



Identifying the paths leading to variation in geographical range size in western North American monkeyflowers

Seema N. Sheth^{1,2*}, Iván Jiménez³ and Amy L. Angert^{1,4}

¹Department of Biology, Colorado State University, Fort Collins, CO 80523, USA,

²Graduate Degree Program in Ecology, Colorado State University, Fort Collins, CO 80523, USA, ³Center for Conservation and Sustainable Development, Missouri Botanical Garden, Saint Louis, MO 63166, USA,

⁴Departments of Botany and Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

ABSTRACT

Aim Closely related species can vary tremendously in size of geographical range, yet the causes of such variation are poorly understood. Prominent hypotheses about range size emphasize effects of niche properties and habitat connectivity via the amount and occupancy of suitable habitat, respectively. Previous studies have examined single hypotheses in isolation; however, we assessed the relative importance of these effects along with their potential interactions, using monkeyflower species (genus *Mimulus*) as a study system.

Location Western North America.

Methods We used primary occurrence data and climatic layers to estimate climatic niche breadth and position (relative to average regional climate), connectivity of climatically suitable habitat, and geographical range size of 72 monkeyflower species. Using path analysis, we then assessed the relative importance of climatic niche properties and connectivity of climatically suitable habitat in explaining variation in the amount and occupancy of climatically suitable habitat, respectively, and in turn, variation in geographical range size.

Results We documented strong support for the hypothesized effects of climatic niche breadth, but not niche position and connectivity of climatically suitable habitat. Amount of climatically suitable habitat explained more variation in range size than occupancy of climatically suitable habitat, with amount and occupancy of suitable habitat together explaining c. 83% of the variation in range size.

Main conclusions To our knowledge, this is the first study to show that climatic niche breadth, via its effects on the amount of climatically suitable habitat, is a strong predictor of geographical range size, thereby improving our understanding of the mechanisms driving species rarity.

Keywords

Extent of occurrence, geographical range size, habitat connectivity, *Mimulus*, niche breadth, niche position, North America, range occupancy, rarity.

*Correspondence: Seema N. Sheth, Department of Biology, Colorado State University, 1878 Campus Delivery, Fort Collins, CO 80523, USA.
E-mail: seema.sheth@colostate.edu

INTRODUCTION

Geographical range size can vary tremendously among species, yet we still do not fully understand the causes of such extreme variation. The present-day distributions of species are contingent upon the geography of speciation and extinction and are influenced by legacies of geological and

climatic history (Mayr, 1963). However, even closely related species with similar biogeographical and evolutionary histories can differ dramatically in range size (Darwin, 1859). Several factors have been implicated in driving variation in range size among species, including body size, species age, properties of ecological niches and species' colonization ability, among others (reviewed in Gaston, 2003), but the

strength and relative importance of each factor in explaining such variation remains uncertain. Understanding the processes that shape the distributions of species can provide important insights into ecological and evolutionary processes, such as dispersal, speciation, extinction and niche evolution, while also improving our ability to prioritize species and areas of conservation concern, forecast species' vulnerability to climate change, and predict the rate and spread of invasive species.

Properties of species' ecological niches, defined as the set of environmental conditions under which the intrinsic rate of increase is non-negative (Chase & Leibold, 2003), may explain differences in range size among species through their effects on the amount of suitable habitat, defined as the geographical area (within a study region) over which the intrinsic rate of increase is non-negative. The niche breadth hypothesis posits that species able to attain non-negative population growth rates across a broad range of environmental conditions tend to achieve larger geographical ranges because they have more suitable habitat than species with narrower niches (Fig. 1a; Hanski, 1982; Brown, 1984). Thus this hypothesis predicts a positive effect of niche breadth on range size via the amount of suitable habitat (Fig. 2a). Another hypothesis emphasizes niche position, which is the location of a species' niche relative to the central tendency of environmental conditions in a study region. Species with a low niche position occupy environmental space that is non-marginal in the sense that it is near the central tendency of environmental conditions within a study region. In contrast, species with a high niche position occupy environmental space that is marginal because it is far from the central tendency of environmental conditions within a study region (Seagle & McCracken, 1986). We emphasize that the definition of niche position adopted here (Fig. 1a) differs from that used in the resource utilization literature (Roughgarden, 1974). The niche position hypothesis proposes that species with a low niche position have more suitable habitat and thus larger range sizes than species with a high niche position (Fig. 1a; Hanski *et al.*, 1993). Hence this hypothesis predicts a negative effect of niche position on range size mediated by a negative effect of niche position on the amount of suitable habitat (Fig. 2a). Niche breadth and position may be negatively related if species with broad niches generalize on environments that are frequent across the study region, and species with narrow niches specialize on environments that are infrequent in the study region, but such a relationship need not exist (Fig. 1b). For example, a species with a narrow niche could specialize on an environment that is abundant in the study region and thus may achieve a large range. If more than one form of rarity is at play (e.g. high niche position and narrow niche breadth), then species may be doubly at risk of extinction.

Variation in species' colonization abilities may also explain differences in geographical range size. The colonization ability hypothesis (Lester *et al.*, 2007) suggests that species with

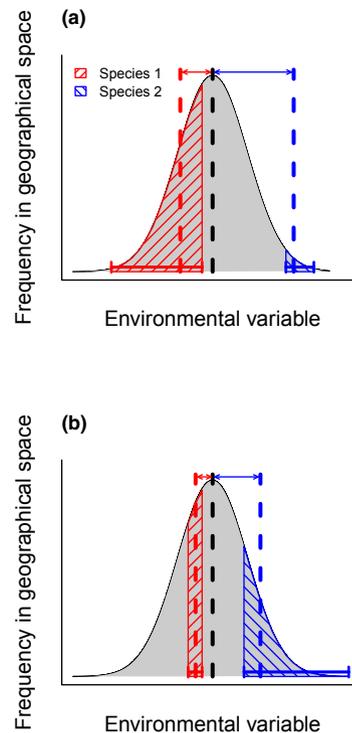
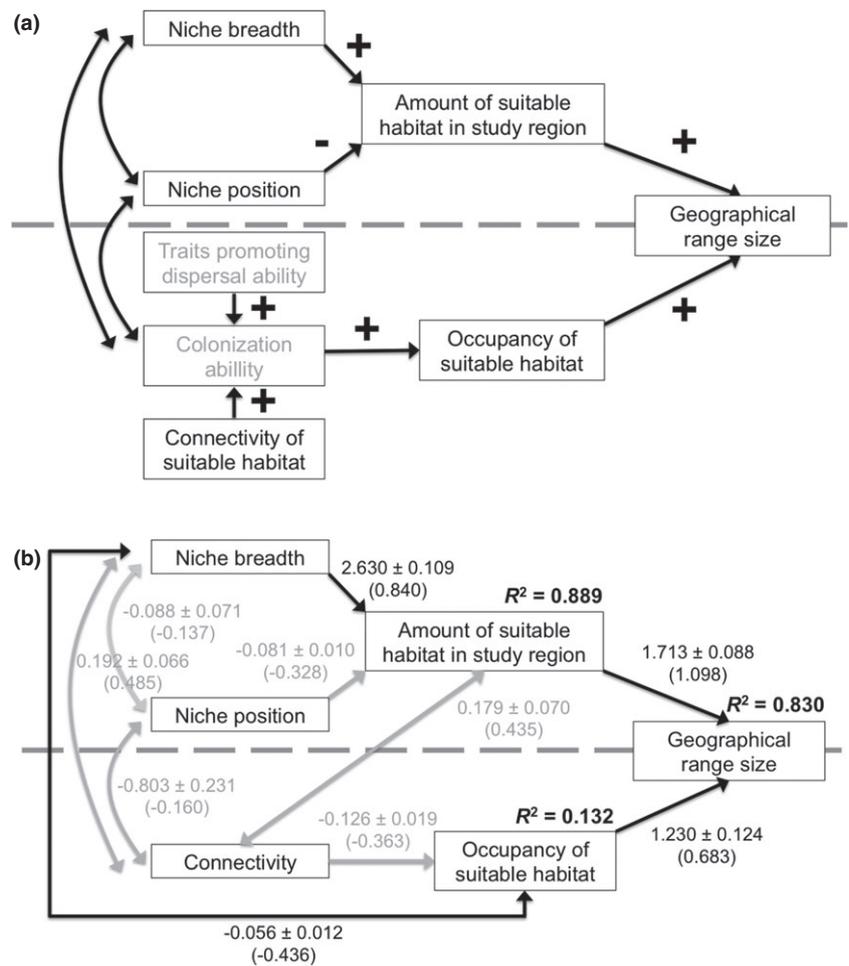


Figure 1 (a) The frequency of environments (grey area) and amount of suitable habitat for two hypothetical species (hatched areas) across a hypothetical study region. Dashed lines indicate the mean environmental values across the study region and across the niche of each species. Arrows indicate the niche position for each species. Species 1 encompasses more variation along the environmental axis (the solid horizontal line immediately above the x-axis) and has more suitable habitat (hatched area) than species 2, a pattern that is consistent with the niche breadth hypothesis. Species 1 also has a mean environmental value that is closer to the average environmental conditions across the study region than the mean environmental value of species 2, so the difference in amount of suitable habitat between species 1 and 2 is also consistent with the niche position hypothesis. The niche breadth and niche position hypotheses predict that increasing amount of suitable habitat increases geographical range size (Fig. 2a). (b) Niche breadth and position need not be negatively correlated, as a species could have a narrow niche and a low niche position (species 1), or a broad niche and a high niche position (species 2).

high colonization ability can become established in more sites and thus achieve larger ranges than species with poor colonization ability, predicting a positive relationship between colonization ability and range size across species. Colonization ability is a product of intrinsic factors, such as traits that affect mobility (e.g. morphological traits such as wing size in insects and seed size in plants) and/or establishment (life-history traits such as propagule number), and extrinsic factors, such as attributes of the landscape that facilitate movement (Lester *et al.*, 2007; Gaston, 2009). A species should have higher extrinsic colonization ability if patches of suitable habitat are well connected rather than

Figure 2 (a) Conceptual diagram of how niche properties (above the dashed line) and colonization ability (below the dashed line) are hypothesized to influence geographical range size. Single-headed arrows indicate causal relationships, double-headed arrows indicate correlations and grey labels correspond to variables that were not measured directly in our study. (b) A modified structural equation model used to assess the relative importance of climatic niche properties (above the dashed line) and connectivity of climatically suitable habitat (below the dashed line) in explaining variation in geographical range size among western North American *Mimulus* species. Unstandardized regression coefficients ± 1 standard error are shown for each arrow, with standardized coefficients in parentheses. Black arrows represent significant path coefficients and grey arrows represent non-significant path coefficients at an $\alpha = 0.05$ significance level according to the test based on the null model (see Fig. 4). R^2 -values above each endogenous variable indicate the amount of variation explained by the model. We transformed niche breadth to the fourth root, and we log-transformed the amount and occupancy of suitable habitat and geographical range size.



highly fragmented (Lester *et al.*, 2007). High connectivity of suitable habitat should facilitate movement and hence site colonization, thereby allowing species to occupy a larger fraction of available suitable habitat and achieve larger ranges (Fig. 2a). Consequently, the colonization ability hypothesis predicts a positive effect of connectivity of species' suitable habitat across the study region on geographical range size, mediated by a positive effect of connectivity of species' suitable habitat on occupancy of suitable habitat (Fig. 2a).

Studies have documented a positive relationship between niche breadth and range size across a variety of taxa and spatial scales (Slatyer *et al.*, 2013) but such a relationship may be an artefact of widely distributed species occupying a broader range of environmental conditions by chance (Gaston, 2003; Davies *et al.*, 2009). While many studies have examined the effects of niche properties or habitat connectivity on range size in isolation (e.g. Lester *et al.*, 2007), few have assessed the relative importance of each in explaining variation in range size among species (but see Hurlbert & White, 2007; Laube *et al.*, 2013) and even fewer have scrutinized the mechanistic pathways by which niche properties or habitat connectivity are hypothesized to affect range size. In

this study, we assessed the relative importance of the mechanistic pathways proposed by the hypotheses outlined above, along with their potential interactions, in a group of closely related species of western North American monkeyflower (genus *Mimulus*, renamed *Erythranthe* in Barker *et al.*, 2012), an emerging model system in ecological and evolutionary studies. We combined occurrence records with climatic variables to estimate range size, climatic niche properties and the connectivity, amount and occupancy of climatically suitable habitat for 72 *Mimulus* species. To disentangle potential artefacts resulting from geographically widespread species occupying greater climatic variation than restricted species simply by chance, we tested all the relationships against those derived from a null model that randomized the location of species' geographical ranges across the study region. To our knowledge, this study is the first to show that climatic niche breadth, via its effects on the amount of suitable habitat, is a strong predictor of geographical range size. Specifically, our results suggest that niche breadth is the best predictor of range size in western North American monkeyflowers, providing strong evidence that species with narrow climatic niches have a limited ability to achieve large ranges.

MATERIALS AND METHODS

Study system

The monkeyflower genus *Mimulus* (Phrymaceae) is a diverse group of wildflowers that occurs worldwide, with *c.* 90 of the global total of 120 species occurring in western North America (Beardsley *et al.*, 2004). *Mimulus* species occupy a wide variety of habitats, including aquatic, alpine, grassland and desert environments, can be herbaceous or woody, annual or perennial, and can exhibit complete outcrossing, obligate selfing or exclusively asexual reproduction (Wu *et al.*, 2008). Because the geographical distributions of *Mimulus* species are well described and largely encompassed within protected lands in western North America, and vary markedly in size (Beardsley *et al.*, 2004; see Fig. S1 in Appendix S1 in Supporting Information), *Mimulus* represents an ideal group for testing hypotheses regarding the variation in range size among species (Wu *et al.*, 2008). Furthermore, there is an existing phylogenetic hypothesis for *Mimulus* allowing for phylogenetically controlled studies (Beardsley *et al.*, 2004; Grossenbacher & Whittall, 2011). A recent taxonomic revision of western North American *Mimulus* (Barker *et al.*, 2012) has proposed primarily nomenclature changes but retained the major patterns of the phylogenetic hypothesis used here, and did not alter our main results (see Appendix S2).

Species occurrence data

To estimate species' geographical distributions, we compiled locality data from herbarium databases of specimen records and our own collections (see Table S1 in Appendix S3), resulting in over 17,000 georeferenced occurrences for 82 species of *Mimulus* that occur in western North America. These locality data are available from figshare (<http://dx.doi.org/10.6084/m9.figshare.1054706>). With the exception of excluding disjunct populations of *M. floribundus* in Arkansas (Nesom, 2012), we estimated the species' known global distributions. To augment the sampling of geographical regions and species for which there were few georeferenced records, we used locality descriptions from herbarium specimen labels to georeference an additional *c.* 500 herbarium specimen records. We removed records with large uncertainty in locality data (e.g. conflict between the description of the collecting locality and the geographical coordinates on specimen labels). Of the 82 species with locality data, 10 were known from fewer than three 5-arc-minute pixels (see details on spatial resolution below), precluding our ability to estimate niche properties, connectivity and range size, resulting in a final sample size of 72 species (see Fig. S2 in Appendix S1).

Climatic niche models

We modelled the climatic niche of each species to estimate its climatic niche breadth and position, and the amount and

connectivity of climatically suitable habitat. Although reducing the niche to only climatic dimensions ignores potential interspecific differences in edaphic specialization, it allows a broad-scale comparison of niche properties across a large number of species at a continental scale. We focused on climatic variables that probably affect the survival and reproduction of *Mimulus* species. Of the 19 climatic variables available from the WorldClim database (Hijmans *et al.*, 2005; <http://www.worldclim.org/>), we selected seven that encompass average and extreme conditions of temperature and precipitation: mean annual temperature, mean diurnal range, temperature annual range, mean temperature of the wettest quarter, annual precipitation, precipitation seasonality, and precipitation of the warmest quarter. These variables were not highly correlated ($r < 0.75$) among 10,000 points placed randomly across the study region, defined as a minimum convex polygon drawn around western North American *Mimulus* species occurrence points and buffered by *c.* 100 km (Fig. 3a). We used climate data at a 5-arc-minute resolution (*c.* 10 km × 10 km) and Albers equal area conic projection of North America to obtain equal-area grid cells, which are better suited for range size calculations and ecological niche models (Elith *et al.*, 2011).

To model the climatic niche of each species, we used the maximum entropy algorithm MAXENT 3.3.3k (Phillips *et al.*, 2006), a machine-learning procedure that only requires presence data and performs well compared with other methods, even for relatively small sample sizes characteristic of rare species (Elith *et al.*, 2006; Pearson *et al.*, 2007). For each species we removed duplicate records from each grid cell. We used MAXENT's default values for the 'regularization multiplier' parameter (= 1), the number of maximum iterations (= 500), the convergence threshold (= 0.00001) and feature types ('auto features'). To quantify climatically suitable habitat for each species, we converted MAXENT's output of continuous suitability values into a binary map based on a threshold of the lowest suitability value among known occurrences (lowest presence threshold; Pearson *et al.*, 2007). This threshold defines climatically suitable grid cells as those that are predicted to be at least as climatically suitable as the lowest suitability value in which a species' presence has been documented, thus eliminating the possibility of omission errors and allowing for the quantification of climatically suitable habitat (Fig. 3b). To evaluate model performance for each species known to occur in ≥ 10 pixels, we built 10 replicate models using the cross-validation approach, by which we randomly split occurrence points into 10 equal-size groups, and ran models 10 times leaving one group out in turn for testing (Elith *et al.*, 2011). For species known to occupy 3–9 pixels, we used a similar cross-validation approach but with two replicates per species instead of 10. To quantify model performance, we obtained the area under the receiver operating characteristic curve (AUC; Fielding & Bell, 1997), which reflects a model's ability to distinguish correctly presence from pseudoabsence (random background

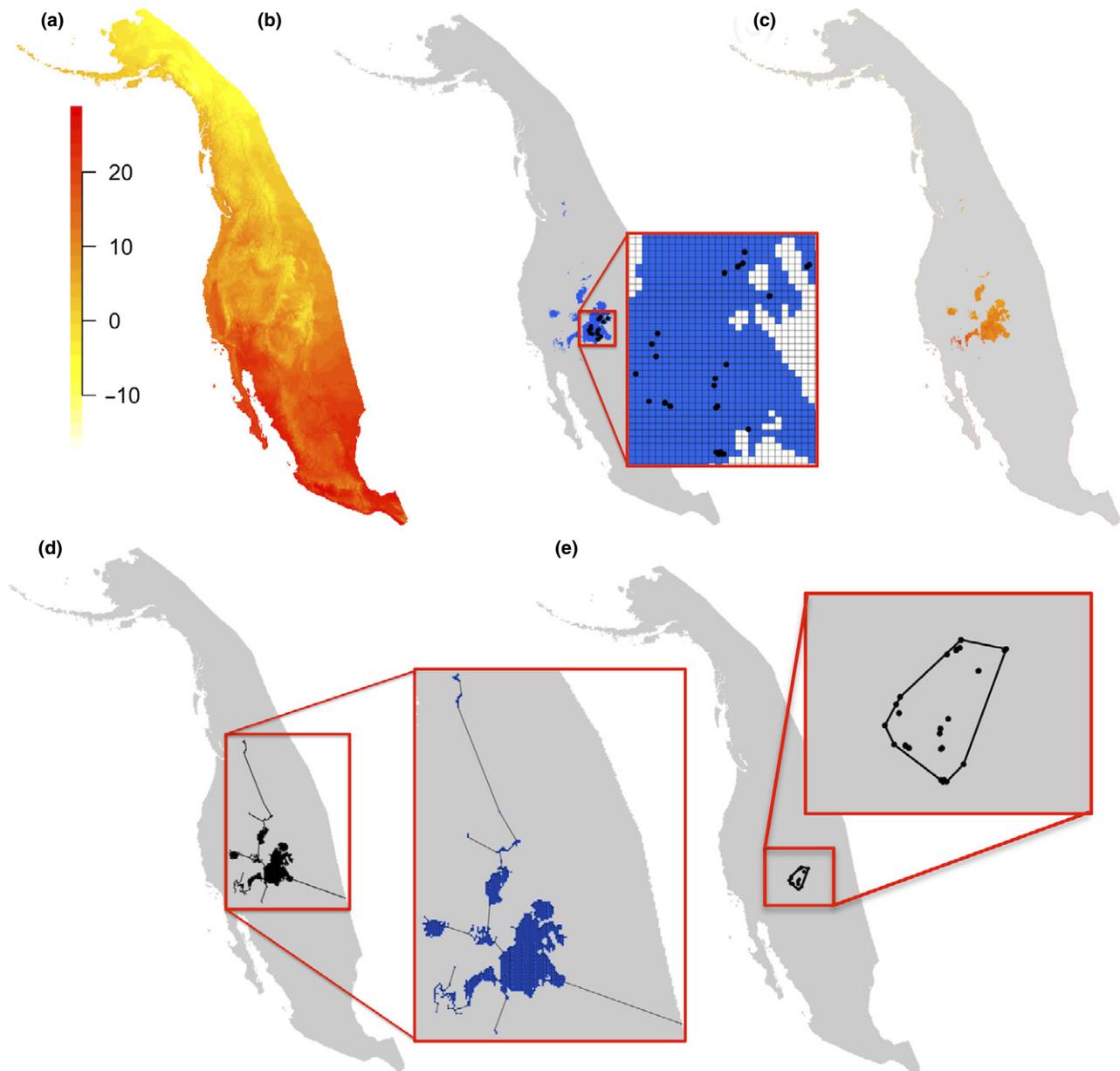


Figure 3 Illustration of how locality records and climatic data were used to estimate range size, climatic niche properties and connectivity of climatically suitable habitat for one species, *Mimulus eastwoodiae*, in western North America. (a) Mean annual temperature ($^{\circ}\text{C}$), one of seven climatic variables used to estimate niche breadth, across the study region in western North America. (b) The proportion of the total number of climatically suitable pixels (shaded in blue; see Materials and Methods for the definition of climatically suitable habitat) that are occupied based on herbarium specimen data (black points). (c) Mean annual temperature across climatically suitable pixels. Niche position is the difference between the mean temperature across the study region (a) and the mean temperature across climatically suitable pixels of a given species (c) but in multivariate climatic space. (d) Minimum spanning tree connecting climatically suitable pixels (used to estimate connectivity). (e) Extent of occurrence based on a minimum convex polygon (black outline) connecting all herbarium specimen data (black points).

points in the study region) for each of the replicated testing datasets (Phillips *et al.*, 2006). AUC ranges from 0 to 1, with AUC = 0.5 suggesting that a model's ability to discriminate presence from pseudoabsence is no better than random. Although MAXENT may have poor performance for species occupying fewer than 5–10 pixels (Wisz *et al.*, 2008), we used it to estimate suitable habitat consistently across all species. MAXENT models built from fewer than 10 occurrences

performed well for most species based on AUC (see Table S2 in Appendix S3). To assess the effects of sample size on estimates of suitable habitat, we randomly subsampled three occupied grid cells from each species 100 times before running MAXENT, and subsequently estimated niche breadth and niche position from this rarefied dataset. When we did so, estimates of niche breadth and position were very similar to those based on all occurrences (niche breadth Pearson's

$r = 0.822$, $P < 0.01$; niche position Spearman's $\rho = 0.817$, $P < 0.01$; see Fig. S3 in Appendix S1), suggesting that MAXENT models of suitable habitat with sample sizes as small as 3 pixels perform reasonably well. Furthermore, results based on analyses excluding species occupying fewer than 10 grid cells were qualitatively similar to those based on all species (see Figs S4 & S5 in Appendix S1).

Climatic niche properties

For each species, we quantified niche breadth as the sum of the variances of standardized climatic variables across climatically suitable grid cells, and niche position as the squared difference between the centroid of the multivariate climatic space encompassed by the entire study region and that of climatically suitable pixels (Fig. 3c; Martin *et al.*, 2008), using the ADEHABITAT 1.8.12 package (Calenge, 2006) in R 3.0.2 (R Core Team, 2013).

Connectivity of climatically suitable habitat

We measured connectivity among climatically suitable pixels by first creating a minimum spanning tree connecting all pixels of climatically suitable habitat for each species in the NNCLUST 2.2 package (Fig. 3d; Lumley, 2010) in R. We then estimated the mean length (in kilometres) of the edges (segments) of the minimum spanning tree, such that shorter distances represent higher connectivity among climatically suitable pixels. We multiplied distances by -1 so that more negative values represented lower connectivity while less negative values represented higher connectivity. We chose this measure of habitat connectivity because it is computationally feasible and need not be influenced by the amount of climatically suitable habitat.

Amount and occupancy of climatically suitable habitat

We estimated the amount of climatically suitable habitat in the study region as the number of climatically suitable pixels for each species (Fig. 3b) using the RASTER 2.1–66 package (Hijmans, 2013) in R. We divided the number of climatically suitable pixels in which each species is known to occur (based on point occurrence data) by the amount of climatically suitable habitat to obtain an estimate of occupancy of suitable habitat (Fig. 3b). This method may underestimate occupancy, but assuming that a species occupies every pixel of climatically suitable habitat within its extent of occurrence (described below) would overestimate occupancy; the true occupancy probably lies somewhere between these two extremes.

Geographical range size

We used three metrics of range size that quantify the overall geographical spread of each species. First, we estimated

global geographical range size as the extent of occurrence, which measures the spatial extent of the areas occupied by a species (Gaston, 1994). We estimated the extent of occurrence by computing the area of a minimum convex polygon in km^2 encompassing the known occurrences of each species (Fig. 3e). Our second and third metrics were the latitudinal and longitudinal extents encompassed by the occurrence points of each species. The three estimates were highly correlated (see Tables S3 & S4 in Appendix S3) and yielded qualitatively similar results, so for simplicity we only present the results for range size estimated as the extent of occurrence. The estimates of geographical range size, climatic niche breadth and position, and connectivity, amount, and occupancy of climatically suitable habitat for each *Mimulus* species are available from figshare (<http://dx.doi.org/10.6084/m9.figshare.1054705>).

Controlling for phylogenetic non-independence

Because *Mimulus* species share a recent evolutionary history and thus may not be statistically independent, we tested the assumption of phylogenetic independence for all explanatory and response variables to determine whether phylogenetically based comparative analyses were needed (Abouheif, 1999). We used the phylogeny published in Grossenbacher & Whittall (2011), which is a Bayesian analysis of nuclear ribosomal ITS and ETS and chloroplast *trnL-F* regions from Beardsley *et al.* (2004), concatenated with chloroplast *rpl16* sequences for the *M. moschatus* alliance (Whittall *et al.*, 2006). Of the 72 *Mimulus* species in our analyses, 68 were sampled in the Grossenbacher & Whittall (2011) phylogeny. We then tested the phylogenetic signal using four widely used methods (Appendix S2). Because tests of phylogenetic signal may fail to detect phylogenetic non-independence, we also performed simple linear regressions on phylogenetically independent contrasts (Felsenstein, 1985) to test for significant relationships between each pair of explanatory and response variables shown in Fig. 2b (Appendix S2). We then compared the results based on contrasts with those based on raw species data.

Path analysis

To evaluate the relative importance of the niche breadth, niche position and colonization ability hypotheses in explaining the variation in range size among species, we created a structural equation model describing a simplified version of Fig. 2a in which we excluded intrinsic dispersal ability and colonization ability and examined the effects of connectivity of climatically suitable habitat on occupancy of climatically suitable habitat (Fig. 2b). Explanatory and response variables were transformed to meet normality assumptions and improve model fit (Fig. 2b). We used the LAVAAN 0.5–15 package (Rosseel, 2012) in R to obtain path coefficients and assessed the significance at $P < 0.05$ for each path in the simplified version of Fig. 2a. Because our data did not meet

the assumption of multivariate normality (multivariate Shapiro–Wilk’s test, $P < 0.001$, obtained in `MVNORMTEST` 0.1–9 package in R; Jarek, 2012), we used the maximum likelihood to estimate model parameters with robust standard errors, and we used a Satorra–Bentler scaled chi-square test statistic to determine whether the covariance matrix observed in our data significantly deviated from that predicted by the structural equation model (Grace, 2006). We ran analyses in a number of different ways, including with and without outliers or taxa undergoing major revision, bootstrapping versus Satorra–Bentler scaled χ^2 and several transformations of variables, and in all cases the observed covariance matrix in our data differed significantly from the model predictions (Satorra–Bentler corrected $\chi^2 = 33.811$, d.f. = 7, $P < 0.01$). Thus we took an exploratory approach by inspecting modification indices to assess which biologically plausible paths needed to be added to achieve an adequate model fit (see Appendix S2; Grace, 2006). With this process, we arrived at a modified model including correlations between the amount of suitable habitat and connectivity, and between niche breadth and occupancy of suitable habitat (Fig. 2b, Appendix S2). The observed covariance matrix from our dataset did not deviate significantly from the modified model (Satorra–Bentler corrected $\chi^2 = 6.883$, d.f. = 5, $P = 0.229$; Fig. 2b). We used estimates of standardized path coefficients and R^2 for each endogenous variable from this resulting modified model to assess the relative importance of each hypothesis.

Null model

Because the geographical ranges of widespread species may encompass greater climatic variation than geographically restricted species simply by chance, observed relationships between range size, niche properties, connectivity and the amount and occupancy of suitable habitat may be artefacts (Gaston, 2003; Davies *et al.*, 2009). To address this issue, we used a null model that randomized the location of species’ geographical ranges across the study region, while preserving the spatial structure of the occurrence data (see details in Appendix S2). We used this null model to create 100 datasets, each including all the variables in the modified structural equation model (Fig. 2b). We then fitted this modified structural equation model to each null dataset. We estimated the 95% confidence interval for the Satorra–Bentler corrected χ^2 and each of the path coefficients derived from the null datasets by calculating percentiles of the distribution of path coefficients. We used one-tailed 95% confidence intervals for the Satorra–Bentler corrected χ^2 because the observed data should have a better model fit than the null datasets, and we used two-tailed 95% confidence intervals for all path coefficients. If a path coefficient from the observed data fell outside the respective confidence interval of the distribution of path coefficients derived from null datasets, we concluded that the observed path coefficient was significantly different from the null model expectation.

RESULTS

Performance of climatic niche models

Cross-validation AUC for test data indicated that MAXENT models performed better than random models for all species, with most species having a mean test AUC > 0.9 across cross-validation replicates (Table S2 in Appendix S3). *Mimus calciphilus* had a mean test AUC < 0.75 , probably because it was only documented in 3 pixels and only one or two occurrences were used to train or test replicate models.

Phylogenetic non-independence

We did not detect a significant phylogenetic signal in niche breadth, niche position, connectivity, amount of suitable habitat or extent of occurrence ($P > 0.05$), with a white noise non-phylogenetic model of evolution having the lowest sample size-corrected Akaike information criterion (AIC_c; Table 1). We detected a weak phylogenetic signal in occupancy of suitable habitat when testing for a phylogenetic signal based on the Abouheif’s test (Abouheif, 1999), with an Ornstein–Uhlenbeck model (Hansen, 1997) having a slightly lower AIC_c than the white noise model of evolution (Table 1). All correlations and simple regressions performed on raw species data were qualitatively similar to those performed on phylogenetically independent contrasts (see Tables S3–S6 in Appendix S3 and Figs S5 & S6 in Appendix S1).

Path analyses and null model

The observed structural equation model test statistic fell outside the 95% confidence interval of the distribution of test statistics derived from null datasets (Fig. 4a), representing a better model fit than the null model expectation. Together, niche breadth and niche position explained nearly 89% of the variation in amount of climatically suitable habitat, with niche breadth explaining more variation than niche position (niche breadth \rightarrow amount of suitable habitat standardized path coefficient = 0.840, versus niche position \rightarrow amount of suitable habitat standardized path coefficient = -0.328 ; Fig. 2b). As predicted, niche breadth had a positive effect and niche position had a negative effect on the amount of suitable habitat (Fig. 2b). While the observed path coefficient from niche breadth to amount of suitable habitat was greater than expected from the null model (Fig. 4b), the observed path coefficient from niche position to amount of suitable habitat was not (Fig. 4c). Contrary to prediction, connectivity had a negative effect on occupancy of suitable habitat, explaining 13% of the variation in occupancy of suitable habitat (Fig. 2b), but the observed path coefficient was not distinguishable from the null model expectation (Fig. 4d).

As predicted, the amount and occupancy of suitable habitat had a positive effect on range size, together explaining 83% of the variation in range size, with the amount of

Table 1 Tests for significant phylogenetic signal in the variables included in the structural equation model (Fig. 2b) for western North American *Mimulus* species.

Variable	Abouheif's C_{mean}	Pagel's λ	Blomberg's K	BM AIC _c	OU AIC _c	WN AIC _c
Niche breadth ^{0.25}	-0.039	0.077	0.039	90.87	-9.60	-11.78
Niche position	0.037	6.641×10^{-5}	0.085	373.52	326.68	325.85
Connectivity	-0.079	6.641×10^{-5}	0.036	378.92	275.36	273.26
Log ₁₀ (amount of suitable habitat)	-0.072	0.033	0.039	247.15	146.73	144.54
Log ₁₀ (occupancy of suitable habitat)	0.171*	6.641×10^{-5}	0.046	172.07	120.15	120.57
Log ₁₀ (geographical range size)	-0.166	6.641×10^{-5}	0.034	317.42	209.58	207.39

*Significant phylogenetic signal at $P < 0.05$.

AIC_c, sample size-corrected Akaike information criterion for Brownian motion (BM), Ornstein-Uhlenbeck (OU) and white noise (WN) models of evolution, with the lowest AIC shown in bold. See Appendix S2 for detailed methods.

suitable habitat explaining more variation than occupancy of suitable habitat (amount of suitable habitat → geographical range size standardized path coefficient = 1.098, versus occupancy of suitable habitat → geographical range size standardized path coefficient = 0.683; Fig. 2b). Furthermore, the path coefficients from both amount and occupancy of suitable habitat to geographical range size were greater than expected from the null model (Fig. 4e,f). Consistent with null model expectations, niche breadth and niche position were not correlated (Figs 2b & 4g). In addition to the predicted relationships, we also detected a positive relationship between niche breadth and connectivity, a negative relationship between niche position and connectivity, and a positive relationship between amount of suitable habitat and connectivity (Fig. 2b), but none of these path coefficients deviated from null model expectations (Fig. 4h-j). We also documented an unexpected negative relationship between niche breadth and occupancy of suitable habitat; this path coefficient was more negative than expected from the null model (Fig. 4k).

DISCUSSION

Despite conspicuous variation in geographical range size among species, few studies have shed light on the relative importance of the multiple mechanisms that may drive such variation. We used western North American monkeyflowers to assess the relative importance of climatic niche properties and connectivity of climatically suitable habitat in determining variation in range size among species. Niche breadth and position explained more than half of the variation in amount of climatically suitable habitat, with niche breadth having a greater effect than niche position. Moreover, the effect of niche breadth on amount of suitable habitat was greater than expected from the null model that randomized geographical ranges across the study region, while the effect of niche position on amount of climatically suitable habitat failed to differ from the null model expectation. The amount of climatically suitable habitat, in turn, explained much of the variation in range size, whereas the occupancy of climatically suitable habitat explained a smaller portion of the variation in range size. These effects of amount and occupancy of climatically

suitable habitat on range size were larger than null model expectations. Our metric of connectivity did not have a positive effect on occupancy of suitable habitat, thereby failing to support one prediction of the colonization ability hypothesis. Finally, the covariance structure of the data yielded an unforeseen relationship between niche breadth and occupancy of suitable habitat, suggesting that niche breadth and occupancy interact to shape the geographical range sizes of western North American monkeyflower species. Below we assess the evidence relevant to each of the mechanistic hypotheses we examined, and draw conclusions about the relative roles of climatic niche breadth and position, connectivity of climatically suitable habitat, and amount and occupancy of climatically suitable habitat, in explaining variation in range size among species.

Climatic niche properties

We found that climatic niche breadth was the strongest predictor of range size in western North American monkeyflowers. Although many studies have detected a positive relationship between niche breadth (or climatic tolerance) and range size (e.g. Pither, 2003), few have corrected for possible effects of range size on estimates of niche breadth, and the extent to which this potential bias has confounded the results of many studies remains unclear. Our study adds to the growing set of results showing that a positive relationship between niche breadth and range size is not artefactual (Fig. 4b; reviewed in Slatyer *et al.*, 2013). Our results contrast with studies finding that niche position is a better predictor of occupancy than niche breadth (Heino & Soininen, 2006; Hurlbert & White, 2007). We found that the relationship between niche position and range size can be explained by a null model that randomly places geographical ranges across the study region (Fig. 4c), in contrast with other findings documenting support for the niche position hypothesis (Gregory & Gaston, 2000; Heino, 2005; Heino & Soininen, 2006; Hurlbert & White, 2007). However, the way we delineated the study region prevented geographically widespread species from having a high niche position. This geometric constraint of species with large ranges might have been alleviated if we had used a larger study region. It may be easier

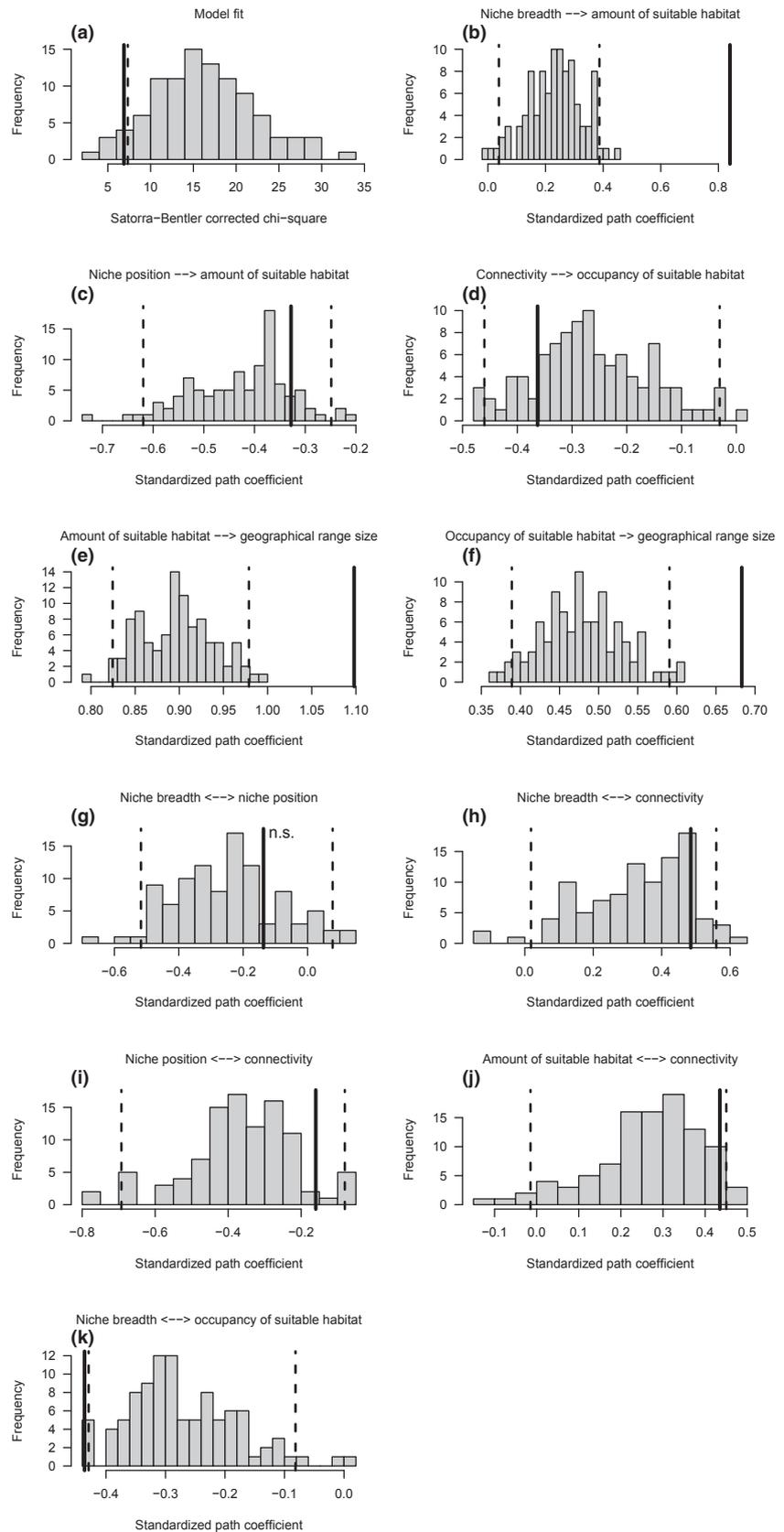


Figure 4 Frequency distributions of (a) Satorra–Bentler corrected chi-square and (b–k) path coefficients derived from 100 null datasets (see Materials and Methods and Appendix S2 for details), with dashed lines representing the respective 95% confidence intervals and the solid lines representing the test statistic or path coefficients derived from observed data for western North American *Mimulus* species. The observed path coefficient denoted as n.s. in (g) was not significantly different from 0 ($P = 0.212$).

to detect an effect of niche position that is not explained by a null model that preserves range size for clades containing mostly allopatric species, because the reference study region

would be quite large. Unlike previous tests of niche breadth and position hypotheses, we have shown that the predicted effects of niche breadth and position on range size are mediated

by the amount of suitable habitat (Fig. 2b), thereby improving our understanding of the mechanisms underlying patterns of variation in range size. Niche breadth and position were not strongly correlated in our final path analysis (Fig. 2b), suggesting that they constitute two rather independent axes of rarity. Nonetheless, our results suggest that Rabinowitz's (1981) form of rarity, in which a species has both a small geographical range and a narrow niche, may be more common than other forms.

Connectivity of climatically suitable habitat

Although occupancy of suitable habitat explained some variation in range size among species, our metric of connectivity was not a good predictor of occupancy of suitable habitat. Bivariate results suggested that the negative effect of connectivity on occupancy of suitable habitat might have been driven by outlier species with high occupancy and low connectivity (see Fig. S5f in Appendix S1). Although we focused on the connectivity of climatically suitable habitat as measured by the average edge length of minimum spanning trees, there are numerous ways to estimate such connectivity (Fortin & Dale, 2005). Furthermore, while we focused on extrinsic climatic factors that affect colonization ability, the diversity of life-history strategies, mating systems, edaphic specializations and habits encompassed by *Mimulus* species (Wu *et al.*, 2008) may outweigh the effects of the connectivity of climatically suitable habitat on the overall colonization ability of species. In the future, it would be beneficial to obtain information on traits associated with species' intrinsic colonization abilities, such as selfing rates (Randle *et al.*, 2009) and seed size (Morin & Chuine, 2006), particularly in light of several examples of dispersal ability being a better predictor of geographical range size than niche properties (Bohning-Gaese *et al.*, 2006; Kristiansen *et al.*, 2009; Blanchet *et al.*, 2013; Laube *et al.*, 2013). There are, however, many circumstances under which dispersal ability need not correlate with geographical range size. For instance, climatically suitable habitat may not be highly fragmented for most *Mimulus* species, as indicated by the small range of connectivity values (see Fig. S5f,h in Appendix S1). Given that occupancy of climatically suitable habitat had a positive effect on range size despite the lack of support for a positive relationship between connectivity and occupancy of climatically suitable habitat, occasional long-distance dispersal events may be more important than connectivity in determining the occupancy of suitable habitat (Lester *et al.*, 2007). However, we emphasize the need to interpret the occupancy results with caution because, without absence data, estimates of occupancy may suffer from collection biases that could potentially result in reshuffling of the relative ranks of estimated occupancy relative to true occupancy (Sheth *et al.*, 2012).

Other determinants of geographical range size

Although we found strong support for the niche breadth hypothesis, our study did not include other potential determinants

of geographical range size and occupancy. In our study, we focused on the climatic niche, but other niche axes, such as edaphic properties, could also influence species distributions. Some studies have documented a positive relationship between species age and range size (e.g. Jablonski, 1987; Webb & Gaston, 2000; Paul *et al.*, 2009), suggesting that over time species are able to fill more of their available niche space and/or adapt and expand into novel niche space, thus achieving broader realized niches and larger range sizes. If species age were driving variation in range size among *Mimulus* species, then the positive relationship we detected between niche breadth and range size could be a result of younger species having narrower niches and thus smaller ranges than older species. Furthermore, the speciation rate within a particular clade could influence the range sizes of species in that clade, such that clades with higher speciation rates may tend to have more species with smaller ranges than clades with low speciation rates (Lester & Ruttenger, 2005). However, if variation in speciation rates among clades within *Mimulus* were driving the variation in range size, then we should have detected a phylogenetic signal in range size as a result of certain clades with high speciation rates having species with small geographical ranges. Recent work suggests that the patterns of niche breadth and range size in western North American *Mimulus* support a budding mode of speciation (Grossenbacher *et al.*, 2014), which may explain the lack of phylogenetic signal in range size.

CONCLUSIONS

In this study we have shown that climatic niche breadth explained more variation in geographical range size among *Mimulus* species than niche position and connectivity of climatically suitable habitat. The results of our study contribute to disentangling the mechanisms underlying patterns of variation in range size among species by providing empirical support for the idea that climatic niche breadth, via its effect on the amount of suitable habitat, drives variation in range size in western North American monkeyflowers, despite other differences among species (e.g. edaphic substrate and mating system) and other causes of range-size variation (e.g. evolutionary and biogeographical history). To understand further the mechanisms underlying the niche breadth hypothesis, experiments assessing whether wide-ranging species have broader niches than narrowly distributed species would be useful. For example, experiments in environmental chambers would allow tests of whether *Mimulus* species with large geographical ranges have broader thermal performance breadths than narrowly distributed relatives. In sum, we have shown that climatic niche breadth influences species' rarity, and thus may constitute a major axis of extinction risk. Consequently, by improving our understanding of the processes driving patterns of rarity, this study increases our ability to assess species' vulnerabilities to extinction.

ACKNOWLEDGEMENTS

We thank M. Wiebush for help with georeferencing; W.C. Funk, R.L. Mueller, J.R. Paul, D. Steingraeber, A. Algar, S. Blanchet and four anonymous referees for comments on earlier versions of this work; K. Hoke for statistical guidance; J. Sobel and N. Fraga for taxonomic insights; D. Grossenbacher and J. Whittall for sharing phylogenetic tree; and R. Moe and the natural history museums and herbaria that provided locality data. This work was supported by National Science Foundation DEB 0950171.

REFERENCES

- Abouheif, E. (1999) A method for testing the assumption of phylogenetic independence in comparative data. *Evolutionary Ecology Research*, **1**, 895–909.
- Barker, W., Nesom, G., Beardsley, P.M. & Fraga, N.S. (2012) A taxonomic conspectus of Phrymaceae: a narrowed circumscription for *Mimulus*, new and resurrected genera, and new names and combinations. *Phytoneuron*, **39**, 1–60.
- Beardsley, P.M., Schoenig, S.E., Whittall, J.B. & Olmstead, R.G. (2004) Patterns of evolution in western North American *Mimulus* (Phrymaceae). *American Journal of Botany*, **91**, 474–489.
- Blanchet, S., Reyjol, Y., April, J., Mandrak, N.E., Rodríguez, M.A., Bernatchez, L. & Magnan, P. (2013) Phenotypic and phylogenetic correlates of geographic range size in Canadian freshwater fishes. *Global Ecology and Biogeography*, **22**, 1083–1094.
- Bohning-Gaese, K., Caprano, T., van Ewijk, K. & Veith, M. (2006) Range size: disentangling current traits and phylogenetic and biogeographic factors. *The American Naturalist*, **167**, 555–567.
- Brown, J.H. (1984) On the relationship between abundance and distribution of species. *The American Naturalist*, **124**, 255–279.
- Calenge, C. (2006) The package “adehabitat” for the R software: a tool for the analysis of space and habitat use by animals. *Ecological Modelling*, **197**, 516–519.
- Chase, J.M. & Leibold, M.A. (2003) *Ecological niches: linking classical and contemporary approaches*. University of Chicago Press, Chicago, IL.
- Darwin, C. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. J. Murray, London.
- Davies, T.J., Purvis, A. & Gittleman, J.L. (2009) Quaternary climate change and the geographic ranges of mammals. *The American Naturalist*, **174**, 297–307.
- Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species’ distributions from occurrence data. *Ecography*, **29**, 129–151.
- Elith, J., 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.
- Felsenstein, J. (1985) Phylogenies and the comparative method. *The American Naturalist*, **125**, 1–15.
- Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.
- Fortin, M.J. & Dale, M.R.T. (2005) *Spatial analysis: a guide for ecologists*. Cambridge University Press, Cambridge, UK.
- Gaston, K.J. (1994) Measuring geographic range sizes. *Ecography*, **17**, 198–205.
- Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.
- Gaston, K.J. (2009) Geographic range limits: achieving synthesis. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1395–1406.
- Grace, J.B. (2006) *Structural equation modeling and natural systems*. Cambridge University Press, Cambridge, UK.
- Gregory, R.D. & Gaston, K.J. (2000) Explanations of commonness and rarity in British breeding birds: separating resource use and resource availability. *Oikos*, **88**, 515–526.
- Grossenbacher, D.L. & Whittall, J.B. (2011) Increased floral divergence in sympatric monkeyflowers. *Evolution*, **65**, 2712–2718.
- Grossenbacher, D.L., Veloz, S.D. & Sexton, J.P. (2014) Niche and range size patterns suggest that speciation begins in small, ecologically diverged populations in North American monkeyflowers (*Mimulus* spp.). *Evolution*, **68**, 1270–1280.
- Hansen, T.F. (1997) Stabilizing selection and the comparative analysis of adaptation. *Evolution*, **51**, 1341–1351.
- Hanski, I. (1982) Dynamics of regional distribution: the core and satellite species hypothesis. *Oikos*, **38**, 210–221.
- Hanski, I., Kouki, J. & Halkka, A. (1993) Three explanations of the positive relationship between distribution and abundance of species. *Species diversity in ecological communities: historical and geographical perspectives* (ed. by R. Ricklefs and D. Schluter), pp. 108–116. University of Chicago Press, Chicago, IL.
- Heino, J. (2005) Positive relationship between regional distribution and local abundance in stream insects: a consequence of niche breadth or niche position? *Ecography*, **28**, 345–354.
- Heino, J. & Soininen, J. (2006) Regional occupancy in unicellular eukaryotes: a reflection of niche breadth, habitat availability or size-related dispersal capacity? *Freshwater Biology*, **51**, 672–685.
- Hijmans, R.J. (2013) *Raster: geographic analysis and modeling*. R package version 2.1–66. Available at: <http://cran.r-project.org/web/packages/raster>.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Hurlbert, A.H. & White, E.P. (2007) Ecological correlates of geographical range occupancy in North American birds. *Global Ecology and Biogeography*, **16**, 764–773.
- Jablonski, D. (1987) Heritability at the species level: analysis of geographic ranges of Cretaceous mollusks. *Science*, **238**, 360–363.

- Jarek, S. (2012) *Package 'mvnrmtest'*. R package version 0.1–9. Available at: <http://cran.r-project.org/web/packages/mvnrmtest/index.html>.
- Kristiansen, T., Svenning, J.-C., Grández, C., Salo, J. & Balslev, H. (2009) Commonness of Amazonian palm (Arecaceae) species: cross-scale links and potential determinants. *Acta Oecologica*, **35**, 554–562.
- Laube, I., Korntheuer, H., Schwager, M., Trautmann, S., Rahbek, C. & Böhning-Gaese, K. (2013) Towards a more mechanistic understanding of traits and range sizes. *Global Ecology and Biogeography*, **22**, 233–241.
- Lester, S.E. & Ruttenberg, B.I. (2005) The relationship between pelagic larval duration and range size in tropical reef fishes: a synthetic analysis. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 585–591.
- Lester, S.E., Ruttenberg, B.I., Gaines, S.D. & Kinlan, B.P. (2007) The relationship between dispersal ability and geographic range size. *Ecology Letters*, **10**, 745–758.
- Lumley, T. (2010) *nnclust: nearest-neighbour tools for clustering*. R package version 2.2. Available at: <http://cran.r-project.org/web/packages/nnclust/index.html>.
- Martin, J., Calenge, C., Quenette, P.-Y. & Allainé, D. (2008) Importance of movement constraints in habitat selection studies. *Ecological Modelling*, **213**, 257–262.
- Mayr, E. (1963) *Animal species and evolution*. Belknap Press, Cambridge, MA.
- Morin, X. & Chuine, I. (2006) Niche breadth, competitive strength and range size of tree species: a trade-off based framework to understand species distribution. *Ecology Letters*, **9**, 185–195.
- Nesom, G.L. (2012) Taxonomy of *Erythranthe* sect. *Mimulosma* (Phrymaceae). *Phytoneuron*, **41**, 1–36.
- Paul, J.R., Morton, C., Taylor, C.M. & Tonsor, S.J. (2009) Evolutionary time for dispersal limits the extent, but not the occupancy of species' potential ranges in the tropical plant genus *Psychotria* (Rubiaceae). *The American Naturalist*, **173**, 188–199.
- 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.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Pither, J. (2003) Climate tolerance and interspecific variation in geographic range size. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 475–481.
- R Core Team (2013) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.
- Rabinowitz, D. (1981) Seven forms of rarity. *The biological aspects of rare plant conservation* (ed. by H. Synge), pp. 205–217. Wiley, New York.
- Randle, A.M., Slyder, J.B. & Kalisz, S. (2009) Can differences in autonomous selfing ability explain differences in range size among sister-taxon pairs of *Collinsia* (Plantaginaceae)? An extension of Baker's Law. *New Phytologist*, **183**, 618–629.
- Rosseel, Y. (2012) lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, **48**, 1–36.
- Roughgarden, J. (1974) Niche width: biogeographic patterns among *Anolis* lizard populations. *The American Naturalist*, **108**, 429–442.
- Seagle, S.W. & McCracken, G.F. (1986) Species abundance, niche position, and niche breadth for five terrestrial animal assemblages. *Ecology*, **67**, 816–818.
- Sheth, S.N., Lohmann, L.G., Distler, T. & Jiménez, I. (2012) Understanding bias in geographic range size estimates. *Global Ecology and Biogeography*, **21**, 732–742.
- Slatyer, R.A., Hirst, M. & Sexton, J.P. (2013) Niche breadth predicts geographical range size: a general ecological pattern. *Ecology Letters*, **16**, 1104–1114.
- Webb, T.J. & Gaston, K.J. (2000) Geographic range size and evolutionary age in birds. *Proceedings of the Royal Society B: Biological Sciences*, **267**, 1843–1850.
- Whittall, J.B., Carlson, M.L., Beardsley, P.M., Meinke, R.J. & Liston, A. (2006) The *Mimulus moschatus* alliance (Phrymaceae): molecular and morphological phylogenetics and their conservation implications. *Systematic Botany*, **31**, 380–397.
- Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H. & Guisan, A. & NCEAS Predicting Species Distributions Working Group (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.
- Wu, C.A., Lowry, D.B., Cooley, A.M., Wright, K.M., Lee, Y.W. & Willis, J.H. (2008) *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity*, **100**, 220–230.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Supplementary figures.

Appendix S2 Detailed methods and results from simple linear regressions.

Appendix S3 Supplementary tables.

BIOSKETCHES

Seema Sheth is interested in understanding the evolutionary and ecological mechanisms underlying species' ecological niches and geographical ranges. This work is part of her dissertation research on the effects of ecological and evolutionary processes on the sizes and limits of geographical ranges.

Iván Jiménez's research focuses on testing theory that attempts to explain spatial patterns of diversity and the size, structure and dynamics of geographical ranges of taxa. He is an associate scientist at the Center for Conservation and Sustainable Development at Missouri Botanical Garden.

Amy Angert studies the ecology and evolution of geographical ranges. She is an assistant professor in the Departments of Botany and Zoology at the University of British Columbia.

Author contributions: S.N.S., I.J. and A.L.A. designed the study; S.N.S. and A.L.A. compiled the data; I.J. designed and coded null distributions; S.N.S. analysed the data; S.N.S. wrote the manuscript with substantial revisions from A.L.A. and I.J.

Editor: Melodie McGeoch