized environmental space than extinct species), (3) had larger geographic ranges than extinct species, and (4) had more suitable area remaining during the Last Glacial Maximum (LGM) than extinct species.

Analysis	Calculation method	≥ Two occurrences				≥ Three occurrences			
		U	z	P-value	r	U	z	P-value	r
FN breadth									
Mean niche model	CPH	1266	1.63	0.05	0.20	966	1.41	0.08	0.18
Mean niche model	MVE	1251	1.51	0.07	0.18	958	1.34	0.09	0.17
RN breadth									
Occurrences	CPH	1402	2.70	0.00	0.33	1022	1.94	0.03	0.25
Occurrences	MVE	1385	2.56	0.00	0.31	1010	1.83	0.03	0.24
LTP niche model	CPH	1363.5	2.40	0.01	0.29	1026.5	1.99	0.02	0.26
LTP niche model	MVE	1337.5	2.19	0.01	0.26	1002.5	1.76	0.04	0.23
Rarefaction	Distance	1302.5	1.92	0.03	0.23	1006.5	1.80	0.04	0.23
Geographic area									
Polygons analyses	Area (km <sup>2</sup> )	1441	3.00	0.00	0.36	1049	2.20	0.01	0.28
Mean model projection	Pixel count	1317.5	2.03	0.02	0.25	1005.5	1.79	0.04	0.23
LGM projection									
Amount of suitable area	Pixel count	1337	2.19	0.01	0.26	1017.5	1.90	0.03	0.25
Sample size		<i>n</i> = 92 (45 extant, 47 extinct)				<i>n</i> = 81 (43 extant, 38 extinct)			

Results are provided from tests using species with ≥ two and ≥ three spatially unique occurrences at 1.25° resolution. The effect size (*r*) is moderate across the analyses, with the exception of the smaller effect sizes for FN breadth. All significant results are still significant when correcting for multiple comparisons applying a Holm–Bonferroni correction for the lowest and second-lowest *P*-value for each estimation method. Abbreviations as follows: CPH, minimum volume complex polyhedron; LTP, least training presence threshold method; mean niche model, mean suitability threshold method; MVE, minimum volume ellipsoid; PCA, principal component analysis.

**Table 2** Univariate binary logistic regression results for each of the parameters.

Analysis	Calculation method	Units	Log transformed?	β	SE β	Wald's z-statistic	P-value	e <sup>β</sup>
FN breadth								
Mean niche model	CPH	PCA volume	Yes	−1.37	0.76	−1.81	0.07	0.25
Mean niche model	MVE	PCA volume	Yes	−1.31	0.78	−1.69	0.09	0.27
RN breadth								
Occurrences	CPH	PCA volume	Yes	−0.31	0.13	−2.45	0.01	0.73
Occurrences	MVE	PCA volume	Yes	−0.27	0.12	−2.23	0.03	0.76
LTP niche model	CPH	PCA volume	Yes	−0.91	0.39	−2.33	0.02	0.40
LTP niche model	MVE	PCA volume	Yes	−0.93	0.47	−2.00	0.05	0.39
Rarefaction	Distance	PCA distance	No	−1.01	0.51	−1.99	0.05	0.36
Geographic area								
Polygons analyses	Area	km <sup>2</sup>	Yes	−0.86	0.31	−2.82	0.00	0.42
Mean model projection	Pixel count	No.	No	−0.01	0.00	−2.22	0.03	0.99
LGM projection								
Amount of suitable area	Pixel count	No.	No	−0.02	0.01	−2.21	0.03	0.98

All analyses were conducted in R for those species with ≥ two spatially unique occurrences. The majority of the variables were log-transformed (indicated by the ‘Log transformed?’ column) in order to conform to the assumption of linearity of the independent variables and the log odds. All previously significant results remained significant if a Holm–Bonferroni correction is applied to the lowest *P*-value for each estimation method, with the exception of suitable area remaining during the Last Glacial Maximum (LGM). Abbreviations as follows: CPH, minimum volume complex polyhedron; FN, fundamental niche proxy; LTP, least training presence threshold method; mean niche model, mean suitability threshold method; MVE, minimum volume ellipsoid; PCA, principal component analysis; RN, realized niche proxy.

mates based on both the occurrence data ( $P = 0.004$  and  $P = 0.005$  for CPH and MVE calculations, respectively) and the LTP niche models ( $P = 0.008$  and  $P = 0.014$  for CPH and MVE calculations, respectively). Moreover, these results are robust to

the bootstrap resampling tests ( $P = 0.028$ ). When niche models are projected to glacial conditions (LGM; ~21 Ka), extinct species lost more suitable area than extant species ( $P = 0.015$ ) (Table 1).

We tested four hypotheses with respect to the FN, RN, geographic range size, and amount of suitable area available during the LGM. All previously-significant tests remain significant for the lowest and second-lowest  $P$ -value for each estimation method applying a Holm–Bonferroni correction (Holm, 1979). We further stress-tested our results by running all analyses using only species with  $\geq$  three spatially unique occurrences at 1.25° resolution. Again, all previously-significant results remain significant using this reduced set of taxa, although significance declines when accounting for multiple comparisons (see ‘ $\geq$  3 occurrences’ results in Table 1). In general, effect sizes for our analyses are moderate or close to moderate (Table 1), with the exception of the FN breadth proxies, which are smaller.

### Binary logistic regressions

We performed both multivariate and univariate binary logistic regressions. The univariate analyses are congruent with the patterns observed in the Mann–Whitney  $U$ -tests: all parameters influence species’ survival, with the exception of the FN proxies (Table 2). All previously-significant results remain significant when a Holm–Bonferroni correction is applied to the lowest  $P$ -value for each estimation method, with the potential exception of suitable area remaining during the LGM.

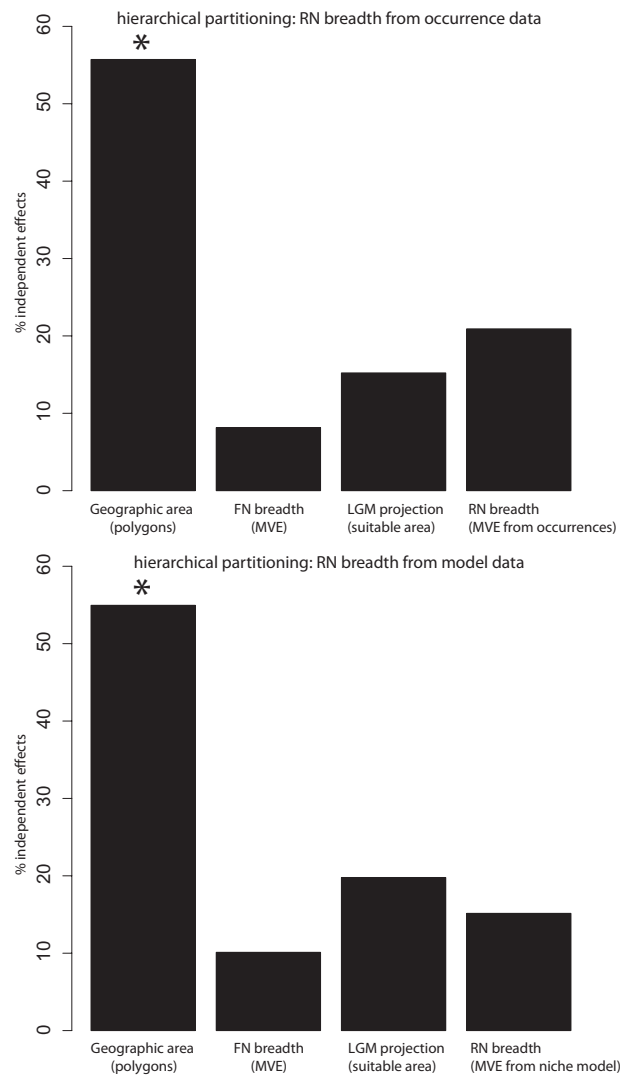
Conversely, analyses considering multiple variables are not significant (Table S7), with the smallest AIC values obtained for the model including only geographic range size (Table S6). The reduced statistical power for the multivariate analyses is likely a function of high multicollinearity among the variables, documented by the Variation Inflation Factor (VIF) in Table S7. Excessive correlation among variables can prevent identification of an optimal set of explanatory variables for a model and increase the likelihood for type II errors.

### Hierarchical partitioning

When such correlation is unpacked using hierarchical partitioning, geographic range size produces the strongest independent effect (55%), regardless of the estimation method used for the four variables (Fig. 2). Following geographic range size, the RN (21%) or the amount of suitable area remaining during the LGM (20%) produces the strongest effect, dependent on whether the RN is measured directly from the geographic imprint of a species or from the niche model results, respectively (Fig. 2). Significant independent effects are obtained only for geographic range size ( $P < 0.05$ ), irrespective of the estimation method used for the four variables. The smallest independent effects are obtained for the FN (< 10%) (Fig. 2).

### DISCUSSION

Our hypothesis that extant species have greater FN breadth proxies than extinct species was not strongly supported. Indeed, FN breadth appears to be loosely decoupled from geographic range size, with only the latter a significant predictor of extinction risk. By contrast, the RN emerged as a predictor of extinc-



**Figure 2** Hierarchical partitioning showing independent effects of niche breadth and range size variables on extinction patterns, expressed as the percentage of the total variance explained. The top panel shows variance explained when the realized niche (RN) was measured directly from the geographic imprint of a species, whereas the bottom panel shows variance explained when the RN is measured from niche model results. These two patterns hold across all other variable permutations. The asterisk denotes significance at the  $P < 0.05$  level. Abbreviations as follows: FN, fundamental niche proxy; LGM, Last Glacial Maximum; LTP, least training presence threshold method; MVE, minimum volume ellipsoid; RN, realized niche proxy.

tion in the univariate analyses, as did geographic range size and the amount of suitable area lost during the LGM. These results suggest that the environmental breadth occupied by a species (i.e. RN) may provide a greater buffer against extinction than potential tolerance limits (i.e. the FN). That is, no significant difference seems to exist in the size of FN breadth proxies between species that are still extant and those that went extinct. Rather, the ability of species to actually fill their potential

suitable area seems to provide extinction resistance, which is a function of both the dispersal ability of a species and biotic interactions.

Of course, the relative effect of these variables on extinction resistance is confounded by their lack of independence. As expected, measures of niche breadth and geographic range size strongly co-vary (Fig. S3). Ultimately, species' geographic ranges are controlled by complex interactions among the FN, dispersal over heterogeneous landscapes, and biotic interactions. Disentangling the effect of these variables is challenging (Peterson *et al.*, 2011; Harnik *et al.*, 2012; Nürnberg & Aberhan, 2013), yet hierarchical partitioning analyses suggest significant effects only for geographic range size and, not, as originally expected, for measures of niche breadth or environmental occupancy.

Taken together, broad geographic range size appears to be the primary variable that confers extinction resistance, but not solely because species with broad ranges have inherently broader environmental tolerances. Being a generalist or specialist *sensu stricto* appears secondary to the unique historical, dispersal, and biotic constraints that dictate a species' occupation of geographic space and suitable environments at a particular time. Therefore, large geographic range size seems to provide extinction resistance primarily as a consequence of buffering: any potential calamity is unlikely to impact the entirety of a species with a large geographic footprint. This argument is congruent with previous studies that find that geographic range plays a primary role in determining extinction, whereas habitat breadth plays a secondary role (e.g. Harnik *et al.*, 2012). This argument also invokes the assumption that species are often in disequilibrium with the environment (Varela *et al.*, 2009) (Fig. S2).

The results herein are congruent with previous studies that find evidence for a positive relationship between proxies of realized niche breadth (Kammer *et al.*, 1997; Nürnberg & Aberhan, 2013) and geographic range size (Kiessling & Aberhan, 2007; Harnik *et al.*, 2012) with taxon longevity and survivorship, and with those that identify population dynamics, dispersal ability, and biotic factors as important for estimating extinction risk (Araújo & Luoto, 2007; Van der Putten *et al.*, 2010; Fordham *et al.*, 2013).

### Study considerations

We focused on the impact of niche breadth and geographic range size on extinction patterns in Pliocene to Recent Western Atlantic mollusks. Other factors, however, may have also dictated extinction patterns. For instance, nutrient declines have been implicated in biotic turnover during this time (Allmon, 2001). Declining nutrient levels may have been associated with the Pliocene formation of the Central American Isthmus and the concomitant changes in oceanic circulation (Allmon, 2001). Although we found that both geographic range size and the RN are predictors of extinction risk, our results do not exclude nutrient decline as a significant factor in observed biotic turnover.

It is important to recognize that, as with any paleontological or modern ENM analysis, sampling biases may lead to incorrect geographic range and niche breadth estimates. The results, however, withstand more stringent rules for species' prevalence ( $\geq$  two and  $\geq$  three occurrences). Sampling bias may also affect the taxa included in our analysis in that species with genuinely small FNs may not be detectable in the fossil record. Exclusion of these species may explain why the FN was not recovered as a significant predictor of extinction risk. That being said, we considered species with both very small and very large geographic range sizes (644 to 691,023 km<sup>2</sup>) and niche volumes (40.6 to 560 and 14.7 to 192 for MVE and CPH calculations, respectively). In general, the frequency distribution for species' prevalence is similar to that for extant taxa (Gaston & He, 2002), suggesting our data are commensurate with neontological data sources (Fig. S1). There was an equal probability of sampling both extinct and extant taxa in Pliocene strata, with no ready explanation for why niche breadth and geographic range size estimates would be biased with respect to taxa that are now extinct or are still extant.

Data from this study derived from one time slice, the mPWP. Analyses that consider only a snapshot of a species' lifetime – including studies focused on the Recent – may be biased by age-area effects. For example, species that originated more recently may have yet to achieve full distributional extent, whereas older species may have artificially shrunken distributions if they were nearing extinction. Spurious results may be obtained if one of the two groups we considered – still-extant and now-extinct species – contained a greater proportion of younger species. No such biases appear to plague our dataset, however, since origination times did not vary dramatically between species that went extinct and those that survived. Studied species originated anywhere from the Miocene to the Pliocene, irrespective of whether they are now extinct or still extant. Furthermore, species that went extinct primarily survived past the late Pliocene and early Pleistocene, indicating their decline was not initiated during the mPWP.

### CONCLUSIONS

Our results aim to assess the factors that contribute to extinction risk by focusing on the Neogene history of mollusks from the Western Atlantic. For the taxa and time period considered, geographic range size appears to play a leading role in dictating survival. In particular, geographic range size serves as a buffer to extinction distinct from species' fundamental tolerance limits. This information may be useful as we consider the massive biodiversity losses predicted in the coming decades (Tewksbury *et al.*, 2008; Barnosky *et al.*, 2012). Macroevolutionarily, the argument that specialist species have higher extinction (and origination) rates still holds (McKinney, 1997; Nürnberg & Aberhan, 2013) but must be tempered by additional factors, such as ease of dispersal to suitable environmental areas and biotic interactions such as competition.

We could not document a highly significant impact of FN proxies on extinction probabilities. Estimation of the FN



remains difficult, and our method probably provides a lower-bound estimate on FNs. In addition, relationships between the FN, RN, and geographic range size are complex, and more information, including that pertaining to ecological parameters such as competition and dispersal abilities, may be required to produce a truly synthetic view of the factors driving such macroevolutionary patterns. Finally, results from other taxa and time periods are needed to assess whether the patterns documented here apply more broadly.

## ACKNOWLEDGEMENTS

The authors thank David Currie, Alberto Jiménez-Valverde, and three anonymous reviewers for improving the quality of this manuscript, and the following individuals for museum collection help: Lauck Ward and Alton Dooley (Virginia Museum of Natural History), Judith Nagel-Myers, Greg Dietl, David Campbell, and Warren Allmon (Paleontological Research Institution), Paul Callomon (The Academy of Natural Sciences of Drexel University), Jessica Utrup (Yale Peabody Museum), and Alex Kittle and Sean Roberts (Florida Museum of Natural History). We thank Alan Haywood (Sellwood Group for Palaeo-Climatology, University of Leeds) for assistance with the paleoclimatic data. We benefited from discussions with the KU ENM Working Group, particularly A. Townsend Peterson and Corinne E. Myers. EES was supported by a Self Graduate Fellowship, an NSF GK12 Fellowship, and a Yale Institute for Biospheric Studies Postdoctoral Fellowship. BSL was supported by NSF Grant #EF-1206757, and SJH was supported by European Research Council funding (FP7/2007–2013)/ERC grant agreement no. 278636.

## REFERENCES

- Allmon, W.D. (2001) Nutrients, temperature, disturbance, and evolution: a model for the late Cenozoic marine record of the western Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **166**, 9–26.
- Angilletta, M. (2009) *Thermal adaptation: a theoretical and empirical synthesis*. Oxford University Press, Oxford.
- Araújo, M.B. & Luoto, M. (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, **16**, 743–753.
- Barnosky, A.D., Hadly, E.A., Bascompte, J. *et al.* (2012) Approaching a state shift in Earth's biosphere. *Nature*, **486**, 52–58.
- Barve, N. & Barve, V. (2014) ENMGadgets: tools for pre and post processing in ENM workflow. *R package version 0.0.5*.
- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810–1819.
- Blois, J.L., Williams, J.W., Fitzpatrick, M.C., Jackson, S.T. & Ferrier, S. (2013) Space can substitute for time in predicting climate-change effects on biodiversity. *Proceedings of the National Academy of Sciences USA*, **110**, 9374–9379.
- Chevan, A. & Sutherland, M. (1991) Hierarchical partitioning. *The American Statistician*, **45**, 90–96.
- Diniz-Filho, J.A.F., Mauricio Bini, L., Fernando Rangel, T., Loyola, R.D., Hof, C., Nogués-Bravo, D. & Araújo, M. (2009) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography*, **32**, 897–906.
- Dudík, M., Schapire, R.E. & Phillips, S.J. (2005) Correcting sample selection bias in maximum entropy density estimation. *Advances in Neural Information Processing Systems*, pp. 323–330. The MIT Press.
- Eldredge, N., Thompson, J.N., Brakefield, P.M., Gavrilets, S., Jablonski, D., Jackson, J.B.C., Lenski, R.E., Lieberman, B.S., McPeck, M.A. & Miller, W.I. (2005) The dynamics of evolutionary stasis. *Paleobiology*, **31**, 133–145.
- 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.
- Fernandez, M.H. & Vrba, E.S. (2005) Macroevolutionary processes and biomic specialization: testing the resource-use hypothesis. *Evolutionary Ecology*, **19**, 199–219.
- Fordham, D.A., Mellin, C., Russell, B.D., Akçakaya, R.H., Bradshaw, C.J.A., Aiello-Lammens, M.E., Caley, J.M., Connell, S.D., Mayfield, S., Shepherd, S.A. & Brook, B.W. (2013) Population dynamics can be more important than physiological limits for determining range shifts under climate change. *Global Change Biology*, **19**, 3224–3237.
- Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.
- Gaston, K.J. & He, F. (2002) The distribution of species range size: a stochastic process. *Proceedings of the Royal Society B: Biological Sciences*, **269**, 1079–1086.
- Gordon, C., Cooper, C., Senior, C.A., Banks, H., Gregory, J.M., Johns, T.C., Mitchell, J.F.B. & Wood, R.A. (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, **16**, 147–168.
- Grinnell, J. (1917) The niche-relationships of the California Thrasher. *The Auk*, **34**, 427–433.
- Harnik, P.G., Simpson, C. & Payne, J.L. (2012) Long-term differences in extinction risk among the seven forms of rarity. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 4969–4976.
- Harris, G. & Pimm, S.L. (2008) Range size and extinction risk in forest birds. *Conservation Biology*, **22**, 163–171.
- Haywood, A.M., Dowsett, H.J., Robinson, M.M., Stoll, D.K., Dolan, A.M., Lunt, D.J., Otto-Bliesner, B. & Chandler, M.A. (2011) Pliocene Model Intercomparison Project (PlioMIP): experimental design and boundary conditions (Experiment 2). *Geoscientific Model Development*, **4**, 571–577.
- Hendricks, J.R., Portell, R.W., Sylva, N., Kittle, B.A., Roberts, S.W., Abdollahian, N. & Lenci, A.M. (2014) *The Neogene Atlas of Ancient Life*. Available at: <http://neogeneatlas.org/> (accessed 2014).

- Holm, S. (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**, 65–70.
- Hooper, H.L., Connon, R., Callaghan, A., Fryer, G., Yarwood-Buchanan, S., Biggs, J., Maund, S.J., Hutchinson, T.H. & Sibly, R.M. (2008) The ecological niche of *Daphnia magna* characterized using population growth rate. *Ecology*, **89**, 1015–1022.
- Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.
- Jablonski, D. (2008) Extinction and the spatial dynamics of biodiversity. *Proceedings of the National Academy of Sciences USA*, **105**, 11528–11535.
- Jones, M.C., Dye, S.R., Fernandes, J.A., Frölicher, T.L., Pinnegar, J.K., Warren, R. & Cheung, W.W.L. (2013) Predicting the impact of climate change on threatened species in UK waters. *PLoS ONE*, **8**, e54216.
- Kammer, T.W., Baumiller, T.K. & Ausich, W.I. (1997) Species longevity as a function of niche breadth: evidence from fossil crinoids. *Geology*, **25**, 219–222.
- Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334–350.
- Kearney, M.R., Simpson, S.J., Raubenheimer, D. & Kooijman, S.A.L. (2013) Balancing heat, water and nutrients under environmental change: a thermodynamic framework. *Functional Ecology*, **27**, 950–966.
- Keith, D.A., Akcakaya, H.R., Thuiller, W., Midgley, G.F., Pearson, R.G., Phillips, S.J., Regan, H.M., Araújo, M.B. & Rebelo, T.G. (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, **4**, 560–563.
- Kiessling, W. & Aberhan, M. (2007) Geographical distribution and extinction risk: lessons from Triassic–Jurassic marine benthic organisms. *Journal of Biogeography*, **34**, 1473–1489.
- Lee, T.M. & Jetz, W. (2011) Unravelling the structure of species extinction risk for predictive conservation science. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 1329–1338.
- Liu, C., White, M. & Newell, G. (2013) Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, **40**, 778–789.
- Mace, G.M., Collar, N.J., Gaston, K.J., Hilton-Taylor, C., Akcakaya, H.R., Leader-Williams, N., Milner-Gulland, E.J. & Stuart, S.N. (2008) Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology*, **22**, 1424–1444.
- McKinney, M.L. (1997) Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annual Review of Ecology and Systematics*, **28**, 495–516.
- Mac Nally, R. (2002) Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables. *Biodiversity and Conservation*, **11**, 1397–1401.
- Mac Nally, R. & Walsh, C. (2004) Hierarchical partitioning public-domain software. *Biodiversity and Conservation*, **13**, 659–660.
- Meseguer, A.S., Lobo, J.M., Ree, R., Beerling, D.J. & Sanmartin, I. (2014) Integrating fossils, phylogenies, and niche models into biogeography to reveal ancient evolutionary history: the case of *Hypericum* (Hypericaceae). *Systematic Biology*, **64**, 215–232.
- Nürnberg, S. & Aberhan, M. (2013) Habitat breadth and geographic range predict diversity dynamics in marine Mesozoic bivalves. *Paleobiology*, **39**, 360–372.
- Owens, H.L., Campbell, L.P., Dornak, L.L., Saupe, E.E., Barve, N., Soberón, J., Ingenloff, K., Lira-Noriega, A., Hensz, C.M., Myers, C.E. & Peterson, A.T. (2013) Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling*, **263**, 10–18.
- Payne, J.L. & Finnegan, S. (2007) The effect of geographic range on extinction risk during background and mass extinction. *Proceedings of the National Academy of Sciences USA*, **104**, 10506–10511.
- Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Peterson, A.T. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.
- Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Nakamura, M. & Araújo, M.B. (2011) *Ecological niches and geographic distributions*. Princeton University Press, Princeton, NJ.
- Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. & Ferrier, S. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, **19**, 181–197.
- Saupe, E.E., Hendricks, J.R., Portell, R.W., Dowsett, H.J., Haywood, A., Hunter, S.J. & Lieberman, B.S. (2014a) Macroevolutionary consequences of profound climate change on niche evolution in marine molluscs over the past three million years. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20141995. DOI: 10.1098/rspb.2014.1995.
- Saupe, E.E., Hendricks, J.R., Townsend, A.T. & Lieberman, B.S. (2014b) Climate change and marine molluscs of the western North Atlantic: future prospects and perils. *Journal of Biogeography*, **41**, 1352–1366.
- Schwartz, M.W., Iverson, L.R., Prasad, A.M., Matthews, S.N. & O'Connor, R.J. (2006) Predicting extinctions as a result of climate change. *Ecology*, **87**, 1611–1615.
- Singarayer, J.S., Valdes, P.J., Friedlingstein, P., Nelson, S. & Beerling, D.J. (2011) Late Holocene methane rise caused by orbitally controlled increase in tropical sources. *Nature*, **470**, 82–85.
- Strubble, D., Broennimann, O., Chiron, F. & Matthysen, E. (2013) Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography*, **22**, 962–970.
- Tewksbury, J.J., Huey, R.B. & Deutsch, C.A. (2008) Putting the heat on tropical animals. *Science*, **320**, 1296–1297.
- Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., de

- Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E. (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T. & Prentice, I.C. (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences USA*, **102**, 8245–8250.
- Todd, J.A., Jackson, J.B.C., Johnson, K.G., Fortunato, H.M., Heitz, A., Alvarez, M. & Jung, P. (2002) The ecology of extinction: molluscan feeding and faunal turnover in the Caribbean Neogene. *Proceedings of the Royal Society B: Biological Sciences*, **269**, 571–577.
- Van der Putten, W.H., Macel, M. & Visser, M.E. (2010) Predicting species distribution and abundance responses to climate change: why it is essential to include biotic interactions across trophic levels. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 2025–2034.
- VanDerWal, J., Shoo, L.P., Graham, C. & Williams, S.E. (2009) Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological Modelling*, **220**, 589–594.
- Varela, S., Rodríguez, J. & Lobo, J.M. (2009) Is current climatic equilibrium a guarantee for the transferability of distribution model predictions? A case study of the spotted hyena. *Journal of Biogeography*, **36**, 1645–1655.
- Varela, S., Lobo, J.M. & Hortal, J. (2011) Using species distribution models in paleobiogeography: a matter of data, predictors and concepts. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **310**, 451–463.
- Wendt, H.W. (1972) Dealing with a common problem in social science: a simplified rank-biserial coefficient of correlation based on the U statistic. *European Journal of Social Psychology*, **2**, 463–465.
- Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, J.A., Nichols, J.D., Campbell Grant, E.H. & Veran, S. (2013) Presence-only

modelling using MAXENT: when can we trust the inferences? *Methods in Ecology and Evolution*, **4**, 236–243.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

**Table S1** Studied bivalve species and associated metadata.

**Table S2** Studied gastropod species and associated metadata.

**Table S3** Formations from which distributional data were derived for studied bivalve species.

**Table S4** Formations from which distributional data were derived for studied gastropod species.

**Table S5** Environmental variables used in the PCA analysis.

**Table S6** Stepwise regression results.

**Table S7** Multivariate binary logistic regression results.

**Figure S1** Histogram of the number of occurrences per species at 1.25° resolution in the mid-Pliocene Warm Period (mPWP).

**Figure S2** Difference between fundamental niche and realized niche volumes in principal component environmental space.

**Figure S3** Spearman  $\rho$  correlation matrix for predictor variables.

## BIOSKETCH

**Erin E. Saupe** is an Institute for Biospheric Studies Postdoctoral Fellow at Yale University. She is interested in how biogeographic processes impact macroevolution. More specifically, her research explores how species respond to environmental changes on both long and short time scales, focusing on the role that species' abiotic environmental preferences play in structuring patterns of speciation, extinction and distribution through time.

Editor: Alberto Jiménez-Valverde