

# Divergence in an obligate mutualism is not explained by divergent climatic factors

William Godsoe<sup>1</sup>, Eva Strand<sup>2</sup>, Christopher Irwin Smith<sup>3</sup>, Jeremy B. Yoder<sup>1</sup>, Todd C. Esque<sup>4</sup> and Olle Pellmyr<sup>1</sup>

<sup>1</sup>Department of Biology, University of Idaho, Moscow, ID 83844, USA; <sup>2</sup>Rangeland Ecology and Management, University of Idaho, Moscow, ID 83844, USA;

<sup>3</sup>Willamette University, Salem, OR 97301, USA; <sup>4</sup>US Geological Survey, Western Ecological Research Centre, 160 N. Stephanie St, Henderson, NV 89074, USA

## Summary

Author for correspondence:

Olle Pellmyr

Tel: +1 208 885 6807

Email: pellmyr@uidaho.edu

Received: 17 February 2009

Accepted: 22 May 2009

*New Phytologist* (2009) **183**: 589–599

doi: 10.1111/j.1469-8137.2009.02942.x

**Key words:** abiotic environment, coevolution, ecological niche modeling, plant–insect interactions, *Tegeticula*, *Yucca*.

• Adaptation to divergent environments creates and maintains biological diversity, but we know little about the importance of different agents of ecological divergence. Coevolution in obligate mutualisms has been hypothesized to drive divergence, but this contention has rarely been tested against alternative ecological explanations. Here, we use a well-established example of coevolution in an obligate pollination mutualism, *Yucca brevifolia* and its two pollinating yucca moths, to test the hypothesis that divergence in this system is the result of mutualists adapting to different abiotic environments as opposed to coevolution between mutualists.

• We used a combination of principal component analyses and ecological niche modeling to determine whether varieties of *Y. brevifolia* associated with different pollinators specialize on different environments.

• *Yucca brevifolia* occupies a diverse range of climates. When the two varieties can disperse to similar environments, they occupy similar habitats.

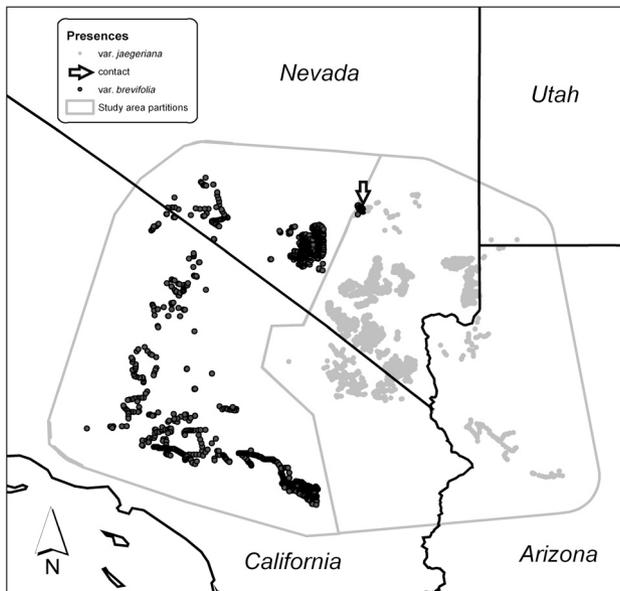
• This suggests that the two varieties have not specialized on distinct habitats. In turn, this suggests that nonclimatic factors, such as the biotic interaction between *Y. brevifolia* and its pollinators, are responsible for evolutionary divergence in this system.

## Introduction

Adaptation to divergent environments creates and maintains biological diversity (Schluter, 2000; Funk *et al.*, 2006), but we know little about the importance of different ecological factors, particularly the relative contributions of the abiotic environment and biotic interactions (Lexer & Fay, 2005; Johnson & Stinchcombe, 2007). It has been hypothesized that coevolution drives ecological divergence in many taxa (Ehrlich & Raven, 1964; Thompson, 2005). The pattern of co-phylogenesis observed at some taxonomic scales in putatively coevolving lineages, such as obligate pollination mutualisms, has offered some support for this contention (Weiblen & Bush, 2002). However, the hypothesis that coevolution produces diversification has rarely been tested against alternative explanations. In particular, there is a lack of tests for the role of ecological divergence caused by the abiotic environment in taxa that are hypothesized to coevolve (Johnson & Stinchcombe, 2007; Nuismer & Gandon, 2008; Toju, 2008). We know that abiotic stresses can have a significant effect on species in these systems (Segraves, 2003), and that many species' boundaries are a result

of environmental differences (Kruckeberg, 1957; Michalak *et al.*, 2001; Lexer & Fay, 2005). It would thus be reasonable to suggest that diversification in these systems is the consequence of adaptation to divergent environments, rather than biotic interactions between mutualists. If this is the case, each taxon should possess different environmental requirements.

We address these questions in a classic example of an obligate mutualism, between *Yucca brevifolia* Engelm. (Joshua tree; Agavaceae) and its yucca moth pollinators (Lepidoptera: Prodoxidae). *Yucca brevifolia* is exclusively pollinated by two sister species of yucca moth, *Tegeticula synthetica* (Riley) and *Tegeticula antithetica* Pellmyr, which speciated whilst interacting with *Y. brevifolia* (Pellmyr & Segraves, 2003; Godsoe *et al.*, 2008). The two moth species differ in overall body size and in ovipositor length, and are parapatrically distributed, with the larger *T. synthetica* occurring in the western half of *Y. brevifolia*'s range and the smaller *T. antithetica* occurring in the eastern part. Interestingly, the two moth species co-occur in a 5-km-wide contact zone in south-central Nevada (Fig. 1). Previous work has suggested that the mutualism between *Y. brevifolia* and its pollinators is mediated by phenotype matching: the female moth



**Fig. 1** Presence of *Yucca brevifolia* var. *brevifolia* and *Y. brevifolia* var. *jaegeriana* and the areas from which we sampled pseudoabsences for each.

uses a blade-like ovipositor to cut into the host flowers and lay her eggs inside the ovary before actively pollinating the flowers (Trelease, 1893), so that her larvae can consume a subset of the developing seeds. Because it must penetrate deeply enough to lay eggs on top of the ovules, but not so deep as to damage the flower, the ovipositor should match the length of the route by which it reaches into the floral ovary (Pellmyr, 2003).

Previous work has documented two varieties of *Y. brevifolia* based on tree architecture and height to first branching (Rowlands, 1978). These taxa are associated with different species of pollinator (Pellmyr & Segraves, 2003; Godsoe *et al.*, 2008), and 12–13 morphological traits examined are significantly different between these varieties (Godsoe *et al.*, 2008). With these data, it is possible to assign an individual tree to the correct variety with 95% accuracy. The varieties are sister taxa (Pellmyr *et al.*, 2007) and have been described as separate species (Lenz, 2007) but, in the absence of an estimate of gene flow based on nuclear markers, we follow previous nomenclatures that treat these entities as varieties (Rowlands, 1978; Baldwin *et al.*, 2002; USDA, 2008).

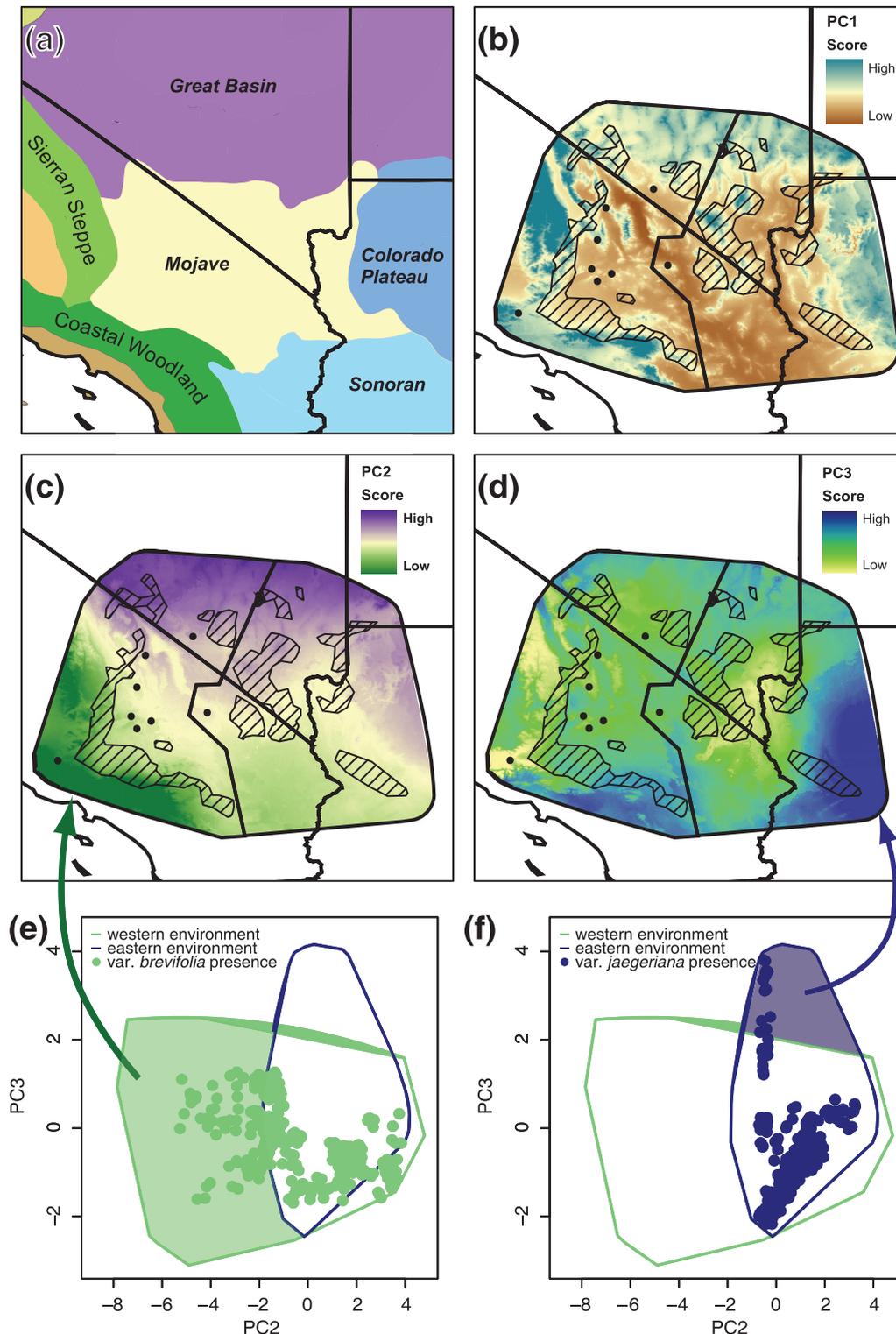
There is a complex history of co-divergence between tree varieties and moth species. The ancestral host plant of the two moth species is *Y. brevifolia* (Godsoe *et al.*, 2008). The distributions of these two varieties parallel those of their pollinators, with *Y. brevifolia* var. *brevifolia* (hereafter var. *brevifolia*) in the west and *Y. brevifolia* var. *jaegeriana* (hereafter var. *jaegeriana*) in the east. Molecular analyses of chloroplast markers indicate that these two varieties are polyphyletic, but diverged significantly earlier than their pollinators (Smith *et al.*, 2008). However, in spite of this history of ancient divergence in the plants,

there are indications of ongoing gene flow between the two varieties (Smith *et al.*, 2008).

The matching between moth and floral phenotypes suggests that the differences between var. *brevifolia* and var. *jaegeriana* may have arisen as a result of coevolutionary interaction with the divergent pollinator species (Godsoe *et al.*, 2008). However, it is also possible that natural selection mediated by the abiotic environment could have produced these differences. To test this alternative hypothesis, we examined whether the two varieties have different climatic requirements. The distribution of *Y. brevifolia* crosses a number of major climatic divisions in the south-western USA. This species occurs primarily in the Mojave desert of the south-western USA, with some populations reaching the Sonoran and Great Basin deserts, the Colorado Plateau semi-desert, as well as montane forest habitats (Fig. 2a). Much of the ecology of plants in this region has been shaped by their ability to survive extreme climates (Bailey, 1995). Moreover, the east–west split between the two varieties of tree has previously been ascribed to a known trend in precipitation across North American deserts (Rowlands, 1978). The Mojave desert receives most of its precipitation in the winter, whereas eastward deserts receive an increasing proportion of summer precipitation, approximating a 50 : 50 summer to winter ratio near the south-eastern distribution limit of *Y. brevifolia* (Brown, 1982). It is thus reasonable to examine whether adaptations to different climates could explain why there are two distinct varieties of *Y. brevifolia* with distinct distributions.

Most plants involved in reciprocally obligate pollination mutualisms are large and long lived (Janzen, 1979; Pellmyr, 2003). For this reason, traditional tests for the role of environmental variation on species' distributions, such as reciprocal transplant experiments, are not typically feasible. This system is, however, an ideal candidate for applying Geographic Information System (GIS)-based methods to test whether these varieties have different climatic requirements (Kozak *et al.*, 2008; Swenson, 2008), because *Y. brevifolia*'s distribution is exceptionally well characterized.

The most popular GIS-based approach is to use ecological niche modeling to test whether the predicted environmental requirements of the taxa are identical or similar (Warren *et al.*, 2008). However, simulation work has shown that some such comparisons can erroneously infer that species have different niches when one species disperses to environments unavailable to the other (W. Godsoe, unpublished). For this reason, we have not focused our investigations on the determination of whether different taxa live in identical environments. Instead, we test whether the two varieties of *Y. brevifolia* have a similar response (pattern of occurrence given important environmental variables) in the environments that both varieties encounter. This is performed in three steps. First, principal component analysis (PCA) is used to determine which environments are available to both varieties. Second, the ability of ecological niche models to predict the presence of each variety of plant



**Fig. 2** A comparison of climates available to each variety of *Yucca brevifolia*. (a) Major ecoregions adjoining the range of *Y. brevifolia*, based on Bailey (1995). (b) Map of scores on principal component 1 (PC1), a surrogate for elevation, and the major source of climatic variation across the Mojave. (c) Map of scores on PC2, a contrast between coastal and continental climates. (d) Map of PC3, a measure of monsoonal precipitation and aseasonal temperature variation. Hatched areas and points on these maps denote the approximate distribution of *Y. brevifolia*. (e, f) Scatter plots of PC2 and PC3, highlighting habitat available to only one variety; (e) presences of *var. brevifolia* with an arrow indicating regions with low loadings on PC2 and the corresponding habitat on the map; (f) presences of *var. jaegeriana*, with an arrow indicating high loadings on PC3 and the corresponding regions on the map.

is assessed. If these models can accurately determine where *Y. brevifolia* occurs, it would suggest that climate is an important determinant of whether environments are suitable for this species. Third, we compared the ability of these distribution models to predict presences in new regions. If there has been ecological divergence caused by physical environmental factors, we should expect a model developed from trees of one variety to be a poor predictor for the presence of the other variety.

## Materials and Methods

### Distribution data assembly

We combined existing survey datasets with extensive, complementary field surveys of nearly all known populations of *Y. brevifolia* to produce a dataset of 5765 presence points [Fig. 1; see also Table S1 (Supporting Information)]. To compensate for variation in sampling density between different areas, we created a stratified subsample of our dataset by overlaying a 5 km × 5 km grid and randomly sampling one presence point from every cell that contained any trees. This produced a final dataset of 319 locations where var. *brevifolia* was present and 417 locations where var. *jaegeriana* was present.

### Climate data

Niche models were based on climate data from 19 WORLDCLIM bioclimatic variables (Hijmans *et al.*, 2005; <http://www.worldclim.org/>) at a spatial resolution of 30 arc seconds (~1 km). We believe that these are the most important abiotic variables to study for *Y. brevifolia* because previous authors have hypothesized that the two varieties specialize on different climatic conditions (Rowlands, 1978). Moreover, climate is believed to be a particularly important determinant of species distributions in the American South-west (Bailey, 1995). Individual populations of *Y. brevifolia* are separated from each other by as much as 50 km. We interpret this as an upper bound on how far away we should expect to find disjunct populations. We thus restricted our analyses to a minimum convex polygon around the range of *Y. brevifolia* plus a 50-km buffer (Fig. 1). We divided our study area into a western region containing presences of var. *brevifolia* and an eastern region containing presences of var. *jaegeriana*. In the south, this division traverses a low-elevation region (Fig. 2b), inhabited by neither variety (Fig. 1). *Yucca brevifolia* disperses over short distances. Seeds are typically dispersed less than 30 m by rodents (Vander Wall *et al.*, 2006), and *Yucca* moths typically disperse pollen over distances of less than 50 m (Marr *et al.*, 2000). We are thus reasonably confident that each variety cannot disperse across this boundary. In the north, there is a great deal of suitable habitat for *Y. brevifolia* and the two varieties come into contact. We are less confident that our division represents a natural barrier in this region but, in the north, similar

environments are available on both sides of this boundary. We thus interpret our division as a reasonable approximation of the environments to which each variety can disperse.

Niche modeling algorithms can produce unreliable extrapolations in environments to which the focal organism cannot disperse (W. Godsoe, unpublished). Thus, we identified regions in which any climatic variable in one portion of the Mojave is more extreme than any value of that variable in the other portion of our study region. These regions represent environments to which one variety cannot disperse, and so we treat inferences of niche models in these regions with caution (Elith & Graham, 2009). It is difficult to visualize all 19 WORLDCLIM variables, and so we have summarized this variation by calculating principal components from 2000 randomly selected locations from the eastern and western Mojave, respectively. Maps generated from these principal components provide a simple way to explore the climates available to both varieties (Fig. 2). In particular, when the color of such a map in one region is more extreme than any shade found in the other region (see, for example, Fig. 2c,d), it indicates that a climate in the first region is more extreme than any climate available in the second. These principal components also provide a convenient way to link the climate data used to model the distribution of *Y. brevifolia* to existing descriptions of environmental variation in the American South-west.

### Niche models

We developed niche models using boosted regression trees (BRT) and maximum entropy (Maxent; Friedman *et al.*, 2000; Phillips *et al.*, 2006), two algorithms that performed well in an extensive comparison of available methods on empirical data (Elith *et al.*, 2006). Unless otherwise noted, we created models by using all of the presences for a particular variety.

BRT is a multi-stepped machine-learning algorithm. The method starts by constructing a decision tree that produces a crude classification of observations into presences and absences. The algorithm builds subsequent trees to explain observations that were poorly predicted in previous steps, and then builds a composite model from a large number of these decision trees, allowing for nonlinear terms, complex interactions and an evaluation of the relative importance of different parameters. We implemented BRT using Elith *et al.*'s (2008) modifications of the *gbm* package in R (R Development Core Team, 2006; Ridgeway, 2006). Our models used presence points for a given variety and an equal number of pseudoabsences selected at random from the appropriate region of the Mojave desert. BRT requires three user-defined parameters that adjust the complexity and stochasticity of the resulting model. We compared the cross-validated area under the receiver operating curve (AUC) scores of models using 27 parameter combinations (three different bagging fractions, three different tree complexities and three different learning rates). AUC scores varied little between different parameter combinations (0.85–0.89).

A relatively fast learning rate of 0.01 and tree complexity of eight produced slightly higher AUC scores for both varieties, and so we used these parameter values in subsequent analyses. A bagging fraction of 0.25 produced the best model for var. *jaegeriana*; a bagging fraction of 0.75 produced the best model for var. *brevifolia*. As a compromise, we used a fraction of 0.5 for subsequent analyses. We estimated the number of trees in the final model with a 10-fold cross-validation procedure (Elith *et al.*, 2008). This resulted in models from a composite of approximately 400 decision trees.

Maxent is a machine-learning algorithm that estimates the probability of occurrence by constraining this probability distribution to resemble the distribution of observed presences and by minimizing the information contained in the residuals (Phillips *et al.*, 2006). We sampled 10 000 background points for analyses of var. *brevifolia* from the western portion of our study area, and 10 000 background points for analyses of var. *jaegeriana* from the eastern portion of our study area. We used the default parameter settings for Maxent version 3.1.0 (logscale output, regularization multiplier = 1, convergence threshold = 0.00001).

### Assessment of model accuracy

We used AUC statistics as our primary measure of model accuracy (Freeman, 2007). This is a nonparametric estimate of a model's ability to distinguish between presences and absences. The AUC metric varies from zero to unity, with a score of unity representing perfect discrimination and a score of 0.5 representing a model that performs no better than random. This metric is threshold independent and so is an appropriate choice for comparing Maxent and BRT, two methods with continuous outputs that vary from zero to unity. We computed cross-validated AUC scores to measure the accuracy of niche models for each variety within its home range. In Maxent, we obtained cross-validated scores by randomly selecting 90% of the observations, and then calculated AUC scores on the remaining 10%. For BRT, we used the R scripts of Elith *et al.* (2008) to create 10 datasets, each using 90% of the available data to create the model and 10% to calculate AUC scores; the final AUC score is the average over all 10 data partitions.

### Comparing predictions of niche models

To test whether the two varieties have similar environmental requirements, we must find a way to compare the predictions of two niche models developed in different regions. Unfortunately, AUC scores will not necessarily reflect the ability of the models to predict locations in new regions (Randin *et al.*, 2006; Peterson *et al.*, 2007). For this reason, we specifically designed a measure of niche similarity that compares the ability of niche models of each variety to predict to a new region. Using the maps from our PCAs, we determined that comparable climates are available approximately 300 km to the east or west of the contact zone. We set aside subsets of this region near the

contact zone for tests of the ability of models developed for each variety to predict presences in adjacent regions with similar climates. For var. *brevifolia*, we created a niche model for the 225 presences further than 200 km from the contact zone. We measured the ability of this model to predict the 94 var. *brevifolia* presences within 200 km of the contact zone. We then used a model of var. *jaegeriana* trees to predict to the same set of presences (Fig. S1A, see Supporting Information), and compared the accuracy of both models using AUC scores. Similarly, we produced a model of the 138 var. *jaegeriana* presences further than 200 km from the contact zone, and then applied it to the 279 var. *jaegeriana* presences within 200 km from the contact zone. Finally, we compared the accuracy of these predictions to that of a model developed for var. *brevifolia* trees (Fig. S1B). Although sample size varied from one model to another, a recent analysis has found that BRT can produce reasonably robust niche models with only 30 presences (Guisan *et al.*, 2007). Our models used at least four times more presences than this minimum.

For these comparisons, we present AUC scores with standard errors computed in the PresenceAbsence package in R (Freeman, 2007). However, we note that this standard error estimate is technically only valid for comparing models against random prediction, and for the particular set of presence and pseudo-absence points used.

This approach for measuring niche similarity has three advantages over existing methods (Peterson *et al.*, 1999; Warren *et al.*, 2008). First, there is some concern that a good niche model in one region may be less accurate in another (Peterson *et al.*, 2007). In the absence of ecological divergence, we might still expect a model developed in one region to produce more accurate predictions in that region than would a model derived elsewhere. Our test mitigates this problem by forcing the models of both varieties to predict in a new region.

Second, this design permits us to specifically compare our models in one region. Niche models produce higher AUC scores in study areas with a great deal of unsuitable habitat than they do in study areas with a great deal of suitable habitat (Lobo *et al.*, 2008). Thus, comparisons of AUC scores from different regions may give the illusion that one model is more accurate than another, even though they may produce equally good predictions when applied to the same study region. We eliminate this problem by restricting our model comparisons to a single study area per test.

Third, our method also has a useful ecological interpretation. If the sharp distribution limits that we observe in sympatry are a consequence of different climatic requirements, a model of one variety should predict very few of the presences for the other variety near the contact zone. For this reason, a model of the wrong variety would be far less accurate than a model of the correct variety. However, if the two varieties have different distributions for some other reason, such as biotic interactions or dispersal limitation, we should expect the distribution models to produce predictions of a comparable quality.

We present two metrics of the ability of niche models of each variety to predict the distribution of the other variety. First, we use the cross-validated AUC score as a measure of the ability of a model to make predictions to the home range of a variety. We compare this score with the ability of the same model to make predictions to the range of the other variety (hereafter referred to as AUC accuracy away from the home range). Second, we test whether BRT models of one variety correctly predict presences of the other variety significantly better than chance (Fig. 3), a frequently reported measure of niche similarity (Peterson *et al.*, 1999; Warren *et al.*, 2008). As our models were fitted to a dataset of 50% presences and 50% pseudoabsences, we identified presences as observations with a probability of occurrence of more than 50%. We then used a  $\chi^2$  test following Peterson *et al.* (1999), with Yates' correction, to determine whether the model predicted a greater number of presences than would be expected by chance.

## Results

### Environmental variation

The first three principal components summarized 88% per cent of the environmental variation. PC1, representing 49% of the variation, contrasts between temperature and precipitation (Table S2, see Supporting Information) and is strongly associated with elevation. PC2, representing 30% of the variation, is negatively correlated with precipitation in the coldest quarter, precipitation in the wettest month and several other traits that are high in regions near the Pacific coast (i.e. the west portion of our study region), and is positively correlated with temperature seasonality and precipitation in the driest month, all of which are properties of interior climates. For this reason, it may be thought of as a contrast between coastal (low values of PC2) and interior (high values of PC2) climates. PC3, representing 9% of the variation, is positively associated with temperature variation and summer precipitation, which are characteristics of the Sonoran desert and Colorado Plateau at the eastern edge of the range of *Y. brevifolia* (Fig. 2a,d). There is much more variation in PC2 in the western Mojave than in the east, and much more variation in PC3 in the eastern Mojave than in the west (Fig. 2c–f). Together, PC2 and PC3 partially represent the previously described east–west gradient from strong winter precipitation in the Mojave desert to strong summer precipitation in the Sonoran desert (Bailey, 1995), although neither shows a steep east–west gradient further north, near the contact zone between var. *brevifolia* and var. *jaegeriana* (Fig. 2c,d). In this region, PC1 exhibits the most variation, but both varieties of *Y. brevifolia* occupy a comparable elevation range in allopatry (Fig. 4b,d). It is in sympatry that we see a sharp deviation from this pattern, where var. *brevifolia* occurs in a region with low scores on this PC, and var. *jaegeriana* occurs in regions with high scores (Fig. 4c).

**Table 1** Area under the receiver operating curve (AUC) scores for ecological niche models of each variety, projected within their range (*Yucca brevifolia* var. *brevifolia* in the western portion of our study area; *Y. brevifolia* var. *jaegeriana* in the eastern portion) and onto the range of the other variety

	Maxent		BRT	
	Home	Away	Home	Away
var. <i>brevifolia</i>	0.832	0.722	0.886	0.758
var. <i>jaegeriana</i>	0.879	0.611	0.831	0.645

BRT, boosted regression trees; Maxent, maximum entropy.

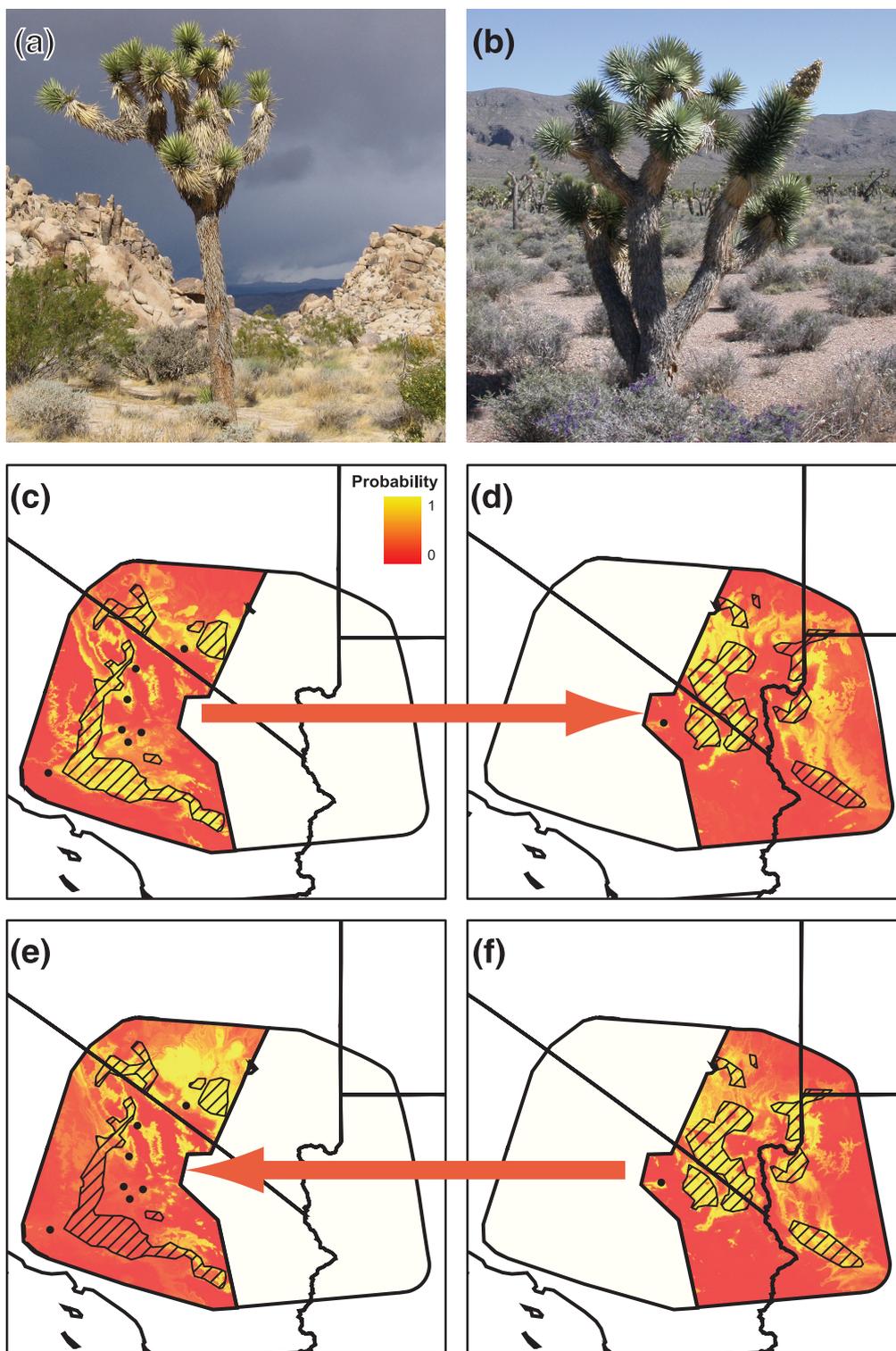
### Niche model accuracy and comparisons

Maxent and BRT each produced models with AUC scores above 0.8 when niche models of each variety were used to predict their own ranges (Table 1). When niche models of one variety were applied to the range of the other, they predicted many presences near the contact zone, but were less accurate in more distant locations (Fig. 3d,e). This pattern was most pronounced when we applied predictions derived from the southern portion of the range of var. *brevifolia* to the prediction of var. *jaegeriana* (AUC = 0.611 and AUC = 0.645 for Maxent and BRT, respectively; Figs 3, S2; Table 1). In that case, the Maxent and BRT models predicted essentially no presences in the southern portion of the range of var. *brevifolia*, producing the lowest AUC scores in the study. The climate in this region has low loadings on PC2; moreover there is no region with an equivalent climate in the eastern Mojave (Figs 2, S2).

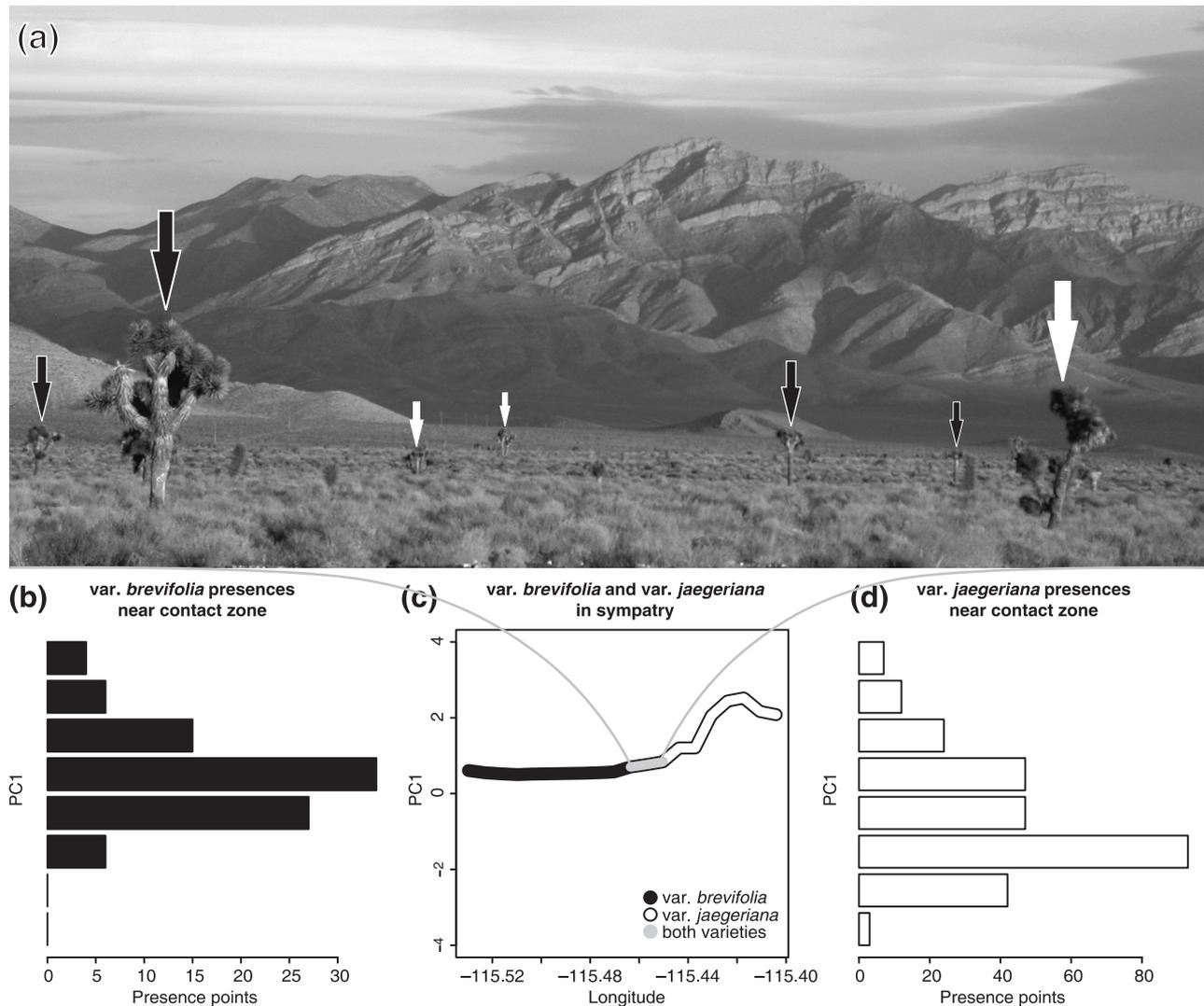
*Yucca brevifolia* var. *jaegeriana* trees near the contact zone were predicted nearly as well by a model derived from var. *brevifolia* as they were by var. *jaegeriana* away from the contact zone (AUC  $\pm$  SE: 0.718  $\pm$  0.02 vs 0.739  $\pm$  0.02). The distribution of var. *brevifolia* near the contact zone was actually predicted better by a model of the distribution of var. *jaegeriana* than it was by a model of var. *brevifolia* away from the contact zone (AUC = 0.748  $\pm$  0.04 vs 0.658  $\pm$  0.04). Across the entire range, BRT produced niche models that performed much better than chance (prediction based on var. *brevifolia* to var. *jaegeriana*,  $P = 1.1e-51$ ; prediction based on var. *jaegeriana* to var. *brevifolia*,  $P = 0.0000578$ ).

## Discussion

Our combined analyses indicate that *Y. brevifolia* is exposed to a diverse array of climates, and that these climates affect which environments are suitable to this species. There is little evidence, however, that trees associated with different pollinators have different ecological requirements. Both the eastern and western partitions of our study area contain locations with climates absent in the other partition (Fig. 2). *Yucca brevifolia*



**Fig. 3** Comparisons of niche models of each variety of *Yucca brevifolia*. (a, b) Illustrations of var. *brevifolia* and var. *jaegeriana*, respectively. It should be noted that var. *jaegeriana* branches much more closely to the ground and that subsequent branches are much more symmetric than those of var. *brevifolia*. (c) The prediction of var. *brevifolia* onto its own range. (d) The prediction of var. *brevifolia* onto the range of var. *jaegeriana*. (e) The prediction of var. *jaegeriana* onto the range of var. *brevifolia*. (f) The prediction of var. *jaegeriana* on its own range. Hatched areas and points denote the approximate distribution of *Y. brevifolia*.



**Fig. 4** A comparison of the climates used by each variety in sympatry and in populations near their contact zone. (a) Photograph of the contact zone with arrows pointing to individual trees of *Yucca brevifolia* var. *brevifolia* (black) and *Y. brevifolia* var. *jaegeriana* (white). The contact zone represents a narrow and relatively homogeneous region, well within the range of environments used by each variety in allopatry. (b, d) Comparisons of the climates used by *var. brevifolia* and *var. jaegeriana* in nearby populations. (c) A transect along the contact zone and the range of principal component 1 (PC1) used by each variety. The peak to the right of this panel corresponds to the mountain range in (a). The y-axes in (b–d) represent scores from –4 to 4 on PC1, a surrogate for elevation and the major source of climatic variation near the contact zone.

*var. brevifolia* occurs along the western edge of the Mojave desert, a region with low loadings on PC2 associated with a more coastal climate. This region includes the transition from desert to montane woodlands (Fig. 2a,c,e; Bailey, 1995). The BRT niche model for *var. jaegeriana* predicts few presences in this region, and the Maxent model predicts none (Fig. S2). Conversely, a few populations of *var. jaegeriana* reach into the edge of the Sonoran desert, a region characterized by summer precipitation; this climate has no equivalent further west. This region has high scores on PC3 (Fig. 2d,f), and niche models of *var. jaegeriana* predict few presences in these regions.

Our niche models indicate that, when eastern and western trees encounter similar climates, they have a similar response.

Not only do models of one variety predict many presences in regions with comparable climates, they actually produce predictions of equal, if not greater, accuracy than models of the correct variety. BRT niche models for each tree variety even produced acceptable predictions across the range of the other variety, and were far more accurate than would be expected by chance. Projections of the niche model of one variety across the entire distribution of the other must be interpreted with caution, but this result at least suggests that, in many places, the climatic requirements of each variety are very similar.

A simple explanation for this pattern would be that moth and tree lineages diverged into allopatric or parapatric populations at some time in the past. We know from PCA that

different climates suitable to *Y. brevifolia* are available in the eastern and western portions of the Mojave. Given this, allopatric lineages should occupy different environments, even in the absence of ecological divergence. Unfortunately, the data available do not allow us to compare the ecological requirements of *Y. brevifolia* across its entire range. Therefore, it is still possible that the tree varieties would perform differently if they were both to encounter environments with high values for PC2 or PC3. To illustrate this, consider an experimental design that corresponds to the extant distribution of *Y. brevifolia*. Imagine if we had planted both varieties in the environments near their contact zone in Nevada, but only one variety in a study site to the west and one in a study site to the east. We could compare the requirements of each variety in Nevada, but it would be impossible to compare their requirements in the other two study sites. By analogy, the extant distribution of the two varieties is similar to a single common garden 'planted' in locations in which the two varieties could encounter each other, but with missing or incomplete information about other locations. This interpretation does, however, depend on our assessment of the environments to which the two varieties of *Y. brevifolia* can disperse. If the two tree varieties really can disperse across the southern Mojave, the absence of var. *jaegeriana* in the west and var. *brevifolia* in the east constitutes strong evidence of ecological differentiation. Nevertheless, we are reasonably confident in our assertion that the trees cannot disperse across this barrier because our distribution models indicate that this region contains a broad area of habitat unsuitable to both varieties, and we know that *Y. brevifolia* disperses over short distances (Vander Wall *et al.*, 2006).

The main conclusion of our study is thus that the two varieties of *Y. brevifolia* have similar climatic requirements across the many environments that they both encounter. This demonstrates that the two varieties have not specialized on distinct climates. In particular, we know that the abrupt boundary between the two varieties in sympatry cannot be explained by adaptation to different climates, as niche models of one variety are reasonably good at predicting presence for the other variety. It is possible that recent contact, or unmeasured abiotic variables, drive this boundary, although the strong performance of our niche models in allopatry suggests that climate is indeed a very important determinant of *Y. brevifolia*'s distribution.

Comparing the predictions of distribution models is a relatively new technique. We have only a limited understanding of the properties of these methods, and so it is important to consider their potential weaknesses. To our knowledge, there are no analyses of the statistical power of these methods. Our comparisons of niche models indicate that the performance of the two varieties is similar, but with current analytical tools we cannot rule out the possibility that there is a biologically significant difference in the performance of the two tree varieties that is too small to be detected with distribution-based methods. We have followed existing conventions for these methods that test for similarity using null models, but recognize

that more theoretical work is needed to develop a nuanced interpretation of distribution-based data.

It is possible that the incorporation of more data, such as Joshua tree population density at different locations, could improve the power of our tests. If the tree varieties have different environmental requirements, we might expect var. *brevifolia* trees to be at low densities in environments with high values of PC3, and var. *jaegeriana* trees to have low densities in regions with low values of PC2. We were unable to collect sufficient density data to test this idea, but the limited information at hand is inconsistent with this hypothesis. Rowlands (1978) analyzed the percentage cover data from *c.* 30 populations. He found that south-eastern populations of var. *jaegeriana* were unusually dense, but that across the rest of the range the density of both varieties increased gradually with elevation. Moreover, the highest percentage cover of var. *brevifolia* was in the south-eastern portion of the range of that variety, a part of the western Mojave with particularly high values of PC3.

Another limitation is that we have not explicitly partitioned out the effects of mutualist-mediated fitness from the abiotic environment. This is an important problem, and one that can now be addressed rigorously, albeit with complex experimental designs that require reciprocal transplants and manipulations of the available mutualists (Nuismer & Gandon, 2008; Piculell *et al.*, 2008). In comparison, our approach is necessarily crude. We found that the two tree varieties had similar climatic requirements, particularly in the northern portion of their range where they encounter one another. This indicates that climatic requirements are insufficient to explain the observed differences in the two taxa. This is an interesting qualitative result, but extensive additional work will be needed to quantify the relative influence of the biotic and abiotic environment.

Aside from the statistical issues, many inferences about the causes of ecological divergence are hindered by our lack of historical information. We know that agents of evolutionary change may have been important in the past, but that subsequent genetic or environmental changes have obscured this fact (Coyne & Orr, 2004). Although there is little evidence of current ecological differentiation, it is possible that the two varieties specialized on different environments in the past, but that these environments do not currently exist (Jackson & Overpeck, 2000). The best way to deal with this problem is to place our observations in a phylogenetic context; in other words, to test whether speciation is associated with pollinator divergence and not with changes in an organism's abiotic requirements. This is not possible in our study because we have only investigated a single example of speciation in the moth lineage and in the plant lineage. The similarity in environmental requirements suggests that adaptation to different climates has not driven ecological divergence in this system. To our knowledge, this represents one of the first tests of this important hypothesis in a coevolving system (but see Toju, 2008). Nonetheless, this result must be interpreted with caution, as a mechanism of divergence may have been very

important in the past, but may be largely irrelevant now (Coyne & Orr, 2004).

We do not yet understand why vegetative traits, such as leaf length and height to first branching, are different between the two varieties. We propose two possible explanations. First, tree architecture may be developmentally linked to floral traits. In particular, branching occurs at the site of previous inflorescences (Rowlands, 1978), and so selection on floral development might indirectly alter other aspects of tree morphology. Second, isolation between the two varieties may have provided the opportunity for genetic divergence. These changes could have resulted from either drift or selection, but they did not alter the ability of the trees to survive in the climates currently available to both varieties. These are important questions which will require a careful investigation of the genetic and physiological basis of these traits.

Obligate mutualisms have offered valuable insights into how ecology shapes the evolution of biological diversity (Weiblen, 2002; Pellmyr, 2003). In particular, we now know that strong interspecific interactions over ecological time-scales are linked with macroevolutionary patterns of diversification. This has been interpreted as evidence that coevolution between mutualists creates and maintains biological diversity, but it has been difficult to determine whether the macroevolutionary pattern is caused by coevolution, or is a side-effect of other ecological or evolutionary forces. Recent work (Godsoe *et al.*, 2008) indicates that the traits that show the greatest degree of divergence between the two varieties of *Y. brevifolia* are tightly associated with pollination, and match the morphology of their respective pollinator. The evidence presented here suggests that the abiotic environment does not explain the pattern of co-diversification we see in this obligate mutualism, which, together with the results of Godsoe *et al.* (2008), provides support for the hypothesis that co-divergence is the result of coevolution.

## Acknowledgements

We thank the US National Park Service, US Fish and Wildlife Service, US Department of Energy, Joshua Tree National Park, Death Valley National Park, Mojave National Preserve, Desert National Wildlife Refuge and The Nevada Test Site for providing access to study sites. W. Kent Ostler helped arrange access to some populations and assisted with data collection. Paul DePrey, Rob Fulton, Ashley Cushman, Tasha LaDoux, Mark Hagan and David Reznick provided logistical support. We thank Gary Reese, Benjamin A. Waitman and Trent M. Draper for technical assistance. Jorge Soberon, Luke Harmon, Jeffrey Evans, Melanie Murphy, Niko Balkenhol, Chris Drummond, Ben Ridenhour and Raymond Dezzani provided valuable comments and suggestions. Jane Elith generously provided her R scripts, and advice on using boosted regression trees. Financial support was provided by National Science Foundation grant DEB-0516841, and the Cooperative Ecosystems Studies Units Network (CESU) of the US National

Park System. Bioinformatics facilities were supported by NIH/NCRR P20RR16448 and P20RR016454. W.G. was supported by a pre-doctoral fellowship from the Canadian National Science and Engineering Research Council and the Foundation for Orchid Research and Conservation. J.B.Y. received support from the Joshua Tree National Park Graduate Student Research Program. Any use of trade, product, or firm names in this publication is for descriptive purposes only and does not imply endorsement by the US government.

## References

- Bailey RG. 1995. *Description of the ecoregions of the United States*, 2nd edn. revised and expanded (1st edn. 1980). *Misc. Publ. No. 1391 (rev.)*. Washington DC, USA: USDA Forest Service.
- Baldwin BG, Boyd S, Ertter BJ, Patterson RW, Rosatti TJ, Wilken D. 2002. *The Jepson desert manual: vascular plants of south eastern California*. Berkeley, CA, USA: University of California Press.
- Brown DE. 1982. *Biotic communities of the American southwest, United States and Mexico*. Tucson, AZ, USA: University of Arizona for Boyce Thompson Southwestern Arboretum.
- Coyne JA, Orr HA. 2004. *Speciation*. Sunderland, MA, USA: Sinauer Associates.
- Ehrlich PR, Raven PH. 1964. Butterflies and plants: a study in coevolution. *Evolution* 18: 586–608.
- Elith J, Graham CH. 2009. Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. *Ecography* 32: 66–77.
- Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A *et al.* 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129–151.
- Elith J, Leathwick JR, Hastie T. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology* 77: 802–813.
- Freeman E. 2007. *Presence-absence: an R package for presence-absence model evaluation*. Ogden, UT, USA: USDA Forest Service, Rocky Mountain Research Station.
- Friedman J, Hastie T, Tibshirani R. 2000. Additive logistic regression: a statistical view of boosting. *Annals of Statistics* 28: 337–407.
- Funk DJ, Nosil P, Etges WJ. 2006. Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proceedings of the National Academy of Sciences, USA* 103: 3209–3213.
- Godsoe W, Yoder JB, Smith CI, Pellmyr O. 2008. Coevolution and divergence in the Joshua tree/yucca moth mutualism. *The American Naturalist* 171: 816–823.
- Guisan A, Zimmerman NE, Elith J, Graham CH, Phillips S, Peterson AT. 2007. What matters for predicting the occurrences of trees: techniques, data, or species' characteristics. *Ecological Monographs* 77: 615–630.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.
- Jackson ST, Overpeck JT. 2000. Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology* 26: 194–220.
- Janzen DH. 1979. How to be a fig. *Annual Review of Ecology and Systematics* 10: 13–51.
- Johnson MTJ, Stinchcombe JR. 2007. An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology and Evolution* 22: 250–257.
- Kozak KH, Graham CH, Wiens JJ. 2008. Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology and Evolution* 23: 141–148.

- Kruckeberg AR. 1957. Variation in fertility of hybrids between isolated populations of the serpentine species, *Streptanthus glandulosus* Hook. *Evolution* 11: 185–211.
- Lenz LW. 2007. Reassessment of *Y. brevifolia* and recognition of *Y. jaegeriana* as distinct species. *Aliso* 24: 97–104.
- Lexer C, Fay MF. 2005. Adaptation to environmental stress: a rare or frequent driver of speciation? *Journal of Evolutionary Biology* 18: 893–900.
- Lobo JM, Jiménez-Valverde A, Real R. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* 17: 145–151.
- Marr DL, Leebens-Mack J, Elms L, Pellmyr O. 2000. Pollen dispersal in *Yucca filamentosa* (Agavaceae): the paradox of self-pollination behavior by *Tegeticula yuccasella* (Prodoxidae). *American Journal of Botany* 87: 670–677.
- Michalak P, Minkov I, Helin A, Lerman DN, Bettencourt BR, Feder ME, Korol AB, Nevo E. 2001. Genetic evidence for adaptation-driven incipient speciation of *Drosophila melanogaster* along a microclimatic contrast in 'Evolution canyon,' Israel. *Proceedings of the National Academy of Sciences, USA* 98: 13 195–13 200.
- Nuismer SL, Gandon S. 2008. Moving beyond common garden and transplant designs: insight into the causes of local adaptation in species interactions. *The American Naturalist* 171: 658–668.
- Pellmyr O. 2003. Yuccas, yucca moths, and coevolution: a review. *Annals of the Missouri Botanical Garden* 90: 35–55.
- Pellmyr O, Seagraves KA. 2003. Pollinator divergence within an obligate mutualism: two yucca moth species (Lepidoptera; Prodoxidae; *Tegeticula*) on the Joshua tree (*Yucca brevifolia*; Agavaceae). *Annals of the Entomological Society of America* 96: 716–722.
- Pellmyr O, Seagraves KA, Althoff DM, Balcazar-Lara M, Leebens-Mack J. 2007. The phylogeny of yuccas. *Molecular Phylogenetics and Evolution* 43: 493–501.
- Peterson AT, Papes M, Eaton M. 2007. Transferability and model evaluation in ecological niche modeling: a comparison of Garp and Maxent. *Ecography* 30: 550–560.
- Peterson AT, Soberon J, Sanchez-Cordero V. 1999. Conservatism of ecological niches in evolutionary time. *Science* 285: 1265–1267.
- Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modeling* 190: 231–259.
- Piculell BJ, Hoeksema JD, Thompson JN. 2008. Interactions of biotic and abiotic environmental factors in an ectomycorrhizal symbiosis, and the potential for selection mosaics. *BMC Biology* 6: 23–34.
- R Development Core Team. 2006. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Randin CF, Dirnbock T, Dullinger S, Zimmerman NE, Zappa M, Guisan A. 2006. Are niche-based species distribution models transferable in space. *Journal of Biogeography* 33: 1689–1703.
- Ridgeway G. 2006. *GBM: Generalized boosted regression models*, R package version 1.5-7. <http://www.I-pensieri.Com/gregr/gbm.Shtml>.
- Rowlands PG. 1978. *The vegetation dynamics of the Joshua tree (Yucca brevifolia Engelm.) in the southwestern United States of America*. PhD thesis, University of California, Riverside, CA, USA.
- Schluter D. 2000. *The ecology of adaptive radiation*. Oxford, UK: Oxford University Press.
- Seagraves KA. 2003. Understanding stability in mutualisms: can extrinsic factors balance the yucca–yucca moth interaction? *Ecology* 84: 2943–2951.
- Smith CI, Godsoe W, Tank S, Yoder JB, Pellmyr O. 2008. Distinguishing coevolution from covariation in an obligate pollination mutualism: asynchronous divergence in Joshua tree and its pollinators. *Evolution* 62: 2676–2687.
- Swenson NG. 2008. The past and future influence of geographic information systems on hybrid zone, phylogeographic and speciation research. *Journal of Evolutionary Biology* 21: 421–434.
- Thompson JN. 2005. *The geographic mosaic of coevolution*. Chicago, IL, USA: University of Chicago Press.
- Toju H. 2008. Fine-scale local adaptation of weevil mouthpart length and camellia pericarp thickness: altitudinal gradient of putative arms race. *Evolution* 62: 1086–1102.
- Trumble W. 1893. Further studies of yuccas and their pollinators. *Annual Reports of the Missouri Botanical Garden* 4: 181–226.
- USDA, NRCS. 2008. *The PLANTS Database*. Baton Rouge, LA, USA: National Plant Data Center, (<http://plants.usda.gov>, last accessed 26 June 2009).
- Vander Wall SB, Esque T, Haines D, Garnett M, Waitman BA. 2006. Joshua tree (*Yucca brevifolia*) seeds are dispersed by seed-caching rodents. *Ecoscience* 13: 539–543.
- Warren DL, Glor RE, Turelli M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* 63: 2868–2883.
- Weiblen GD. 2002. How to be a fig wasp. *Annual Review of Entomology* 47: 299–330.
- Weiblen GD, Bush GL. 2002. Speciation in fig pollinators and parasites. *Molecular Ecology* 11: 1573–1578.

## Supporting Information

Additional supporting information may be found in the online version of this article.

**Fig. S1** A comparison of the ability of niche models to predict to new regions.

**Fig. S2** A comparison of the predictions of niche models and available climates.

**Table S1** Distribution records of *Yucca brevifolia*

**Table S2** Principal component loadings for environmental variation in the range of *Yucca brevifolia*

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.