



International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report

IR-13-035

Evolutionary branching in complex landscapes

Benjamin C. Haller
Rupert Mazzucco (mazzucco@iiasa.ac.at)
Ulf Dieckmann (dieckmann@iiasa.ac.at)

Approved by

Pavel Kabat
Director General and Chief Executive Officer

June 2015

Interim Reports on work of the International Institute for Applied Systems Analysis receive only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work.

Title: **Evolutionary branching in complex landscapes**

Type: Article (E-Article)

Benjamin C. Haller^{1*}, Rupert Mazzucco^{2**}, Ulf Dieckmann^{2***}

¹Department of Biology and Redpath Museum, McGill University, 859 Sherbrooke Street West, Montreal, Quebec, Canada H3A 0C4

²Evolution and Ecology Program, International Institute for Applied Systems Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria

*Corresponding author: Haller, B.C. (benjamin.haller@mail.mcgill.ca)

**Email: mazzucco@iiasa.ac.at

***Email: dieckmann@iiasa.ac.at

Keywords: spatial environmental heterogeneity, local adaptation, individual-based model, negative frequency-dependent selection, ecological speciation, evolutionary refugia

List of elements for expanded online edition: Online Appendix: Supplemental methods and results; Tables A1–A3; Figures A1–A15.

Figures to print in color: All figures (E-Article)

Abstract

Divergent adaptation to different environments can promote speciation, and it is thus important to consider spatial structure in models of speciation. Earlier theoretical work, however, has been limited to particularly simple types of spatial structure – linear environmental gradients and spatially discrete metapopulations – leaving unaddressed the effects of more realistic patterns of landscape heterogeneity, such as nonlinear gradients and spatially continuous patchiness. To elucidate the consequences of such complex landscapes, we adapt an established spatially explicit individual-based model of evolutionary branching. We show that branching is most probable at intermediate levels of various types of heterogeneity, and that different types of heterogeneity have, to some extent, additive effects in promoting branching. In contrast to such additivity, we find a novel refugium effect in which refugia in hostile environments provide opportunities for colonization, thus increasing the probability of branching in patchy landscapes. Effects of patchiness depend on the scale of patches relative to dispersal. Providing a needed connection to empirical research on biodiversity and conservation policy, we introduce empirically accessible spatial environmental metrics that quantitatively predict a landscape's branching propensity.

Introduction

We have only partial answers to questions such as why there are so many species in the world, and why some clades are so much more speciose than others (Hutchinson 1959; Schluter 2000; Coyne and Orr 2004; Butlin et al. 2012). A central question in speciation theory concerns the cause of speciation: in particular, which factors (environmental, ecological, phenotypic, genetic) promote speciation, and which factors hinder it? According to the ecological speciation model (Schluter 2001), ecology drives speciation: the process of speciation begins with adaptation to different environments by different populations or subpopulations. From this perspective, evolutionary branching in an ecological trait is the first step toward speciation; reproductive isolation follows, either as a byproduct of ecological divergence or as a result of reinforcement (Schluter 2001).

Empirical support for the ecological speciation model has come from studies in which the adaptive traits that exhibit divergence between different populations also contribute to reproductive barriers (Rundle and Nosil 2005; Hendry 2009; Schluter 2009). Recent theoretical work has also demonstrated that local adaptation in a heterogeneous environment, driven by local competition for resources, can lead to speciation in both sexual and asexual populations (e.g., Doebeli and Dieckmann 2003; Gavrillets and Vose 2005; Birand et al. 2012).

Complex landscapes

Theoretical models of divergence as a result of local adaptation to different environments necessarily consider some type of spatial heterogeneity. Two types of spatial models have dominated: metapopulation models, in which populations inhabit two or more discrete patches that differ in some way (e.g., Doebeli and Ruxton 1997; Meszéna et al. 1997; Geritz et al. 1998;

Kisdi and Geritz 1999; Day 2000; Gavrillets et al. 2000a; Geritz and Kisdi 2000; Parvinen and Egas 2004; Gavrillets and Vose 2005; Thibert-Plante and Hendry 2009; Birand et al. 2012), and spatially continuous models of linear environmental gradients (e.g., Kirkpatrick and Barton 1997; Doebeli and Dieckmann 2003; Mizera and Meszina 2003; Leimar et al. 2008; Heinz et al. 2009; Ispolatov and Doebeli 2009; Payne et al. 2011; Irwin 2012), although a few models have tried to bridge the gap between these approaches (Sutter and Kawecki 2009; Debarre and Gandon 2010).

Real environments are more complex than these simple cases. Patchy environments are not always discrete, as metapopulation models assume; rather, spatial environmental heterogeneity is often continuous (Manning et al. 2004; Fischer and Lindenmayer 2006). On the other hand, this continuous environmental heterogeneity is often not linear, as linear gradient models assume; rather, spatial environmental heterogeneity may be pronounced in some areas and minimal in others. The effects of more realistically complex spatial heterogeneity on adaptive divergence and speciation have not been explored in theoretical models. Given the intimate feedbacks between spatial heterogeneity, dispersal, and local adaptation, we expect such heterogeneity to be important.

Here we model the dynamics of evolutionary branching due to local competition and dispersal in complex heterogeneous landscapes. These landscapes combine continuously varying patchiness and a nonlinear ecological gradient, thus reconciling and extending the “continuous linear gradient” and “discrete metapopulation” approaches of previous models. Environmental heterogeneity in our model has two components (Figure 1; Online Appendix: *Landscape generation*). The first component is an environmental gradient, described by both a linear coefficient, here called the gradient slope, and a quadratic coefficient, here called the gradient

curvature. The second component is spatially continuous patchiness, described by its amplitude and its spatial scale (autocorrelation length). Together, these components produce a wide variety of landscapes with patterns of heterogeneity reminiscent of real environments. In the following four sections we present four hypotheses regarding the likelihood that evolutionary branching will occur in such complexly heterogeneous landscapes.

Intermediate heterogeneity maximizes branching propensity

Although negative frequency-dependent selection due to competition can cause divergence even in the absence of environmental heterogeneity (Dieckmann and Doebeli 1999; Doebeli and Dieckmann 2000), some degree of spatial environmental heterogeneity must exist for local adaptation to produce divergence as a result of external environmental effects (Doebeli and Dieckmann 2003). In this way, heterogeneity promotes branching by providing divergent selection among environments. However, if heterogeneity is too pronounced, maladapted colonizer populations will be unable to persist long enough to adapt, and so organisms will be restricted to those areas to which they are already well-adapted. An intermediate level of heterogeneity is therefore expected to produce maximal divergence among locally adapted populations, and thus to maximally promote evolutionary branching. Doebeli and Dieckmann (2003) demonstrated this phenomenon in a model involving spatial heterogeneity due to a linear environmental gradient; furthermore, their study suggested that the stringency of this optimum depends on dispersal distance, with short-range dispersal allowing branching to occur over a broader range of gradient slopes (their Figure 3b). We hypothesize that this finding will generalize to other types of heterogeneity; specifically, evolutionary branching will be maximized also at an intermediate nonlinear gradient curvature, and at an intermediate amplitude

of spatially continuous patchiness, with the stringency of these optima increasing with increasing dispersal distance.

Additivity of heterogeneity types

If different types of heterogeneity have similar effects on evolutionary branching, each with its own intermediate optimum that maximizes branching, then such different types might be expected to have additive effects. For example, if the amount of heterogeneity due to gradient slope is already optimal, such that adding more heterogeneity by increasing the slope decreases the likelihood of branching, then it seems reasonable to think that adding other types of heterogeneity, such as gradient curvature, would also decrease branching. Furthermore, this decrease in branching might be compensated for by reducing the gradient slope as curvature is increased. We thus hypothesize that the maximal branching propensity should be achievable through an essentially additive mixture of different types of heterogeneity (gradient slope, gradient curvature, patchiness). An increase in one type would be compensated by a decrease in another type, to the extent that such additivity applies.

Refugium effect

As explained above (*Additivity of heterogeneity types*), different types of heterogeneity may be additive in their effects on branching; however, we expect one major exception to this general principle of additivity. This exception arises because the lethality of a given high level of environmental heterogeneity, and its concomitant negative effect on evolutionary branching, might be lower if some of that heterogeneity is unevenly distributed in the form of spatially continuous patchiness. In particular, such patchiness might provide refugia in areas of an environmental gradient that would otherwise be too lethal to colonize. These refugia might

promote evolutionary branching through several complementary mechanisms. First, a refugium might provide an ecologically distinct patch of sufficient size to allow diversification *in situ*, despite the harshness of the underlying steep gradient. Second, a refugium might act as a spatial stepping stone, providing a clement habitat patch that could act as a proximate source for repeated colonization attempts into the surrounding inhospitable environment (Havel et al. 2005; Thomas et al. 2012; Travis et al. 2012). Third, a refugium might act as an adaptive stepping stone, providing an environment intermediate between the clement home range and a destination too inhospitable to colonize directly; the refugium would thus allow partial adaptation to occur prior to further range expansion (Havel et al. 2005; Lombaert et al. 2010; Heinicke et al. 2011; Westley et al. 2013).

For all these reasons, we hypothesize that spatially continuous patchiness should promote evolutionary branching particularly strongly when combined with an environmental gradient so steep as to ordinarily hinder branching. In other words, a positive interaction between patchy heterogeneity and steep gradients should exist, mitigating or even opposing the additivity that would otherwise be expected between these different types of heterogeneity.

Autocorrelation length and dispersal

The effects of patchiness previously discussed – both the refugium effect and, more generally, the promotion of branching due to patchiness – should depend on the size of patches relative to the typical dispersal distance. In particular, a population that disperses over a large area containing many patches would adapt to conditions over the whole area, whereas populations with shorter dispersal distance might divergently adapt to local conditions within particular patches (Hovestadt et al. 2001; Bolker 2003; Snyder and Chesson 2003; Bolker 2010; Bonte et al. 2010; Hanski and Mononen 2011; Richardson 2012). If dispersal distance is so small relative

to patch size that populations effectively never disperse outside their local patch, however, then colonization of the larger landscape would be prevented, and the heterogeneity of the landscape would no longer promote divergence. We thus hypothesize that the ratio of the autocorrelation length (ACL) of environmental heterogeneity to the typical dispersal distance should affect the likelihood of evolutionary branching. As this ratio increases, the scale of patchiness relative to dispersal should increasingly promote evolutionary branching by allowing local adaptation to individual patches, up to some optimum ratio at which branching is maximized. Beyond that optimum, branching should be hindered due to a decrease in the colonization of new patches.

Empirical landscapes and evolutionary branching

Together, these four hypotheses suggest that metrics of heterogeneity could be taken from a real landscape and used to quantitatively predict its propensity for branching. To test the feasibility of this idea, we define simple, empirically accessible metrics (“realized-landscape metrics”) that describe the salient features of our modeled landscapes. Focusing on ecological divergence – the first step of ecological speciation – by considering evolutionary branching in asexual populations, we use these metrics to quantitatively predict the branching propensity of our landscapes. This method, which could be easily applied to real landscapes using, e.g., satellite data and known biodiversity patterns, provides the beginnings of a bridge between theoretical and empirical work on biodiversity and speciation. We end by discussing the implications of these findings to conservation and management.

Methods

Model summary

To test our hypothesis, we constructed a spatially explicit, stochastic, continuous-time individual-based model, summarized here (for the full model description, see Online Appendix: *Model description*). All parameters of the model are given in Table 1.

Our model is derived from the asexual model of Doebeli and Dieckmann (2003), who explored the propensity for evolutionary branching in two-dimensional landscapes in which the optimum phenotype varied across space following a linear environmental gradient. With the present model, we explore further types of spatial environmental heterogeneity through the addition of a nonlinear environmental gradient and continuously varying environmental patchiness (Introduction: *Complex landscapes*). The heterogeneity of the landscape is governed by four model parameters: s , the linear slope of the environmental gradient; c , the nonlinear “curvature” of the gradient; l_g , the autocorrelation length of the continuously varying patchiness; and a , the amplitude of that patchiness (Figure 1; Online Appendix: *Landscape generation*). Generated landscapes are ultimately described through a tabulated function $u_0(x, y)$ specifying the optimum phenotype at each location. Landscapes are always periodic in the y -direction, while in the x -direction we model four possible boundary conditions: stopping, reflecting, and absorbing, with their usual meanings, and reprising, which entails redrawing invalid locations until a valid location is obtained. The topology of the landscapes is thus cylindrical.

The environmental landscape is inhabited by a population of individuals, with their abundance regulated by a maximum carrying capacity density K_0 . Individuals are each characterized by a quantitative ecological trait u and a spatial location (x, y) . Their fitness

depends on their degree of local adaptation as determined by the difference between u and the locally optimum phenotype $u_0(x, y)$. Additionally, the fitness of each individual is decreased through competition from the other individuals. The strength of competition depends on both the spatial and the phenotypic distances between the focal individual and its competitors. The death rate of each individual is governed by its fitness, while the birth rate is constant for all individuals.

Individuals reproduce asexually, with each offspring inheriting its ecological trait nearly faithfully from its parent. Dispersal in the model is natal; offspring are displaced from their parents' location by a random distance.

The full model description (Online Appendix: *Model description*) discusses these and other aspects of the model in detail. A movie of the model is provided as Movie 1 (available on Dryad, <http://dx.doi.org/10.5061/dryad.43cj7/1>).

Model realizations

A total of 1,500,000 realizations of the model were generated. One set of 300,000 realizations was generated for each of the four boundary conditions, and an additional set of 300,000 realizations was generated with reprising boundary conditions to serve as an independent test dataset. Each realization was supplied with random values for the five governing parameters s , c , l_g , a , and v , with the value for each parameter drawn from a uniform distribution spanning the range of values for that parameter (Table 1).

For each realization, a landscape was generated (Online Appendix: *Landscape generation*) and analyzed to determine its standard deviation σ , skewness γ_1 , kurtosis γ_2 , and autocorrelation length l_r , together called the “realized-landscape metrics” (Online Appendix:

Landscape analysis). The initial state was then constructed and its evolution was traced for 5000 generations. A histogram of the ecological trait values in the population was recorded every 10 generations (Online Appendix: *Model description*, OBSERVABLES). Branching of the population was identified from these histograms as a stable divergence from the initially unimodal phenotypic distribution into a bi- or multi-modal distribution. Finally, a full census of the population was recorded at the end of the realization. A typical model realization is shown in Figure 2.

Statistical analysis

For analysis of the results of these realizations, binomial (logistic) Generalized Additive Models (GAMs; Wood 2006) were fitted. The independent variables in these GAMs were the scaled parameters s , c , l_g , a , and v . The dependent variable was taken to be whether the realization branched, encoded as 0 (not branched) or 1 (branched). This choice was motivated by the clear unimodality of branching times; realizations typically either branched early or did not branch at all (Online Appendix: *Distribution of branching times*). Nested models including interactions of the independent variables were constructed (Table A1, Figure A1). In the full model, the hypothesis of an intermediate optimum for each type of heterogeneity, with an interaction of that optimum with dispersal, was represented by $s*v$, $c*v$, and $a*v$ terms. The hypotheses of additivity between heterogeneity types and of deviation from additivity due to the refugium effect was represented by the three-way interaction $s*c*a$ and its component interactions $s*c$, $s*a$, and $c*a$. Finally, the hypothesis regarding the importance of the ratio of the dispersal distance to the autocorrelation length (ACL) was represented by l_g*v .

Similarly, logistic GAMs were also fitted using the realized-landscape metrics σ , γ_1 , γ_2 , l_r , and v , with the same dependent variable (not branched vs. branched). Nested models including interactions of the independent variables were constructed (Table A2, Figure A2). In the full model, the hypothesis that an intermediate optimum for the overall realized-landscape heterogeneity would exist, and would be affected by dispersal distance, was represented by $\sigma * v$, and the hypothesis regarding the importance of the ratio of the dispersal distance to the realized-landscape ACL was represented by $l_r * v$.

The Bayesian Information Criterion (BIC; Schwarz 1978) was used to select the best model from each of these two sets of nested models. Several other metrics were also calculated for each model: the Nagelkerke R^2 , a standard substitute for R^2 in the context of logistic regression (Nagelkerke 1991); C , the area under the Receiver Operating Characteristic curve (Metz 1978); D , the Discrimination coefficient (Tjur 2009); and the proportion of realizations in which the outcome (branched or unbranched) was correctly predicted by the model.

To fit the GAM models, we used the `gam()` function in the `mgcv` package of R (version 1.7-20; Wood 2006). Tensor product smooths, `te()`, were used for both variables and interactions to ensure proper nesting (S. Wood, pers. comm.). Restricted maximum likelihood (REML) was used for fitting because it is the best method with a binary dependent variable (S. Wood, pers. comm.). Apart from these choices, default values were used for all fitting parameters.

Visualizations of two-dimensional slices through the five-dimensional GAM spaces were generated using a modified version of the `vis.gam()` function of `mgcv`. These slices show the predicted probability of branching for given combinations of all five independent variables of the GAM (given that three variables are held constant across a slice while varying the other two). Since only these slices are shown, caution must be used in interpreting the patterns observed.

However, results presented are qualitatively robust to variation of the parameters fixed for each slice except where otherwise noted.

To confirm that our hypotheses encompassed all important interactions among the parameters varied, we conducted *ex post facto* data exploration using logistic Generalized Linear Models (GLMs). GLMs were used instead of GAMs for this purpose because fitting GAMs that included every possible interaction, up to the full five-way interaction of all parameters, was computationally infeasible. This exploration indicated that other interactions were of very small effect size, and the prediction rate and Nagelkerke R^2 for these GLMs was never as high as for the best GAMs; results from these GLMs are thus not presented here.

Results

A dataset containing the parameter values, realized-landscape metrics, and outcomes for all realizations is published on Dryad (Haller et al. 2013, <http://dx.doi.org/10.5061/dryad.43cj7>), with R code to fit the GAMs presented. End-of-realization censuses and per-generation results comprise approximately 500 GB, and the generated landscapes comprise several thousand GB; because these sizes greatly exceed Dryad's 10 GB dataset limit, online provision of these data was not possible.

Reprising boundaries constitute a middle ground between the extremes of stopping boundaries, which mostly promote branching by generating disruptive selection, and absorbing boundaries, which generally inhibit branching by generating stabilizing selection (Mazzucco et al., unpublished manuscript). For this reason, reprising boundaries were used for all main results, to minimize the influence of the boundary condition on the results. Results for the other

boundary conditions are qualitatively similar, and are presented in the Online Appendix: *Effects of boundary conditions* (Table A3 and Figures A7–A10).

Generalized additive models

Of the 300,000 main realizations using reprising boundary conditions, 64,858 (21.6%) exhibited evolutionary branching. We fit two nested sets of models to the results of these realizations, using, in addition to the dispersal parameter v , either the landscape-generating parameters (s , c , l_g , a ; Table A1, Figure A1) or the realized-landscape metrics (σ , γ_1 , γ_2 , l_r ; Table A2, Figure A2), as described in Methods: *Statistical analysis*.

The best GAM model using the landscape-generating parameters (henceforth “GAM_g”) was the full model including all interactions predicted by our hypotheses,

$$\textit{branched} \sim s + c + l_g + a + v + s * v + c * v + a * v + l_g * v + s * c + s * a + c * a + s * c * a .$$

GAM_g, with 13 terms and 167.0 effective degrees of freedom, provided a fit with a Nagelkerke R^2 of 0.642, and correctly predicted the outcome of 89.5% of realizations, as compared to the null model prediction rate of 78.4% (Table A1). GAM_g also correctly predicted 89.6% of the outcomes of the 300,000 realizations in the separate test dataset, indicating that overfitting did not occur.

The best GAM using the realized-landscape metrics (henceforth “GAM_r”) was also the full model including all predicted interactions,

$$\textit{branched} \sim \sigma + \gamma_1 + \gamma_2 + l_r + v + \sigma * v + l_r * v .$$

GAM_r, with only 7 terms and 49.3 effective degrees of freedom, provided a fit with a Nagelkerke R^2 of 0.681, and correctly predicted the outcome of 90.6% of realizations, as

compared to the null model prediction rate of 78.4% (Table A2). GAM_r also correctly predicted 90.7% of the test dataset outcomes, again indicating that overfitting did not occur.

Although GAM_g and GAM_r were chosen on the basis of their BIC scores, the other model metrics assessed (Nagelkerke R^2 , C , D , prediction accuracy; see Methods: *Statistical analysis*) were also optimized by this choice (Tables A1 and A2, Figures A1 and A2). For further discussion see Online Appendix: *Comparison of nested GAM models*.

Intermediate heterogeneity maximizes branching propensity

For all types of heterogeneity explored, branching propensity was maximized at some particular level of heterogeneity, showing a “hump-shaped” relationship. In particular, an intermediate slope was found to maximally promote branching (optimum $s \approx 0.75$; Figure 3A), and likewise for an intermediate curvature (optimum $c \approx 0.25$; Figure 3B). An intermediate patchiness amplitude similarly appeared to maximally promote branching (optimum $a \approx 2.75$; Figure 3C), but this effect depended on the boundary conditions, being most apparent with absorbing boundaries (Figure A7C) and least apparent for stopping boundaries (Figure A7O). For reflecting and reprising boundaries, the optimum value appeared to be close to the edge of the parameter space explored (Figure A7G and A7K). Branching was also maximally promoted at an intermediate level of overall landscape heterogeneity, as measured by the standard deviation of the realized landscape, σ (optimum $\sigma \approx 2.0$ for $v \lesssim 0.5$; Figure 3D). The probability of branching always decreased with increasing dispersal distance, but dispersal distance had little effect on the optimal values of s , c , and a (Figures 3A–C). The optimum σ decreased with increasing dispersal distance, reaching zero for $v \gtrsim 0.5$ (Figure 3D), although this effect varied somewhat among boundary conditions (Figure A7).

Additivity and the refugium effect

The different types of heterogeneity also interacted in their effects on branching, making the complete picture somewhat more complex. Slope and curvature showed simple additivity (Figure 4A). This was not the case, however, for the combination of slope and patchiness amplitude (Figure 4B); here, as amplitude increased, the sensitivity to slope decreased, indicating less than full additivity between slope and amplitude. This effect was even more pronounced for curvature and amplitude (Figure 4C); here for $c \gtrsim 0.4$ the probability of branching actually increased with increasing amplitude, up to a point, indicating a net positive interaction between large curvature and amplitude.

Autocorrelation length and dispersal

Landscapes with a larger realized autocorrelation length l_r produced a higher probability of branching (Figure 5). An interaction between the effects of l_r and dispersal distance ν was also observed: the branching probability decreased more strongly with increasing ν when l_r was larger (Figure 5). The generating autocorrelation length l_g had little effect (see Online Appendix: *Landscape-generating parameters versus realized-landscape metrics*).

Other results

We had no specific hypotheses regarding the role that the realized-landscape skewness, γ_1 , and kurtosis, γ_2 , would play. For this reason, although substantial and significant effects were observed for both (Table A2 and Figure A10), they are presented in the Online Appendix: *Effects of skewness and kurtosis*.

Analysis using the alternative cluster-based assessment of branching (Online Appendix: *Model description*, OBSERVABLES) produced results very similar to those presented, which suggests that our findings are robust to variations in the method used to assess branching (Online Appendix: *Alternative assessment of branching*).

Branching times for realizations that branched were typically less than 1000 generations, following an approximately lognormal distribution, and showed little correlation with the predicted probabilities of branching given by the GAM models (Figure A11, Online Appendix: *Distribution of branching times*). Branching was thus generally binary: it happened early or not at all, for most realizations. Nevertheless, some effects of parameters on branching time were observed (Figures A12 and A13, Online Appendix: *Correlations with branching times*).

Although branches occasionally went extinct, merging of branches, or “reverse speciation” (Seehausen 2006), was not observed. Furthermore, the phenotype of unbranched lineages often drifted over time, while branched lineages appeared to be stabilized by competition and typically exhibited evolutionary stasis after arriving at equilibrium. These and other qualitative observations are presented further in Figures A14 and A15 and in Online Appendix: *Patterns of evolutionary branching*.

Discussion

Here we have shown that complex spatial heterogeneity affects evolutionary branching in several ways previously unexplored by theoretical models. An intermediate optimum level of environmental heterogeneity exists that maximally promotes branching, and such an intermediate optimum also appears to exist for particular types of heterogeneity – linear gradients, nonlinear gradients, and spatially continuous patchiness. The effects of these different types of

heterogeneity are additive to some extent, but linear gradients of an optimal slope are more effective in promoting branching than are the other types of heterogeneity investigated. A refugium effect also causes a departure from additivity of different types of heterogeneity because patchiness can break up an otherwise hostile environment, leading to a higher branching propensity for landscapes than would otherwise be expected given the overall level of heterogeneity present. The scale of patchiness relative to the dispersal distance of organisms is also important: large patch size relative to dispersal distance isolates populations in distinct areas, promoting local adaptation.

Furthermore, our results allow us to quantitatively and accurately predict the branching propensity of heterogeneous landscapes through empirical metrics that might often be easily obtained. This method could be used to predict the branching propensity of real landscapes. The effects of complex spatial environmental heterogeneity might help explain some contentious observations, such as the higher biodiversity of the tropics, differences in the speciosity of parapatric sister clades, the causes of adaptive radiation (or the lack thereof) in different environments, and the evolutionary effects of differences in dispersal behavior among species. This connection between our model and real landscapes also has potential implications for conservation.

Intermediate heterogeneity maximizes branching propensity

Previous research indicates that an intermediate environmental gradient slope maximally promotes evolutionary branching (Doebeli and Dieckmann 2003; Heinz et al. 2009): very shallow slopes provide little opportunity for adaptive divergence, while very steep slopes are likely to be lethal for dispersers and provide little habitable space in which a colony might settle.

Our model confirmed this result (Figure 3A), with good agreement to previous findings (Doebeli and Dieckmann 2003, their Figure 3b).

Going beyond previous results, we also observed an intermediate optimum for the degree of curvature of an environmental gradient (Figure 3B), and for the overall measured heterogeneity of the landscape (Figure 3D). These findings support our hypothesis that an intermediate optimum level should exist for other types of heterogeneity (Introduction: *Intermediate heterogeneity maximizes branching propensity*). In fact, the argument that steep slopes are lethal for colonizers readily generalizes to other kinds of heterogeneity, since the area of suitable habitat for each species shrinks as heterogeneity increases (Allouche et al. 2012). The existence of an intermediate optimum for the amplitude of spatially continuous patchiness depended on the boundary condition, however, and was not always apparent (Figure A7). This may be because a very high level of patchiness can have both negative effects on speciation, by hindering colonization, and also positive effects on speciation, by providing obstacles that increase the geographic isolation of populations and thus promote divergence (Golestani et al. 2012). Separating these competing effects of heterogeneity would be an interesting direction for future research.

Additivity and the refugium effect

If there is an optimum overall level of spatial environmental heterogeneity, as discussed in the previous section, then it might be natural to expect particular types of heterogeneity – gradient slope, gradient curvature, and patchiness amplitude, in this study – to be additive in their effects (Introduction: *Additivity of heterogeneity types*). This effect was observed for the interaction of slope and curvature (Figure 4A). If both are too low, then the optimum level of heterogeneity is not attained, and branching is hindered. Increasing either one to its optimal level maximizes

branching due to that type (although curvature is less effective than slope in promoting branching even at its optimum level, presumably because curvature can only optimize a fraction of the whole landscape for branching). Finally, mixtures of both slope and curvature exhibit nearly straight, diagonal isoclines indicating a high degree of additivity. This supports our hypothesis that the overall magnitude of heterogeneity determines the branching propensity of the landscape, to a first approximation. However, other heterogeneity types showed more complex interactions, as we now discuss.

We also hypothesized that a refugium effect should cause a deviation from this additivity, for the combination of patchiness with an extreme gradient due to either slope or curvature. This effect is strongest for curvature, since increasing patchiness amplitude actually increases the branching propensity for all magnitudes of curvature substantially greater than the optimum curvature (Figure 4C). The effect is less pronounced for slope; here, increasing amplitude merely has less strong of an effect than would be expected from pure additivity, and at very steep slopes increasing amplitude has almost no effect at all (Figure 4B). The refugium effect, then, causes a positive, synergistic interaction between high curvature and patchiness, but merely acts to partially mitigate the expected additivity between high slope and patchiness. This difference might be because, without patchiness, only a small section of a curved gradient is likely to be hospitable; if patchiness can open up the entire landscape to colonization, the refugium effect may thus be quite large.

Because of the refugium effect, the probability of branching (let us here call this p) increased substantially with an increase in patchy heterogeneity in some cases – from less than 85% to more than 95% for the largest curvatures explored in our realizations with reprising boundaries (Figure 4C), and even more with absorbing and stopping boundaries (Figure A8). However, the

probability that branching will not occur, $1-p$, shows a proportionately much larger effect, falling from more than 15% to less than 5%. The odds ratio, $p/(1-p)$, is a standard metric that unifies these two perspectives; using it, we can see that in the best case the refugium effect more than tripled the odds in favor of branching, from about 5.5:1 to 19:1. Such a large change in odds could produce very large effects on the net biodiversification rate in empirical scenarios in which branching is not otherwise very likely to occur.

Autocorrelation length and dispersal

We also hypothesized that the effects of patchiness – both the refugium effect and the overall promotion of branching by patchy heterogeneity – should depend on the relative scales of dispersal and patchiness (Hovestadt et al. 2001; Bolker 2003; Snyder and Chesson 2003; Bolker 2010; Bonte et al. 2010; Hanski and Mononen 2011). Specifically, as the ratio of the autocorrelation length (ACL) of patchiness to the typical dispersal distance increases, local adaptation to individual patches should be increasingly favored (up to a point), promoting evolutionary branching (Introduction: *Autocorrelation length and dispersal*). This hypothesized interaction between ACL and dispersal was observed for the realized ACL, l_r (Figure 5). For small values of l_r , patches are too small for local adaptation to be possible, and so the dispersal distance is relatively unimportant. For larger l_r , however, local adaptation becomes possible as long as dispersal distance is sufficiently short-range; the longer the dispersal distance, the less likely branching is. For very large l_r and short-range dispersal, the isoclines flatten out, indicating an insensitivity to l_r ; patches in this case are large enough relative to dispersal that making them even larger has little effect on branching, because populations are already able to fully adapt to their local conditions. These findings demonstrate that for a patch to provide a

good site for local adaptation, it must be much larger than the individual dispersal distance, both because many individuals in a population will not be located at the center of the patch, and because the flat bottom of the patch that provides a relatively constant environment is much smaller than the overall patch size; for much of the radius of a patch, the environment is changing, with a gradient from inside-patch to outside-patch conditions.

This interaction between dispersal distance and ACL was not observed for the generating ACL, l_g (Figure A3B); indeed, l_g had almost no effect. This was expected, as the realized ACL is not highly correlated with the generating ACL due to stochasticity and system size constraints (Figure A3A), and because it also depends on the other landscape-generating parameters. For these reasons, l_g contains much less information than l_r , and is a poor predictor of evolutionary branching (Online Appendix: *Landscape-generating parameters versus realized-landscape metrics*).

Empirical implications

We have suggested that empirical studies might use our realized-landscape metrics to measure the heterogeneity of real landscapes, and then test the predictions of our model against the actual biodiversity of clades endemic to that landscape. This exciting opportunity to confront theory with data is, however, not without challenges.

Sampling a landscape at high resolution may be non-trivial for some environmental metrics not provided by existing resources such as satellite images (e.g., soil mineral concentrations; Yost et al. 2012). However, many environmental variables are typically correlated (e.g., temperature, altitude, and rainfall), and in some cases sophisticated methods exist for predicting unknown environmental variables across a landscape using such correlated proxies (McKenzie

and Ryan 1999). Studies might also compare the predictive value of all available metrics, or might compress multiple metrics into a single axis of environmental variation, using Principal Components Analysis (PCA). PCA might be particularly appropriate when local adaptation to a suite of correlated environmental variables is believed to have occurred; however, it should be used with caution, as the axis of greatest environmental variation might not correspond to the axis of variation to which a particular clade adapted.

Furthermore, clades should be chosen that match the biology modeled here (e.g., unbiased natal dispersal, asexual reproduction); further theoretical exploration will be needed before other radiations can be studied within this framework (see *Future directions*). We suggest that adaptation of asexual or clonal plants and fungi to serpentine soil outcrops might present an attractive opportunity, due to the many independent instances of speciation due to local adaptation to serpentine soils in heterogeneous landscapes (Brady et al. 2005; Harrison and Rajakaruna 2011). Experimental evolution of microbial systems in heterogeneous culture media might also be a worthwhile approach (Rainey and Travisano 1998; Bailey and Kassen 2012). There may also be cases in which historical data, perhaps from paleoclimatological reconstructions and dated phylogenies, are sufficiently complete to test our model's predictions for cases such as post-glacial radiations (Linder 2008) or adaptation to different islands (Givnish 2010). Finally, natural heterogeneity also commonly varies over time; to test our model's predictions, a landscape would have to be chosen that has been relatively invariant over the time in which speciation occurred (or this model would have to be generalized to spatiotemporal patterns of heterogeneity; see *Future directions*).

Natural heterogeneity might exhibit patterns not modeled here. For example, sudden transitions in habitat type may be superimposed on less extreme heterogeneity (e.g., ocean-to-

land transitions); the autocorrelation length may vary depending upon the spatial axis or the position in space (e.g., “basin-and-range” topography); or the landscape structure may not be reducible to a single autocorrelation length due to hierarchical structure (Kolasa et al. 2012). In such cases, model predictions could still be generated by replacing our generated landscapes by a representation of the actual landscape.

If environmental conditions that promote biodiversification also promote the maintenance of biodiversity (Rosenzweig 2001; Rosenzweig 2003), then our results have implications for conservation policy. Specifically, the loss of, or alteration of, patterns of environmental heterogeneity may lead to “reverse speciation” or the loss of locally adapted taxa (Seehausen et al. 1997; Templeton et al. 2001; Seehausen 2006; Seehausen et al. 2008; Crispo et al. 2011; De León et al. 2011; Vonlanthen et al. 2012). One implication is that even if a large reserve is ideal habitat for a given species, the variation within that species may not be maintained unless environmental heterogeneity is preserved to provide divergent selective pressures. A second implication is that refugia may provide stepping-stones, both spatially and adaptively (Introduction: *Refugium effect*), allowing more rapid and effective colonization and adaptation than would otherwise be possible in the face of anthropogenic disturbances such as climate change (Havel et al. 2005; Thomas et al. 2012). Indeed, assisted migration programs might do well to consider introducing species into carefully chosen refugia within a larger hostile environment, rather than into a larger area of compatible habitat, to maximally accelerate adaptation toward anticipated future conditions. The plausibility of this approach is illustrated by the importance of heterogeneity, refugia, and adaptation in the spread of invasive species (Havel et al. 2005; Lombaert et al. 2010; Heinicke et al. 2011; Tingley et al. 2012; Westley et al. 2013). Given this, our results might also help to predict biological invasions (Kolar and Lodge 2001),

extending approaches such as environmental niche modeling (Herborg et al. 2007) by adding consideration of dispersal and evolution.

Future directions

There are many future directions for this research that we expect to be fruitful. Because the present model is asexual with natal dispersal, it applies most strongly to selfing and asexual plants; extending the model to the sexual case would reveal the importance of gene flow due to hybridization (Dieckmann and Doebeli 1999; Doebeli and Dieckmann 2003). The need to establish assortative mating in the sexual case can hinder speciation, particularly when mate choice is based upon an ecologically neutral marker trait (Felsenstein 1981; Dieckmann and Doebeli 1999; Servedio et al. 2011). However, this negative effect of sexual reproduction on speciation has been shown to be mitigated greatly by effects of spatial distance and environmental heterogeneity (Doebeli and Dieckmann 2003). We expect this to apply also for complex spatial heterogeneity, mediated by the interaction between the relative spatial scales of heterogeneity and dispersal.

Similarly, the addition of different types of dispersal, such as non-natal dispersal, conditional dispersal, and habitat preference, would allow an exploration of the effects of complex spatial heterogeneity with dispersal behaviors more typical of animal species (Ronce 2007; Edelaar et al. 2008; Clobert et al. 2009; Payne et al. 2011; Webster et al. 2012). Allowing the evolution of dispersal rate (Doebeli and Ruxton 1997; Mathias et al. 2001), distance (Heinz et al. 2009), and kernel shape (Hovestadt et al. 2001; Bolker 2010) would also be worthwhile; these factors might interact with complex spatial heterogeneity in interesting ways, perhaps modifying the dynamics of the refugium effect with an effect of the dispersal kernel shape. Steep environmental gradients have previously been shown to promote the evolution of short dispersal distance (Heinz et al.

2009), and we expect that the same would likely apply to other types of heterogeneity, but this might inhibit branching on some landscapes, if it means that the refugium effect is diminished by decreased dispersal into refugia.

Our results indicate that branching, when it occurs, typically occurs immediately; however, some landscapes promote branching in a delayed fashion, perhaps because the population first has to attain a favorable spatial configuration on the landscape before branching can occur (see Online Appendix, *Distribution of branching times* and *Correlations with branching times*). This result underscores the need for a theory linking spatial structure to the expected waiting time to branching, a question that has barely begun to be explored (Orr and Orr 1996; Gavrilets et al. 2000b).

Allowing temporal environmental heterogeneity, in addition to spatial environmental heterogeneity, would connect this work to previous studies related to environmental change (e.g., Pease et al. 1989; Bürger and Lynch 1995; Kopp and Hermisson 2007; North et al. 2011), with particular relevance for predicting the effects of climate change for species occupying realistically heterogeneous landscapes. Finally, only one ecological dimension has been examined in this work; an exploration of the consequences of complex spatial heterogeneity in multiple ecological dimensions is needed (Nosil and Harmon 2009; Guillaume 2011; Birand et al. 2012). We hope to pursue these topics in future research.

Spatial heterogeneity, by producing divergent selective regimes that promote diversification, is an important driver of speciation. A complete understanding of the effects of complex, realistic patterns of spatial heterogeneity is therefore fundamental to an understanding of the origins of biodiversity. The model and results we have presented here represent a step toward that goal.

Acknowledgements

We thank Andrew Hendry for discussions, support, and comments on a previous version of the manuscript, and Georg Heinze for helpful discussions. BCH is supported by a National Science Foundation Graduate Research Fellowship under Grant No. 1038597. This research was conducted as part of BCH's participation in the Young Scientists Summer Program (YSSP) at the International Institute for Applied Systems Analysis (IIASA), with support from the National Academy of Sciences, thanks to NSF Grant No. OISE-738129. UD received financial support from the European Commission, the European Science Foundation, the Austrian Science Fund, the Austrian Ministry of Science and Research, and the Vienna Science and Technology Fund. We also thank T. Day, B. Bolker, and two anonymous reviewers for their valuable suggestions.

Literature Cited

- Allouche, O., M. Kalyuzhny, G. Moreno-Rueda, M. Pizarro, and R. Kadmon. 2012. Area-heterogeneity tradeoff and the diversity of ecological communities. *Proceedings of the National Academy of Sciences of the United States of America* 109:17495–17500.
- Bailey, S. F., and R. Kassen. 2012. Spatial structure of ecological opportunity drives adaptation in a bacterium. *American Naturalist* 180:270–283.
- Birand, A., A. Vose, and S. Gavrillets. 2012. Patterns of species ranges, speciation, and extinction. *American Naturalist* 179:1–21.
- Bolker, B. 2010. Evolution of dispersal scale and shape in heterogeneous environments: A correlation equation approach, Pages 231–249 in S. Cantrell, C. Cosner, and S. Ruan, eds. *Spatial Ecology*. Chapman & Hall/CRC, Boca Raton, FL.
- Bolker, B. M. 2003. Combining endogenous and exogenous spatial variability in analytical population models. *Theoretical Population Biology* 64:255–270.
- Bonte, D., T. Hovestadt, and H. J. Poethke. 2010. Evolution of dispersal polymorphism and local adaptation of dispersal distance in spatially structured landscapes. *Oikos* 119:560–566.
- Brady, K. U., A. R. Kruckeberg, and H. D. Bradshaw. 2005. Evolutionary ecology of plant adaptation to serpentine soils. *Annual Review of Ecology Evolution and Systematics* 36:243–266.
- Bürger, R., and M. Lynch. 1995. Evolution and extinction in a changing environment: A quantitative-genetic analysis. *Evolution* 49:151–163.

- Butlin, R., A. Debelle, C. Kerth, R. R. Snook, L. W. Beukeboom, R. F. C. Cajas, W. Diao et al. 2012. What do we need to know about speciation? *Trends in Ecology & Evolution* 27:27–39.
- Clobert, J., J. F. Le Galliard, J. Cote, S. Meylan, and M. Massot. 2009. Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology Letters* 12:197–209.
- Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Inc., Sunderland, Massachusetts.
- Crispo, E., J.-S. Moore, J. A. Lee-Yaw, S. M. Gray, and B. C. Haller. 2011. Broken barriers: Human-induced changes to gene flow and introgression in animals. *Bioessays* 33:508–518.
- Day, T. 2000. Competition and the effect of spatial resource heterogeneity on evolutionary diversification. *American Naturalist* 155:790–803.
- De León, L. F., J. A. M. Raeymaekers, E. Bermingham, J. Podos, A. Herrel, and A. P. Hendry. 2011. Exploring possible human influences on the evolution of Darwin's finches. *Evolution* 65:2258–2272.
- Debarre, F., and S. Gandon. 2010. Evolution of specialization in a spatially continuous environment. *Journal of Evolutionary Biology* 23:1090–1099.
- Dieckmann, U., and M. Doebeli. 1999. On the origin of species by sympatric speciation. *Nature* 400:354–357.
- Doebeli, M., and U. Dieckmann. 2000. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *American Naturalist* 156:S77–S101.
- . 2003. Speciation along environmental gradients. *Nature* 421:259–264.

- Doebeli, M., and G. D. Ruxton. 1997. Evolution of dispersal rates in metapopulation models: Branching and cyclic dynamics in phenotype space. *Evolution* 51:1730–1741.
- Edelaar, P., A. M. Siepielski, and J. Clobert. 2008. Matching habitat choice causes directed gene flow: A neglected dimension in evolution and ecology. *Evolution* 62:2462–2472.
- Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or Why are there so few kinds of animals? *Evolution* 35:124–138.
- Fischer, J., and D. B. Lindenmayer. 2006. Beyond fragmentation: The continuum model for fauna research and conservation in human-modified landscapes. *Oikos* 112:473–480.
- Gavrilets, S., R. Acton, and J. Gravner. 2000a. Dynamics of speciation and diversification in a metapopulation. *Evolution* 54:1493–1501.
- Gavrilets, S., H. Li, and M. D. Vose. 2000b. Patterns of parapatric speciation. *Evolution* 54:1126–1134.
- Gavrilets, S., and A. Vose. 2005. Dynamic patterns of adaptive radiation. *Proceedings of the National Academy of Sciences of the United States of America* 102:18040–18045.
- Geritz, S. A. H., and E. Kisdi. 2000. Adaptive dynamics in diploid, sexual populations and the evolution of reproductive isolation. *Proceedings of the Royal Society B-Biological Sciences* 267:1671–1678.
- Geritz, S. A. H., E. Kisdi, G. Meszner, and J. A. J. Metz. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology* 12:35–57.
- Givnish, T. J. 2010. Ecology of plant speciation. *Taxon* 59:1326–1366.

- Golestani, A., R. Gras, and M. Cristescu. 2012. Speciation with gene flow in a heterogeneous virtual world: Can physical obstacles accelerate speciation? *Proceedings of the Royal Society B-Biological Sciences* 279:3055–3064.
- Guillaume, F. 2011. Migration-induced phenotypic divergence: The migration-selection balance of correlated traits. *Evolution* 65:1723–1738.
- Haller, B. C., R. Mazzucco, and U. Dieckmann. 2013. Data from: Evolutionary branching in complex landscapes. *Dryad Data Repository* doi:10.5061/dryad.43cj7
- Hanski, I., and T. Mononen. 2011. Eco-evolutionary dynamics of dispersal in spatially heterogeneous environments. *Ecology Letters* 14:1025–1034.
- Harrison, S. P., and N. Rajakaruna. 2011. *Serpentine: The Evolution and Ecology of a Model System*. University of California Press, Berkeley, CA.
- Havel, J. E., C. E. Lee, and M. J. Vander Zanden. 2005. Do reservoirs facilitate invasions into landscapes? *Bioscience* 55:518–525.
- Heinicke, M. P., L. M. Diaz, and S. B. Hedges. 2011. Origin of invasive Florida frogs traced to Cuba. *Biology Letters* 7:407–410.
- Heinz, S. K., R. Mazzucco, and U. Dieckmann. 2009. Speciation and the evolution of dispersal along environmental gradients. *Evolutionary Ecology* 23:53–70.
- Hendry, A. P. 2009. Ecological speciation! Or the lack thereof? *Canadian Journal of Fisheries and Aquatic Sciences* 66:1383–1398.
- Herborg, L.-M., C. L. Jerde, D. M. Lodge, G. M. Ruiz, and H. J. MacIsaac. 2007. Predicting invasion risk using measures of introduction effort and environmental niche models. *Ecological Applications* 17:663–674.

- Hovestadt, T., S. Messner, and H. J. Poethke. 2001. Evolution of reduced dispersal mortality and ‘fat-tailed’ dispersal kernels in autocorrelated landscapes. *Proceedings of the Royal Society of London Series B-Biological Sciences* 268:385–391.
- Hutchinson, G. E. 1959. Homage to Santa Rosalia, or Why are there so many kinds of animals? *American Naturalist* 93:145–159.
- Irwin, D. E. 2012. Local adaptation along smooth ecological gradients causes phylogeographic breaks and phenotypic clustering. *American Naturalist* 180:35–49.
- Ispolatov, J., and M. Doebeli. 2009. Diversification along environmental gradients in spatially structured populations. *Evolutionary Ecology Research* 11:295–304.
- Kirkpatrick, M., and N. H. Barton. 1997. Evolution of a species’ range. *American Naturalist* 150:1–23.
- Kisdi, E., and S. A. H. Geritz. 1999. Adaptive dynamics in allele space: Evolution of genetic polymorphism by small mutations in a heterogeneous environment. *Evolution* 53:993–1008.
- Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: Predicting invaders. *Trends in Ecology & Evolution* 16:199–204.
- Kolasa, J., C. R. Allen, J. Sendzimir, and C. A. Stow. 2012. Predictions and retrodictions of the hierarchical representation of habitat in heterogeneous environments. *Ecological Modelling* 245:199–207.
- Kopp, M., and J. Hermisson. 2007. Adaptation of a quantitative trait to a moving optimum. *Genetics* 176:715–719.

- Leimar, O., M. Doebeli, and U. Dieckmann. 2008. Evolution of phenotypic clusters through competition and local adaptation along an environmental gradient. *Evolution* 62:807–822.
- Linder, H. P. 2008. Plant species radiations: Where, when, why? *Philosophical Transactions of the Royal Society of London B Biological Sciences* 363:3097–3105.
- Lombaert, E., T. Guillemaud, J.-M. Cornuet, T. Malausa, B. Facon, and A. Estoup. 2010. Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *Plos One* 5.
- Manning, A. D., D. B. Lindenmayer, and H. A. Nix. 2004. Continua and Umwelt: Novel perspectives on viewing landscapes. *Oikos* 104:621–628.
- Mathias, A., E. Kisdi, and I. Olivieri. 2001. Divergent evolution of dispersal in a heterogeneous landscape. *Evolution* 55:246–259.
- Mazzucco, R., M. Doebeli, and U. Dieckmann. (unpublished manuscript). The influence of habitat boundaries on evolutionary branching along environmental gradients.
- McKenzie, N. J., and P. J. Ryan. 1999. Spatial prediction of soil properties using environmental correlation. *Geoderma* 89:67–94.
- Meszéna, G., I. Czibula, and S. Geritz. 1997. Adaptive dynamics in a 2-patch environment: A toy model for allopatric and parapatric speciation. *Journal of Biological Systems* 5:265–284.
- Metz, C. E. 1978. Basic principles of ROC analysis. *Seminars in Nuclear Medicine* 8:283–298.
- Mizera, F., and G. Meszena. 2003. Spatial niche packing, character displacement and adaptive speciation along an environmental gradient. *Evolutionary Ecology Research* 5:363–382.

- Nagelkerke, N. J. D. 1991. A note on a general definition of the coefficient of determination. *Biometrika* 78:691–692.
- North, A., J. Pennanen, O. Ovaskainen, and A. L. Laine. 2011. Local adaptation in a changing world: The roles of gene-flow, mutation, and sexual reproduction. *Evolution* 65:79–89.
- Nosil, P., and L. J. Harmon. 2009. Niche dimensionality and ecological speciation, Pages 127–154 *in* R. Butlin, J. R. Bridle, and D. Schluter, eds. *Speciation and Patterns of Diversity*. Cambridge University Press, Cambridge, UK.
- Orr, H. A., and L. H. Orr. 1996. Waiting for speciation: The effect of population subdivision on the time to speciation. *Evolution* 50:1742–1749.
- Parvinen, K., and M. Egas. 2004. Dispersal and the evolution of specialisation in a two-habitat type metapopulation. *Theoretical Population Biology* 66:233–248.
- Payne, J. L., R. Mazzucco, and U. Dieckmann. 2011. The evolution of conditional dispersal and reproductive isolation along environmental gradients. *Journal of Theoretical Biology* 273:147–155.
- Pease, C. M., R. Lande, and J. J. Bull. 1989. A model of population growth, dispersal and evolution in a changing environment. *Ecology* 70:1657–1664.
- Rainey, P. B., and M. Travisano. 1998. Adaptive radiation in a heterogeneous environment. *Nature* 394:69–72.
- Richardson, J. L. 2012. Divergent landscape effects on population connectivity in two co-occurring amphibian species. *Molecular Ecology* 21:4437–4451.
- Ronce, O. 2007. How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology Evolution and Systematics* 38:231–253.

- Rosenzweig, M. L. 2001. Loss of speciation rate will impoverish future diversity. *Proceedings of the National Academy of Sciences of the United States of America* 98:5404–5410.
- . 2003. Reconciliation ecology and the future of species diversity. *Oryx* 37:194–205.
- Rundle, H. D., and P. Nosil. 2005. Ecological speciation. *Ecology Letters* 8:336–352.
- Schluter, D. 2000. *The Ecology of Adaptive Radiation: Oxford Series in Ecology and Evolution*. Oxford University Press, Oxford.
- . 2001. Ecology and the origin of species. *Trends in Ecology & Evolution* 16:372–380.
- . 2009. Evidence for ecological speciation and its alternative. *Science* 323:737–741.
- Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6:461–464.
- Seehausen, O. 2006. Conservation: Losing biodiversity by reverse speciation. *Current Biology* 16:R334–R337.
- Seehausen, O., G. Takimoto, D. Roy, and J. Jokela. 2008. Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular Ecology* 17:30–44.
- Seehausen, O., J. J. M. van Alphen, and F. Witte. 1997. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* 277:1808–1811.
- Servedio, M. R., G. S. Van Doorn, M. Kopp, A. M. Frame, and P. Nosil. 2011. Magic traits in speciation: ‘magic’ but not rare? *Trends in Ecology & Evolution* 26:389–397.
- Snyder, R. E., and P. Chesson. 2003. Local dispersal can facilitate coexistence in the presence of permanent spatial heterogeneity. *Ecology Letters* 6:301–309.
- Sutter, M., and T. J. Kawecki. 2009. Influence of learning on range expansion and adaptation to novel habitats. *Journal of Evolutionary Biology* 22:2201–2214.
- Templeton, A. R., R. J. Robertson, J. Brisson, and J. Strasburg. 2001. Disrupting evolutionary processes: The effect of habitat fragmentation on collared lizards in the Missouri Ozarks.

- Proceedings of the National Academy of Sciences of the United States of America
98:5426–5432.
- Thibert-Plante, X., and A. P. Hendry. 2009. Five questions on ecological speciation addressed with individual-based simulations. *Journal of Evolutionary Biology* 22:109–123.
- Thomas, C. D., P. K. Gillingham, R. B. Bradbury, D. B. Roy, B. J. Anderson, J. M. Baxter, N. A. D. Bourn et al. 2012. Protected areas facilitate species' range expansions. *Proceedings of the National Academy of Sciences of the United States of America* 109:14063–14068.
- Tingley, R., B. L. Phillips, M. Letnic, G. P. Brown, R. Shine, and S. J. E. Baird. 2012. Identifying optimal barriers to halt the invasion of cane toads *Rhinella marina* in arid Australia. *Journal of Applied Ecology*.
- Tjur, T. 2009. Coefficients of determination in logistic regression models—a new proposal: The coefficient of discrimination. *American Statistician* 63:366–372.
- Travis, J. M. J., K. Mustin, K. A. Barton, T. G. Benton, J. Clobert, M. M. Delgado, C. Dytham et al. 2012. Modelling dispersal: An eco-evolutionary framework incorporating emigration, movement, settlement behaviour and the multiple costs involved. *Methods in Ecology and Evolution* 3:628–641.
- Vonlanthen, P., D. Bittner, A. G. Hudson, K. A. Young, R. Muller, B. Lundsgaard-Hansen, D. Roy et al. 2012. Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature* 482:357–U1500.
- Webster, S. E., J. Galindo, J. W. Grahame, and R. K. Butlin. 2012. Habitat choice and speciation. *International Journal of Ecology* 2012:1–12.

- Westley, P. A. H., E. J. Ward, and I. A. Fleming. 2013. Fine-scale local adaptation in an invasive freshwater fish has evolved in contemporary time. *Proceedings of the Royal Society B-Biological Sciences* 280.
- Wood, S. N. 2006. *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC, Boca Raton, Florida.
- Yost, J. M., T. Barry, K. M. Kay, and N. Rajakaruna. 2012. Edaphic adaptation maintains the coexistence of two cryptic species on serpentine soils. *American Journal of Botany* 99:890–897.

Tables

Table 1: Model parameters and landscape metrics with symbols, value/range/formula, and units

	Symbol	Value	Units
Initial number of individuals	N_{init}	100	–
Maximum carrying capacity density; scales the number of individuals in the system	K_0	500	$1/L^2$
Standard deviation of carrying capacity density; scales the reduction in carrying capacity density due to the difference of an individual's ecological character from the locally optimal ecological character $u_0(x, y)$	σ_K	0.3	E
Standard deviation of the spatial competition function; scales the reduction in the strength of competition as the spatial distance between individuals increases	σ_s	0.19	L
Scaled standard deviation of the phenotypic competition function	c_p	5	–
Standard deviation of the phenotypic competition function; scales the reduction in the strength of competition as the phenotypic distance between individuals increases	σ_c	$c_p \cdot \sigma_K$	E
Probability of a mutation occurring	μ_m	0.005	–
Standard deviation of the mutation effect size	σ_m	0.05	E
Per capita birth rate; scales time in the model	b	1	T^{-1}
Scaled environmental gradient slope	s	0.0–2.0	–

Scaled environmental gradient curvature	c	0.0–1.0	–
Scaled environmental patchiness autocorrelation length	l_g	0.05–3.0	–
Scaled environmental patchiness amplitude	a	0.0–3.0	–
Scaled standard deviation of natal dispersal distance	v	0.01–3.0	–
Environmental gradient slope	S	$s \cdot \sigma_K / \sigma_s$	E/L
Environmental gradient curvature	C	$c \cdot \sigma_K / \sigma_s^2$	E/L ²
Environmental patchiness autocorrelation length	L_g	$l_g \cdot \sigma_s$	L
Environmental patchiness amplitude	A	$a \cdot \sigma_K$	E
Standard deviation of natal dispersal distance; scales the spatial distance offspring land from their parent	V	$v \cdot \sigma_s$	L
Realized-landscape heterogeneity scaled standard deviation	σ	$\sqrt{\mu_2} / \sigma_K$	–
Realized-landscape heterogeneity skewness	γ_1	μ_3 / σ^3	–
Realized-landscape heterogeneity excess kurtosis	γ_2	$\mu_4 / \sigma^4 - 3$	–
Realized-landscape heterogeneity scaled autocorrelation length	l_t	See text ^a	–

Note: Units are expressed using the symbols L (length), E (ecological phenotype), T (time), and – (dimensionless). The symbols μ_2 , μ_3 , and μ_4 represent the second, third, and fourth moments, respectively, of the distribution of values in the realized landscape^a.

^asee Online Appendix: *Landscape analysis*.

Figures

Figure 1: Examples of generated landscapes. We illustrate the effect of the four landscape-generating parameters on the pattern of spatial heterogeneity. A–D: Small values of parameters s , c , a , and l_g (scaled slope, scaled curvature, scaled amplitude, and scaled generating autocorrelation length), respectively. E–H: Larger values of the same parameters, so that each column shows the effect of a low versus high value of one parameter. Colors indicate locally optimal ecological trait values, ranging from white (low) to dark green (high). Black lines show locally optimal ecological trait values across one horizontal transect of each landscape. For purposes of comparison, all panels are generated with the same random number generator seed so that they are based on the same stochastic spatial structure. For each column, the other landscape-generating parameters are held to zero (except that to show the effect of the scaled amplitude, a particular autocorrelation length must be specified).

Figure 2: One realization of the model. A–C: A census of all individuals in the model after 50 (A), 2000 (B), and 5000 (C) generations (circles), superimposed on the generated landscape, which includes both a spatial gradient and spatially continuous patchiness. Background colors indicate locally optimal ecological trait values, while circle colors indicate the actual trait values of individuals, both ranging from white (low) to dark green (high). D: A plot of ecological trait values (y axis) through time. For each time, the plot shows the distribution of the ecological trait values of the individuals extant at that time in the model. Red lines indicate time points for panels A–C. Evolutionary branching can be observed at several points, and three branches exist at the end of execution, two of which are phenotypically similar; note also that one lineage went extinct after approximately 2600 generations.

Figure 3: Predicted branching probabilities from GAMs, showing effects of dispersal distance and different heterogeneity types. We show that an intermediate level of heterogeneity typically maximizes the branching probability. Panels show contour plots of the branching probability as a function of the scaled dispersal distance v and A: scaled slope s ; B: scaled curvature c ; C: scaled amplitude a ; and D: scaled landscape standard deviation σ . Red (white) indicates predicted low (high) branching probability (see color scale at right). The probability of branching is maximized at an intermediate value of s , c , and perhaps a (see text), regardless of the dispersal distance; an intermediate value of σ also maximizes branching for $v \lesssim 0.5$. Results shown are for reprising boundaries; results for other boundary conditions are qualitatively similar (Online Appendix: *Effects of boundary conditions*, and Figure A7). Other parameters: A–C: $s = 0$, $c = 0$, $a = 0$, and $l_g = 0.15$, D: $\gamma_1 = 0$, $\gamma_2 = 0$, and $l_r = 0.15$. Contour lines below 10% and above 90% are placed to best show the contours of the data in each panel.

Figure 4: Predicted branching probabilities from GAM_g , showing interactions between heterogeneity types. We show additive effects, and deviations therefrom due to the refugium effect, between scaled slope s , scaled curvature c , and scaled amplitude a . Panels show the branching probability as a function of A: s and c , demonstrating additivity between the two; B: s and a , demonstrating non-additive effects between the two, and a weak mitigation of the negative effect of steep slope on branching when combined with large amplitude; and C: c and a , demonstrating non-additive effects between the two, and a strong mitigation of the negative effect of high curvature on branching when combined with large amplitude. Red (white) indicates predicted low (high) branching probability (see color scale at right). Results shown are for reprising boundaries; results for other boundary conditions are qualitatively similar (Online

Appendix: *Effects of boundary conditions*, and Figure A8). Other parameters: $s = 0$, $c = 0$, $a = 0$, $l_g = 0.15$, and $v = 0.075$. Contour lines above 90% are placed to best show the contours of the data in each panel.

Figure 5: Predicted branching probabilities from GAM_r , showing interaction between dispersal distance and realized-landscape autocorrelation length (ACL). Contours show branching probability as a function of the scaled realized autocorrelation length l_r and the scaled dispersal distance v , demonstrating the importance of the relative scale of the two. Red (white) indicates predicted low (high) branching probability (see color scale at right). Branching probability decreases with increasing v , but increases with increasing l_r ; v and l_r also interact, with v having less effect when l_r is small. Results shown are for reprising boundaries; results for other boundary conditions are qualitatively similar (Online Appendix: *Effects of boundary conditions*, and Figure A9). Other parameters: $\sigma = 2.0$, $\gamma_1 = 0$, and $\gamma_2 = 0$. Contour lines above 90% are placed to best show the contours of the data.

Movie 1: Movie of the model realization depicted in Figure 2. Frames show the population census at every tenth generation, superimposed on the generated landscape, which includes both a spatial gradient and spatially continuous patchiness. Background colors indicate locally optimal ecological trait values, while circle colors indicate the trait values of individuals, both ranging from white (low) to dark green (high). Dryad, <http://dx.doi.org/10.5061/dryad.43cj7/1> (34.9 MB).

Figure 1

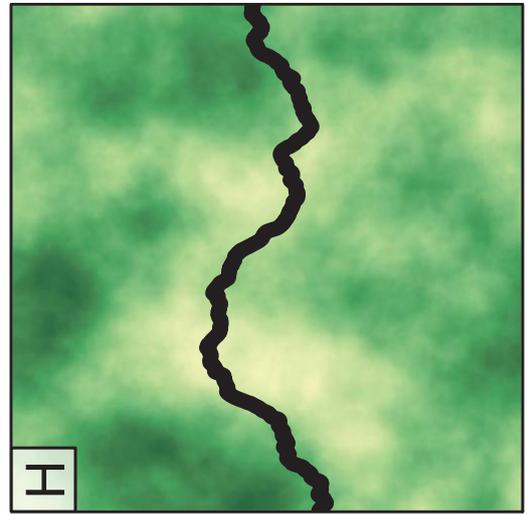
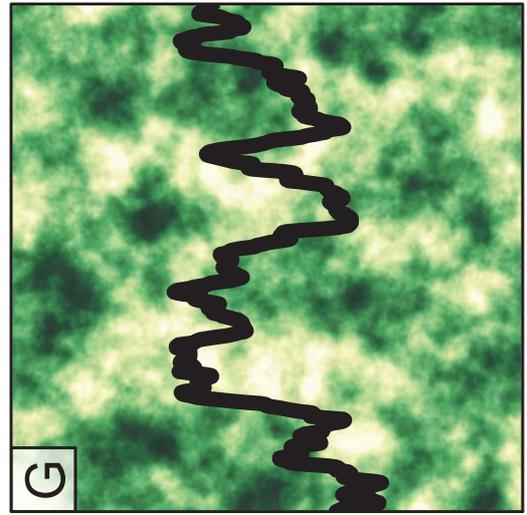
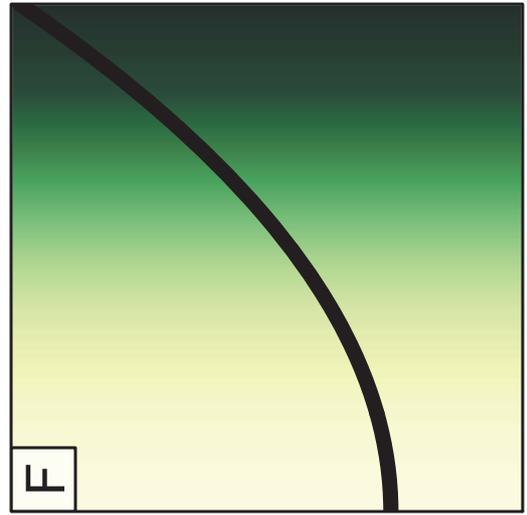
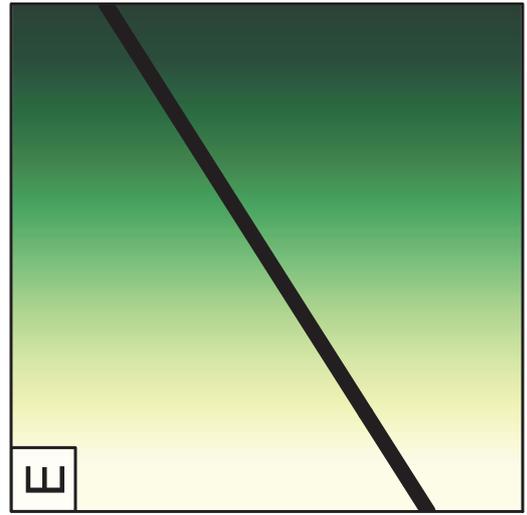
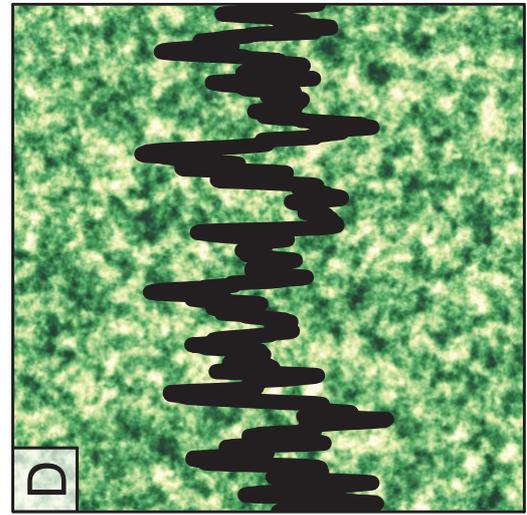
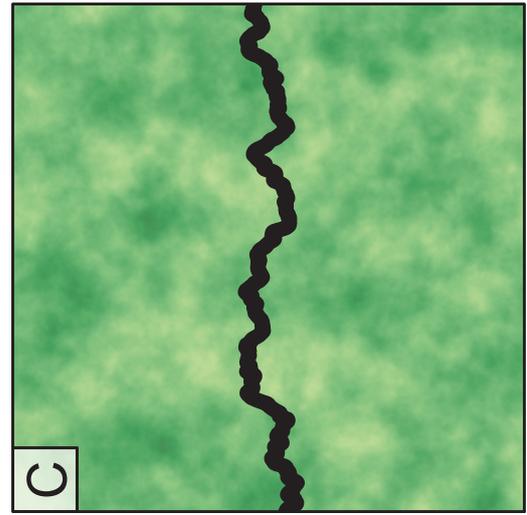
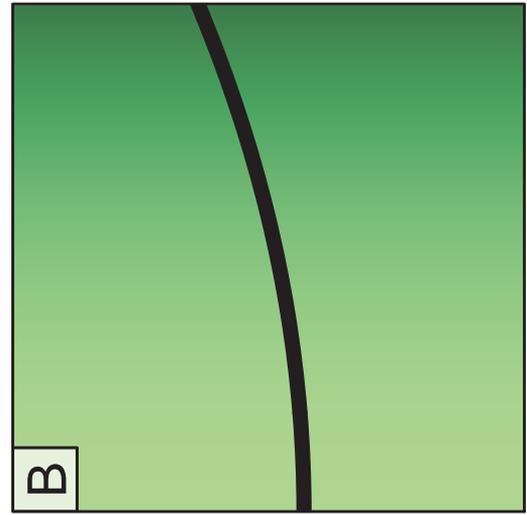
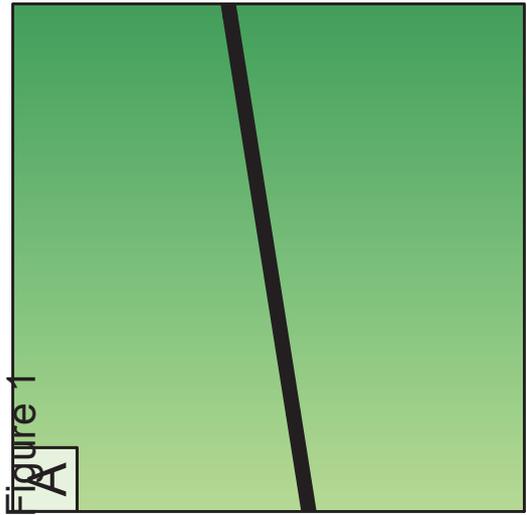
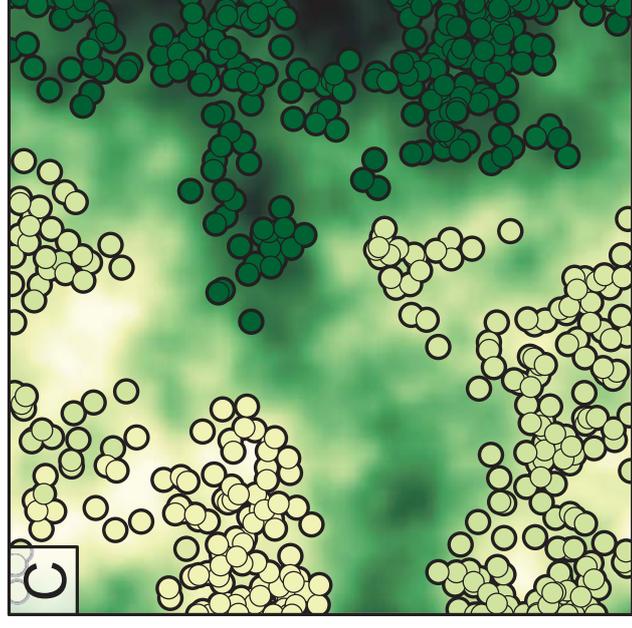
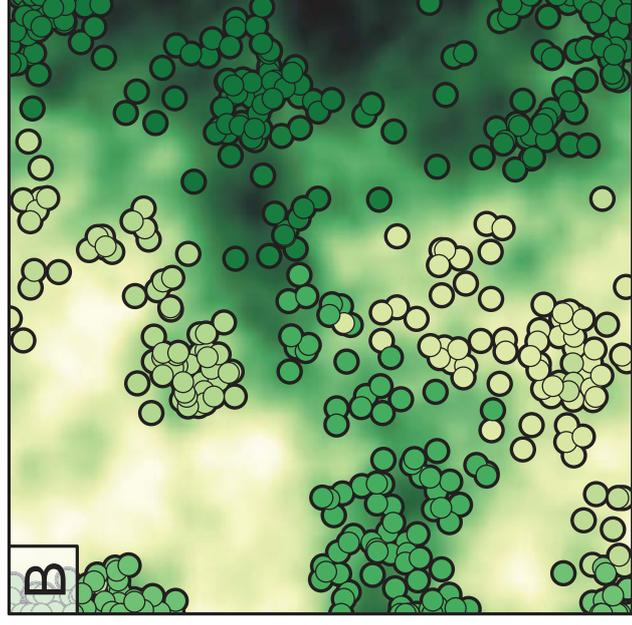
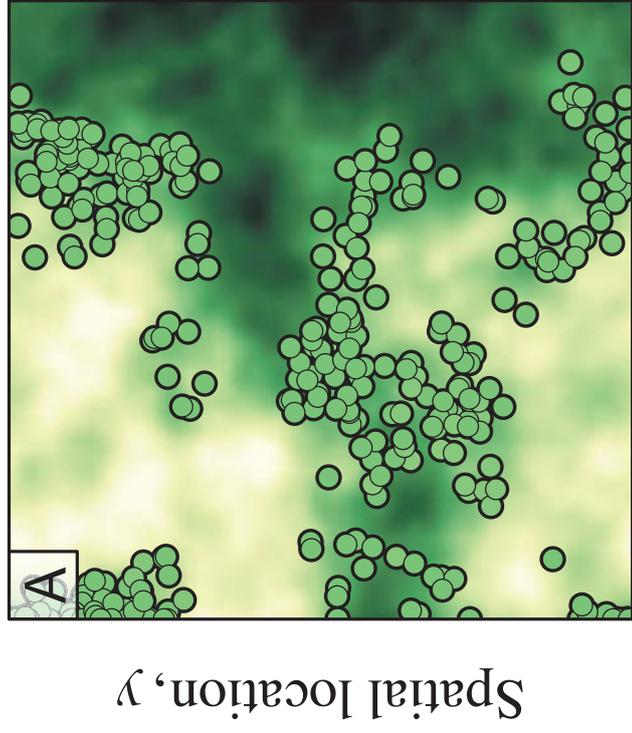


Figure 2

Spatial location, x



Spatial location, y

Ecological trait value, u

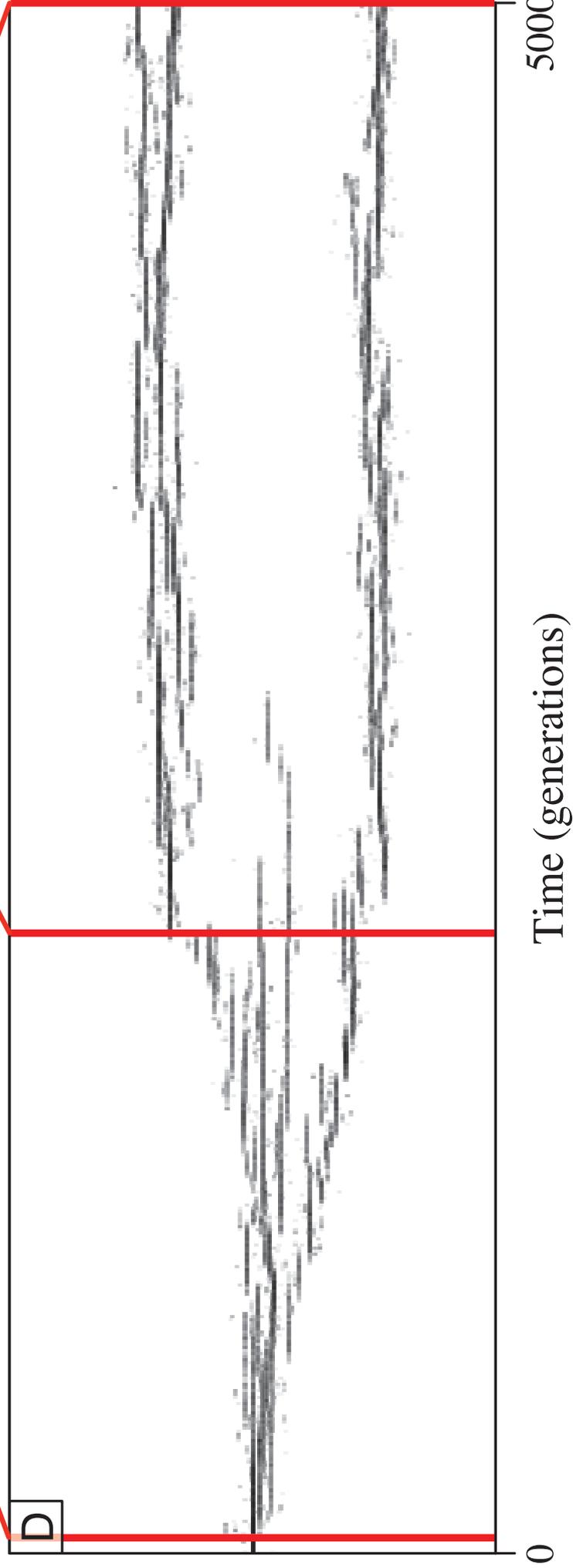


Figure 3

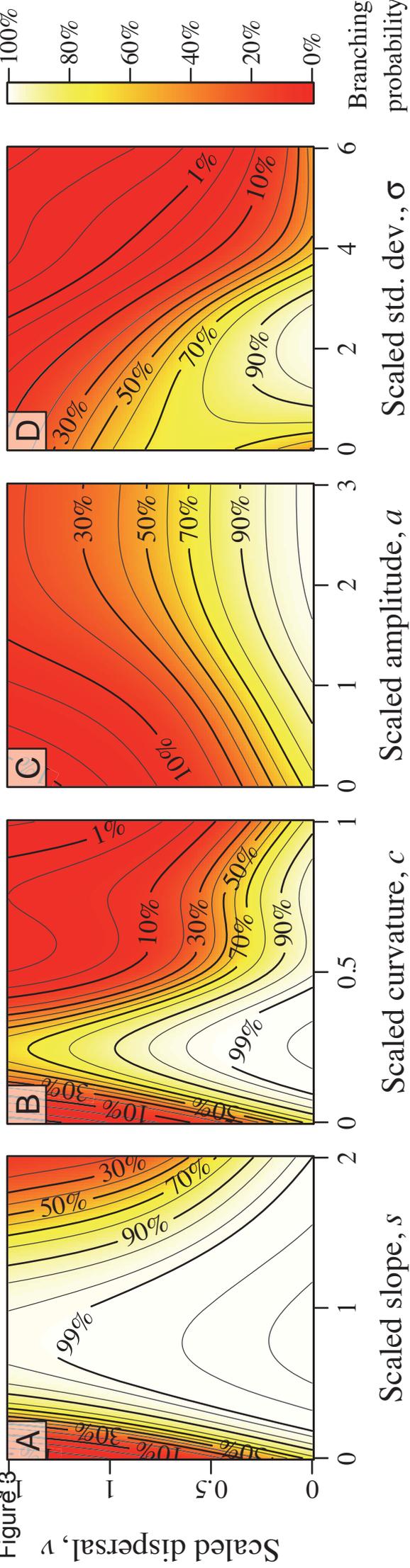
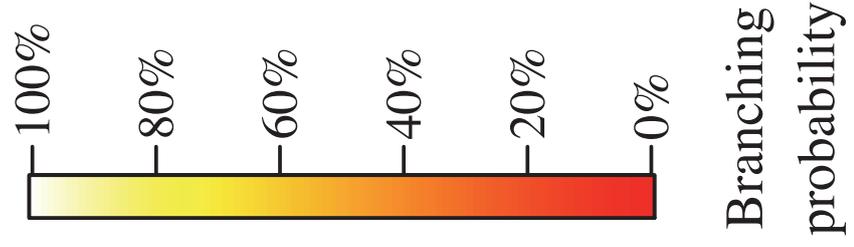
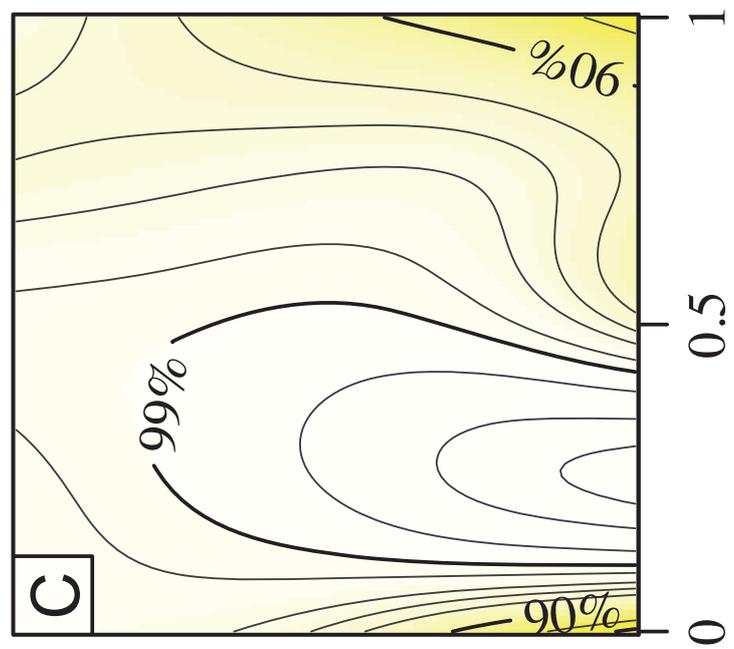
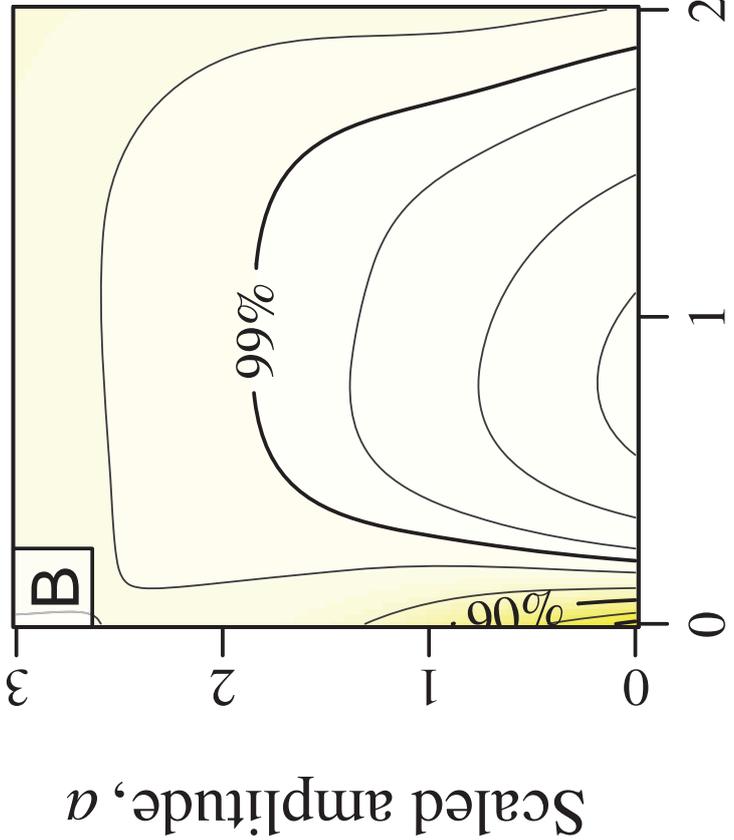
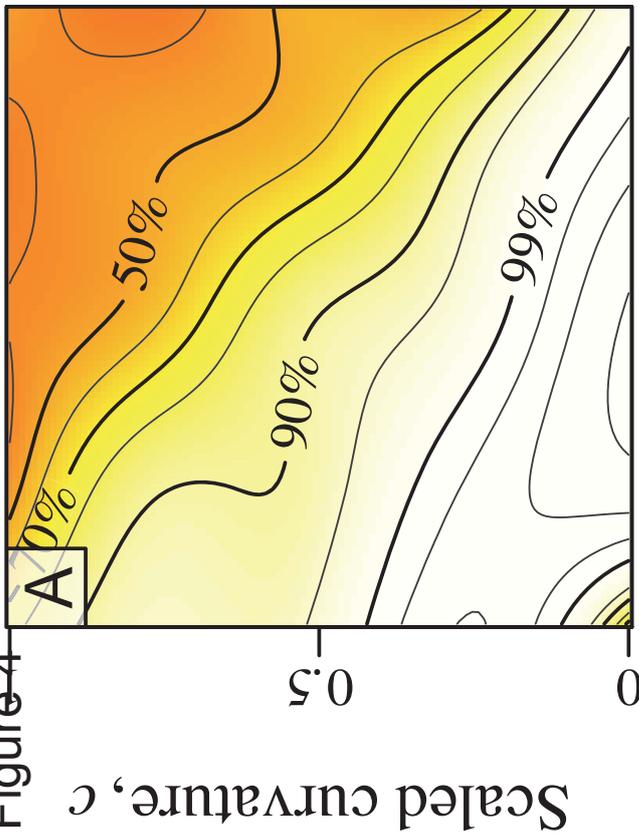


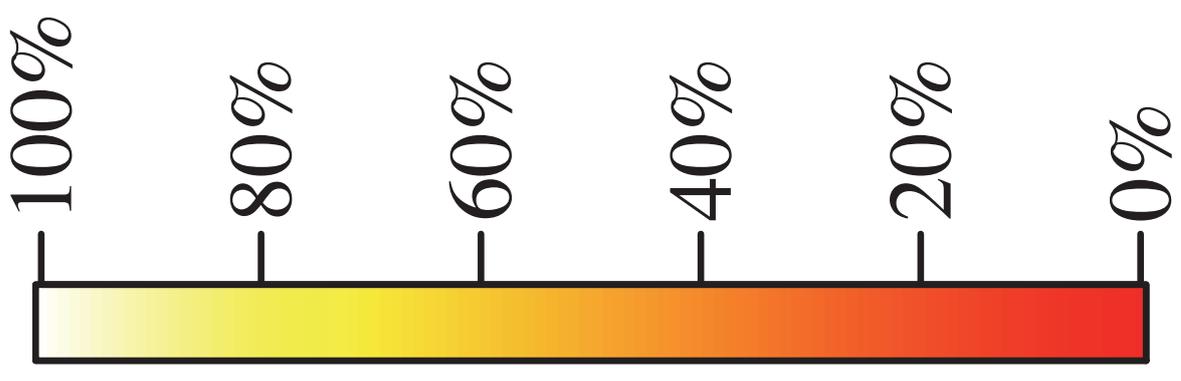
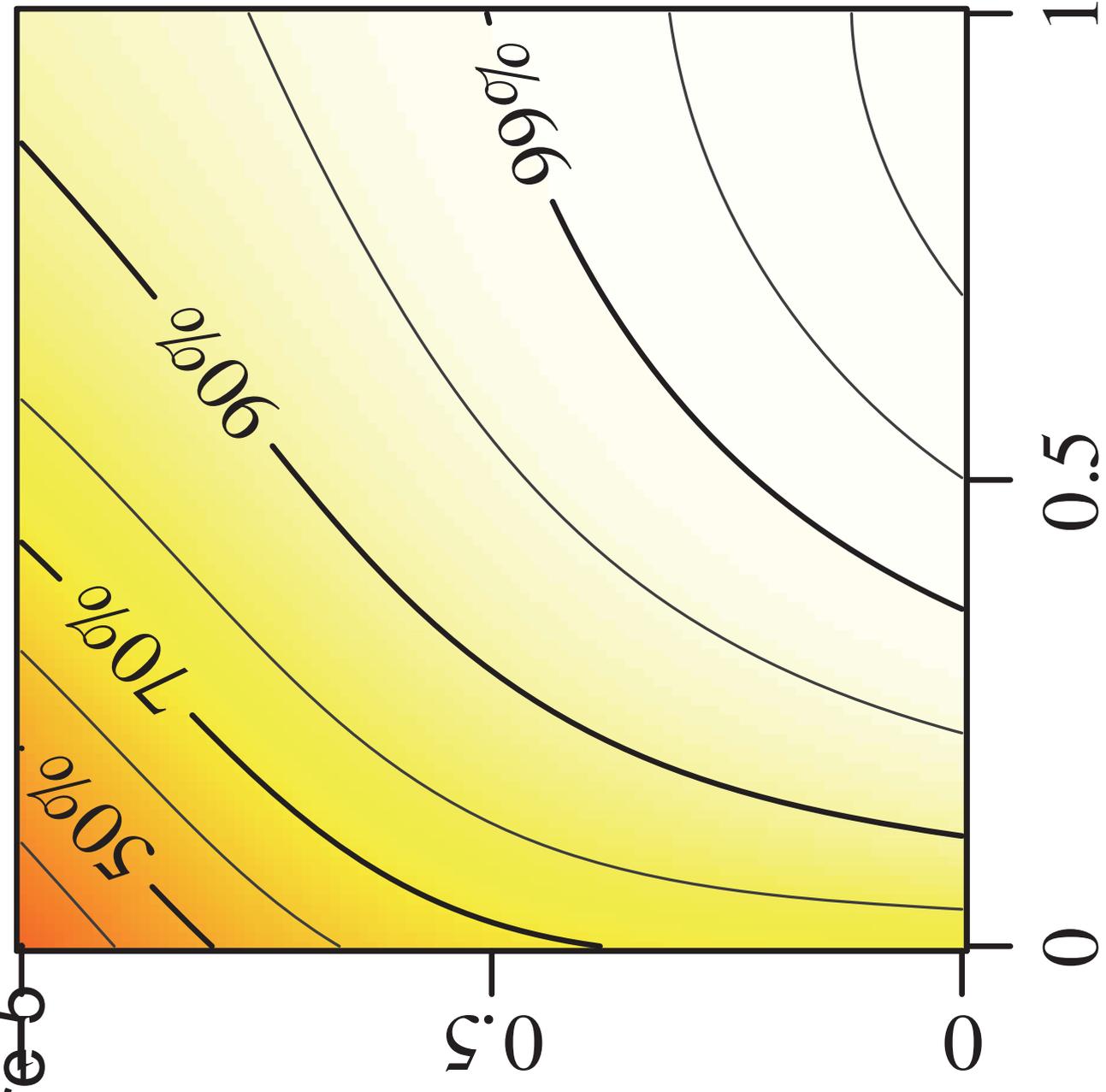
Figure 4



Branching probability

Figure 5

Scaled dispersal, v



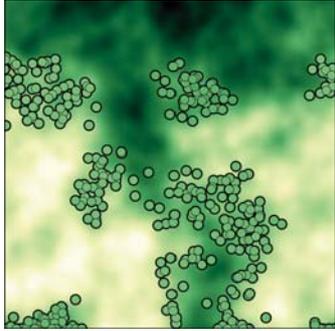
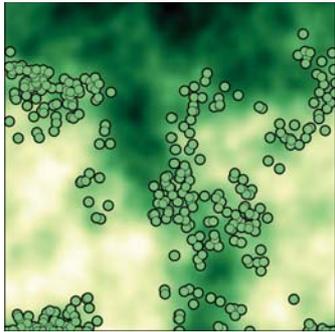
Branching probability

Scaled realized ACL, l_r

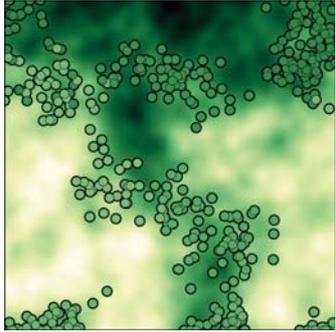
Movie 1 still image

200-6 KODAK

0



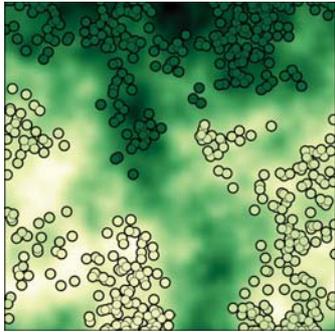
0A



1

2

GOLD 200-6



00

00A

0

0A

1

1A

2

2A

Online Appendix: Supplemental methods and results

for Evolutionary branching in complex landscapes

Benjamin C. Haller,¹ Rupert Mazzucco,² and Ulf Dieckmann²

1. Department of Biology and Redpath Museum, McGill University, 859 Sherbrooke Street West, Montreal, Quebec, Canada H3A 0C4;

2. Evolution and Ecology Program, International Institute for Applied Systems Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria

Model description

To test our hypotheses, we constructed an individual-based model, described below following guidelines suggested for the standardized description of such models (Grimm et al. 2006). The model description is also intended to follow closely that of Heinz et al. (2009), itself a descendant of Doebeli and Dieckmann (2003). A brief summary of the model is provided in the main text (see Methods, *Model summary*).

MODEL OVERVIEW

The purpose of this model is to understand the effects of complex spatial environmental heterogeneity on the process of speciation. The model is related to the spatially explicit individual-based stochastic model of Doebeli and Dieckmann (2003), restricted to the asexual case of that model. Our model replaces the linear spatial environmental gradient of Doebeli and Dieckmann (2003) with a complex landscape composed of a nonlinear environmental gradient and continuously varying environmental patchiness. As a consequence of the individual-based approach, stochasticity manifests in several aspects of our model, including variation in the landscape, variation in the initial state, demographic stochasticity, random dispersal, and random mutations during reproduction, as described in the following sections.

ENVIRONMENTAL LANDSCAPE AND STATE VARIABLES

Individuals inhabit a two-dimensional continuous landscape. The landscape is constant in time; its spatial heterogeneity is composed of a gradient in the x -direction, and of patchiness that varies continuously in both the x - and y -directions (see LANDSCAPE HETEROGENEITY). Individuals in this landscape are characterized by their location (x, y) ,

with $0 \leq x, y \leq 1$, and by their ecological character u , which affects their adaptation to the local environment. The state of the model at any point in time is thus fully specified by triplets (x_i, y_i, u_i) for all individuals i in the population ($i = 1, \dots, N$, where N is the current population size). Individuals are affected both by the landscape and by the other individuals in the model through their death rate (see DEATH EVENTS).

LANDSCAPE HETEROGENEITY

The carrying capacity density K at a location (x, y) for ecological character u is defined as

$$K(u, x, y) = K_0 N_{\sigma_x} (u - u_0(x, y)),$$

where K_0 is the maximum carrying capacity density, $N_{\sigma}(z) = \exp(-z^2/2\sigma^2)$ denotes a Gaussian function of standard deviation σ , and $u_0(x, y)$ denotes the optimal ecological character value (the value that produces the greatest carrying capacity density) at the given location; $u_0(x, y)$ thus describes the environmental landscape and is defined as

$$u_0(x, y) = \alpha + Sx + \frac{1}{2}Cx^2 + \Phi_{L_g, A}(x, y).$$

In this expression, the centering constant α is always selected after landscape generation so as to yield a mean value of 0.5 for $u_0(x, y)$ over the landscape as a whole, while S and C are model parameters defining the slope and curvature, respectively, of the environmental gradient in the x -direction. C represents the maximum environmental gradient due to curvature (at $x = 1$); with a quadratic curvature term, the coefficient $\frac{1}{2}C$ thus represents the average environmental gradient due to curvature. The function $\Phi_{L_g, A}(x, y)$ specifies the autocorrelated noise that describes the environmental patchiness, which is further characterized by the autocorrelation length L_g and the patchiness amplitude A . The function Φ is constructed as the convolution of

Gaussian white noise and a two-dimensional filter function constructed to produce the desired autocorrelation length L_g . The model parameters S , C , L_g , and A thus together govern the generation of the landscape. For further details on the generation procedure, see *Landscape generation*.

A survey of some landscapes generated by this model, and of the effects of parameters on the generated landscape, is shown in Figure 1.

BOUNDARY CONDITIONS

In the x -direction (the direction of the environmental gradient), we tested four boundary conditions: stopping, reflecting, reprising, and absorbing. These boundary conditions differ in the action taken when a natal dispersal location (see BIRTH EVENTS AND DISPERSAL) is drawn outside of the valid range $0 \leq x \leq 1$. Stopping boundaries replace values $x < 0$ or $x > 1$ by values 0 or 1, respectively, causing dispersers to stop at the edge of the landscape. Reflecting boundaries replace values $x < 0$ or $x > 1$ by values $-x$ or $2 - x$, respectively, as if dispersers bounced off of the edge of the landscape (repeatedly, if necessary, until the location is valid). Reprising boundaries redraw invalid locations until a valid location is obtained, as if dispersers avoid areas beyond the landscape's edge. Absorbing boundaries, finally, remove the individual from the population if an invalid location is drawn; dispersal beyond the edge of the landscape is in this case lethal.

In the y -direction, for which only stochastic environmental heterogeneity exists, periodic boundary conditions were utilized. Periodic boundaries replace values $y < 0$ or $y > 1$ by values $y + 1$ or $y - 1$, respectively (repeatedly, if necessary, until the location is valid). The landscape was generated in such a manner as to guarantee that it meshed seamlessly across the periodic y boundaries (see *Landscape generation*). The landscape thus described is cylindrical in topology for all four boundary conditions used. Periodic boundaries (or cline-periodic boundaries; Heinz et al. 2009) could not be used in the x -direction because the nonlinear environmental gradient (see LANDSCAPE HETEROGENEITY) would cause discontinuity in the landscape.

PHENOTYPE RANGES AND INITIALIZATION

Initially, the model contains N_{init} individuals with state (x_i, y_i, u_i) for each individual i , where x_i and y_i are drawn from a uniform distribution between 0 and 1, and $u_i = 0.5$ (matching the mean value of $u_0(x, y)$ over the landscape). No bounds are enforced by the model on values of the ecological character u thereafter.

This initial state distributes individuals across the entire landscape, although the poorly adapted individuals typically die almost immediately. Biologically, this distribution might represent an unusual event such as a storm that introduces individuals, propagules, or seeds across a landscape, or might represent a previously homogeneous landscape suddenly transformed into heterogeneity by a catastrophic event such as fire, flooding, or volcanism. However, additional test realizations (not shown) suggest that the final outcome of realizations does not depend strongly upon this initial distribution. With all but the shortest dispersal distances, the landscape is quickly colonized and the initial distribution has little effect. With very short-range dispersal, movement across the landscape is very slow and the waiting time to branching can thus be much longer, but the final outcome is similar. Given sufficiently long runtimes, then, we expect that different initial conditions would not substantially affect our results, other than by broadening the branching time distribution.

INTERACTIONS

Individuals interact only through local competition for resources. The strength of competition felt by focal individual i is the sum of the competitive impacts of all other individuals in the system, and is expressed as the effective number of individuals competing with the focal individual,

$$n_{\text{eff}}(u_i, x_i, y_i) = \left(2\pi\sigma_s^2\right)^{-1} \sum_{j=1, j \neq i}^N N_{\sigma_c}(u_j - u_i) N_{\sigma_s}(x_j - x_i) N_{\sigma_s}(y_j - y_i).$$

The strength of competition decreases with increasing spatial distance, as described by a Gaussian function with standard deviation σ_s , the spatial competition radius. The strength of competition also decreases with increasing

ecological character difference between the individuals, as described by a Gaussian function with standard deviation σ_c , the phenotypic competition width. The parameters σ_s and σ_c can therefore be thought of as scaling factors for the impact of spatial differences and phenotypic differences, respectively, on the strength of competition (Doebeli and Dieckmann 2003). The normalization factor $(2\pi\sigma_s^2)^{-1}$ ensures that a spatially uniform population is regulated to a given overall environmental carrying capacity regardless of the value of σ_s . Interactions span the periodic y -boundary; the shortest distance between individuals in the cylindrical landscape is used. As the competition felt by an individual increases, that individual's death rate increases, as described below (DEATH EVENTS).

DEATH EVENTS

The death rate d_i of each individual i is computed as $n_{\text{eff}}(u_i, x_i, y_i)/K(u_i, x_i, y_i)$, where $n_{\text{eff}}(u_i, x_i, y_i)$ denotes the strength of competition felt by individual i (see INTERACTIONS), and $K(u_i, x_i, y_i)$ denotes the carrying capacity density experienced by individual i (see LANDSCAPE HETEROGENEITY). Competition in this model therefore produces soft selection (Wallace 1975), and cannot lead to extinction. When an individual dies, it is removed from the population.

BIRTH EVENTS AND DISPERSAL

The birth rate of every individual i is equal and constant: $b_i = b$. When an individual gives birth, an offspring individual is generated asexually and added to the population. We model natal dispersal as the only form of movement; the spatial coordinates x and y of the offspring are thus drawn from normal distributions with standard deviation V and means equal to the parental values x_i and y_i , and then the boundary conditions are applied as described in BOUNDARY CONDITIONS. The ecological character u of the offspring is nearly faithfully inherited from the parent, except that with probability μ_m a mutational offset, drawn from a normal distribution with mean 0 and standard deviation σ_m , is added to the parental value u_i to represent a mutation event during reproduction.

Since the birth rate b is constant, we scale time by $1/b$. In $1/b$ units of time every individual is expected to produce one offspring, and if the

population size is at equilibrium, the same number of individuals is expected to die in that period of time. For this reason, we refer to $1/b$ units of time as a "generation".

PROCESS OVERVIEW AND SCHEDULING

Overlapping generations of individuals are implemented in the model using a continuous-time birth–death process implemented following the so-called Direct Method of Gillespie (1976). First, the current birth rate b_i and current death rate d_i of each individual i are determined by the state of the model, as described above (DEATH EVENTS; BIRTH EVENTS AND DISPERSAL). From these, the population-level birth and death rates are calculated as sums over the individual rates,

$$B = \sum_{i=1}^N b_i \text{ and } D = \sum_{i=1}^N d_i$$

respectively. The overall event rate is then calculated as $E = B + D$, and the waiting time to the next event is drawn from an exponential distribution with mean $1/E$. With probability B/E this is a birth event, and an individual i is chosen for reproduction with probability b_i/B . Otherwise, with probability D/E , it is a death event, and an individual i is chosen for removal with probability d_i/D . At the end of the event, the model state is updated, and the birth and death rates of all individuals are recalculated.

OBSERVABLES

The $u_0(x, y)$ values that define the landscape were recorded on generation of the landscape, at the start of each realization of the model. Each generated landscape was analyzed to extract metrics describing its salient features: the landscape's scaled standard deviation σ , skewness γ_1 , excess kurtosis γ_2 , and autocorrelation length l_r (see *Landscape analysis*).

Every 10 generations during each realization a histogram of the ecological character values of the population was recorded. At the end of each realization, the final state (x_i, y_i, u_i) of all individuals was recorded. Following the recommendations of Grimm (2002) regarding the importance of visual debugging, the landscape and the state of all individuals were also observable graphically during model runs. An example of the information available for each realization is presented in Figure 2, and a movie showing an example realization is provided as

Movie 1 (available in the Dryad data repository, <http://dx.doi.org/10.5061/dryad.43cj7/1>).

Branching was identified by a heuristic analysis of the histograms of ecological character values recorded during the realization (for example, Figure 2D). The initial distribution of ecological character values was unimodal, since we set $u_i = 0.5$ for all individuals. Branching was defined as divergence into a visibly bi- or multi-modal distribution, established not only at the end of the realization, but for a substantial period of time (3000 generations) prior to the end of the realization, to demonstrate the stability of the outcome. An alternative assessment of branching based on cluster analysis was also conducted (see *Alternative assessment of branching*). Finally, for realizations considered to have branched, the branching time was determined (see *Distribution of branching times*).

PARAMETERS

Parameters of the model, including the metrics taken from the realized model landscapes, are shown in Table 1. The parameter σ_s defines a fundamental length scale of the model, because distance in space affects model dynamics through the strength of competition scaled by σ_s . Similarly, the parameter σ_K defines a fundamental ecological phenotype scale of the model, because distance in phenotype affects model dynamics through the carrying capacity density, scaled by σ_K . For this reason, following Doebeli & Dieckmann (2003), we rescaled all other parameters with units involving length or ecological phenotype, using σ_s and σ_K , to derive dimensionless scaled parameters. The unscaled parameters V , S , C , L_g , and A were transformed according to the formulas shown (Table 1) to produce the scaled parameters v , s , c , l_g , and a . These scaled parameters were varied among realizations of the model.

Values for the parameters that were not varied (N_{init} , K_0 , σ_K , σ_s , μ_m , σ_m , b) were taken from the equivalent parameters of Doebeli & Dieckmann (2003) so as to allow easy comparison of results between these models. The exception is c_p , chosen here to explore dynamics in the limit of phenotypically indiscriminate competition (large c_p). In this limit, negative frequency-dependent selection would not cause branching in the absence of spatial heterogeneity; this choice thus guarantees that all

observed branching is due to the effects of spatial heterogeneity (Doebeli and Dieckmann 2003).

MODEL IMPLEMENTATION

The model was implemented in Objective-C using the Cocoa object-oriented toolkit (Mac OS X 10.6.4; Apple Inc., <http://www.apple.com>), except for the scheduling module, landscape construction module, and branching analysis module, all of which were written in the R programming language (version 2.11.1–2.15.1; R Development Core Team, <http://www.r-project.org>).

Comparison of nested GAM models

The nested models based on the landscape-generating parameters s , c , a , and l_g , in addition to the dispersal distance v , are compared in Table A1 and Figure A1. Similarly, the nested models based on the realized-landscape metrics σ , γ_1 , γ_2 , and l_r , in addition to v , are compared in Table A2 and Figure A2.

Intercept-only models correctly predicted the outcome of 78.4% of realizations, since that percentage of realizations did not branch (Table A3); these models simply predict that branching will never occur. The prediction rate observed in other models should thus be evaluated relative to this baseline.

A model including only the generating ACL, l_g , only marginally improved on the intercept-only null model (Table A1, model #1 vs. #0); similarly, a model with the interaction $l_g * v$ only marginally improved on a model without that interaction (Table A1, model #8 vs. #6). However, both terms did improve the BIC score, and were therefore retained. Other terms were more clearly of non-negligible effect size, according to both BIC and Nagelkerke R^2 (change in prediction rate is not a good indicator of effect size with a binary outcome, because a term can improve the certainty of predictions without changing their direction).

Binary predictions (not branched vs. branched) from the best GAM from each set of nested models (GAM_g and GAM_r) agreed in 94.8% of realizations (Figure A5).

The results presented in this section are for reprising boundary conditions; qualitative results were the same for other boundary conditions.

Table A1: Comparison of nested GAM models predicting the outcome of a realization (not branched versus branched) from the scaled landscape-generating parameters s , c , a and l_g and the scaled dispersal distance v , for reprising boundary conditions.

Model	Terms	EDF	C	R^2	D	Match	ΔBIC
0. (Intercept)	0	1.0	.500	.000	.000	.784	159104
1. l_g	1	3.2	.507	.000	.000	.784	159098
2. a	1	4.2	.522	.002	.001	.784	158821
3. s	1	4.9	.625	.052	.034	.784	148851
4. c	1	4.9	.703	.130	.088	.784	132781
5. v	1	5.0	.829	.386	.288	.845	72896
6. $s + c + a + l_g + v$	5	20.0	.922	.600	.493	.882	11585
7. [6] + $c * v$	6	33.0	.923	.601	.493	.882	11473
8. [6] + $l_g * v$	6	33.6	.923	.601	.494	.883	11350
9. [6] + $s * a$	6	33.9	.923	.604	.496	.884	10641
10. [6] + $s * v$	6	34.8	.924	.604	.496	.884	10629
11. [6] + $a * v$	6	36.2	.925	.609	.502	.885	8910
12. [6] + $s * c$	6	36.0	.926	.613	.506	.886	7792
13. [6] + $c * a$	6	36.4	.928	.617	.511	.887	6368
14. [7–13]	12	112.7	.933	.637	.533	.894	753
15. [14] + $s * c * a$	13	167.0	.935	.642	.537	.895	0

Note: The Terms column shows the number of terms in the model, while the EDF column gives the estimated degrees of freedom used. C is the area under the Receiver Operating Characteristic curve (Metz, 1978), R^2 is the Nagelkerke generalized R^2 (Nagelkerke, 1991), D is the Discrimination coefficient (Tjur, 2009), Match is the fraction of cases in which the model prediction matched the observed binary outcome, and ΔBIC is the Bayesian Information Criterion relative to the best model (Schwarz, 1978). Better models are indicated by high values, except for ΔBIC , for which a low score is better. Models are sorted by descending ΔBIC ; the best model according to BIC (the “GAM_g” model) is shown in bold. All terms in all models are significant with $P < 0.001$. See Figure A1 for a graphical depiction of the best ten models.

Table A2: Comparison of nested GAM models predicting the outcome of a realization (not branched versus branched) from the scaled realized-landscape metrics σ , γ_1 , γ_2 , and l_r and the scaled dispersal distance v , for reprising boundaries.

Model	Terms	EDF	C	R^2	D	Match	ΔBIC
0. (Intercept)	0	1.0	.500	.000	.000	.784	173875
1. γ_2	1	5.0	.579	.022	.014	.784	169695
2. γ_1	1	5.0	.600	.032	.021	.784	167732
3. l_r	1	5.0	.611	.040	.026	.784	165999
4. σ	1	5.0	.760	.213	.149	.794	129258
5. v	1	5.0	.829	.386	.288	.845	87667
6. $\sigma + \gamma_1 + \gamma_2 + l_r + v$	5	20.8	.943	.666	.564	.899	4795
7. [6] + $\sigma * v$	6	35.5	.943	.668	.568	.900	4055
8. [6] + $l_r * v$	6	33.4	.946	.677	.577	.904	1087
9. [7–8]	7	49.3	.947	.681	.582	.906	0

Note: The Terms column shows the number of terms in the model, while the EDF column gives the estimated degrees of freedom used. C is the area under the Receiver Operating Characteristic curve (Metz, 1978), R^2 is the Nagelkerke generalized R^2 (Nagelkerke, 1991), D is the Discrimination coefficient (Tjur, 2009), Match is the fraction of cases in which the model prediction matched the observed binary outcome, and ΔBIC is the Bayesian Information Criterion relative to the best model (Schwarz, 1978). Better models are indicated by high values, except for ΔBIC , for which a low score is better. Models are sorted by descending ΔBIC ; the best model according to BIC (the “GAM_r” model) is shown in bold. All terms in all models are significant with $P < 0.001$. See Figure A2 for a graphical depiction of the best four models.

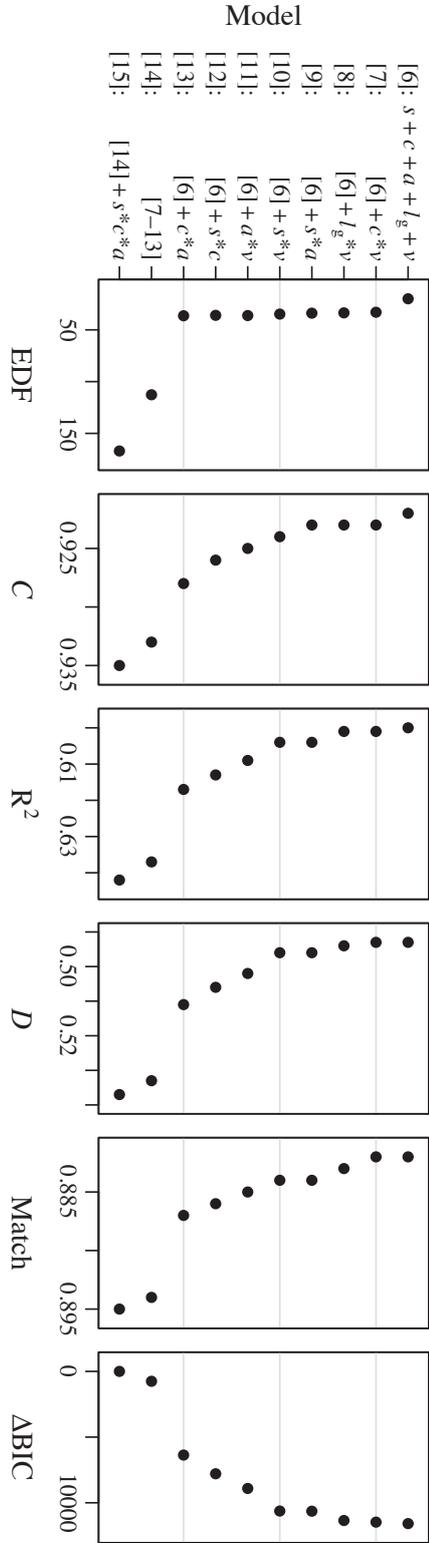


Figure A1: Comparison of the best (ABIC < 15000) nested GAMs predicting the outcome of a realization (not branched versus branched) from the scaled landscape-generating parameters s , c , a and l_g and the scaled dispersal distance v , for reprising boundary conditions. This figure is a graphical depiction of models 6–15 in Table A1. EDF: the estimated degrees of freedom used. C : the area under the Receiver Operating Characteristic curve (Metz, 1978). R^2 : the Nagelkerke generalized R^2 (Nagelkerke, 1991). D : the Discrimination coefficient (Tjur, 2009). Match: the fraction of cases in which the model prediction matched the observed binary outcome. ABIC: the Bayesian Information Criterion, relative to that of the best model (Schwarz, 1978). Better models are indicated by high values, except for ABIC, for which a low value is better. All terms in all models are significant with $P < 0.001$. Note that all goodness-of-fit statistics are strongly correlated, and have perfect rank correlation.

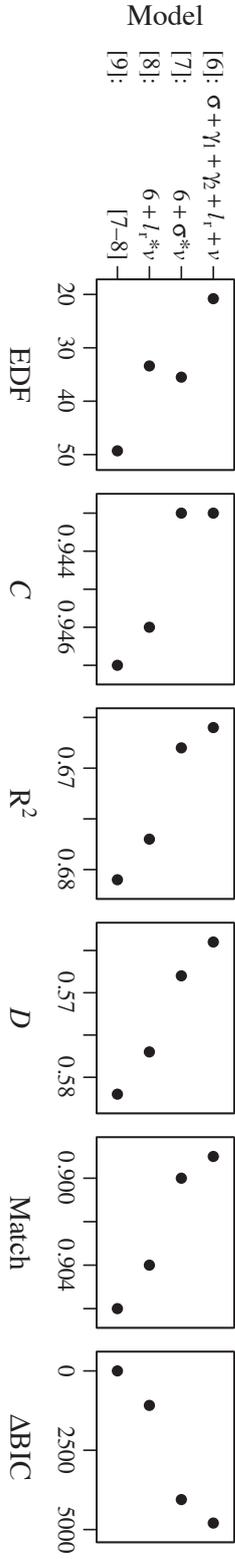


Figure A2: Comparison of the best (ABIC < 15000) nested GAMs predicting the outcome of a realization (not branched versus branched) from the scaled realized-landscape metrics σ , γ_1 , γ_2 , and l , and the scaled dispersal distance v , for reprising boundary conditions. This figure is a graphical depiction of models 6–9 in Table A2. EDF: the estimated degrees of freedom used. C : the area under the Receiver Operating Characteristic curve (Metz, 1978). R^2 : the Nagelkerke generalized R^2 (Nagelkerke, 1991). D : the Discrimination coefficient (Tjur, 2009). Match: the fraction of cases in which the model prediction matched the observed binary outcome. ABIC: the Bayesian Information Criterion, relative to that of the best model (Schwarz, 1978). Better models are indicated by high values, except for ABIC, for which a low value is better. All terms in all models are significant with $P < 0.001$. Note that all goodness-of-fit statistics are strongly correlated, and have perfect rank correlation.

Table A3: Comparison of results for the four boundary conditions.

	Absorbing	Reprising	Reflecting	Stopping
Branched (%)	18.5	21.6	21.9	24.0
Branched \bar{s}	0.814	0.804	0.795	0.769
Branched \bar{c}	0.358	0.341	0.340	0.336
Branched \bar{a}	1.416	1.449	1.473	1.497
Branched \bar{l}_g	1.555	1.542	1.525	1.524
Branched \bar{v}	0.552	0.732	0.728	0.787
Branched $\bar{\sigma}$	2.856	2.770	2.763	2.714
Branched $\bar{\gamma}_1$	0.268	0.248	0.248	0.246
Branched $\bar{\gamma}_2$	-0.784	-0.759	-0.744	-0.717
Branched \bar{l}_r	1.131	1.115	1.104	1.083
Clustered (%)	19.6	23.0	23.2	25.4
Avg. clusters	2.64	2.63	2.65	2.66
GAM _g R ²	0.700	0.642	0.651	0.658
GAM _g pred.	0.918	0.895	0.895	0.887
GAM _r R ²	0.728	0.681	0.689	0.689
GAM _r pred.	0.926	0.906	0.906	0.897

Note: Branched (%) is the percentage of realizations of the 300,000 realizations conducted for the given boundary condition that resulted in evolutionary branching according to our phenotypic history analysis. The Branched \bar{s} , \bar{c} , \bar{a} , \bar{l}_g , \bar{v} , $\bar{\sigma}$, $\bar{\gamma}_1$, $\bar{\gamma}_2$, and \bar{l}_r are the mean value of that parameter or metric, among only those realizations that branched. Clustered (%) is the percentage of realizations that resulted in evolutionary branching according to alternative cluster analysis method of detecting branching; “Avg. clusters” is the average number of clusters produced, among only those realizations that produced two or more clusters. GAM_g results are for the full GAM model fitted using the model parameters, as in Table A1 model #15; the Nagelkerke R² and the proportion of realization outcomes correctly predicted are given. Similarly, GAM_r results are for the full GAM fitted using the landscape metrics, as in Table A2 model #9.

Landscape-generating parameters versus realized-landscape metrics

Although the generating ACL l_g and its interaction $l_g * v$ had very little effect (Table A1 and Figure A3B, and previous section), the realized ACL l_r and its interaction $l_r * v$ had a substantial effect (Table A2 and Figure 5). The generating and realized ACLs are correlated when l_g is very small, because very small-scale patchiness can be accurately reproduced with little distortion due to stochasticity, but when l_g is larger the correlation breaks down, because large-scale patchiness is stochastic at a scale commensurate with the scale of the landscape itself; the realized ACL of the landscape may therefore be quite different from the generating ACL (Figure A3A). Since the environmental gradient slope and curvature also affect the realized ACL, that is another reason why l_g and l_r are generally poorly

correlated; for clarity, however, Figure A3A shows only values for landscapes with no slope or curvature, so that it depicts solely the loss of correlation for large values of l_g due to stochasticity in the landscape generation and system size effects.

The superiority of l_r to l_g in predicting branching is one aspect of the larger phenomenon that GAM_r, the best GAM based on the realized-landscape metrics, substantially outperformed GAM_g, the best model based on the landscape-generating parameters (Table A1 model #15 vs. Table A2 model #9). This was the case even though GAM_g contained 13 terms and used 167.0 effective degrees of freedom, compared to 7 terms and 49.3 effective degrees of freedom for GAM_r. For large l_g , stochasticity caused considerable variation not only in l_r , but also in the realized-landscape metrics σ , γ_1 , and γ_2 , for given values of the landscape-generating parameters s , c ,

a , and l_g (Figure A4). Unsurprisingly, given this, the realized-landscape metrics are a better predictor of the effect of the landscape than are the landscape-generating parameters; the realized landscape is what the individuals in the model actually occupy, and so its properties are what matters. However, it is perhaps surprising that the simple metrics we used proved so effective; one might ponder whether different metrics could be even better. In any case, this result suggests that caution must be used whenever modeling heterogeneous landscapes, to be sure that the effect of the realized pattern of heterogeneity, and not solely the effect of the landscape-generating parameters, is analyzed, since the two may be quite different – an issue that has not, to our knowledge, been raised previously.

The realized-landscape metrics spanned a wide range of types of heterogeneity, although negative skewness, positive kurtosis, and short ACLs were relatively infrequent (Figure A6). Although this means that there is less data in some regions of the parameter space, the very large number of realizations analyzed mitigates this problem, except where specifically noted.

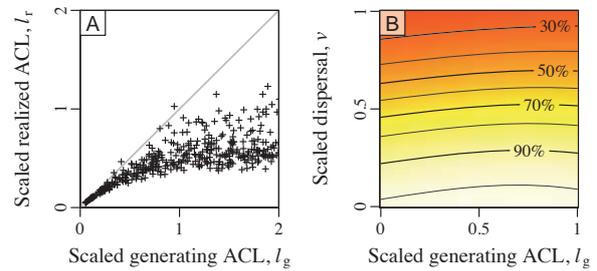


Figure A3: Comparison of the scaled generating autocorrelation length (ACL) l_g to the scaled realized ACL l_r . A: Scatter plot of l_g versus l_r for 500 randomly chosen landscapes; the gray line shows the 1:1 line at which $l_g = l_r$. The deviation of l_g from l_r illustrates that autocorrelation lengths that are large relative to the spatial scale of the landscape can be unrepresentable because they correspond to spatial features approaching or exceeding the scale of the landscape itself. B: Predicted branching probabilities from GAM_g , showing the effects of l_g and the scaled dispersal distance v ; compare the weak relationships shown here (little effect of l_g and little interaction between l_g and v) to the stronger effects from GAM_r shown in Figure 5. The color scale is as shown in Figure 5. Results shown are for reprising boundaries; results for other boundary conditions are qualitatively similar (not shown). Other parameters: $s = 0$, $c = 0$, $a = 1.5$. Contour lines above 90% are placed to best show the contours of the data.

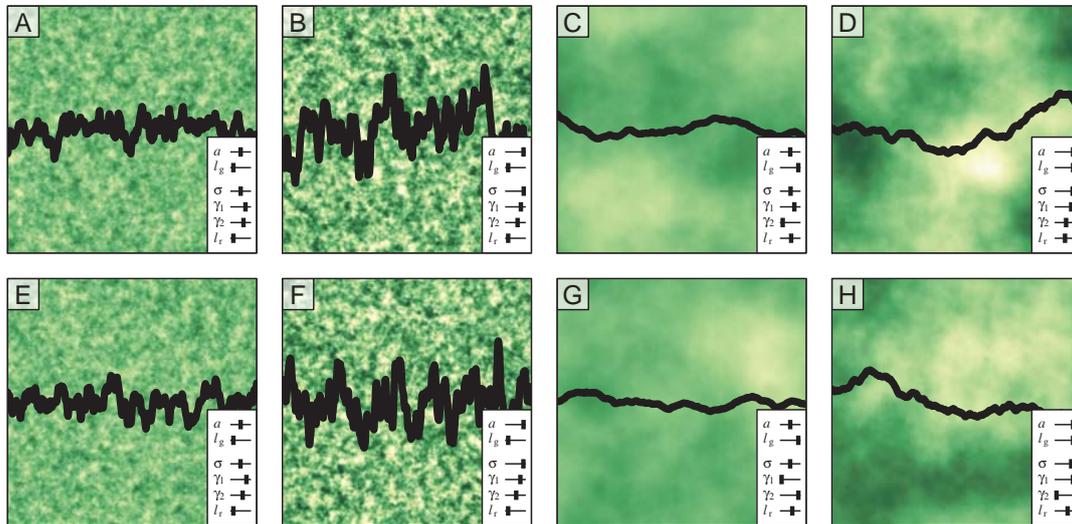


Figure A4: Examples of generated landscapes. These examples provide an illustration of the variability in realized-landscape metrics resulting from different types of spatially continuous patchiness. Vertical pairs of panels used the same landscape-generating parameters (scaled amplitude a and scaled generating ACL l_g ; $s = 0$ and $c = 0$ for all panels), but used different random number generator seeds and thus differ. Sliders in each panel show landscape-generating parameters (a , l_g) and realized-landscape metrics (σ , γ_1 , γ_2 , l_r) on an arbitrary scale. Colors indicate locally optimal ecological trait values, ranging from white (low) to dark green (high). Black lines show locally optimal ecological trait values across one horizontal transect of each landscape. Panels A, B, E, F used short l_g while panels C, D, G, H used long l_g . Panels A, C, E, G used intermediate a while panels B, D, F, H used large a . These results illustrate that landscapes that share generating parameters will also share similar realized metrics if l_g is small, but when l_g is large the realized metrics can be quite divergent because of the stochastic patterns at large spatial scales in the landscapes generated.

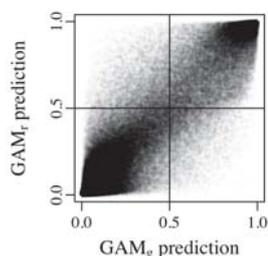


Figure A5: Correlation between the predictions of the two GAM models, GAM_g and GAM_r . A total of 300,000 points are plotted; each point represents one realization, showing the predicted probability of branching from the GAM_g model (x-axis) and the GAM_r model (y-axis). The lower left quadrant, in which both models predict no branching, contains 78.9% of realizations, and the upper right quadrant, in which both models predict branching, contains 15.9% of realizations; the two models thus made the same prediction in 94.8% of realizations. The upper left quadrant, in which GAM_r predicts branching but GAM_g does not, contains 2.8% of realizations, while the lower right quadrant, in which GAM_r does not predict branching but GAM_g does, contains 2.4% of realizations. Note also that for the quadrants in which the models agree, both models tend to be relatively certain about their predictions (predicted probabilities near 0.0 or 1.0), while predictions are generally much less certain for the quadrants in which the models disagree (predicted probabilities near 0.5). Results shown are for reprising boundaries; results are very similar for other boundary conditions, with small differences in the proportion of realizations predicted to branch (see Table A3).

Effects of boundary conditions

Some basic metrics comparing the four tested boundary conditions are shown in Table A3. Boundary condition effects generally followed the findings of previous research (Mazzucco et al., unpublished manuscript). In particular, the stopping boundary condition produced the highest branching propensity due to its effect of generating disruptive selection, while the absorbing boundary condition produced the lowest branching propensity due to its generation of stabilizing selection. Similarly, absorbing boundaries produced branching only for relatively short dispersal distances, because long-range dispersal was more likely to be lethal, while stopping boundaries produced branching even for large dispersal distances, because long-range dispersal predictably stopped individuals at the edge of the landscape. Effects of the boundary condition

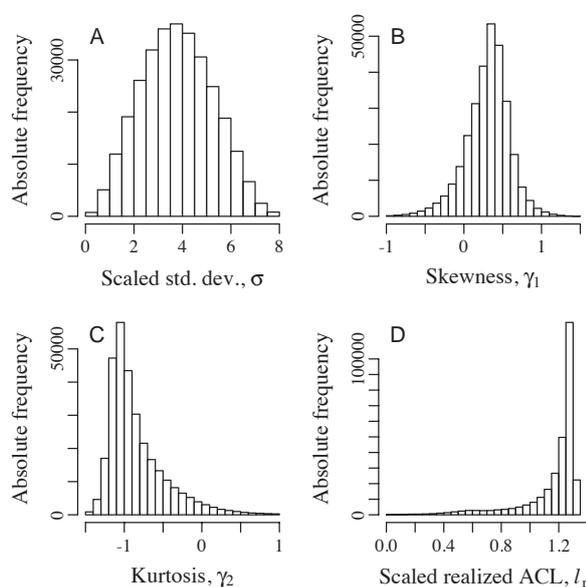


Figure A6: Distributions of the realized-landscape metrics. A: Scaled standard deviation, σ . B: Skewness, γ_1 . C: Excess kurtosis, γ_2 . D: Scaled realized ACL, l_r . A small number of outliers are omitted in panels A–C to allow the remaining data to be better displayed. Distributions shown encompass all landscapes generated in realizations utilizing reprising boundary conditions; since landscape generation is independent of the boundary conditions chosen, this is a representative sample. Note that although the landscape-generating parameters were drawn from uniform distributions (Methods: *Model realizations*), the realized-landscape metrics were not uniformly distributed, and so the realized metric parameter space was not uniformly sampled. Nevertheless coverage was extensive due to the large number of realizations conducted; more than 1000 realizations had $l_r < 0.2$, for example.

on the distribution of branching times followed the same pattern, with the modal time to branching being longest with absorbing boundaries and shortest with stopping boundaries, although the differences were slight (see *Distribution of branching times* and *Correlations with branching times*).

Reprising boundaries were chosen for our main results as a middle ground, minimizing the influence of the landscape boundary and thus best approximating the effect expected on a landscape of infinite extent. Stopping boundaries cause an accumulation of individuals near the boundary, whereas absorbing boundaries cause a deficit of individuals near the boundary, and these density effects influence the branching propensity. Reprising boundaries do not bias the branching propensity through such density effects, and thus insights from

this case are likely to apply generally – even when the density effects of other boundary conditions are added in. Reflecting boundaries would have been a similarly reasonable choice, and showed very similar effects in our analysis. Minimization of boundary effects is preferable for theoretical purposes, to maximally reveal the effects of model parameters without any obscuring influence from the model’s assumptions; however, when comparing our model results to a particular biological system, the boundary condition should be chosen that best matches the empirical dispersal characteristics of the system.

The effects of different types of heterogeneity on branching are qualitatively similar under different boundary conditions (Figure A7). Interestingly, the intermediate optimum patchiness amplitude is most apparent for absorbing boundaries and least apparent for stopping boundaries. A secondary, weaker optimal magnitude of curvature is also observed for all boundary conditions except absorbing; this may reflect the degree of curvature that produces an optimal difference between the environments at the extreme left and right edges of the landscape, since the landscape edges can be exploited as niches with non-absorbing boundary conditions (particularly stopping boundaries, for which this secondary optimum is strongest).

The interactions between different types of heterogeneity are also qualitatively similar under different boundary conditions (Figure A8). The particular contour line shape that indicates the refugium effect (Results: *Additivity and the refugium effect*) appears strongest with reflecting boundaries for the interaction of slope and amplitude, and strongest with reprising boundaries for the interaction of curvature and amplitude, but it is visible for all boundary conditions; this suggests that the refugium effect is general.

The interaction between realized ACL and dispersal distance is shown for different boundary conditions in Figure A9. The insensitivity of branching to the dispersal distance for short realized ACL is most apparent for stopping boundaries, and least apparent for absorbing boundaries; this may be a simple result of dispersal distance being least important for stopping boundaries (since dispersal often stops at the edge regardless of the “intended” dispersal distance), and most important for

absorbing boundaries (since dispersing too far is lethal). In any case, the qualitative pattern of interaction between the realized ACL and the dispersal distance is similar for all boundary conditions.

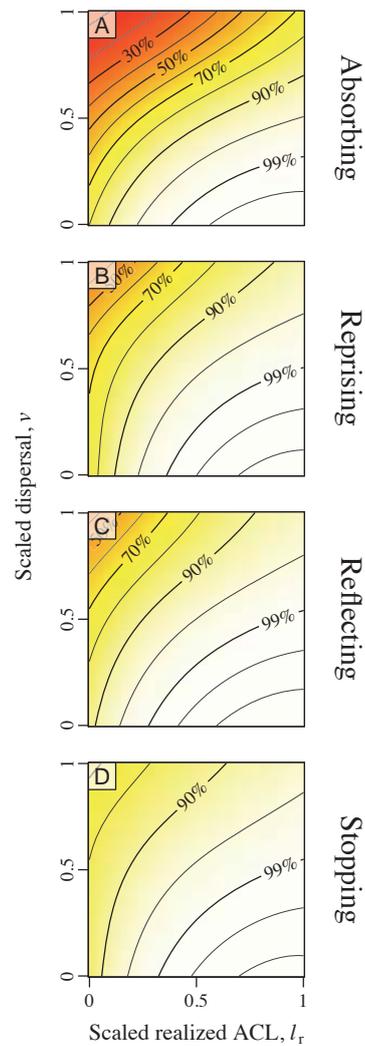


Figure A9: Predicted branching probabilities from fitted GAMs, showing a comparison among boundary conditions in their interaction between realized autocorrelation length (ACL) and dispersal distance. Each row of this figure follows Figure 5, but is plotted for a particular boundary condition; the second row, for reprising boundaries, thus matches Figure 5 exactly. Contour lines show the predicted branching probability; the color scale is as shown in Figure 5. Branching is always promoted by large realized ACL and by short dispersal distance. An interaction between v and l_r exists for all boundary conditions, causing a decrease in the effect of v for small l_r , but this interaction varies in strength among boundary conditions.

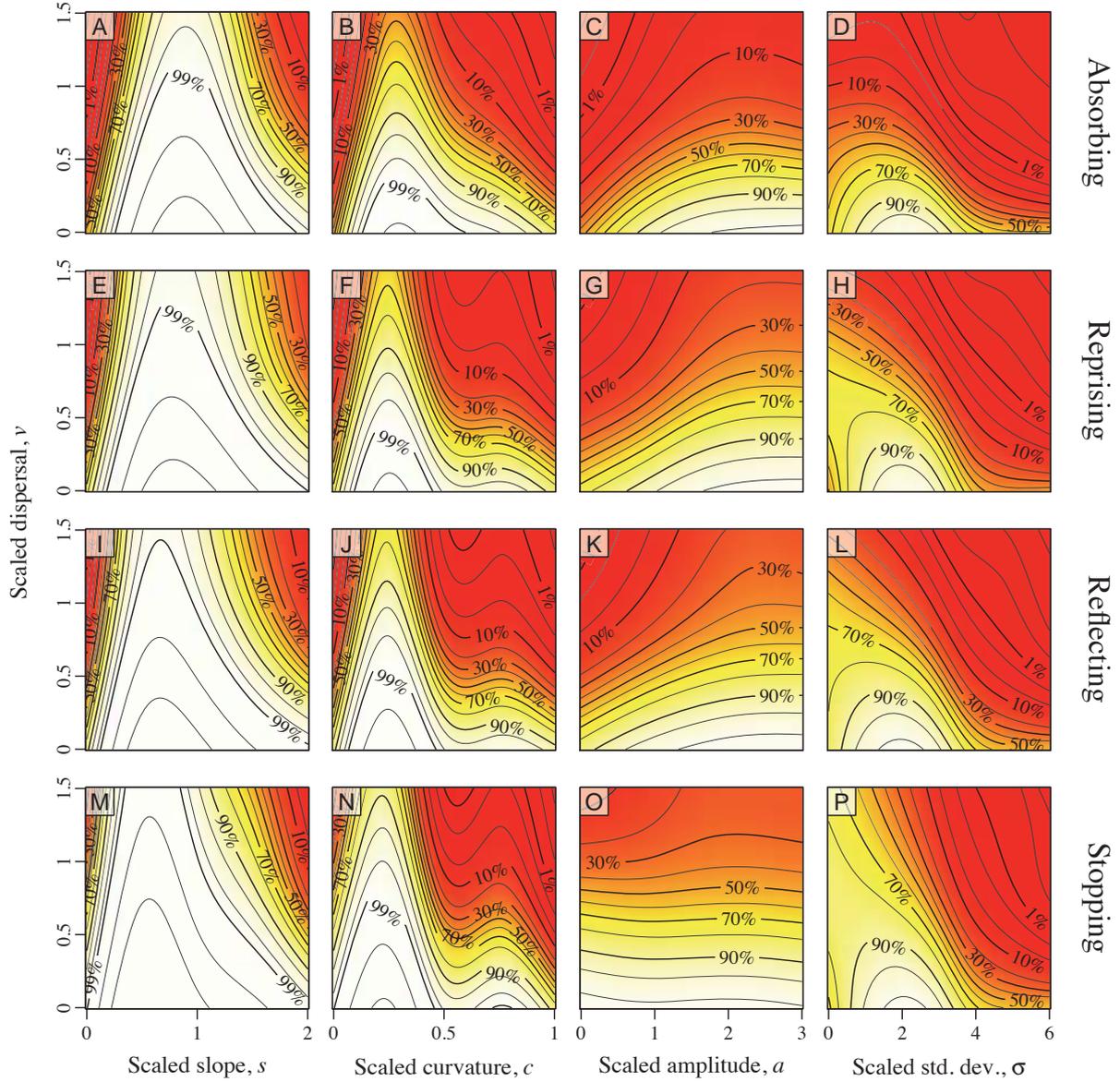


Figure A7: Predicted branching probabilities from fitted GAMs, showing a comparison among boundary conditions in the effects of scaled dispersal distance and different heterogeneity types. Each row of this figure follows Figure 3, but is plotted for a particular boundary condition; the second row, for reprising boundaries, thus matches Figure 3 exactly. Contour lines show the predicted branching probability; the color scale is as shown in Figure 3. The probability of branching is maximized at an intermediate value of s and c for all boundary conditions, although this optimum value varies among boundary conditions. Similarly, an intermediate optimum for σ exists for all boundary conditions, but typically only for smaller values of v . For a , an intermediate optimum that maximizes branching clearly exists for absorbing boundaries, and may also exist for reflecting and reprising boundary conditions (but since the optimum in these cases is close to the edge of the parameter space explored, this is uncertain); no well-defined intermediate optimum for a appears to exist for stopping boundaries.

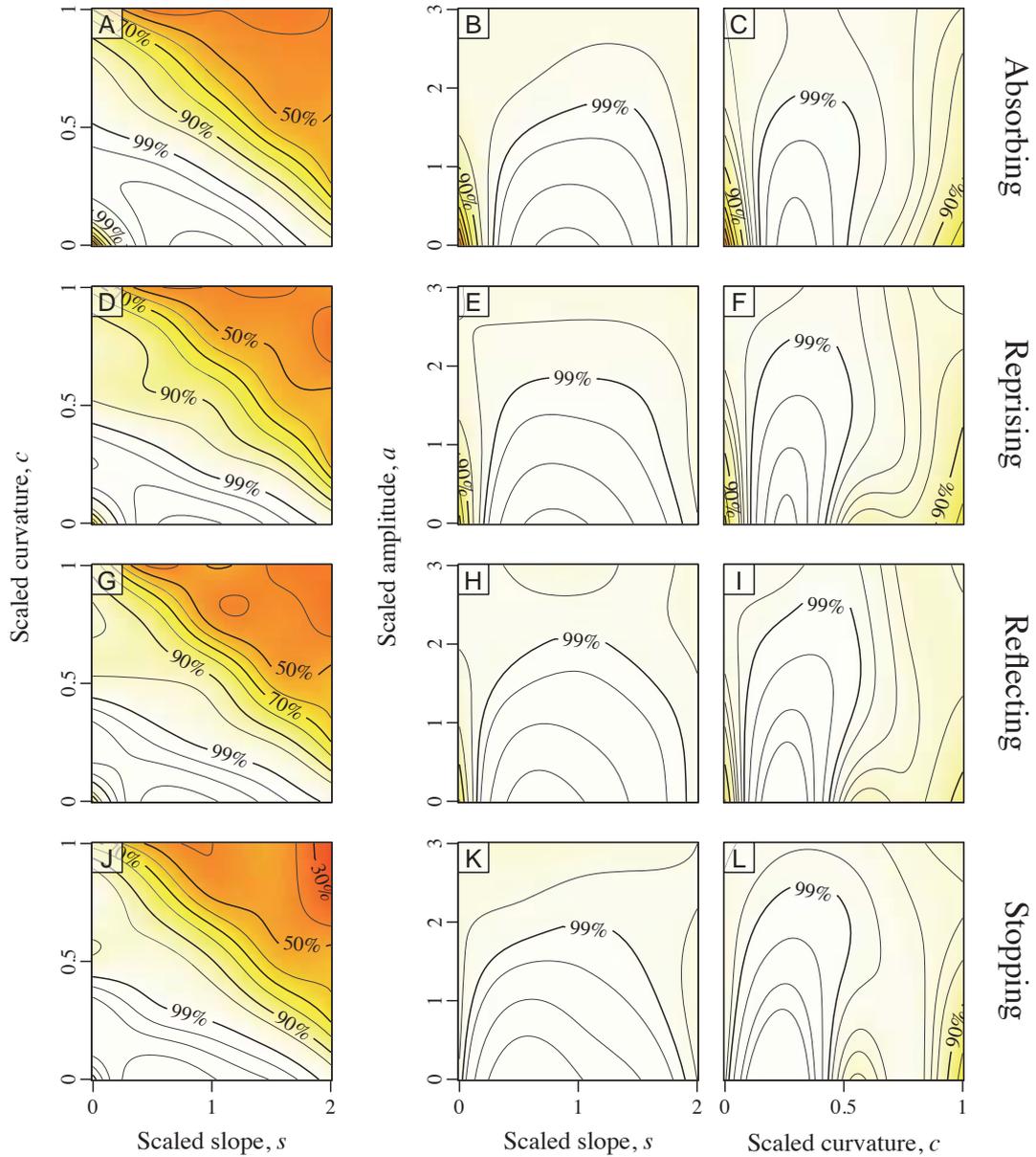


Figure A8: Predicted branching probabilities from fitted GAMs, showing a comparison among boundary conditions in their interactions between different heterogeneity types. Each row of this figure follows Figure 4, but is plotted for a particular boundary condition; the second row, for reprising boundaries, thus matches Figure 4 exactly. Contour lines show the predicted branching probability; the color scale is as shown in Figure 4. Additivity of the effects of s and c is apparent for all boundary conditions, as is the lack of additivity of s and a , and of c and a , due to the refugium effect. The refugium effect, which causes mitigation of the negative effect on branching of large s and large c due to increasing amplitude a , is generally visible; however, the effect varies in strength, and is always stronger with c than with s .

Effects of skewness and kurtosis

The bias towards positive skewness values among our generated landscapes (Figure A6B) is due to the fact that curvature, as implemented in our model as increasing gradient slope from left to right, leads to positive skewness. For this reason, the causes of positive skewness and negative skewness in our model are not symmetric; positive skewness is caused by both patchiness and curvature, while negative skewness is caused by patchiness alone. These causes would not segregate in this manner for real environments; empirically, the meaning of the sign of the skewness depends simply on the ordination of environmental values (whether a cold environment is given a value less than, or greater than, a hot environment, for example). Since the effect of skewness was found to be reasonably symmetric around zero (Figure A10), however, this concern can likely be ignored. In general, then, our results indicate that large skewness, whether caused by gradient nonlinearity or by an uneven pattern of patchiness, hinders branching; the optimal level of skewness appears to be close to zero (Figure A10).

Most kurtosis values were negative, with a peak near -1 (Figure A6C), indicating that platykurtic distributions of environmental conditions were much more common than leptokurtic distributions. This is unsurprising, since an environmental gradient is inherently platykurtic, resembling a uniform distribution (kurtosis -1.2) much more than a normal distribution (kurtosis 0), not to mention a more leptokurtic distribution such as a Laplace distribution (kurtosis 3). To a first approximation, then, the landscape kurtosis may be thought of as representing the degree to which the pattern of heterogeneity of the landscape is dominated by a linear gradient that produces a uniform distribution of environmental values. It is interesting, therefore, that more negative values of kurtosis hindered branching (Figure A10). We tentatively interpret this as a different angle on the same refugium effect discussed previously (Discussion: *Additivity and the refugium effect*), since it shows that a linear gradient with no other heterogeneity (more negative kurtosis values) is less effective in promoting branching than more complexly heterogeneous landscapes.

Interestingly, then, our results suggest that a normal distribution of environmental values (with

both skewness and kurtosis close to zero) may be close to optimal for evolutionary branching. The spatial distribution of those values is, of course, also important, however, and so broad conclusions regarding the optimal pattern of environmental heterogeneity for the promotion of evolutionary branching will require further investigation.

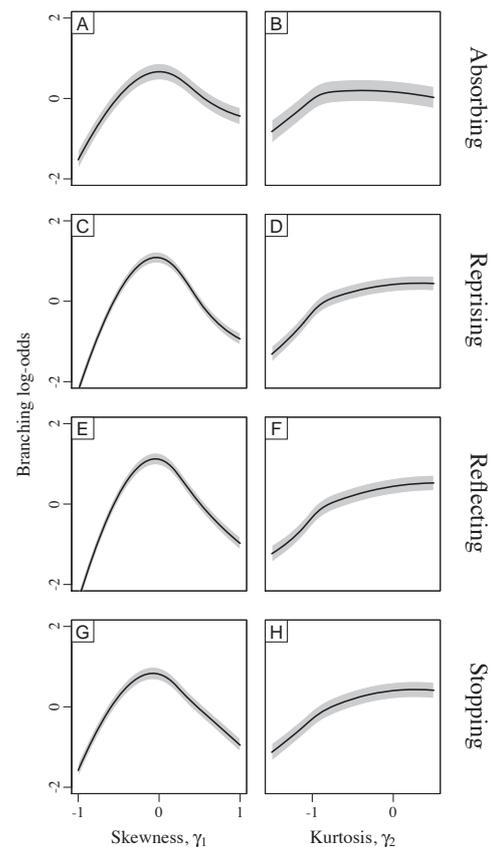


Figure A10: Predicted branching propensities from fitted GAMs, showing a comparison among boundary conditions in the effects of skewness and excess kurtosis. Branching propensity is shown as the branching log-odds (the logit of the branching probability); positive log-odds indicate relative promotion of branching, while negative log-odds indicate relative hindrance of branching. The curve in each panel is re-centered on a mean y value of zero, removing the absolute differences in branching propensity among boundary conditions, for ease of interpretation. Gray shaded bands indicate standard error confidence intervals. An intermediate skewness maximally promotes branching for all boundary conditions. The branching propensity generally increases with increasing kurtosis, although this effect saturates and even, for absorbing boundary conditions, declines at high kurtosis values. Other parameter values are unspecified, since no interaction terms involving skewness and kurtosis are included in the GAM models.

Alternative assessment of branching

As described previously (*Model description, OBSERVABLES*), a heuristic analysis of the distribution of ecological trait values over the final 3000 generations was used to determine whether evolutionary branching had occurred. An alternative assessment of branching was also conducted using a Gaussian mixture model cluster analysis performed on the ecological character values from the end-of-realization census data, with the `Mclust()` function of the R package `mclust` (Fraley and Raftery 2003). A Bayesian prior was used on the variance of the cluster analysis, with a `scale` parameter value of 0.5 for the `priorControl()`, determined by trial and error to provide a good detection of clustering as compared to a visual determination.

For the detection of branching, the heuristic method and the cluster analysis method produced similar results, except that cluster analysis, since it considered only the pattern of phenotypes in the final generation of a realization, overestimated the branching rate slightly (Table A3). Despite their very different mathematical bases, the level of agreement between the two methods was quite high, giving some assurance that our results were not a consequence, in some way, of our method of assessing branching. Furthermore, GAM models fit using the cluster analysis branching assessment provided a close match to the GAM models presented.

One advantage of the cluster analysis method is that it produces not just a binary assessment of whether branching did or did not occur, but a count of the number of clusters. This ranged from 1 to 9, for the realizations conducted; typically, however, even those realizations that did branch produced only 2 or 3 clusters (Table A3), with the number of clusters produced following roughly a negative exponential distribution across all realizations (not shown). Interestingly, the number of clusters observed was negatively correlated with branching times, suggesting that landscapes that can support a larger number of species also produce more rapid evolutionary branching (see *Correlations with branching times*). Ordinal multinomial GLM fits (not shown) indicated that the factors causing a high branching propensity are the same as the factors

causing a large number of clusters, suggesting that adaptive radiation in our model is simply adaptive branching “writ large”; the two metrics are simply different views on the propensity for adaptive diversification.

Distribution of branching times

For each realization that was classified as branched, we also determined the branching time using a heuristic algorithm that looked backwards from the end of the realization until finding a generation without a bimodal (or multimodal) phenotypic distribution. Since the phenotypic distribution was saved every ten generations (see *Model description, OBSERVABLES*), the branching time was then estimated as the generation found, plus five (intermediate between the last unbranched and the first branched generations). This determination was generally robust for a wide variety of phenotypic histories, although a small number (~2%) could not be assigned a well-defined branching time.

The distribution of branching times for branched realizations using reprising boundary conditions is shown in Figure A11A. Most realizations that branched did so quite early – within the first ~1000 generations – as indicated by the strong modal peak of the distribution. The distribution also has a long tail that asymptotically approaches a non-zero constant value. This non-zero offset probably represents realizations that branched only after establishing a favorable spatial configuration; however, there is currently no good theory for spatial branching processes that would allow us to quantify this contribution (see Discussion: *Future directions*). Finally, the bump at the end of the distribution, comprised of realizations evaluated to have branched at close to 5000 generations, represents noise in the detection of branching times caused by stochastic demographic effects. We fitted lognormal and exponential curves with a vertical offset added, using branching times less than 4500 generations to exclude the final bump (Figure A11B). The exponential fit is poor, suggesting that branching is not a Poisson process; the lognormal fit is clearly better, and indeed, fits the distribution quite closely. In fact this result is expected, since branching is the result of many independent random events that combine multiplicatively (mutation,

dispersal, survival). Wakano and Iwasa (2013) showed similar results with a substantially different model (their Figure 9).

The predicted probability of branching had no effect on the modal branching time (Figures A11C and A11D). This indicates that branching is essentially a binary process once a favorable spatial configuration is reached: it happens early or does not happen at all. However, the mean and median branching times were negatively correlated with the predicted probability of branching (Figures A11C and A11D). This was due to a decreased frequency of early branching for realizations with low predicted branching probabilities, and thus an increased influence of the late-branching realizations over the mean and median branching time (see the next section, *Correlations with branching times*).

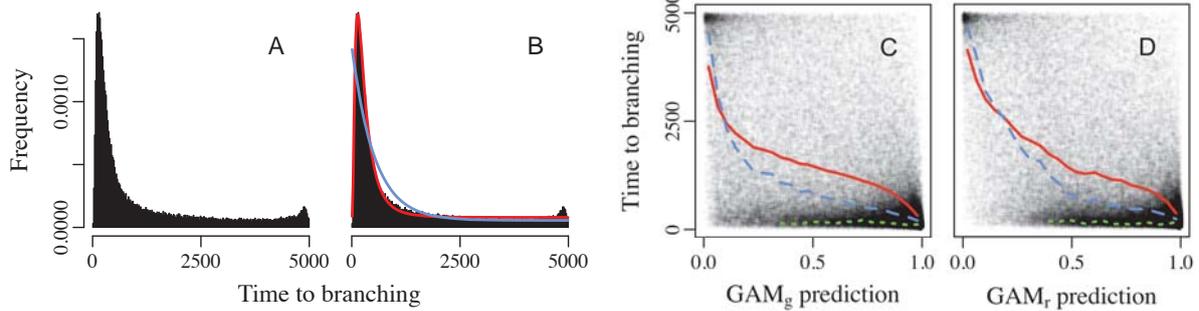
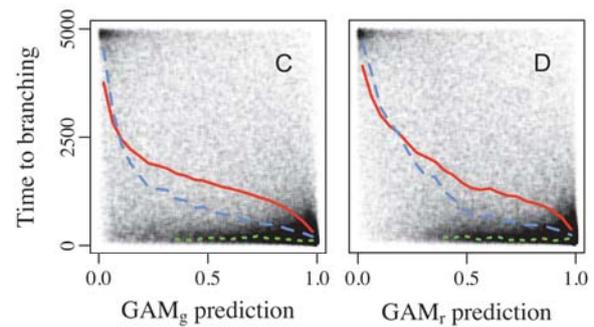


Figure A11: Distribution of branching times for branched realizations, and its relationship to the predicted probability of branching. A: The distribution of branching times for branched realizations, showing a strong peak indicating that branching typically occurs in the first ~1000 generations, with a long tail asymptotically approaching a non-zero limit, and a final bump (see text). B: The same distribution, overlaid with best fit lognormal (red) and exponential (blue) curves that include a height offset parameter. Curves were fitted to the branching times < 4500, to exclude the final bump (since it is an analysis artifact; see text). C–D: Scatter plots showing the prediction of GAM_g (C) and GAM_r (D) versus the branching time for branched realizations. Red solid, blue dashed, and green dotted curves respectively show the mean, median, and mode of 25 bins of equal x span; the mode is shown only when the early-branching peak is well-defined. The negative relationship between predicted branching probability and branching time shown by the mean and median curves is due to a progressive shift from early branching with a consistent modal branching time unrelated to the predicted probability (predicted branching probabilities > ~0.4) to less frequent early branching and thus a higher proportion of realizations with longer branching times (the long tail; predicted branching probabilities < ~0.4), and then into a regime in which the rare detection of branching is usually an analysis artifact (the final bump; predicted branching probabilities < ~0.1). Results shown are for reprising boundaries; qualitative results were the same for other boundary conditions (see text).

Correlations with branching times

Mean and median branching times correlated with some parameters and metrics; the modal branching time, however, was nearly invariant (Figure A12). These qualitative differences between the mean and median compared to the mode are due to changes in the proportion of realizations that branched early

(the modal peak of the branching time distribution) versus realizations that branched later (the long tail of the distribution; see *Distribution of branching times*). Large mean/median branching times may thus perhaps be understood as indicating that a high proportion of realizations had to establish a favorable spatial configuration prior to branching (see *Distribution of branching times*). The effects



presented are for reprising boundaries; results were qualitatively identical for other boundary conditions with the exception of an interaction between the boundary condition and the effect of dispersal distance, discussed below.

Increasing slope increased the mean/median branching time (Figure A12A), as did increasing curvature (Figure A12B) and increasing realized-landscape heterogeneity (Figure A12E). It is worth noting that the mean/median branching time was not

minimized at an intermediate value of these parameters (in contrast to Results: *Intermediate heterogeneity maximizes branching propensity*); this is another view on the result that branching time and branching probability are not closely related phenomena in our model. Interestingly, increasing patchiness amplitude had no effect on mean/median branching time, unlike the other types of heterogeneity (Figure A12C).

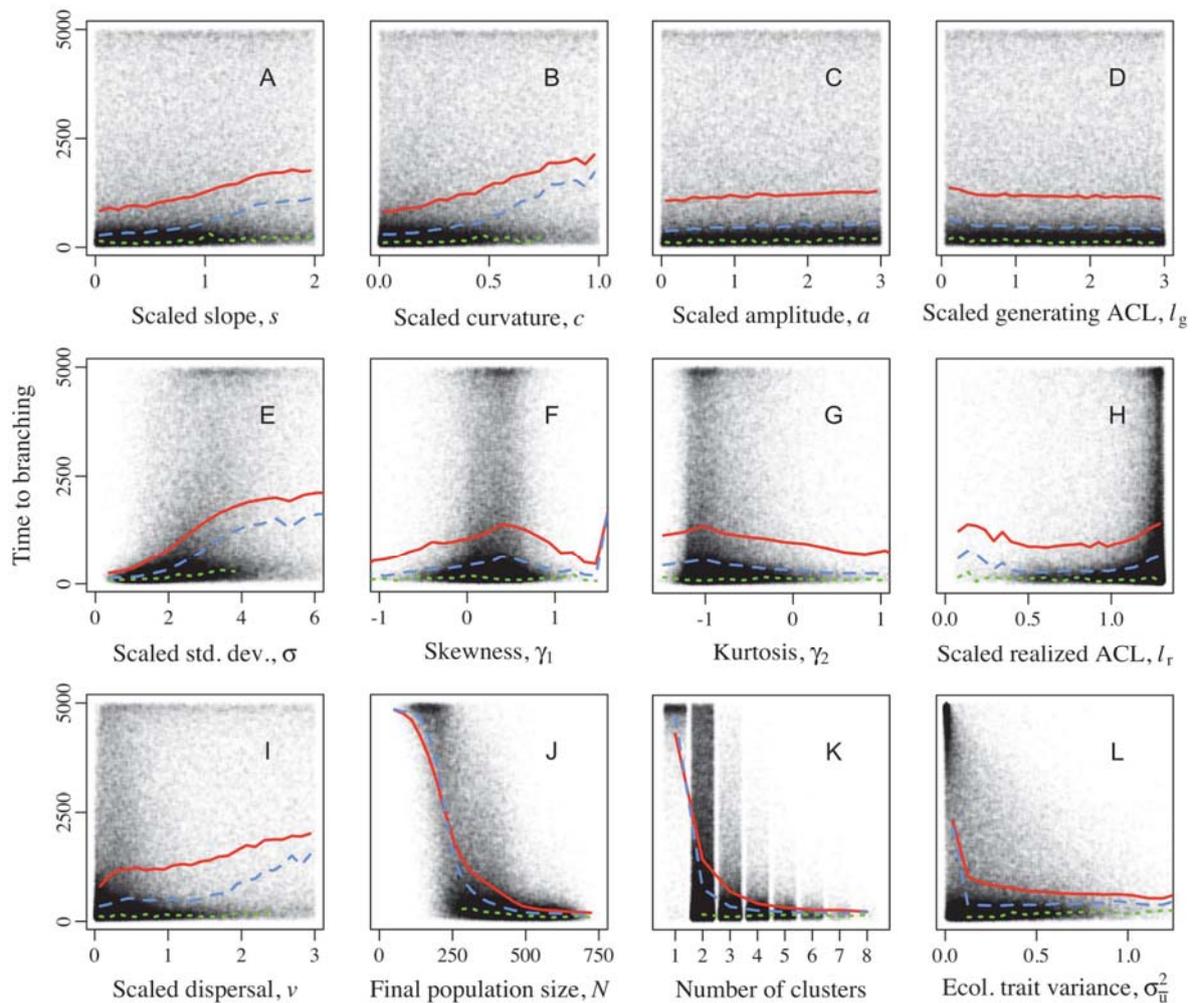


Figure A12: Correlations between branching times for branched realizations and landscape-generating parameters (top row), realized-landscape metrics (middle row), and other parameters and metrics (bottom row). Red solid, blue dashed, and green dotted curves respectively show the mean, median, and mode of 25 bins of equal x span (8 bins, for panel K); the mode is shown only when the early-branching peak is well-defined. The position of the early-branching peak was nearly constant in all panels, but positive relationships were observed between the mean/median branching time and several parameters and metrics: slope (A), curvature (B), realized-landscape standard deviation (E) and autocorrelation length (H), and dispersal distance (I). Horizontal jitter was added to panel K to separate points. Results shown are for reprising boundaries; qualitative results were the same for other boundary conditions, except that the strength of the effect of dispersal (panel I) varied (see text).

Large realized ACL increased the mean/median branching time (Figure A12H), presumably because the scale of heterogeneity was then so large that it often took populations a long time simply to reach a part of the environment that was ecologically different enough to produce branching. The generating ACL had no effect (Figure A12D), as might be expected (see *Landscape-generating parameters versus realized-landscape metrics*).

The mean/median branching time increased with increasing dispersal (Figure A12I). This effect was weak with reprising boundaries (Figure A12I) and reflecting boundaries, stronger with absorbing boundaries, and entirely absent with stopping

boundaries (not shown). This variation appears to be driven by the interaction between dispersal and the boundary condition (see *Effects of boundary conditions*). In particular, long-distance dispersal with stopping boundaries does not hinder branching; whether dispersal is short or long, realizations that are going to branch typically do so early (see *Distribution of branching times*), and so the relationship between dispersal distance and mean/median branching time is flat. With absorbing boundaries, long-distance dispersal is lethal and thus early branching is hindered; late branching, however, can still (very rarely) occur.

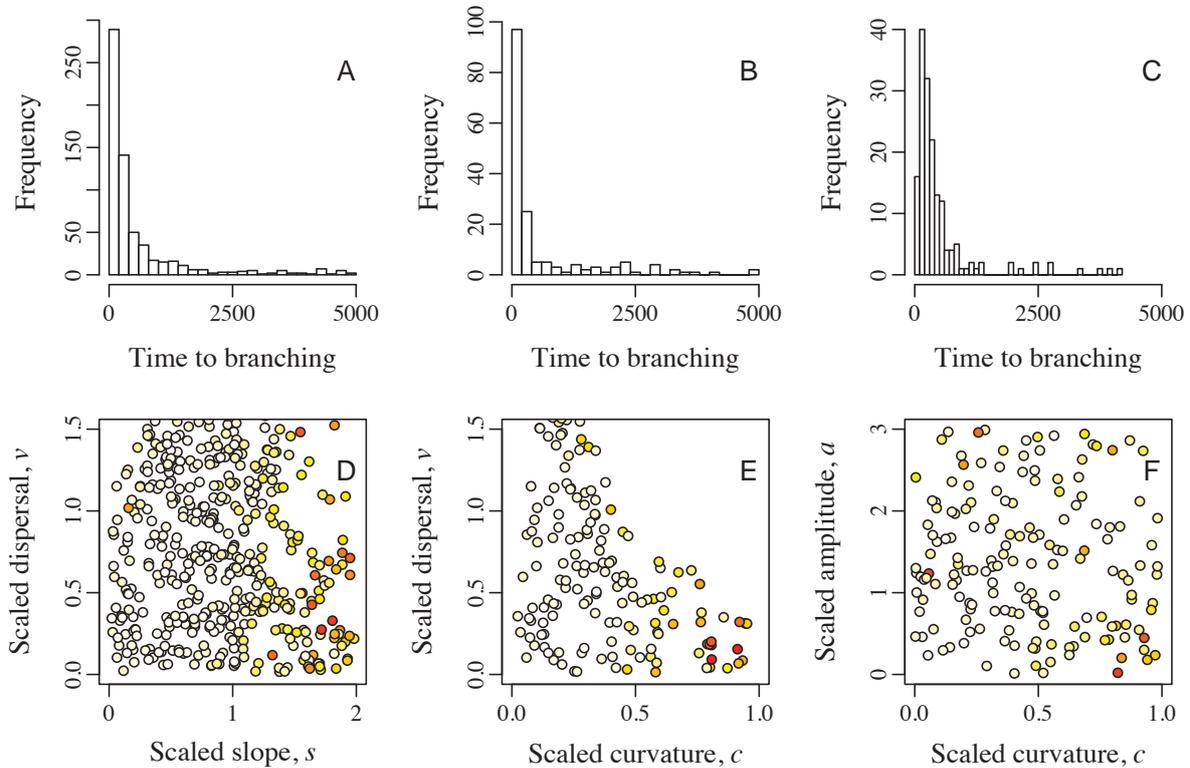


Figure A13: Alternate view of the parameter-region slices shown in Figures 3A (left column), 3B (middle column), and 4C (right column), showing branching times for branched realizations near those slices. A–C: distributions of branching times for each set of realizations, showing similarity to the overall distribution shown in Figure A11A. D–F: each set of realizations plotted following the axes of Figures 3A, 3B, and 4C; colors indicate branching time, from white (0 generations) to red (5000 generations). Comparison to Figures 3A, 3B, and 4C shows that intermediate probabilities of branching do not generally correspond to long branching times, although a weak association may exist in panel F. Instead, large slope and/or large curvature appear to cause long branching times (see also Figure A12). Parameter ranges for these plots: A/D: $a < 0.1$, $c < 0.1$, 631 realizations; B/E: $a < 0.1$, $s < 0.1$, 166 realizations; C/F: $s < 0.2$, $l_g < 0.3$, $\nu < 0.225$, 166 realizations (parameters not listed were unconstrained to maximize the number of realizations plotted). Results shown are for reprising boundaries; qualitative results are the same for other boundary conditions.

Finally, short mean/median branching times were associated with large population size (Figure A12J), a large number of phenotypic clusters (Figure A12K), and a large ecological trait variance (Figure A12L). These associations might appear to be almost tautological: the faster branching is, the more clusters are produced at equilibrium, the more variance there is in the ecological trait, and the larger population size can be (due to relaxed competition resulting from greater phenotypic variance). However, it could have been otherwise; if it were true here that “slow and steady wins the race”, then realizations with long branching times might have ultimately produced the most variation. Our result suggests instead that radiations producing a large final clade size should be expected to proceed more rapidly than smaller radiations. This may fit with empirical results, since some of the largest adaptive radiations known are believed to have occurred in a short time, such as the radiations of cichlids in African lakes (Seehausen 2002), of plants in the Andes (Hughes and Eastwood 2006) and in the Cape flora of South Africa (Richardson et al. 2001), and other cases (Schluter 2000); however, the generality of this pattern is not clear.

Figure A13 shows these results regarding branching times in the light of three particular cases, following three figures in the main text (Figures 3A, 3B, and 4C). Comparison of Figure A13 to those figures shows again that branching time and branching probability were unrelated in our model, but that branching time increased with increasing environmental gradient slope and curvature.

Complex landscapes

Some previous work has been done regarding the generation of complexly heterogeneous landscapes. A method similar to ours was briefly outlined by Cliff and Ord (1973; Appendix 1) but was not fully specified. Keitt (2000) describes a variety of methods for landscape generation; our method is essentially a special case of one of these. Similar ideas have also been developed in other fields, such as materials science (Hu and Tonder 1992; Wu 2000; Uchidate et al. 2010). However, these papers are not generally accessible to non-mathematicians; we thus felt there would be value in presenting our landscape generation algorithm explicitly.

A few previous models have looked at ecological and/or evolutionary dynamics on complex two-dimensional landscapes. However, the landscapes of these studies differ from ours in several important ways. In some models, the landscape consists of only two types of terrain, habitable and uninhabitable (e.g., Hiebeler 2000; Bonte et al. 2010), or two discrete resource types distributed heterogeneously (e.g., North et al. 2011b; Karonen 2012). In others, the environment does not change smoothly from one type to another, but instead exhibits abrupt transitions between types (e.g., Birand et al. 2012). Commonly, the landscapes modeled are considered to vary in suitability in some absolute sense, producing areas that are more habitable and areas that are less habitable, rather than allowing local adaptation to the different environment types to occur (e.g., North and Ovaskainen 2007; Stoddard 2010; North et al. 2011a). As far as we are aware, our model is unique in its incorporation of both spatially and ecologically continuous complex spatial heterogeneity that determines the locally optimal phenotype rather than the habitability *per se*.

Landscape generation

Given parameters S , C , L_g , and A , the landscapes used in our model are constructed in a six-step process. These steps are explained here.

First, a patchiness matrix with resolution sufficient to represent patchiness with autocorrelation length L_g (confirmed by post-analysis of the generated landscape) is filled with random values drawn from a Gaussian distribution with mean 0 and standard deviation A .

Second, this patchiness matrix is convoluted with a filter function (a modified Bessel function of the second kind, order 0) using Fourier transforms. This yields the desired periodic properties for the y -direction; in the x -direction the patchiness matrix is padded prior to convolution to remove the correlation that would otherwise result between the left and right edges.

Third, an environmental gradient matrix is constructed, of the same resolution as the patchiness matrix, expressing the linear and quadratic gradients as a function $e(x) = Sx + \frac{1}{2}Cx^2$.

Fourth, the final patchiness matrix and the environmental gradient matrix are added together to produce the preliminary landscape matrix.

Fifth, the preliminary landscape matrix is adjusted by adding a constant $\alpha = 0.5 - \mu_{\text{env}}$, where μ_{env} is the mean of the preliminary landscape matrix; this re-centers the matrix to have a mean value of 0.5, so that the initial population is adapted to the mean environmental value (see *Model description, PHENOTYPE RANGES AND INITIALIZATION*).

Sixth and last, bilinear interpolation is used to construct the canonical landscape matrix. The size of this matrix is constant, 512 by 512, for ease of implementation. This is a sufficiently high resolution to represent with high fidelity all landscapes generated by the values of S , C , L_g , and A used in this study. Values of $u_0(x, y)$ were obtained from the canonical landscape matrix, scaled to exactly span the dimensions $0 \leq x, y \leq 1$ of the landscape.

Landscape analysis

The canonical landscape matrix (see previous section, *Landscape generation*) was analyzed to derive metrics regarding the heterogeneity of the landscape. The scaled standard deviation σ , skewness γ_1 , and excess kurtosis γ_2 were computed across all the values in the canonical landscape matrix, thus characterizing the distribution of environmental values represented in the landscape without regard to their spatial distribution. The standard deviation was scaled by σ_{κ} (Table 1) in order to yield a dimensionless metric of the environmental heterogeneity, σ .

Additionally, a metric of the characteristic spatial scale of heterogeneity in the landscape was computed. First, the autocorrelation function (ACF) for the canonical landscape matrix was computed via the Wiener-Khinchin Theorem, using Fast Fourier Transforms (FFTs); specifically, the autocorrelation function is $F^{-1}(F(E)\overline{F}(E))$, where F denotes the Fourier transform, F^{-1} the inverse transform, \overline{F} the complex conjugate, and E the canonical landscape matrix. To eliminate the periodicity in the x -direction implicitly imposed by the Fourier transform, flipped copies of the landscape were pasted to its left and right edge. Given the ACF (normalized to a maximum value of 1), the autocorrelation length (ACL) was estimated as the

lag at which the ACF first crossed below the threshold value $1/e$ (using linear interpolation to estimate the exact position of crossing, between the last lag above the threshold and the first lag below it). The lag was converted to the model coordinate system by dividing by 512, and was then divided by σ_{κ} to yield the scaled realized ACL, l_r , of the landscape. This scaled realized ACL is a dimensionless metric comparable to the scaled generating ACL, l_g .

Several other ACL estimation methods were tried, including fitting an exponential curve to the ACF, and using estimates of Moran's I and Geary's c . Although these metrics were all strongly correlated with l_r , they produced less accurate predictions of the branching propensity of the landscape, indicating that they did not capture as much useful information as the $1/e$ crossing point. The $1/e$ crossing point was therefore used to estimate the realized ACL, l_r , in all subsequent analyses.

Patterns of evolutionary branching

In the analysis presented, we have focused on the effects of model parameters on the probability of branching during a realization. We have also presented data regarding the timing of branching (see *Distribution of branching times* and *Correlations with branching times*), and have briefly considered the number of branches generated (see *Alternative assessment of branching* and *Correlations with branching times*). Except for the example realization shown in Figure 2, however, we have thus far left unaddressed more qualitative questions: What does the typical pattern of branching in a realization look like? What do landscapes that cause branching, or do not cause branching, look like? How stable is the branching in a typical realization? What does the spatial distribution of branched and unbranched populations look like? What, if anything, differentiates landscapes that allow a single branching event from landscapes that allow larger adaptive radiations? To answer these questions, we have chosen a small selection of unbranched (Figure A14) and branched (Figure A15) realizations that illustrate different types of observed model dynamics.

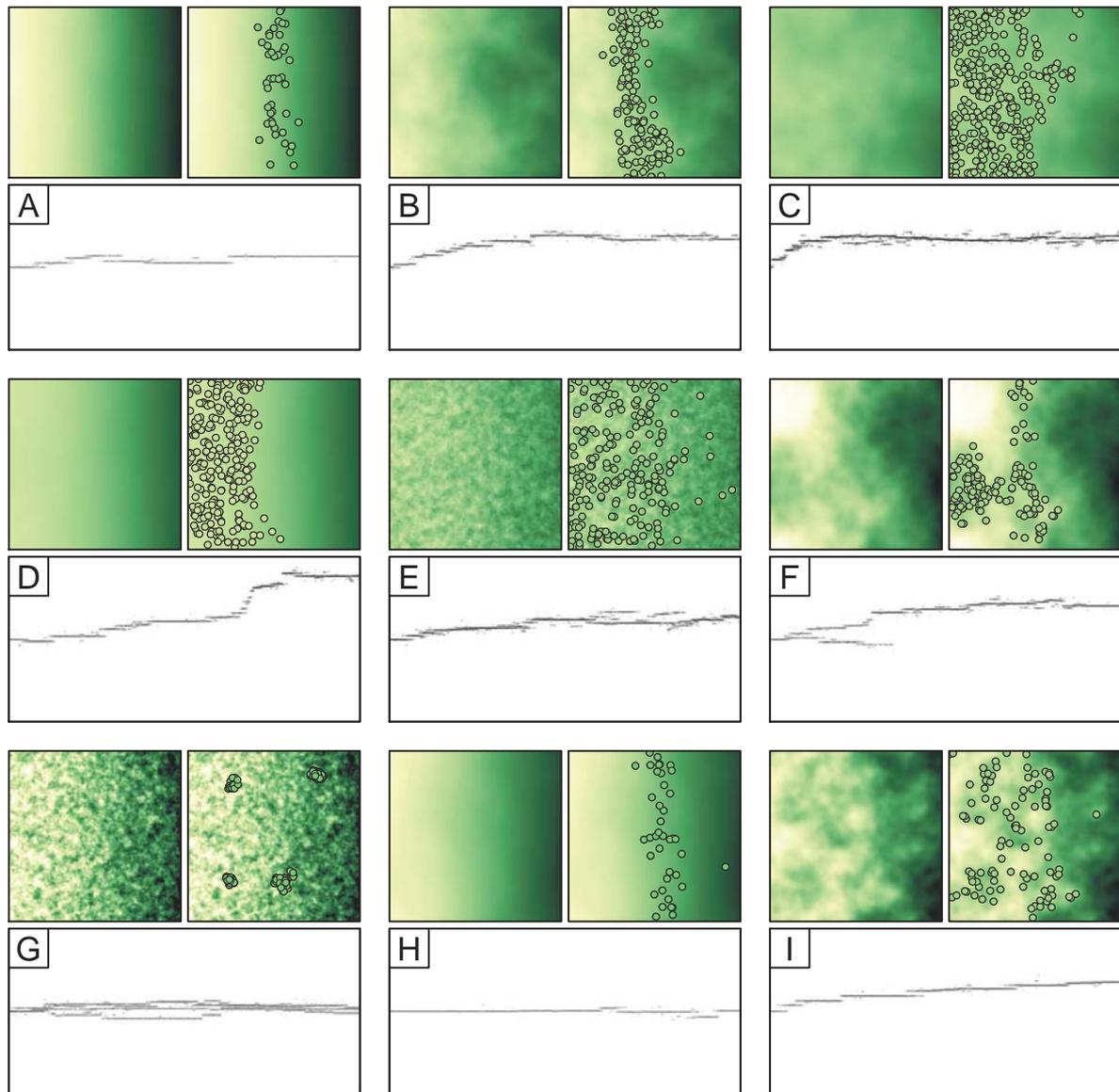


Figure A14: A selection of realizations that did not branch. Each panel is composed of three subpanels: the landscape (upper left), the end-of-realization population census (upper right), and the evolutionary history (bottom). Landscape colors indicate locally optimal ecological trait values, while circle colors indicate the actual trait values of individuals, both ranging from white (low) to dark green (high). The evolutionary history shows ecological trait values (y-axis) through time (from the beginning of the realization, on the left, to the end at 5000 generations, on the right). A: a steep gradient can inhibit branching; note the small population size due to low mean fitness. B: a shallower gradient allows a larger population size and higher genetic variability. C: a relatively flat landscape with mild patchiness allows even larger population size and genetic variability. D: curvature pushes the population to the left, toward the shallower end of the gradient. E: fine-grained heterogeneity promotes branching, but branches may be ephemeral. F: strong patchiness promotes branching, but it can be difficult to persist; note how the population's distribution fits to the shape of the landscape. G: very short-range dispersal creates highly localized clusters, but with extreme heterogeneity even this may be insufficient to allow branching. H: A steep gradient hinders branching; note the lone colonist on the right, which will surely die. I: Strong patchiness also hinders branching, but refugia may mitigate this; note the lone colonist on the right, which might establish a foothold and found a new branch.

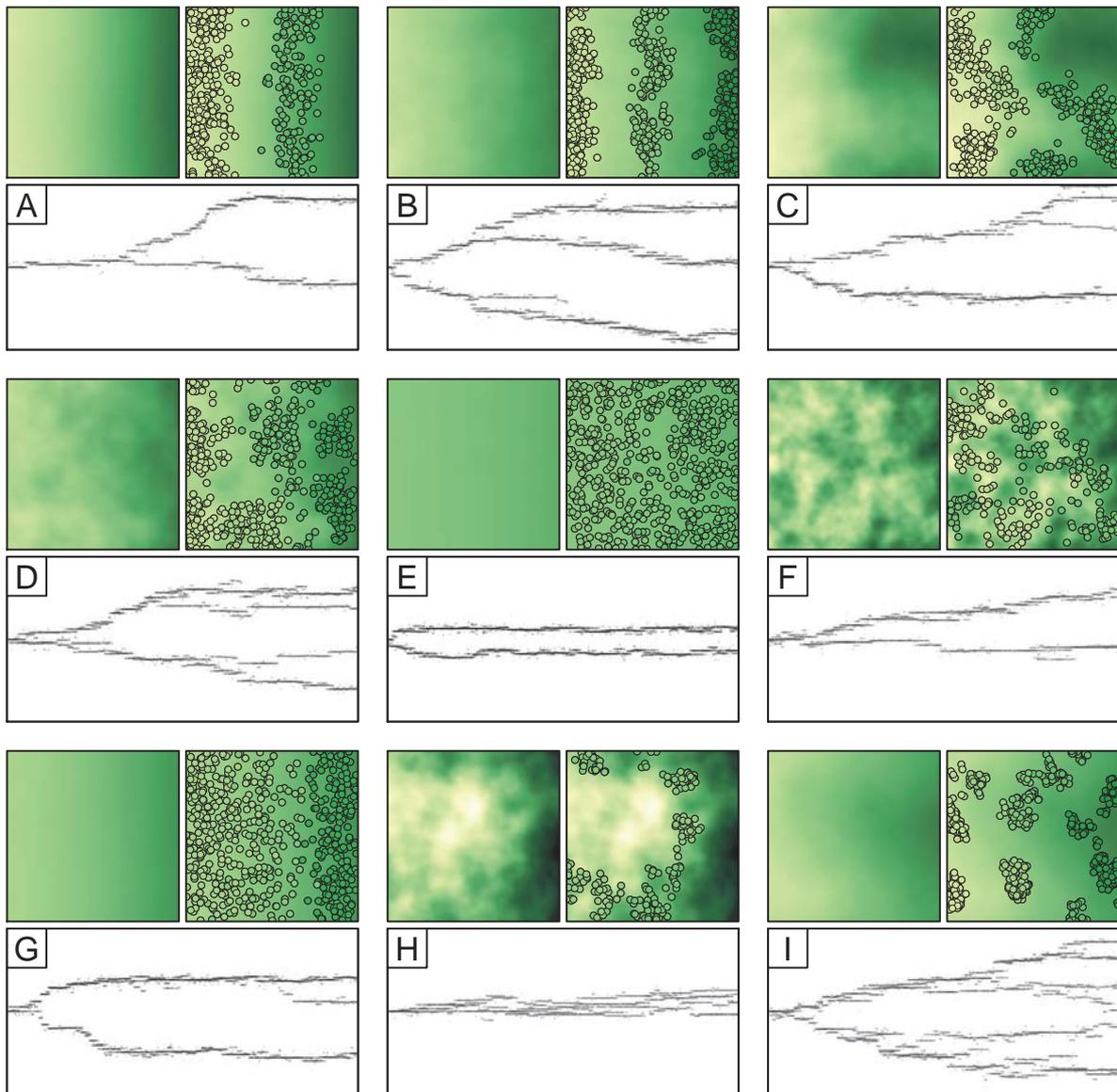


Figure A15: A selection of realizations that branched. Each panel is composed of three subpanels: the landscape (upper left), the end-of-realization population census (upper right), and the evolutionary history (bottom). Landscape colors indicate locally optimal ecological trait values, while circle colors indicate the actual trait values of individuals, both ranging from white (low) to dark green (high). The evolutionary history shows ecological trait values (y -axis) through time (from the beginning of the realization, on the left, to the end at 5000 generations, on the right). A: simple branching on a gradient; note that branching was somewhat delayed. B: sometimes several branches can become established; however, note the transient branch that went extinct. C: large-scale patchiness creates local gradients that can promote branching. D: medium-scale patchiness can support several locally-adapted groups. E: close to pure competitive branching (subdivision of a homogeneous resource maintained by character displacement due to competition); a very slight gradient is present, however. F: strong patchiness can subdivide the landscape into interleaved, fragmented habitat types. G: branching with long-range dispersal on a curved gradient; note that the flatter part of the gradient supports a larger population, allowing further branching to occur (although it may not persist). H: very strong patchiness can restrict the population to a single zone, just as with a very steep gradient. I: A rare example of extreme adaptive radiation, the product of an optimally intermediate degree of heterogeneity, a short dispersal distance, and luck.

Beyond remarks on individual realizations (see captions, Figure A14 and A15), some general observations can be made. Perhaps the broadest observation is that branching is a one-way process in this model; “reverse speciation” (Seehausen 2006) does not occur, although branches occasionally go extinct. This is likely for two reasons. First, the collapse of incipient branching due to non-assortative mating does not occur since this model is asexual; this might be expected to occur in a sexual version of the model, until the strength of assortative mating grew sufficiently strong to prevent hybridization (Felsenstein 1981; Dieckmann and Doebeli 1999; Doebeli and Dieckmann 2003). Second, the heterogeneity in this model is temporally invariant, and thus causes a constant selective pressure towards diversification; temporal variation in environmental heterogeneity, on the other hand, might be expected to lead to more ephemeral divergence (Seehausen 2006).

Another common pattern is that of evolutionary stasis (Estes and Arnold 2007) apart from branching and extinction (“punctuated equilibrium”, Eldredge and Gould 1972; Gould and Eldredge 1977; Gould and Eldredge 1993). The phenotype of unbranched lineages often wanders over time due to drift (e.g., Figures A14B, A14D, A14I), particularly on landscapes with weak heterogeneity that provide some variety in local conditions, but not so much variation (e.g., Figures A14A, A14H) as to constrain evolution through strong selection. Branched lineages, on the other hand, more typically reach an equilibrium and exhibit stasis thenceforth (Figure A15) – although reaching equilibrium may take a long time, as populations jockey for position on both the physical landscape and on an adaptive landscape with peaks that shift due to frequency-dependent selection. Once equilibrium is reached, each branch exerts selection on every other branch, due to competition, and this stabilizes the system. Destabilization occurs, however, when a new branch arises, causing other branches to evolve away to avoid competitive pressure (e.g., Figures A15C, A15D). Similarly, when a branch goes extinct this may remove previously existing competitive pressures, allowing other branches to stop diverging (e.g., Figure A14F) or even evolve toward the now-unoccupied niche (e.g., Figure A15B). Nevertheless,

after character displacement has equilibrated, stasis is generally again observed.

These qualitative patterns are not here developed to the point of rigorously shown results; rather we offer them as interesting observations, in the spirit of “natural philosophy”, to spur further inquiry into less-studied aspects of the process of diversification in biological systems.

Literature Cited

- Birand, A., A. Vose, and S. Gavrilets. 2012. Patterns of species ranges, speciation, and extinction. *American Naturalist* 179:1–21.
- Bonte, D., T. Hovestadt, and H. J. Poethke. 2010. Evolution of dispersal polymorphism and local adaptation of dispersal distance in spatially structured landscapes. *Oikos* 119:560–566.
- Cliff, A. D., and J. K. Ord. 1973. *Spatial autocorrelation: Monographs in spatial and environmental systems analysis*. Pion Limited, London.
- Dieckmann, U., and M. Doebeli. 1999. On the origin of species by sympatric speciation. *Nature* 400:354–357.
- Doebeli, M., and U. Dieckmann. 2003. Speciation along environmental gradients. *Nature* 421:259–264.
- Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: An alternative to phyletic gradualism, Pages 82–115 in T. J. M. Schopf, ed. *Models in Paleobiology*. Freeman, Cooper & Co., San Francisco.
- Estes, S., and S. J. Arnold. 2007. Resolving the paradox of stasis: Models with stabilizing selection explain evolutionary divergence on all timescales. *American Naturalist* 169:227–244.
- Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or Why are there so few kinds of animals? *Evolution* 35:124–138.
- Fraley, C., and A. E. Raftery. 2003. Enhanced model-based clustering, density estimation, and discriminant analysis software: MCLUST. *Journal of Classification* 20:263–286.
- Gillespie, D. T. 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* 22:403–434.
- Gould, S. J., and N. Eldredge. 1977. Punctuated equilibria: The tempo and mode of evolution reconsidered. *Paleobiology* 3:115–151.
- Gould, S. J., and N. Eldredge. 1993. Punctuated equilibrium comes of age. *Nature* 366:223–227.
- Grimm, V. 2002. Visual debugging: A way of analyzing, understanding and communicating bottom-up simulation models in ecology. *Natural Resource Modeling* 15:23–38.
- Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard et al. 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198:115–126.
- Heinz, S. K., R. Mazzucco, and U. Dieckmann. 2009. Speciation and the evolution of dispersal along environmental gradients. *Evolutionary Ecology* 23:53–70.
- Hiebeler, D. 2000. Populations on fragmented landscapes with spatially structured heterogeneities: Landscape generation and local dispersal. *Ecology* 81:1629–1641.
- Hu, Y. Z., and K. Tonder. 1992. Simulation of 3-D random rough surface by 2-D digital filter and Fourier analysis. *International Journal of Machine Tools & Manufacture* 32:83–90.
- Hughes, C., and R. Eastwood. 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences of the United States of America* 103:10334–10339.
- Karonen, I. 2012. Stable trimorphic coexistence in a lattice model of spatial competition with two site types. *Journal of Theoretical Biology* 295:77–85.
- Keitt, T. H. 2000. Spectral representation of neutral landscapes. *Landscape Ecology* 15:479–493.
- Mazzucco, R., M. Doebeli, and U. Dieckmann. (unpublished manuscript). The influence of habitat boundaries on evolutionary branching along environmental gradients.
- North, A., S. Cornell, and O. Ovaskainen. 2011a. Evolutionary responses of dispersal distance to landscape structure and habitat loss. *Evolution* 65:1739–1751.
- North, A., and O. Ovaskainen. 2007. Interactions between dispersal, competition, and landscape heterogeneity. *Oikos* 116:1106–1119.
- North, A., J. Pennanen, O. Ovaskainen, and A. L. Laine. 2011b. Local adaptation in a changing world: The roles of gene-flow, mutation, and sexual reproduction. *Evolution* 65:79–89.
- Richardson, J. E., F. M. Weitz, M. F. Fay, Q. C. B. Cronk, H. P. Linder, G. Reeves, and M. W. Chase. 2001. Rapid and recent origin of species richness in the Cape flora of South Africa. *Nature* 412:181–183.
- Schluter, D. 2000. *The Ecology of Adaptive Radiation*. Oxford Series in Ecology and Evolution. Oxford University Press, Oxford.
- Seehausen, O. 2002. Patterns in fish radiation are compatible with Pleistocene desiccation of Lake Victoria and 14 600 year history for its cichlid species flock. *Proceedings of the Royal Society B-Biological Sciences* 269:491–497.
- . 2006. Conservation: Losing biodiversity by reverse speciation. *Current Biology* 16:R334–R337.
- Stoddard, S. T. 2010. Continuous versus binary representations of landscape heterogeneity in spatially-explicit models of mobile populations. *Ecological Modelling* 221:2409–2414.
- Uchida, M., T. Shimizu, and A. Iwabuchi. 2010. Data generation techniques of three-dimensional surface roughness for tribology simulations. *Journal of Japanese Society of Tribologists* 55:375–380.
- Wakano, J. Y., and Y. Iwasa. 2013. Evolutionary Branching in a Finite Population: Deterministic Branching vs. Stochastic Branching. *Genetics* 193:229–241.
- Wallace, B. 1975. Hard and soft selection revisited. *Evolution* 29:465–473.
- Wu, J. J. 2000. Simulation of rough surfaces with FFT. *Tribology International* 33:47–58.