The effects of ecological selection on species diversity and trait distribution: predictions and an empirical test

NIV DEMALACH,1,2,5 PO-JU KE,1,3,4 AND TADASHI FUKAMI1

1Department of Biology, Stanford University, Stanford, California 94305 USA
2Institute of Plant Sciences and Genetics in Agriculture, Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 7610001 Israel
3Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544 USA
4Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei 10617 Taiwan

Citation: DeMalach, N., P.-J. Ke, and T. Fukami. 2022. The effects of ecological selection on species diversity and trait distribution: predictions and an empirical test. Ecology 103(1):e03567. 10.1002/ecy.3567

Abstract. Ecological selection is a major driver of community assembly. Selection is classified as stabilizing when species with intermediate trait values gain the highest reproductive success, whereas selection is considered directional when fitness is highest for species with extreme trait values. Previous studies have investigated the effects of different selection types on trait distribution, but the effects of selection on species diversity have remained unclear. Here, we propose a framework for inferring the type and strength of selection by studying species diversity and trait distribution together against null expectations. We use a simulation model to confirm our prediction that directional selection should lead to lower species diversity than stabilizing selection despite a similar effect on trait community-weighted variance. We apply the framework to a mesocosm system of annual plants to test whether differences in species diversity between two habitats that vary in productivity are related to differences in selection on seed mass. We show that, in both habitats, species diversity was lower than the null expectation, but that species diversity was lower in the more productive habitat. We attribute this difference to strong directional selection for large-seeded species in the productive habitat as indicated by trait community-weighted mean being higher and community-weighted variance being lower than the null expectations. In the less productive habitat, we found that community-weighted variance was higher than expected by chance, suggesting that seed mass could be a driver of niche partitioning under such conditions. Altogether, our results suggest that viewing species diversity and trait distribution as interrelated patterns driven by the same process, ecological selection, is helpful in understanding community assembly.

Key words: annual plants; community-weighted mean; community-weighted variance; competition; environmental filtering; functional diversity; functional traits; seed mass; seed size; species pool; species richness; trait selection.

INTRODUCTION

A major goal of ecology is to understand the mechanisms behind patterns of species diversity and trait distribution (Grime 1979, Tilman 1982, Huston 1994, Chesson 2000, Chase and Leibold 2013, HilleRisLambers et al. 2012). The theory of ecological communities (Vellend 2010, 2016) argues that variation among communities results from four high-level processes: ecological selection, dispersal, ecological drift, and speciation.

Ecological selection, the biotic and abiotic filtering of species from the species pool of potential colonizers, has been a major focus of studies on species diversity (Cornwell and Ackerly 2009, Kraft et al. 2015, Vellend 2016, Germain et al. 2018). Selection can be characterized by its strength, which increases with increasing differences in reproductive success among species (ecological fitness sensu Vellend [2016]). Additionally, selection is characterized by type: it is stabilizing when species with intermediate trait values gain the highest reproductive success, whereas it is directional when species with extreme trait values gain the highest fitness (Fig. 1). Originally, this classification was proposed for the evolutionary selection of phenotypes within populations but it is currently applied for the ecological selection of traits within communities (Shipley 2010, Vellend 2016, Loranger et al. 2018) as we do here.

Both types of selection are thought to reduce community-weighted variance (CWV) of the traits being selected but have different effects on community-weighted mean (CWM; Cornwell and Ackerly 2009, Rolfauser and Pucheta 2017, Loranger et al. 2018). The CWM should be similar to the mean of the trait...
distribution of the species pool under stabilizing selection, whereas it should vary under directional selection. Species diversity is expected to decrease with increasing selection strength because larger differences in fitness lead to faster exclusion (Vellend 2016). Nonetheless, the relationships between selection type and species diversity have not been fully investigated.

Since species diversity and trait distribution are both driven by ecological selection, we argue that they should be studied together within the same framework. We therefore propose to characterize selection attributes (i.e., strength and type) by simultaneously studying species diversity and trait distribution against null expectations built from the species pool. Specifically, let us assume the same species pool is shared across different communities, which could be different sites or habitat types within a heterogeneous landscape. Additionally, let us assume that trait differences among species reflect competitive hierarchy instead of niche partitioning, i.e., ecological fitness decreases with increasing trait distance from the optimal value. These simplifying assumptions follow previous models of trait selection in ecological communities (Shipley 2010, Loranger et al. 2018) and their consequences are discussed toward the end of the manuscript.

Under the above assumptions, one would have evidence that selection had taken place if species diversity was lower than the null expectation (Fig. 2A). In addition, the hypothesis that a specific trait has been under selection would be supported if the observed CWV was lower than the null expectation (Fig. 2B). We can also infer whether the selection was stabilizing or directional based on the deviation of the observed CWM from the null expectation (Fig. 2C). Stronger directional selection is expected to increase the deviation between observed CWM and the null expectation.

We investigated the proposed framework using a simulation model and applied it to an experimental case study. Our simulation model aims to verify the logic of the framework, i.e., that CWM and CWV can indicate the type and strength of ecological selection, respectively. Another aim of the model was to investigate the effects of different selection types on species diversity. Specifically, we expected that when the trait distribution of the species pool is unimodal, directional selection would lead to lower species diversity than stabilizing selection because there are fewer species with extreme values.

To demonstrate the utility of our framework for understanding community assembly, we reanalyzed data from a mesocosm experiment of annual plant communities (Ron et al. 2018, DeMalach et al. 2019). In that system, the same set of species was sown under different levels of resource availability, which enabled characterizing the selection processes. In the analysis presented here, we test the theoretical prediction (DeMalach and Kadmon 2018) that under low resource availability selection on seed size should be weak and stabilizing while under high resource availability it should be strong and directional due to asymmetric light competition that favors large-seeded species. Together, the simulation and the case study show that the selection mechanisms can be identified only when trait distribution and species diversity are measured simultaneously.
METHODS

Simulation model

Our spatially implicit model describes population dynamics in a meta-community comprised of $n$ local communities. Competition occurs within each local community, and the local communities are connected by dispersal. For simplicity, the model assumes that the local communities have a fixed size and that there is no overlap among generations, as in annual species. In each time step, proportion $D$ of the community arrives from other local communities (hereafter dispersers), proportion $I$ arrives from outside the meta-community (hereafter immigrants), and the rest are descendants of individuals from the local community (hereafter residents).

Among each of the residents, the probability of belonging to species ($P_i$) in timestep $t+1$ is determined by the following equation:

$$P_i(t+1) = \frac{\omega_i f_i(t)}{\sum_{j=1}^{S} \omega_j f_j(t)}$$

(1)

where $f_i(t)$ is the frequency of species $i$ in the local community (in the previous time step), $\omega_i$ is its ecological fitness (mean reproductive success), and $S$ is the number of species in the species pool (potential colonizers). A similar probabilistic rule applies for dispersers with the only difference being that meta-community frequency is used instead of the local community (i.e., they have an equal chance to arrive from all local communities). All species have the same (extremely low) probability to arrive as immigrants from the species pool.

Fig. 2. A scheme for inferring selection attributes based on the comparison between observed patterns and null expectations. (A) If species diversity is lower than the null expectation then selection has occurred (the larger the difference the stronger the selection). (B) If community-weighted variance (CWV) of a particular trait is lower than the null expectation it implies a trait-specific selection for that trait (the larger the difference the stronger the selection). (C) If community-weighted mean (CWM) of a particular trait differ from the null expectation it implies a directional selection. Alternatively, if CWM is similar, it implies a stabilizing selection (assuming the previous step has shown that CWV is lower than expected).
The ecological fitness of each species is determined by its specific trait value $\delta$, based on the following Lorentzian function:

$$ω_i = \frac{1}{1 + \theta(\delta_{\text{best}} - \delta_i)^2},$$

where $\delta_{\text{best}}$ represents the optimal trait value and $\theta$ determines the strength of the selection, i.e., the degree of the fitness differences for a given trait distance. When $\theta = 0$, all species have equal fitness while increasing $\theta$ intensifies fitness differences. This Lorentzian function was used to restrict $ω_i$ to be positive for all values of selection strength. The trait values ($\delta_i$) in the species pool were assumed to be normally distributed ($\delta_{\text{mean}}$ and $\delta_{\text{SD}}$ are the mean and the SD of this distribution). For simplicity, the simulation focused on two scenarios, representing two extremes of a continuum. In the first, the simulation focused on two scenarios, representing the mean and the SD of this distribution. For simplicity, being measured (Appendix S1: Fig. S10).

**alternative fitness function** (Appendix S1: Figs. S7–S9), where $\omega_i$ can be expressed as a constant (Appendix S1: Figs. S11). 

The ecological fitness of each species is determined by finding the value in the pool (hereafter directional selection). In the second, it was equal to the highest value in the pool (hereafter directional selection).

All local communities started from a uniform abundance distribution of all species. We ran the model for 5,000 timesteps. Visual inspection suggests that communities reached equilibrium by approximately 3,000 timesteps (Appendix S1: Figs. S1–S3). We conducted three simulation runs for each parameter combination that we investigated (differences were minor). All the results represent the means of the three simulation runs, averaged from time step 3,000 to 5,000. The description of parameters and their values in the simulation are found in Table 1. In Appendix S1, we tested the robustness of the model by investigating different assumptions about the trait distribution of the species pool (Appendix S1: Figs. S4–S6), alternative fitness function (Appendix S1: Figs. S7–S9), and a scenario when the trait being selected is not the trait being measured (Appendix S1: Fig. S10).

For each simulation run, we calculated CWM and CWV. Additionally, we calculated species diversity under two scales (local communities and metacommunity), in terms of species richness and inverse Simpson index (hereafter Simpson diversity).

### Table 1. Parameters of the simulation model.

| Parameter | Description | Value(s) |
|-----------|-------------|----------|
| $S$       | species pool size (number of species) | 100 |
| $N$       | local community size (number of individuals) | 1,000 |
| $n$       | number of local communities | 10 |
| $\delta_{\text{mean}}$ | trait mean (arbitrary units) | 0 |
| $\delta_{\text{SD}}$ | the standard deviation of the trait (arbitrary units) | 0.5 |
| $\theta$ | selection strength (dimensionless) | $10^{-3}$ to $10^{3}$ |
| $D$       | the proportion of dispersed individuals (arbitrary units) | $10^{-3}$ |
| $I$       | the proportion of immigrants from the species pool | $5 \times 10^{-4}$ |

**Mesocosm experiment**

We applied the framework to a mesocosm experiment of annual plants growing in two habitats varying in soil depth (55 cm and 18 cm) and therefore productivity (Ron et al. 2018, DeMalach et al. 2019), hereafter referred to as the productive and the less productive habitat, respectively. We focused on seed mass selection patterns because previous analysis (DeMalach et al. 2019) has shown that it is the main predictor for abundance patterns along natural and experimental soil depth gradients (other measured traits were found to be insignificant).

A detailed description of the experimental system is found in the original papers (Ron et al. 2018, DeMalach et al. 2019). Briefly, the experiment was conducted at the botanical gardens of the Hebrew University of Jerusalem in Israel and consisted of nine artificial plant communities for each soil depth category. The mesocosm communities were established within metal containers with an area of $1 \times 1 \text{ m}$. In December 2011, 51 annual species were sown in equal density (200 seeds per species, a total of $51 \times 200 = 10,200$ seeds per container). The species emerging in each container were let to grow and interact for five successive years (2011–2016) following their germination. All containers were blocked against dispersal (using mesh nets), which enabled interpreting all patterns as consequences of selection and drift only.

At the (experimental) species pool level, seed mass was log-normally distributed (Appendix S1: Fig. S11, Table S1) and therefore our analysis was based on log$_{10}$-transformed seed mass data (as in most analyses of seed mass patterns). We used abundance data from the fifth growing season in a fixed quadrat of $25 \times 25 \text{ cm}$ at the center of each container for calculating species diversity (species richness and inverse Simpson’s index) and seed mass patterns (CWM and CWV).

To test whether selection occurred (regardless of which trait was selected for, Fig. 2A) we compared the species diversity patterns with a simulation model of drift dynamics (i.e., $\theta = 0$) based on the specific parameter values of the mesocosm system (Appendix S1: Table S2). Although the experiment included 51 species, in the drift model species pool size was set to 47 based on the number of species blooming during the first year. This conservative assumption was used to avoid finding false evidence for selection based on technical artifacts (e.g., non-viable seeds) of the experiment. The estimation of community size was based on the mean number of individuals measured in each container multiplied by 16 (the ratio between the sampled area and the total community). Initial composition was assumed to be a random sample from a multinomial distribution where all species have the same chance to be sampled (since sowing density was equal). To generate a distribution of outcomes we ran the drift model 1,000 times (for each iteration, we sample from a multinomial distribution where all species have the same chance to be sampled (since sowing density was equal)).

Based on our simulation results (Appendix S1: Fig. S10), a decrease in CWV could be driven by a selection...
acting on a different uncorrelated trait than the specific traits under investigation. While one could use dynamic simulations to determine whether the selection was specifically related to seed mass, this would require imposing assumptions from the theoretical model on the empirical data (e.g., the specific function relating fitness to traits). Instead, for CWM and CWV we used a trait-shuffling approach to generate the null expectation. We used the observed species abundance distribution from the mesocosm communities, assigned for each species a random seed mass value from the “species pool” (i.e., the seed mass values of the sown species), and calculated the average of the nine communities in each treatment. Then, we compared the observed patterns of CWM and CWV to the expectation from 10,000 different randomizations.

**Results**

**Simulation**

The model supports the assumptions of our framework that the CWM is mainly determined by selection type and CWV is mainly determined by selection strength (Fig. 3). The CWM differs from the mean of the species pool only under directional selection while the CWV is almost unaffected by selection type.

As expected, species diversity decreases with increasing selection strength (Fig. 3). However, under any given level of selection strength, diversity is lower under directional selection. These results are robust to the scale (local scale vs. meta-community scale) and the diversity indices (Simpson diversity vs. species richness). Moreover, transient dynamics are qualitatively similar to equilibrium results (Appendix S1: Figs. S1–S3).

The lower diversity under directional selection is driven by two mechanisms operating in the same direction. First, as the species pool’s trait distribution is normally distributed, species with intermediate traits have more similar fitness, which reduces extinction rate and enhances diversity under a given level of colonization. Accordingly, the difference between the two selection types is smaller under a uniform trait distribution in the species pool (Appendix S1: Figs. S4–S6). Still, diversity is higher under stabilizing selection, even under a uniform species pool trait distribution because of boundary constraints (similar to the mid-domain effect; Letten et al. 2013), where only under stabilizing selection species with traits close to the optimum are found in both sides of the optimum. Our findings seem general and not restricted to the function we have chosen, because similar results were observed under alternative fitness function (Appendix S1: Figs. S7–S9) and other alternative assumptions (Appendix S1: Figs. S1–S10).

**Mesocosm experiment**

Species diversity was lower in the productive habitat (Fig. 4). In both habitats, species diversity was lower than the null expectation generated from the drift model, indicating that selection occurred. In accordance with our hypothesis, CWV was lower and CWM was higher than the null expectation in the productive habitat, implying directional selection for large-seeded species. In the less productive habitat, CWM was similar to the null expectation. Contrary to our hypothesis of stabilizing selection, CWV in the less productive habitat was higher (rather than lower) than the null expectation.

**Discussion**

Recently, it was proposed that characterizing the type and strength of ecological selection could improve our understanding of the drivers of community assembly (Shipley 2010, Vellend 2016, Loranger et al. 2018). Here we sought to infer the underlying selection characteristics by measuring species diversity, CWM, and CWV in the local community and trait distribution in the species pool. We demonstrated the utility of our framework by analyzing data from a mesocosm experiment that included two habitats that varied in their productivity. By comparing species diversity to the null expectation, we showed that selection took place in both habitats but species loss was more severe in the productive habitat. The differences in species diversity probably arose from strong directional selection in seed mass that occurred only in the productive habitat. Below, we discuss our findings, elaborate on the simplifying assumptions of our approach, and highlight their implications.

**Interpretation of the mesocosm experiment**

We sought to explain patterns of species diversity in the mesocosm experiment based on seed mass selection. Previous analysis of this system (DeMalach et al. 2019) has shown that seed mass is the main predictor of species habitat preferences; this raised the hypothesis that species diversity patterns in that system are related to seed size selection. Here, we tested this hypothesis by comparing the observed patterns of CWM, CWV, and species diversity to the null expectation. Specifically, we predicted strong directional selection under high productivity and weaker stabilizing selection under low productivity. These predictions were driven by a resource competition model (DeMalach and Kadmon 2018) suggesting that large-seed species (that produce large seedlings) would be favored under high productivity where light competition was more intense.

In both habitats, selection took place, as indicated by species diversity being lower than the null expectation (Fig. 2). As we predicted, in the productive habitat, CWM was higher and CWV was lower than the null expectation, implying strong and directional selection for large seed mass. This strong directional selection was probably one of the main drivers of the low diversity under these conditions.
In the less productive habitat, we found that, contrary to our prediction of stabilizing selection, CWV was higher (rather than lower) than the null expectation. If CWV was not different from the null expectation, the interpretation would be that traits other than seed mass were selected for (Fig. 2B). However, the finding that CWV was higher than the null expectation suggests some kind of niche partitioning, where species with more distant trait values are more likely to coexist (limiting similarity sensu Macarthur and Levins [1967]).
The maintenance of limiting similarity in seed mass is often explained by a trade-off between higher fecundity of small-seeded species and higher stress tolerance of large-seeded species (Muller-Landau 2010, D’Andrea and O’Dwyer 2021). This tradeoff enables the coexistence of species varying in seed mass when there is spatial heterogeneity in microhabitat quality and a positive correlation between seed mass and stress tolerance. However, in this system, spatial heterogeneity was minimized (Ron et al. 2018) and there was no evidence of higher stress tolerance of large-seeded species (DeMalach et al. 2019). We, therefore, attribute the coexistence of plants with different seed masses to a competition-fecundity trade-off that enables small-seeded species to grow in small microsites left unoccupied by the less fecund large-seeded species (Rees and Westoby 1997, Geritz et al. 1999, Coomes and Grubb 2003).

Simplifying assumption of the simulation model

Our model assumes that ecological fitness is affected by a single trait. However, our framework is not limited to this simplistic assumption. For multiple correlated traits that are difficult to disentangle, ordination techniques could be used as commonly done for sets of leaf traits in plant ecology (the leaf economic spectrum; Diaz et al. 2016). Furthermore, our framework can be applied also for multiple uncorrelated traits by using a multidimensional trait space where directionality is characterized by the distance between trait values and the
of species with traits close to CWM increases species diversity but its effect on CWV may not be positive. In some cases, there are some minor effects of selection type on CWV (Appendix S1: Figs. S4, S6, S9). We speculate that these differences are driven by different geometric constraints of stabilizing selection.

Implications

Our framework and model present new predictions regarding the relationship between species diversity and functional diversity across environmental gradients (Mayfield et al. 2010, Cadotte et al. 2011, Rapacciuolo et al. 2019). When using CWV as the indicator for functional diversity, our model predicts that the two aspects of diversity will be positively correlated if diversity patterns along an environmental gradient are mainly determined by changes in selection strength (i.e., species diversity and CWV vary in the same direction with increasing selection strength). However, we expect a weaker correlation if diversity patterns along an environmental gradient are determined mostly by changes in selection type (e.g., a transition from stabilizing selection to directional selection affects only species diversity but not CWV; Fig. 3).

It may seem that there is a clear dichotomy between trait-based and species-based approaches (Shipley et al. 2016). However, explanations for ecological patterns often involve selection that affects both species diversity and trait distribution (e.g., Grime 1979, Tilman 1982). Previous studies have shown the benefits of studying species and traits together (Fukami et al. 2005, Pavoine and Bonsall 2011). Our study corroborates that simultaneous investigation of species diversity and trait distribution helps to understand community assembly even in a system where both species diversity (Ron et al. 2018) and trait distribution (DeMalach et al. 2019) are thoroughly investigated. We hope that our proposed approach will assist in shedding light on the underlying selection processes in many other communities.

Acknowledgments

We thank Ronen Ron and Ronen Kadmon for contributing data from their mesocosm experiment. Karen Abbott, Matthieu Barbier, Gili Greenbaum, Susan Harrison, an anonymous reviewer, and the members of the community ecology group at Stanford University provided constructive comments on earlier drafts. This work was supported by the Rothschild fellowship (N. DeMalach), the Terman Fellowship of Stanford University (T. Fukami), and the studying abroad scholarship of the Ministry of Education of Taiwan (P.-J. Ke). Author contributions: N. DeMalach, P.-J. Ke, and T. Fukami conceived and designed the framework and analyses; N. DeMalach performed the simulation and empirical analyses; N. DeMalach wrote the first draft; all authors substantially contributed to the writing of the manuscript.

Literature Cited

Botta-Dukát, Z. 2005. Rao’s quadratic entropy as a measure of functional diversity based on multiple traits. Journal of Vegetation Science 16:533–540.
