Supplementary Material to "Character displacement in the presence of multiple trait differences: Evolution of the storage effect in germination and growth."

Nicholas Kortessis[1,2,*] and Peter Chesson[1,3]
[1] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA.
[2] Department of Biology, University of Florida, Gainesville, Florida 32611, USA.
[3] Department of Life Sciences and Center for Global Change Biology, National Chung Hsing University, Taichung 402, Taiwan.
*Corresponding author. Email: n.kortessis@ufl.edu

Table of Contents

S1. The Johnson S_B distribution

The Johnson S_B distribution (Johnson and Kotz 1994, pg. 34) is a distribution, $X \in (\xi, \gamma + \xi)$ which gives a standard normal distribution, Z, under the following transformation,

$$Z = \frac{1}{\sigma}\left[\ln\left\{\frac{X-\xi}{\gamma+\xi-X}\right\} - \mu\right]. \tag{S1}$$

The parameters have the following restrictions: ξ and μ are both real numbers, $\gamma > 0$, and, by convention, $\sigma > 0$. The parameters μ and σ are the mean and standard deviation of a normal distribution, $\ln\{(X-\xi)/(\gamma + \xi - X)\} = \mu + \sigma Z$. The parameters ξ and γ are the minimum and range of the domain of X, respectively. Our use of the S_B is to describe temporal fluctuations in the yearly germination fraction, which is bounded between the values 0 and 1. Hence, $\xi = 0$ and $\gamma = 1$ for our purposes, and we define $G = (X - \xi)/\gamma$. Rewriting (S1) in terms of E_G and G yields

$$E_G \equiv \mu + \sigma Z = \ln\left(\frac{G}{1-G}\right), \tag{S2}$$

which states that there is a class of distributions G that is normally distributed on the log-odds scale. Rearranging (S2) yields the formulation of the Johnson S_B distribution used in the main text (4) for G,

$$G = \frac{e^{E_G}}{1 + e^{E_G}}, \tag{S3}$$

where $G \in (0,1)$. The behavior of G is thus determined entirely by the normal distribution E_G, which has two parameters μ and σ.

The probability density function for G can be found via transformation of a normally distributed random variable E_G. First, note that the probability density function of a transformed variable $G = h(X)$ of a random variable X is

$$f_G(g) = \frac{dh^{-1}}{dg} f_X\left(h^{-1}(g)\right), \tag{S4}$$

where f_X is the density function for random variable X and h is a monotone function. Equation (S4) is common in texts on probability, but we briefly derive it here. First, recognize that a continuously differentiable distribution function for a random variable G, F_G, is by definition

$$F_G(g) = P(G \leq g) = \int_{-\infty}^{g} f_G(z)dz. \tag{S5}$$

By the fundamental theorem of calculus,

$$f_G(g) = \frac{d}{dg} F_G. \tag{S6}$$

From (S5), it follows that

$$\frac{d}{dG}F_G = \frac{d}{dG}P(G \leq g) \qquad \text{by (S5)}$$

$$= \frac{d}{dh(X)}P\left(h(X) \leq g\right) \qquad \text{by def. of } G$$

$$= \frac{d}{dh(X)}P\left(X \leq h^{-1}(g)\right)$$

$$= \frac{dX}{dh} \cdot \frac{d}{dX}P\left(X \leq h^{-1}(g)\right) \qquad \text{by the chain rule}$$

$$= \frac{dX}{dh} \cdot f_X\left(h^{-1}(g)\right)$$

$$= \frac{dh^{-1}}{dG} f_X\left(h^{-1}(g)\right). \tag{S7}$$

The penultimate line follows again from the fundamental theorem of calculus and the final line follows from the definition of $G = h(X)$. Using (S7) in (S6) yields (S4).

For the germination distribution G, h is the inverse logit function of E_G, where $G = h(E_G) = \exp(E_G)/(1+\exp(E_G))$. Thus, $E_G = h^{-1}(G) = \ln\{G/(1-G)\}$ and

$$\frac{dh^{-1}}{dg} = \frac{d}{dg}\ln\left\{\frac{g}{1-g}\right\} \tag{S8}$$
$$= \left[g(1-g)\right]^{-1}.$$

In addition,

$$f_{E_G}\left(h^{-1}(g)\right) = \frac{1}{\sqrt{2\pi\sigma_{E_G}^2}}\exp\left\{-\frac{\left(\ln\left\{\frac{g}{1-g}\right\}-\mu_{E_G}\right)^2}{2\sigma_{E_G}^2}\right\} \tag{S9}$$

because $E_G \sim N(\mu_{E_G}, \sigma_{E_G}^2)$.

Using expressions (S8) and (S9) in (S4) yields the pdf for G:

$$f_G(g) = \frac{1}{g(1-g)\sqrt{2\pi\sigma_{E_G}^2}}\cdot\exp\left\{-\frac{\left(\ln\left\{\frac{g}{1-g}\right\}-\mu_{E_G}\right)^2}{2\sigma_{E_G}^2}\right\} \tag{S10}$$

for $g \in (0,1)$. Equation (S10) is the probability density function for the Johnson S_B distribution describing germination when the underlying environmental variable has mean μ_{E_G} and variance $\sigma_{E_G}^2$. Equation (S10) was used to create the density curves in Figure 1 of the main text.

S2. Writing correlations between environmental responses as differences in θs.

There are two correlations used in the main text. The first is the between-species correlation in environmental responses, $\mathrm{Corr}(E_{G_1}, E_{G_2}) = \cos(\theta_{G_1} - \theta_{G_2})$. To derive this correlation, first note that $\mathrm{Corr}(E_{G_1}, E_{G_2}) = \mathrm{Cov}(E_{G_1}, E_{G_2})/\sqrt{\mathrm{Var}(E_{G_1})\mathrm{Var}(E_{G_2})}$. Using expressions for E_G from Table 1 in the main text, the covariance is

$$\mathrm{Cov}\left(E_{G_1}, E_{G_2}\right) = \mathrm{Cov}\left(\mu_{E_G} + \sigma_{E_G}\left(e_{11}X_1 + e_{12}X_2\right), \mu_{E_G} + \sigma_{E_G}\left(e_{21}X_1 + e_{22}X_2\right)\right)$$
$$= \sigma_{E_G}^2\,\mathrm{Cov}\left(e_{11}X_1 + e_{12}X_2, e_{21}X_1 + e_{22}X_2\right)$$
$$= \sigma_{E_G}^2\left[e_{11}e_{21}\,\mathrm{Var}\left(X_1\right) + 2\left(e_{11}e_{21} + e_{12}e_{22}\right)\mathrm{Cov}\left(X_1, X_2\right) + e_{12}e_{22}\,\mathrm{Var}\left(X_2\right)\right]. \tag{S11}$$

We assume X_1 is independent of X_2 and so $\mathrm{Cov}(X_1, X_2) = 0$. Furthermore, X_1 and X_2 are both standard normal, meaning $\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = 1$. Thus (S11) simplifies to

$$\mathrm{Cov}\left(E_{G_1}, E_{G_2}\right) = \sigma_{E_G}^2\left(e_{11}e_{21} + e_{12}e_{22}\right). \tag{S12}$$

Substituting in the angular definitions for the environmental sensitivities, $e_{j1} = \sin(\theta_{G_j})$ and $e_{j2} = \cos(\theta_{G_j})$, into the term in parentheses in (S12) yields

$$\mathrm{Cov}\left(E_{G_1}, E_{G_2}\right) = \sigma_{E_G}^2\left(e_{11}e_{21} + e_{12}e_{22}\right)$$
$$= \sigma_{E_G}^2\left(\sin(\theta_{G_1})\sin(\theta_{G_2}) + \cos(\theta_{G_1})\cos(\theta_{G_2})\right). \tag{S13}$$

Now, we use the trigonometric identity

$$\sin(x)\sin(y)+\cos(x)\cos(y)=\cos(x-y),\tag{S14}$$

in (S13) to yield

$$\text{Cov}\left(E_{G_1},E_{G_2}\right)=\sigma_{E_G}^2\cos\left(\theta_{G_1}-\theta_{G_2}\right).\tag{S15}$$

Now, since $\text{Var}(E_{G_1})=\text{Var}(E_{G_2})=\sigma_{E_G}^2$, placing (S15) in the definition of a correlation yields our final result:

$$\begin{aligned}\text{Cov}\left(E_{G_i},E_{G_j}\right)&=\frac{\text{Cov}\left(E_{G_1},E_{G_2}\right)}{\sqrt{\text{Var}(E_{G_1})\text{Var}(E_{G_2})}}\\&=\frac{\sigma_{E_G}^2\cos\left(\theta_{G_1}-\theta_{G_2}\right)}{\sqrt{\left(\sigma_{E_G}^2\right)^2}}\\&=\cos\left(\theta_{G_1}-\theta_{G_2}\right)\end{aligned}\qquad.\tag{S16}$$

The same derivation applies to correlations between vigor responses of the two species.

The other correlation of use in the main text is the within-species correlation of germination and vigor responses, $\text{Corr}(E_{G_j},E_{V_j})=\rho\cos(\theta_{G_j}-\theta_{V_j})$, our measure of predictive germination. The definitions of E_G and E_V imply that

$$\begin{aligned}\text{Cov}\left(E_{G_j},E_{V_j}\right)&=\text{Cov}\left(\mu_{E_G}+\sigma_{E_G}\left(e_{j1}X_1+e_{j2}X_2\right),\mu_{E_V}+\sigma_{E_V}\left(v_{j1}Z_1+v_{j2}Z_2\right)\right)\\&=\sigma_{E_G}\sigma_{E_V}\text{Cov}\left(e_{j1}X_1+e_{j2}X_2,v_{j1}Z_1+v_{j2}Z_2\right)\\&=\sigma_{E_G}\sigma_{E_V}\sum_{k=1}^{2}\sum_{l=1}^{2}e_{jk}v_{jl}\text{Cov}\left(X_k,Z_l\right)\end{aligned}\qquad.\tag{S17}$$

Following the assumptions in the main text, $\text{Cov}(X_k, Z_k) = \rho$ for k = 1,2, and $\text{Cov}(X_k, Y_l) = 0$ for $k \neq l$. Hence, (S17) simplifies to

$$\begin{aligned}\text{Cov}\left(E_{G_j},E_{V_j}\right)&=\sigma_{E_G}\sigma_{E_V}\rho\left(e_{j1}v_{j1}+e_{j2}v_{j2}\right)\\&=\sigma_{E_G}\sigma_{E_V}\rho\left(\sin(\theta_{G_j})\sin(\theta_{V_j})+\cos(\theta_{G_j})\cos(\theta_{V_j})\right),\end{aligned}\tag{S18}$$

where we have substituted in the angular representations of environmental sensitivities in the last line. Using the same trigonometric identity (S14), (S18) simplifies to

$$\text{Cov}\left(E_{G_j},E_{V_j}\right)=\sigma_{E_G}\sigma_{E_V}\rho\cos\left(\theta_{G_j}-\theta_{V_j}\right).\tag{S19}$$

The definition of a correlation is $\text{Corr}(E_{G_j},E_{V_j})=\text{Cov}(E_{G_j},E_{V_j})/\sqrt{\text{Var}(E_{G_j})\text{Var}(E_{V_j})}$ and since $\text{Var}(E_{G_j})=\sigma_{E_G}^2$ and $\text{Var}(E_{V_j})=\sigma_{E_V}^2$, the correlation can be written using (S19) as

$$\begin{aligned}\text{Corr}\left(E_{G_j},E_{V_j}\right)&=\frac{\text{Cov}\left(E_{G_j},E_{V_j}\right)}{\sqrt{\text{Var}(E_{G_j})\text{Var}(E_{V_j})}}\\&=\frac{\sigma_{E_G}\sigma_{E_V}\rho\cos\left(\theta_{G_j}-\theta_{V_j}\right)}{\sigma_{E_G}\sigma_{E_V}}\\&=\rho\cos\left(\theta_{G_j}-\theta_{V_j}\right)\end{aligned}\qquad,\tag{S20}$$

which is Eqn (10) in the main text.

S3. Constructing the environmental responses

A description of the modeling problem

Our model includes variable germination and vigor for each species. To investigate potential correlations between these variables, we model them as normal on some scale. Normal distributions are chosen because it is straightforward to model correlations of normally distributed variables. The normal variables are $E_G = \log\{G/(1-G)\}$ and $E_V = \log(V)$, which we term environmental responses. The set of all environmental responses $\mathbf{E} = (E_{G1}, E_{G2}, E_{V1}, E_{V2})^T \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is a 4-d multivariate normal distribution. Multivariate normal distributions are completely defined by the vector of means, $\boldsymbol{\mu}$, and the covariance matrix, $\boldsymbol{\Sigma}$. From Eqns (6)-(7) in the main text, the vector of means is $\boldsymbol{\mu} = (\mu_{EG}, \mu_{EG}, \mu_{EV}, \mu_{EV})^T$ and the diagonal elements of $\boldsymbol{\Sigma}$ (the variances of each marginal distribution of $\mathbf{E}$).

With the means and variances of the multivariate normal specified, all that remains to be specified are the off-diagonal elements of $\boldsymbol{\Sigma}$, which are the covariances between each of the random variables in $\mathbf{E}$. Instead of the covariance matrix, we consider the correlation matrix, $\mathbf{R}$, which has all the same information of the covariance matrix when combined with the variances because $\boldsymbol{\Sigma} = \mathbf{DRD}$, where $\mathbf{D} = \mathrm{diag}(\sigma_{EG}, \sigma_{EG}, \sigma_{EV}, \sigma_{EV})$ is the diagonal matrix of standard deviations. The correlation matrix, $\mathbf{R}$, is symmetric because $\mathrm{Corr}(U,W) = \mathrm{Corr}(W,U)$ for any random variables U and W and has values of 1 along its diagonal. There are 6 unique elements of $\mathbf{R}$ for a 4-d multivariate normal (and $n(n-1)/2$ unique elements for an n-dimensional multivariate normal). These are

     (i) $\mathrm{Corr}(E_{G1},E_{G2})$,   (ii) $\mathrm{Corr}(E_{G1},E_{V1})$,
     (iii) $\mathrm{Corr}(E_{G1},E_{V2})$,   (iv) $\mathrm{Corr}(E_{G2},E_{V1})$,
     (v) $\mathrm{Corr}(E_{G2},E_{V2})$,   and   (vi) $\mathrm{Corr}(E_{V1},E_{V2})$.

The goal is to write a model where correlations (i)-(v) change as both species germination responses evolve ($\mathrm{Corr}(E_{V1},E_{V2})$ is fixed by assumption that vigor does not evolve).

Covariance matrices must be positive-definite

An important consideration when constructing covariance matrices (and correlation matrices) is that all elements of the matrices must be internally consistent with one another. Internal consistency is guaranteed if the covariance matrix is positive-definite. A matrix $\mathbf{R}$ is said to be positive-definite if it satisfies the inequality, $\mathbf{a}^T\mathbf{Ra} > 0$, for any real-valued column vector $\mathbf{a}$. Importantly, all the elements of $\mathbf{R}$ cannot be independently varied; some correlations between variables are constrained given specified correlations between other variables.

For an example of a matrix entries that are not internally consistent (and so it not positive-definite), consider the slightly simpler problem of 3 normal variables W_1, W_2, and W_3 with nonzero, finite variances. Assume W_1 and W_2 have correlation ρ_1 and W_2 and W_3 also have correlation ρ_1. The remaining entry of the correlation matrix is the correlation between W_1 and W_3, i.e., $\mathrm{Corr}(W_1,W_3) = x$, for $-1 \leq x \leq 1$. The correlation matrix $\mathbf{R}_W$ in this case is

$$\mathbf{R}_W = \begin{pmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & x \\ \rho_1 & x & 1 \end{pmatrix}. \tag{S21}$$

Might we choose any value of x for this matrix? No, because some values of x are not consistent with the prior assumptions that W_2 is correlated with W_1 and W_3. To find what values of x allow $\mathbf{R}_W$ to be positive semi-definite, we use Sylvester's criterion, which states that a real, symmetric matrix is positive semi-definite if all leading principal minors are nonnegative (Gilbert 1991). In other words, the determinant of all upper left sub square matrices of $\mathbf{R}_W$ are nonnegative and that the full matrix $\mathbf{R}_W$ has a nonnegative determinant.

Applying Sylvester's Criterion to $\mathbf{R}_W$ reveals that $\mathbf{R}_W$ is positive-definite if $x^2 - 2x\rho_1 \leq 1 - 2\rho_1^2$. Clearly, the value of ρ_1 constrains the feasible values of x. For example, when $\rho_1 = 1$, the condition on x is $(x - 1)^2 \leq 0$, which can only be satisfied for $x = 1$. Similarly, when $\rho_1 = -1$, the condition on x is $(x + 1)^2 \leq 0$, which again is only satisfied for $x = -1$. In fact, the value of x can only be independently chosen when $\rho_1 = 0$, in which case $\mathbf{R}_W$ is positive definite if $x^2 \leq 1$, which is satisfied for all feasible values of x, i.e., for any feasible correlation.

This example shows that arbitrarily choosing some correlation values of $\mathbf{R}$ constrains other possible correlations in $\mathbf{R}$ and may fix the values of some correlations. One could keep track of these constraints, but it would be analytically quite involved for the 4 environmental responses we have in our model.

An approach to create any positive-definite matrix
Instead of specifying the correlations and ensuring that the covariance matrix is positive-definite, an alternative approach is to construct a self-consistent multivariate normal distribution from a number of independent standard normal distributions. Johnson and Kotz (1972b, Chapter 35) shows that any multivariate normal distribution, $\mathbf{U} \sim \mathcal{N}(\boldsymbol{\mu}_U, \boldsymbol{\Sigma}_U)$, where $\mathbf{U} = (U_1, U_2, U_3, \ldots, U_n)^T$, can be constructed from n independent standard normal variables, $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$, according to the following linear transformation

$$\mathbf{U} = \boldsymbol{\mu}_U + \mathbf{HX}, \tag{S22}$$

where $\boldsymbol{\mu}_U$ is a column vector of means, $\mathbf{X} = (X_1, X_2, X_3, \text{and } X_4)^T$ is a column vector of i.i.d. standard normal random variables, and $\mathbf{H}$ is an n x n matrix for which $\mathbf{HH}^T = \boldsymbol{\Sigma}_U$. Eqn (1.2) states that any n-dimensional multivariate random variable, with any arbitrary covariance matrix, can be constructed from a linear combination of n independent standard normal variables. In fact, n independent standard normal variables are necessary and sufficient to generate $\mathbf{U}$. Use fewer than four and one marginal distribution of $\mathbf{U}$ is statistically identical to another. Use more than four and the model is overdetermined.

To describe environmental responses in the model, we have taken the approach of constructing the multivariate distribution of environmental responses by way of a linear combination of independent random variables. Since we are interested in the correlation structure, we need to specify the components of $\mathbf{H}$ that do not depend on the variances. We do this by writing $\mathbf{H} = \mathbf{DK}$. The matrix $\mathbf{K}$ contains all the information present in the correlation matrix, which can be seen by substituting the identity $\boldsymbol{\Sigma} = \mathbf{HH}^T$ into the definition for the correlation matrix, $\mathbf{R} = \mathbf{D}^{-1}\boldsymbol{\Sigma}\mathbf{D}^{-1}$, which yields

$$\begin{aligned}\mathbf{R} &= \mathbf{D}^{-1}\mathbf{HH}^T\mathbf{D}^{-1} \\ &= \mathbf{D}^{-1}(\mathbf{DK})(\mathbf{DK})^T\mathbf{D}^{-1} \\ &= (\mathbf{D}^{-1}\mathbf{D})(\mathbf{KK}^T)(\mathbf{DD}^{-1}) \\ &= \mathbf{KK}^T\end{aligned}. \tag{S23}$$

Hence, $\mathbf{E}$ can be written as

$$\mathbf{E} = \boldsymbol{\mu} + \mathbf{DKX}, \tag{S24}$$

where $\boldsymbol{\mu}$ contains the means of the marginal distributions, $\mathbf{D}$ contains the standard deviations of the marginal distributions, $\mathbf{K}$ contains constants determining the correlation structure, and $\mathbf{X}$ is a column vector of 4 i.i.d. standard normal random variables. For our model,

$$\mathbf{K} = \begin{pmatrix} e_{11} & e_{12} & 0 & 0 \\ e_{21} & e_{22} & 0 & 0 \\ v_{11}\rho & v_{12}\rho & v_{21}\sqrt{1-\rho^2} & v_{12}\sqrt{1-\rho^2} \\ v_{21}\rho & v_{22}\rho & v_{21}\sqrt{1-\rho^2} & v_{22}\sqrt{1-\rho^2} \end{pmatrix}, \tag{S25}$$

which can be written much more compactly in the block matrix form

$$\mathbf{K} = \begin{pmatrix} \mathbf{e} & 0 \\ \rho\mathbf{v} & \sqrt{1-\rho^2}\,\mathbf{v} \end{pmatrix}, \tag{S26}$$

with $\mathbf{e}$ being the 2x2 matrix with elements e_{ij} and $\mathbf{v}$ being the 2x2 matrix with elements v_{ij}. The correlation matrix is then

$$\mathbf{R} = \mathbf{K}\mathbf{K}^T = \begin{pmatrix} \mathbf{e}\mathbf{e}^T & \rho\mathbf{e}\mathbf{v}^T \\ \rho\mathbf{v}\mathbf{e}^T & \mathbf{v}\mathbf{v}^T \end{pmatrix}. \tag{S27}$$

An important note here is one extra step is required in order to derive the expressions for environmental responses in the main text. The step is to write $Z_1 = \rho X_1 + \sqrt{1-\rho^2}\,X_3$ and $Z_2 = \rho X_2 + \sqrt{1-\rho^2}\,X_4$. Using these two definitions for Z_1 and Z_2 in Eqn (7) of the main text yields the system of equations (1.3) above.

As $\mathbf{R}$ is a correlation matrix, the diagonal elements are 1, which implies that $e_{i1}^2 + e_{i2}^2 = 1$ and $v_{i1}^2 + v_{i2}^2 = 1$ for $i = 1,2$. Since we have two parameters to describe a single value, one element in each row of $\mathbf{e}$ and $\mathbf{v}$ is redundant. In effect, each matrix can be described by 2 values, one for each row. The choice of the polar coordinate representation works well in this case. Now, we have 4 parameters to describe the elements of $\mathbf{e}$ and $\mathbf{v}$ (θ_{G1}, θ_{G2}, θ_{V1}, and θ_{V2}) and 5 parameters overall (θ_{G1}, θ_{G2}, θ_{V1}, θ_{V2}, and ρ) to describe 6 correlations.

Tactically speaking, the approach here ensures that the model is internally consistent for all parameter values and lets the covariances emerge as the parameters change. Furthermore, the parameters here are meaningful in terms of different components of the multivariate distribution. The parameters μ and σ control the means and variances, respectively, of the marginals of $\mathbf{E}$ (and the median and spread of G and V). The θ parameters control the correlations between species and within species, once coupled with ρ.

What if there are more (or fewer) than 2 environments affecting germination and vigor?
We have described the random variables X and Z as representing unspecified environmental factors that affect germination and plant size, respectively. Describing X and Z as environmental factors gives some biological meaning to the idea of sensitivities evolving. However, one may wonder whether 2 is the appropriate number of environmental factors. Biologically speaking, it seems unlikely that germination responds to more than just two environmental factors. On the other hand, one may ask why not do the same but with one environmental factor, which is simpler, if the same results could be derived.

Johnson and Kotz (1972) provides us with the answer to both in the demonstration that any n-dimensional multivariate normal can be constructed from n i.i.d. standard normal random variables. Adding more variables would lead to redundancy in the description of the system. Having fewer would limit the types of covariances that can be described by the model. Furthermore, we might want to assume that the environmental variables are correlated in some other way, perhaps by assuming that X_1, X_2, X_3, and X_4 are correlated in some way. In doing so, we could always rewrite such correlated variables as a linear function transformation of 4 i.i.d. random variables. As such, any correlated set of variables can be described by the model we have, which is sufficient to describe all possible covariances between 4 multivariate normal variables.

Finally, we might consider only a single environmental variable to be important. In this case, our model again is sufficient to describe this situation. Imagine that the germination response is a reaction norm to temperature at the time of rainfall, $T \sim \mathcal{N}(\mu_T, \sigma^2_T)$, such that we can describe the germination response with a linear regression against temperature. In that case, $\hat{E}_j(t) = \beta_j^0 + \beta_j T(t)$ where β_j^0 and β_j are

regression coefficients. For two species, the environmental responses follow a bivariate normal distribution $(\hat{E}_1, \hat{E}_2)$ with means $E[\hat{E}_j] = \beta_j^0 + \beta_j \mu_T$ and covariance matrix,

$$\Sigma_{\hat{E}} = \sigma_T^2 \begin{pmatrix} \beta_1^2 & \beta_1\beta_2 \\ \beta_1\beta_2 & \beta_2^2 \end{pmatrix}.$$

As before, we can always rewrite this distribution as a linear combination of two independent standard normal random variables, according to (S22). Note however in this case that the correlations between species, $\mathrm{Corr}(\hat{E}_1, \hat{E}_2)$ is limited to take values ± 1 depending on whether β_1 and β_2 have the same sign. This temperature model is thus more limited than the one we use.

How do the multiple transformations of X affect the resulting distributions of G and V?
A reasonable question is whether information is lost in the multiple transformations from X to G and V. The process for generating G and V is as follows

1. Start with 4 independent standard normal random variables, $\mathbf{X} = (X_1, X_2, X_3, X_4)$.
2. Construct $\mathbf{E}$ with a linear transformation (S24).
3. Transform the first two marginals of $\mathbf{E}$ (E_{G1}, E_{G2}) by applying the inverse log-odds transformation (eqn 4 of the main text, sometimes called the logistic transform) to create G_1 and G_2.
4. Transform the second two marginals of $\mathbf{E}$ (E_{V1}, E_{V2}) by taking the exponential of the marginals (eqn 5 of the main text) to create V_1 and V_2.

The result is that we have two germination distributions, G_1 and G_2, that are each S_B and two vigor distributions, V_1 and V_2, that are each log normal. The first two central moments of the marginals V_j are $E[V_j] = \exp(\mu_{EV} + \sigma^2_{EV}/2)$ and $\mathrm{Var}(V_j) = [\exp(\sigma^2_{EV}) - 1]\exp(2\mu_{EV} + \sigma^2_{EV})$ and are completely determined by means and variances of E_V.

Johnson (1994) gives some important relationships of the S_B distribution that we use to model G, although they give no closed form expressions for statistical properties of G. We can, however, have good approximations to the first two central moments of G. Approximating G to third order about $G^* = \exp(\mu_{EG})/[1+\exp(\mu_{EG})]$, its median, and taking expectations yields

$$E[G] = G^* \left[1 + (1-G^*)\left(\frac{1}{2} - G^*\right)\sigma^2_{E_G} \right] + o(\sigma^3),$$

where $y = o(\sigma)$ is used in the sense that $y/\sigma \to 0$ as $\sigma \to 0$. We can also approximate G to second order about its median and take the variance to find that

$$\mathrm{Var}(G) = G^{*2}(1-G^*)^2 \sigma^2_{E_G} + o(\sigma^2).$$

These properties would be no different from constructing any log-normal or S_B distributions. And you can see that the first two moments are described by parameters which are unrelated to any evolving characters in the model. However, we need a way to correlate the different environmental responses to have correlations between all combinations of G_1, G_2, V_1, and V_2. That is where the transformation from $\mathbf{X}$ to $\mathbf{E}$ is helpful because there is no obvious way to correlate Gs and Vs without such an underlying structure.

S4. Derivation of the selection gradient under the assumption of small environmental variance.

To find an expression for the growth rate, we follow the general approach of Chesson (1994) and write the growth rate, $g(E_G, E_V, C) \equiv r$, as a function of germination on the log-odds scale, vigor on the log-scale, and competition on the log scale. Thus, the growth rate of a species j can be written as

$$g_j\left(E_{G_j}, E_{V_j}, C\right) = \ln\left\{ s + y_j e^{E_{G_j} + E_{V_j} - C} \right\} - \ln\left\{ 1 + e^{E_{G_j}} \right\}, \tag{S28}$$

where E_{G_j} and E_{V_j} are given by equations (6) and (8), respectively, in the main text, and $C = \ln\{1 + \alpha G_1 V_1 N_1 + \alpha G_2 V_2 N_2\}$. The approach follows a Taylor expansion in three variables around an equilibrium point, i.e., a point $\left(E_{G_j}^*, E_{V_j}^*, C^*\right)$ for which $g_j\left(E_{G_j}^*, E_{V_j}^*, C^*\right) = 0$. The expansion is limited to terms $O(\sigma^2)$ where σ is a parameter representing the variation in E_G and E_V and is assumed small (detailed below).

Following Holt and Chesson (2014), we choose $C^* = \frac{1}{2}\left(\bar{C}_i^* + \bar{C}_j^*\right)$, where $\bar{C}_i^*$ is the solution to $g_j\left(\bar{E}_G, \bar{E}_V, \bar{C}_j^*\right) = 0$, which in this model formulation is

$$\bar{C}_j^* = \ln\left\{\frac{y_j \bar{G}^* \bar{V}^*}{1 - s\left(1 - \bar{G}^*\right)}\right\}. \tag{S29}$$

In (S29), $\bar{G}^* = \exp(\bar{E}_G)/\left(1 + \exp(\bar{E}_G)\right)$ and $\bar{V}^* = \exp(\bar{E}_V)$ are the values of G and V evaluated for the mean environmental values, $\bar{E}_{G_j}$ and $\bar{E}_{V_j}$, respectively. These are the same for both species as the means of the environmental responses are assumed the same for both species. We choose $E_{V_j}^* = E_V^* = \mu_{E_V}$, the mean of the vigor responses. The choices of E_V^* and C^* fix the value of $E_{G_j}^*$ to be the solution to $g_j\left(E_{G_j}^*, E_V^*, C^*\right)$, which from (S28) is

$$E_{G_j}^* = \ln\left\{\frac{1 - s}{y_j e^{E_V - C^*} - 1}\right\}. \tag{S30}$$

Using the fixed point $(E_{G_j}^*, E_V^*, C^*)$, the Taylor expansion of g to second order terms is

$$r_j = g^{(1,0,0)*}\Delta E_{G_j} + g^{(0,1,0)*}\Delta E_{V_j} + g^{(0,0,1)*}\Delta C +$$

$$+ \frac{1}{2}g^{(2,0,0)*}\Delta E_{G_j}^2 + \frac{1}{2}g^{(0,2,0)*}\Delta E_{V_j}^2 + \frac{1}{2}g^{(0,0,2)*}\Delta C^2$$

$$+ g^{(1,1,0)*}\Delta E_{G_j}\Delta E_{V_j} + g^{(1,0,1)*}\Delta E_{G_j}\Delta C + g^{(0,1,1)*}\Delta E_{V_j}\Delta C, \tag{S31}$$

where $\Delta X = X - X^*$ and $g^{(i,j,k)*}$ is the $i + j + k^{\text{th}}$ partial derivative of g with respect to the associated variables (E_G, E_V, C) evaluated at the equilibrium point $\left(E_{G_j}^*, E_V^*, C^*\right)$. Expressions for the derivatives are given in Table S1, which are all written in terms of a species' sensitivity to competition, $\beta_j = 1 - s(1 - G_j^*)$, where $G_j^* = \exp(E_{G_j}^*)/\left(1 + \exp(E_{G_j}^*)\right)$.

Table S1. Expressions for the coefficients of the Taylor approximation (D4). Note that $\beta_j = 1 - s(1 - \exp(E_j^*)/(1 + \exp(E_j^*)))$.

First Order Derivatives	Expression	Second Order Derivatives	Expression	Cross Partial Derivatives	Expression
$g^{(1,0,0)*}$	$\beta_j - G_j^*$	$g^{(2,0,0)*}$	$(1-s)(1-\beta_j)$	$g^{(1,1,0)*}$	$\beta_j(1-\beta_j)$
$g^{(0,1,0)*}$	β_j	$g^{(0,2,0)*}$	$\beta_j(1-\beta_j)$	$g^{(1,0,1)*}$	$-\beta_j(1-\beta_j)$
$g^{(0,0,1)*}$	$-\beta_j$	$g^{(0,0,2)*}$	$\beta_j(1-\beta_j)$	$g^{(0,1,1)*}$	$-\beta_j(1-\beta_j)$

To control the approximation, we introduce the critical assumption of small environmental fluctuations by writing

$$\Delta E_{G_j} \equiv E_{G_j} - E_{G_j}^* = O(\sigma)$$

(S32)

$$\Delta E_{G_j} \equiv E_{V_j} - E_V^* = O(\sigma),$$

for all j where $y = O(\sigma)$ is defined to mean $|y| < K\sigma$ for some constant K. The assumptions in (S32) mean that fluctuations in the environmental parameters about their equilibrium values are bounded within a range of relatively small values. While this may appear restrictive, insights from the small variance approximation prove useful in understanding results from much larger fluctuations. Further, we assume that $E(\Delta E_{G_j}) = O(\sigma^2)$ and $E(\Delta E_{V_j}) = O(\sigma^2)$. It follows these assumptions and (S32) that

$E(\Delta E_{G_j}^2) = \mathrm{Var}(E_{G_j}) + o(\sigma^2)$ and $E(\Delta E_{V_j}^2) = \mathrm{Var}(E_{V_j}) + o(\sigma^2)$, where $y = o(\sigma)$ means $y/\sigma \to 0$ as $\sigma \to 0$.

Chesson (1994) demonstrates that fluctuations in C are likewise bounded for this model, such that $C - C^* = O(\sigma)$ and $E(C - C^*) = O(\sigma^2)$, which also means that $E[(C - C^*)^2] = \mathrm{Var}(C) + o(\sigma^2)$.

Taking expectation of expression (S31), substituting in the expressions in Table S1, and following the small variance assumptions in the previous paragraph, expression (S31) can now be written as

$$\bar{r}_j \approx \left(\beta_j - G_j^*\right)E(\Delta E_{G_j}) - \beta_j E(\Delta C)$$

$$+\frac{1}{2}(1-s)(1-\beta_j)\mathrm{Var}(E_G) + \frac{1}{2}\beta_j(1-\beta_j)\left[\mathrm{Var}(E_V) + \mathrm{Var}(C)\right]$$

$$+\beta_j(1-\beta_j)\left[\mathrm{Cov}\left(E_{G_j}, E_{V_j}\right) - \mathrm{Cov}\left(E_{G_j}, C\right) - \mathrm{Cov}\left(E_{V_j}, C\right)\right],$$

(S33)

where $\approx$ means the approximation omits all terms $o(\sigma^2)$. Note that the term $E(\Delta E_V)$ has been omitted because $E(\Delta E_V) = E(E_V) - E_V^* = E(E_V) - \mu_{EV} = 0$ and that the species subscripts for $\mathrm{Var}(E_G)$ and $\mathrm{Var}(E_V)$ have been dropped in accordance with the assumptions in the main text that species have equal variance.

Selection

The approximation for fitness (S33), once applied to an invader, gives the invasion rate of a mutant phenotype θ'_{G_j}. Note however, the value C, which takes on different meanings in different contexts. For a mutant phenotype, it is assumed to be rare enough so as not to contribute to C. Thus, C only includes competition from the resident phenotype of the species when in allopatry and of both species when in sympatry. Given that the mutant phenotype does not affect C, mutant long-term growth rates can be approximated as

$$\bar{r}'_j \approx \left(\beta_j - G_j^*\right)E(\Delta E'_{G_j}) - \beta_j E(\Delta C)$$

$$+\frac{1}{2}(1-s)(1-\beta_j)\mathrm{Var}(E_G) + \frac{1}{2}\beta_j(1-\beta_j)\left[\mathrm{Var}(E_V) + \mathrm{Var}(C)\right]$$

$$+\beta_j(1-\beta_j)\left[\mathrm{Cov}\left(E'_{G_j}, E_{V_j}\right) - \mathrm{Cov}\left(E'_{G_j}, C\right) - \mathrm{Cov}\left(E_{V_j}, C\right)\right],$$

(S34)

where again $C = \ln\{1 + \alpha G_1 V_1 N_1 + \alpha G_2 V_2 N_2\}$ and E'_{G_j} is the environmental response of the invader.

Differentiating (S34) with respect to the invader trait θ'_{G_j} yields the selection differential. The trait θ'_{G_j} only affects patterns of variation in G and so has no effect on mean or variance terms. Thus, the derivative of each of the first 5 terms of (S34) w.r.t. θ'_{G_j} are zero. Further, since a mutant phenotype does not contribute to C, $d\mathrm{Cov}(E_{V_j}, C)/d\theta'_{G_j} = 0$. Thus, the derivative of (S34) is

$$\frac{d\bar{r}'_j}{d\theta'_{G_j}} \approx \beta_j(1-\beta_j)\left[\frac{d}{d\theta'_{G_j}}\mathrm{Cov}\left(E'_{G_j}, E_{V_j}\right) - \frac{d}{d\theta'_{G_j}}\mathrm{Cov}\left(E'_{G_j}, C\right)\right].$$

(S35)

Defining $\gamma_j \equiv \beta_j(1 - \beta_j)$ yields expression (12) in the main text.

S5. Results of evolution using the selection gradient under special cases

In section S4, we derived the result that the selection gradient can be approximated by the expression

$$\frac{d\bar{r}'_j}{d\theta'_{G_j}} \approx \gamma_j \left[\frac{d}{d\theta'_{G_j}} \text{Cov}\left(E'_{G_j}, E_{V_j}\right) - \frac{d}{d\theta'_{G_j}} \text{Cov}\left(E'_{G_j}, C\right) \right] \tag{S36}$$

where E'_{G_j} is the environmental response of the mutant phenotype of species j. Result (S36) can be used to predict evolutionary dynamics in some special cases. The cases here include (*i*) the single species case, (*ii*) the case of large variation of germination relative to vigor, i.e., $\sigma_{EV}/\sigma_{EG} \to 0$, (*iii*) the inverse case of large vigor variation relative to germination, i.e., $\sigma_{EG}/\sigma_{EV} \to 0$, and (*iv*) the case of no cue, i.e., $\rho = 0$.

The first term in (S36) can be expressed exactly (eqn 13 of the main text). However, the second term does not have an exact, closed form expression. In each of the special cases, we approximate the Cov(E,C) term in (S33) for "small" environmental fluctuations.

The single species case

In the single species case, we drop the species subscripts and note that C is

$$C(t) = \ln\{1 + \alpha G(t)V(t)N(t)\}. \tag{S37}$$

We cannot derive a closed form expression for Cov(E,C), but instead approximate it using the linear approximation for competition, $C \approx C^* + \partial C/\partial E_G(E_G - E_G^*) + \partial C/\partial E_V(E_V - E_V^*)$, which leads to the following expression for Cov(E,C):

$$\text{Cov}\left(E'_G, C\right) \approx \text{cov}\left(E'_G, p^*\left[E_G\left(1 - G^*\right) + E_V\right]\right)$$

$$= p^*\sigma_{E_G}^2 \left[\left(1 - G^*\right)\text{corr}\left(E'_G, E_G\right) + \frac{\sigma_{E_V}}{\sigma_{E_G}}\text{corr}\left(E'_G, E_V\right)\right] \tag{S38}$$

$$= p^*\sigma_{E_G}^2 \left[\left(1 - G^*\right)\cos\left(\theta'_G - \theta_G\right) + \frac{\sigma_{E_V}}{\sigma_{E_G}}\rho\cos\left(\theta'_G - \theta_V\right)\right],$$

where $p^* = E[\partial C/\partial E_V]$ evaluated at equilibrium E_G^*, E_V^*. The value p^* is bounded between 0 and 1, approaching 1 in the limit of large population biomass and 0 in the limit of zero biomass. Now, the derivative of CovEC with respect to the invader trait, θ'_G, evaluated at the resident value, θ_G, is

$$\left. \frac{d}{d\theta'_G}\text{Cov}\left(E'_G, C\right)\right|_{\theta'_G = \theta_G} \approx p^*\sigma_{E_G}^2 \left[\left(1 - G^*\right)\frac{d}{d\theta'_G}\cos\left(\theta'_G - \theta_G\right) + \frac{\sigma_{E_V}}{\sigma_{E_G}}\rho\frac{d}{d\theta'_G}\cos\left(\theta'_G - \theta_V\right)\right]_{\theta'_G = \theta_G}$$

$$= -p^*\sigma_{E_G}^2 \left[\left(1 - G^*\right)\sin\left(\theta'_G - \theta_G\right) + \frac{\sigma_{E_V}}{\sigma_{E_G}}\rho\sin\left(\theta'_G - \theta_V\right)\right]_{\theta'_G = \theta_G}, \tag{S39}$$

$$= -p^*\sigma_{E_G}\sigma_{E_V}\rho\sin\left(\theta_G - \theta_V\right)$$

where we have taken that $\sin(\theta'_G - \theta_G)|_{\theta_{G'} = \theta_G} = \sin(0) = 0$ and so it is absent in the final line of (S39). In the single-species case, intraspecific competition is not appreciably different for mutants and residents and so it has not affect on selection. This property holds for any model where competition is maximized for species with the same trait value and mutation are small in effect.

The derivative of Cov(E_G, E_V) is

$$\frac{d}{d\theta_G'}\text{cov}\left(E_G',E_V\right)\bigg|_{\theta_G'=\theta_G} \approx \sigma_{E_G}\sigma_{E_V}\rho\frac{d}{d\theta_G'}\cos\left(\theta_G'-\theta_V\right)\bigg|_{\theta_G'=\theta_G}$$

$$= -\sigma_{E_G}\sigma_{E_V}\rho\sin\left(\theta_G-\theta_V\right).$$

(S40)

Placing expressions (S39) and (S40) in (S36) yields an expression for the selection gradient,

$$\frac{d\overline{r}'}{d\theta_G'}\bigg|_{\theta_G'=\theta_G} \approx -\gamma\left(1-p^*\right)\sigma_{E_G}\sigma_{E_V}\rho\sin\left(\theta_G-\theta_V\right).$$

(S41)

Solutions to the selection gradient are candidate evolutionary equilibria, which may be attracting or repelling. The sine function has infinitely many solutions that are integer multiples of π. However, given the restriction on phenotype space that $-\pi/2 \leq \theta_G \leq \pi/2$ and $-\pi/2 \leq \theta_V \leq \pi/2$, there is only one solution to equation (S38), $\theta_G = \theta_V$ except in the limiting case that $\theta_V = \pm\pi/2$, in which case there are two solutions, $\theta_G = \pm\pi/2$. The solution $\theta_G = \theta_V$ corresponds to the germination phenotype that maximizes predictive germination of the species. In the limiting case where $\theta_V = \pm\pi/2$, the solution $\theta_G = \theta_V$ also maximizes predictive germination at $\text{Corr}(E_G, E_V) = \rho$, but the solution $\theta_G = -\theta_V$ corresponds to minimizing predictive germination at $\text{Corr}(E_G, E_V) = 0$.

Stability of these equilibria are given by the sign of the (S41) above and below each equilibrium point. The terms γ, $1 - p^*$, σ_{E_G}, σ_{E_V}, and ρ are positive. Hence, the sign of (S41) is given by the sign of $-\sin(\theta_G - \theta_V)$. Take first the case of $-\pi/2 < \theta_V < \pi/2$. For all values of $\theta_G < \theta_V$, (S41) is positive. For all values of $\theta_G > \theta_V$, (S41) is negative. Hence, the equilibrium $\theta_G = \theta_V$ is globally attractive. Now for special case of $\theta_V = \pm\pi/2$, the solution $\theta_G = \theta_V$ is a local attractor according to the argument above. However, this is not a global attractor because the solution $\theta_G \mp \pi/2$ is an equilibrium, albeit a repellor. To see how it is a repellor, consider the derivative of (S41), which is proportional to $-\cos(\theta_G - \theta_V)$. This derivative at the solutions $\theta_G \mp \pi/2$ for $\theta_V \pm \pi/2$ is $-\cos(\pm\pi) = 1$, indicating an unstable equilibrium.

Large germination variation relative to vigor variation with equal vigor
In the second case, we consider large variation in germination relative to vigor, i.e., $\sigma_{EV}/\sigma_{EG} \to 0$. Without loss of generality, we evaluate selection for species 1 and then note that it is straightforward to exchange subscripts to derive selection on species 2.

Here it is valuable to write C explicitly as functions of the rescaled environmental variables:

$$C = \ln\left\{1 + \alpha N_1 \frac{e^{E_{G_1}+E_{V_1}}}{1+e^{E_{G_1}}} + \alpha N_2 \frac{e^{E_{G_2}+E_{V2}}}{1+e^{E_{G_2}}}\right\}.$$

(S42)

Just as in the single species case, there is no exact expression for CovEC, and so we use the linear approximation for C,

$$C \approx C^* + p_1^*\left[\left(1-G_1^*\right)\left(E_{G_1}-E_{G_1}^*\right)+\left(E_{V_1}-E_{V_1}^*\right)\right] + p_2^*\left[\left(1-G_2^*\right)\left(E_{G_2}-E_{G_2}^*\right)+\left(E_{V_2}-E_{V_2}^*\right)\right],$$

(S43)

where $p_j^* = \text{E}[\partial C/\partial E_{Vj}]$ evaluated at E_{Gj}^* and E_V^* and $0 \leq p_j^* \leq 1$ for $j = 1,2$. Placing (S43) in for C in $\text{Cov}(E'_{G1}, C)$ yields

$$\text{Cov}\left(E_{G_1}',C\right) \approx \text{Cov}\left(E_{G_1}', p_1^*\left[\left(1-G_1^*\right)E_{G_1}+E_{V_1}\right]+p_2^*\left[\left(1-G_2^*\right)E_{G_2}+E_{V_2}\right]\right)$$

$$= \overline{p}_1^*\left[\left(1-G_1^*\right)\text{Cov}\left(E_{G_1}',E_{G_1}+E_{V_1}\right)\right] + \overline{p}_2^*\left[\left(1-G_2^*\right)\text{Cov}\left(E_{G_1}',E_{G_2}+E_{V_2}\right)\right].$$

(S44)

This covariance can be rewritten as

$$\text{Cov}\left(E_{G_1}',C\right) = \sigma_{E_G}^2\left\{\begin{array}{l}\left[\left(1-G_1^*\right)\overline{p}_1^*\text{Corr}\left(E_{G_1}',E_{G_1}\right)+\left(1-G_2^*\right)\overline{p}_2^*\text{Corr}\left(E_{G_1}',E_{G_2}\right)\right] \\ +\dfrac{\sigma_{E_V}}{\sigma_{E_G}}\left[\overline{p}_1^*\text{Corr}\left(E_{G_1}',E_{V_1}\right)+\overline{p}_2^*\text{Corr}\left(E_{G_1}',E_{V_2}\right)\right]\end{array}\right\}.$$

(S45)

Now using the last line of (S45) and the equation $\text{Cov}\left(E_{G_1}, E_{V_1}\right) = \sigma_{E_G}^2 \left[\dfrac{\sigma_{E_V}}{\sigma_{E_G}} \text{Corr}\left(E_{G_1}, E_{V_1}\right) \right]$ in the selection

gradient (S36) for species 1 yields:

$$\frac{d\overline{r_1'}}{d\theta_{G_1}'} \approx \gamma_1 \frac{d}{d\theta_{G_1}'} \left[\text{Cov}\left(E_{G_1}', E_{V_1}\right) - \text{Cov}\left(E_{G_1}', C\right) \right]$$

$$= \gamma_1 \sigma_{E_G}^2 \frac{d}{d\theta_{G_1}'} \left[\begin{array}{l} \dfrac{\sigma_{E_V}}{\sigma_{E_G}} \left\{ \text{Corr}\left(E_{G_1}', E_{V_1}\right) - \overline{p}_1^* \text{Corr}\left(E_{G_1}', E_{V_1}\right) - \overline{p}_2^* \text{Corr}\left(E_{G_1}', E_{V_2}\right) \right\} \\ - \left[\left(1 - G_1^*\right) \overline{p}_1^* \text{Corr}\left(E_{G_1}', E_{G_1}\right) + \left(1 - G_2^*\right) \overline{p}_2^* \text{Corr}\left(E_{G_1}', E_{G_2}\right) \right] \end{array} \right]. \qquad \text{(S46)}$$

Upon taking the limit as $\sigma_{E_V} / \sigma_{E_G} \to 0$, (S46) reduces to

$$\frac{d\overline{r_1'}}{d\theta_{G_1}'} \approx -\gamma_1 \sigma_{E_G}^2 \frac{d}{d\theta_{G_1}'} \left[\left(1 - G_1^*\right) \overline{p}_1^* \text{Corr}\left(E_{G_1}', E_{G_1}\right) + \left(1 - G_2^*\right) \overline{p}_2^* \text{Corr}\left(E_{G_1}', E_{G_2}\right) \right], \qquad \text{(S47)}$$

which includes only the effects of germination on seedling competition. When the model only includes the component of selection reflecting competition, trait evolution maximizes trait divergence.

To show that maximum trait divergence results, we rewrite (S47) explicitly in terms of the germination trait, θ_G, and the vigor trait, θ_V. Replacing the covariance terms with their trait representations yields

$$\frac{d\overline{r_1'}}{d\theta_{G_1}'} \approx -\gamma_1 \sigma_{E_G}^2 \frac{d}{d\theta_{G_1}'} \left[\left(1 - G_1^*\right) \overline{p}_1^* \cos\left(\theta_{G_1}' - \theta_{G_1}\right) + \left(1 - G_2^*\right) \overline{p}_2^* \cos\left(\theta_{G_1}' - \theta_{G_2}\right) \right]. \qquad \text{(S48)}$$

The derivative $d\cos(x - y)/dx = -\sin(x - y)$. Hence, (S48) can be rewritten as

$$\frac{d\overline{r_1'}}{d\theta_{G_1}'} \approx \gamma_1 \sigma_{E_G}^2 \frac{d}{d\theta_{G_1}'} \left[\left(1 - G_1^*\right) \overline{p}_1^* \sin\left(\theta_{G_1}' - \theta_{G_1}\right) + \left(1 - G_2^*\right) \overline{p}_2^* \sin\left(\theta_{G_1}' - \theta_{G_2}\right) \right]. \qquad \text{(S49)}$$

Following the assumptions of adaptive dynamics, we evaluate (S49) at the resident trait value, $\theta_{G_1}' = \theta_{G_1}$,

which means that $\sin\left(\theta_{G_1}' - \theta_{G_1}\right) = \sin\left(\theta_{G_1} - \theta_{G_1}\right) = \sin(0) = 0$. Hence, (S49) simplifies to

$$\left. \frac{d\overline{r_1'}}{d\theta_{G_1}'} \right|_{\theta_{G_1}' = \theta_{G_1}} \approx \gamma_1 \sigma_{E_G}^2 \left(1 - G_2^*\right) \overline{p}_2^* \sin\left(\theta_{G_1} - \theta_{G_2}\right). \qquad \text{(S50)}$$

The selection gradient given by equation (S50) has three solutions, $\theta_{G_1} = \theta_{G_2}$ and $\theta_{G_1} - \theta_{G_2} = \pm\pi$. Again, all

parameters γ_1, $(1 - G_2^*)$, $\overline{p}_2^*$, and $\sigma_{E_G}^2$ are positive and so the sign of (S50) is determined by the sign of

$\sin(\theta_{G_1} - \theta_{G_2})$. The solution $\theta_{G_1} = \theta_{G_2}$ is unstable because (S50) is negative for $\theta_{G_1} < \theta_{G_2}$ and positive for

$\theta_{G_1} > \theta_{G_2}$, both of which increase differences in θs between species. These differences between species are

increased until either solution ($\theta_{G_1} - \theta_{G_2} = \pm\pi$) is reached, both of which are stable. This is the

characteristic of classical character displacement but for temporal niche partitioning (see Abrams et al 2013) and accurately predicts simulation results of greater evolved differences between species in their germination traits with decreasing σ_{E_V} at fixed σ_{E_G} (figs 4a,c).

The analysis is based only on the invasion rates of mutant relative to resident. However, it is not guaranteed that all phenotype values as resident have positive population densities, which is to say that they coexist with the competitor. As shown in section S6, the storage effect promoting coexistence in this case is monotonically decreasing in corr(E_{G1}, E_{G2}). Since corr(E_{G1}, E_{G2}) is a monotonically decreasing function of $|\theta_{G_1} - \theta_{G_2}|$, eqns (S56) – (S60) show that both species' invader growth rates also increase with $|\theta_{G_1} - \theta_{G_2}|$. This means that, if species can coexist for some initial trait difference, then they will also coexist as selection proceeds, since both species invasions rates increase with increasing $|\theta_{G_1} - \theta_{G_2}|$

caused by selection. However, if mutual invasibility is not satisfied for initial trait differences, then the selection gradient given by expression (S50) does not apply because $\bar{p}_i^* = 0$ for the species with $\bar{r}_i / \beta_i < 0$. Selection in this case simplifies to the single species case considered in section S5 "The single species case."

Large vigor variation relative to germination variation and equal vigor
Here, we assume that species have the same vigor responses, in which case $V_1(t) = V_2(t) = V(t)$ for all t. To investigate this case, we rewrite the selection gradient in (S46) as

$$\frac{d\bar{r}_1'}{d\theta_{G_1}'} \approx \gamma_1 \sigma_{E_V} \sigma_{E_G} \cdot \frac{d}{d\theta_{G_1}'} \left[\left(1 - \bar{p}_1^* - \bar{p}_2^*\right) \mathrm{Corr}\left(E_{G_1}', E_V\right) - \frac{\sigma_{E_G}}{\sigma_{E_V}} \left[\left(1 - G_1^*\right) \bar{p}_1^* \mathrm{Corr}\left(E_{G_1}', E_{G_1}\right) + \left(1 - G_j^*\right) \bar{p}_2^* \mathrm{Corr}\left(E_{G_1}', E_{G_2}\right) \right] \right]$$

(S51)

where the subscripts of E_V are dropped because the two species have identical vigor responses. In the limit as $\sigma_{E_G} / \sigma_{E_V} \to 0$, then the selection gradient (S51) simplifies to

$$\frac{d\bar{r}_i'}{d\theta_{G_1}'} \approx \gamma_i \left(1 - \bar{p}_1^* - \bar{p}_2^*\right) \sigma_{E_V} \sigma_{E_G} \rho \frac{d}{d\theta_{G_1}'} \cos\left(\theta_{G_1}' - \theta_V\right)$$
$$= -\gamma_i \left(1 - \bar{p}_1^* - \bar{p}_2^*\right) \sigma_{E_V} \sigma_{E_G} \rho \sin\left(\theta_{G_1}' - \theta_V\right)$$

(S52)

Equation (S52) is nearly identical to the long-term growth rate for an invading type in the single species case (S41), except that now the term $(1 - p^*)$ is replaced with a measure of the amount to which competition from both species reduces selection for vigor, $(1 - \bar{p}_1^* - \bar{p}_2^*)$. The conclusions from the single species case apply in this case because both $(1 - p^*)$ and $(1 - \bar{p}_1^* - \bar{p}_2^*)$ are bounded between 0 and 1. To see why, note that the sum $\bar{p}_1^* + \bar{p}_2^*$ can be written as

$$\bar{p}_1^* + \bar{p}_2^* = \frac{\alpha\left(N_1 G_1 V_1 + N_2 G_2 V_2\right)}{1 + \alpha\left(N_1 G_1 V_1 + N_2 G_2 V_2\right)},$$

(S53)

and so is bounded above by 1. As shown in section S5 "The single species case," selection in the single species case favors maximization of the correlation between germination and vigor, i.e., $|\theta_{G1} - \theta_V| \to 0$. Here, the same conclusion applies. However, since both species have identical vigor responses, evolution leads to identical germination responses of the species as well, i.e., $|\theta_{G1} - \theta_{G2}| \to 0$. This result predicts the trend of smaller difference in germination traits with increasing σ_{E_V} from simulations (figs 4a,c).

Uncorrelated germination and vigor ($\rho = 0$).
This case is the simplest to investigate. Regardless of the species' trait values, $\mathrm{Cov}(E_G, E_V) = 0$ when $\rho = 0$. Hence, the density-independent component of selection is absent from the selection gradient. Furthermore, the correlation between E_G and E_V in the competition term is zero. This means the selection gradient is identical to previous case considered "Large germination variation relative to vigor variation with equal vigor." In that section, species with initial trait difference $|\theta_{G1} - \theta_{G2}|$ that satisfies mutual invasibility diverge under selection. This result accurately predicts the simulation results in figs 4c and 6a that maximum divergence occurs when ρ is small, regardless of other parameter values, even under large environmental variation ($\sigma_{E_G} = \sigma_{E_V} = 1$ for simulations in figs 4 and 6).

S6. Deriving and quantifying species coexistence mechanisms
Deriving coexistence mechanisms
To identify mechanisms affecting coexistence of two species, we follow the general partitioning method of the low-density, long-term growth rate of a species as outlined in Chesson (1994). The approximation

(S33) for the low-density growth rate can be used in an invasion analysis, which we use here to determine species coexistence in terms of stochastic boundedness (ala Chesson and Ellner 1989 and Schreiber et al 2011). One species, termed the invader and labelled with subscript i, is assumed to be at zero density and grows in the presence of a resident, with dynamics given by its stationary population fluctuations and labelled with subscript k. Operationally, this is done by removing the influence of the invader from C but retaining the effects of the resident. Following previous notation, $C^{\{-i\}}$ designates competition with species i set to zero density, which means that $C^{\{-i\}} = \ln\{1 + \alpha G_k V_k N_k\}$.

In section S4, the long-term growth rate of a species i at low density, $\bar{r}_i$, is derived and given in equation (S33):

$$
\begin{aligned}
\bar{r}_i \approx & \left(\beta_i - G_i^*\right) E(\Delta E_{G_i}) - \beta_i E(\Delta C^{\{-i\}}) \\
& + \frac{1}{2}(1-s)(1-\beta_i)\mathrm{Var}(E_G) + \frac{1}{2}\beta_i(1-\beta_i)\left[\mathrm{Var}(E_V) + \mathrm{Var}(C^{\{-i\}})\right] \\
& + \beta_i(1-\beta_i)\left[\mathrm{Cov}\left(E_{G_i}, E_{V_i}\right) - \mathrm{Cov}\left(E_{G_i}, C^{\{-i\}}\right) - \mathrm{Cov}\left(E_{V_i}, C^{\{-i\}}\right)\right].
\end{aligned}
\tag{S54}
$$

To determine mechanisms of coexistence between species, comparisons of the components of long-term growth between species uncover ways in which species differ that allow for or undermine coexistence. Chesson (2018) reviews the necessity for scaling components of growth rates by factors related to species life history speed. The scaling factor in the annual plant model here is $\beta_j = 1 - s(1 - G_j^*) = 1 - s(1 + \exp(E_{Gj}^*))^{-1}$, which represents the fraction of seed in the seed bank at the beginning of the season that is lost prior to new seed entering the seed bank. This quantity is the average fraction of the population that turns over across years. Since s is assumed to be the same between species, differences between species in β_j are determined by differences in $E_{G_j}^*$, (S29), which are caused by species differences in y.

We use these scaling factors to compare components of the growth rates of resident and invader species. Since the resident long-term growth rate is assumed to satisfy $\bar{r}_k = 0$, the invader long-term growth rate can be written as

$$
\frac{\bar{r}_i}{\beta_i} = \frac{\bar{r}_i}{\beta_i} - \frac{\bar{r}_k}{\beta_k} .
\tag{S55}
$$

Following Holt and Chesson (2014), to first order, $(\beta_i - G_i^*)[E(E_{G_i}) - E_{G_i}^*] - \beta_i(\bar{C}_i^* - C^*) \approx 0$. Therefore, we substitute $\beta_i(\bar{C}_i^* - C^*)$ for $(\beta_i - G_i^*)E(\Delta E_{G_i})$ in (S54) and use (S54) in (S55) to yield

$$
\frac{\bar{r}_i}{\beta_i} \approx \left(\kappa_i - \kappa_k\right) + \Delta J_i + \Delta I_{G_i} + \Delta I_{V_i} ,
\tag{S56}
$$

where

$$
\begin{aligned}
\kappa_i - \kappa_k = & \ln y_i - \ln y_k + \frac{1}{2}(\beta_k - \beta_i)\left[\frac{(1-s)}{\beta_i \beta_k}\mathrm{Var}(E_G) + \mathrm{Var}(E_V)\right] \\
& + \frac{\gamma_i}{\beta_i}\mathrm{Cov}\left(E_{G_i}, E_{V_i}\right) - \frac{\gamma_k}{\beta_k}\mathrm{Cov}\left(E_{G_k}, E_{V_k}\right)
\end{aligned}
\tag{S57}
$$

$$
\Delta J_i = \frac{1}{2}(\beta_i - \beta_k)\mathrm{Var}\left(C^{\{-i\}}\right)
\tag{S58}
$$

$$
\Delta I_{G_i} = \frac{\gamma_k}{\beta_k}\mathrm{Cov}\left(E_{G_k}, C^{\{-i\}}\right) - \frac{\gamma_i}{\beta_i}\mathrm{Cov}\left(E_{G_i}, C^{\{-i\}}\right)
\tag{S59}
$$

15

$$\Delta I_{V_i} = \frac{\gamma_k}{\beta_k}\text{Cov}\left(E_{V_k}, C^{\{-i\}}\right) - \frac{\gamma_i}{\beta_i}\text{Cov}\left(E_{V_i}, C^{\{-i\}}\right) \ . \tag{S60}$$

Expression (S56) shows that the invasion rate of a species can be written as a sum of four components, given in expression (S57) – (S60). Expression (S57) represents species i's average fitness advantage over its competitor. If positive, species i excludes species k in the scenario of no coexistence (Chesson 2018). In the absence of any coexistence mechanisms, the species with larger κ excludes the other. (S58) gives an expression for the relative nonlinearity coexistence mechanism, which is traditionally denoted by $-\Delta N$, but modified here to more directly demonstrate the potential for it to boost the recovery rate of species. Furthermore, we use J to represent the influence of Jensen's Inequality via the nonlinear averaging that relative nonlinearity quantifies. (S59) gives an expression for the storage effect coexistence mechanism from germination variation and (S60) gives an expression for the storage effect coexistence mechanism from vigor variation. Storage effects stabilize competition between species when environmental responses between species are less than correlated.

The main text considers three scenarios. In the first, species have the same seed yield ($y_1 = y_2 = y$) and identical vigor traits, i.e., $\theta_{V_1} = \theta_{V_2} = \theta_V$, which means $E_{V_1} = E_{V_2} = E_V$. Since species have the same y, $\beta_1 = \beta_2 = \beta$. Therefore, $\gamma_1 = \gamma_2 = \gamma$ and species average fitness difference is $\kappa_1 - \kappa_2 = (1-\beta)\text{Cov}(E_{G_1} - E_{G_2}, E_V)$. Hence, species only differ in average fitness if based on differences in predictive germination. The only stabilizing mechanisms in this case are storage effects from germination (ΔI_G). Identical y and identical E_V for both species mean that $\Delta J = 0$ and $\Delta I_V = 0$, respectively.

The main text also considers the case of identical vigor responses but unequal seed yield, $y_1 \neq y_2$. Species average fitness inequality in this case is

$$\kappa_i - \kappa_k = \ln y_i - \ln y_k + \frac{1}{2}(\beta_k - \beta_i)\left[\frac{(1-s)}{\beta_i\beta_k}\text{Var}(E_G) + \text{Var}(E_V)\right] + \text{Cov}\left((1-\beta_i)E_{G_i} - (1-\beta_k)E_{G_k}, E_V\right). \tag{S61}$$

Both predictive germination and differences in seed yield contribute to average differences between species. Furthermore, the storage effect mechanism from vigor has asymmetrical effects between species. Since $\gamma_k/\beta_k - \gamma_i/\beta_i = (1-\beta_k) - (1-\beta_i) = \beta_i - \beta_k$, the storage effect from vigor can be written as

$$\Delta I_{V_i} = (\beta_i - \beta_k)\text{Cov}\left(E_V, C^{\{-i\}}\right). \tag{S62}$$

Note the similar structure of (S62) to (S58). Like relative nonlinearity, ΔI_V is asymmetrical between species in this case and so facilitates invasion of one species and disadvantages another. This can be seen from (S62) because the sign of $\text{Cov}(E_V, C) > 0$ regardless of the identity of the invading species and because the sign of $\beta_i - \beta_j$ changes with species. The species with larger β relatively benefits from ΔI_V. The species with larger y has smaller β meaning that the competitive dominant has $\Delta I_V < 0$ while the subordinate has $\Delta I_V > 0$. Hence, ΔI_V reduces species average differences; it acts as an equalizing mechanism. A similar argument holds for ΔJ owing to the similar mathematical structure and the fact that $\text{Var}(C) > 0$.

The last case considered in the text is one of unequal vigor traits ($E_{V1} \neq E_{V2}$) but identical seed yield ($y_1 = y_2 = y$). Both storage effects contribute to stabilizing competition in this case. Relative nonlinearity is zero due to equal seed yield. Furthermore, species average fitness differences are

$$\kappa_1 - \kappa_2 = (1-\beta)\left[\text{Cov}(E_{G_1}, E_{V_1}) - \text{Cov}(E_{G_2}, E_{V_2})\right]. \tag{S63}$$

Hence, species average fitness differences are proportional to species differences in predictive germination. These results are summarized in Table 3 of the main text.

Quantifying stabilizing mechanisms and fitness inequalities

To measure the strength of all stabilizing mechanisms, we take the average over both species of their invader growth rates, i.e., $1/2\sum_{i=1,2} E[r_i/\beta_i]$. This provides a measure of all stabilizing mechanisms because the average recovery rate can be written as

$$\frac{1}{2}\left(\frac{\bar{r}_1}{\beta_1}+\frac{\bar{r}_2}{\beta_2}\right)=\frac{1}{2}(\kappa_1-\kappa_2)-\frac{1}{2}(\kappa_1-\kappa_2)+\overline{\Delta J}+\overline{\Delta I_G}+\overline{\Delta I_V}$$

$$=\overline{\Delta J}+\overline{\Delta I_G}+\overline{\Delta I_V}\equiv\bar{A}. \tag{S64}$$

Equation (S64) shows that the sum of all community average mechanisms is given directly by the average of the growth rates of each species, when evaluated at low density. The amount to which a species 1's recovery rate departs from this average value is

$$\frac{\bar{r}_1}{\beta_1}-\bar{A}=(\kappa_1-\kappa_2)+\left(\Delta J_1+\overline{\Delta J}\right)+\left(\Delta I_{G_1}-\overline{\Delta I_G}\right)+\left(\Delta I_{V_1}-\overline{\Delta I_V}\right)$$

$$=\kappa_1+\frac{1}{2}\left(\Delta J_1+\overline{\Delta J}\right)+\frac{1}{2}\left(\Delta I_{G_1}-\overline{\Delta I_G}\right)+\frac{1}{2}\left(\Delta I_{V_1}-\overline{\Delta I_V}\right) \tag{S65}$$

$$-\kappa_2+\frac{1}{2}\left(\Delta J_1+\overline{\Delta J}\right)+\frac{1}{2}\left(\Delta I_{G_1}-\overline{\Delta I_G}\right)+\frac{1}{2}\left(\Delta I_{V_1}-\overline{\Delta I_V}\right).$$

Note that $\Delta X_1-\overline{\Delta X}=\overline{\Delta X}-\Delta X_2$ because

$$\Delta X_1-\overline{\Delta X}=\Delta X_1-\frac{(\Delta X_1+\Delta X_2)}{2}$$

$$=\frac{(\Delta X_1-\Delta X_2)}{2}$$

$$=\frac{(\Delta X_1+\Delta X_2)-2\Delta X_2}{2}$$

$$=\frac{(\Delta X_1+\Delta X_2)}{2}-\Delta X_2$$

$$=\overline{\Delta X}-\Delta X_2. \tag{S66}$$

Define $\delta X_j=\Delta X_j-\overline{\Delta X_j}$ as the difference of a species stabilizing mechanism from the community average. Rewritting (S65) using the identity (S66) and the definition δX_j yields

$$\frac{\bar{r}_1}{\beta_1}-\bar{A}=\kappa_1+\frac{1}{2}\left(\Delta J_1+\overline{\Delta J}\right)+\frac{1}{2}\left(\Delta I_{G_1}-\overline{\Delta I_G}\right)+\frac{1}{2}\left(\Delta I_{V_1}-\overline{\Delta I_V}\right)$$

$$-\kappa_2-\frac{1}{2}\left(\Delta J_1+\overline{\Delta J}\right)-\frac{1}{2}\left(\Delta I_{G_1}-\overline{\Delta I_G}\right)-\frac{1}{2}\left(\Delta I_{V_1}-\overline{\Delta I_V}\right)$$

$$=\left(\kappa_1+\frac{1}{2}\delta J_1+\frac{1}{2}\delta I_{G_1}+\frac{1}{2}\delta I_{V_1}\right)-\left(\kappa_2'+\frac{1}{2}\delta J_2+\frac{1}{2}\delta I_{G_2}+\frac{1}{2}\delta I_{V_2}\right)$$

$$=\kappa_1'-\kappa_2'. \tag{S67}$$

where $\kappa_j' = \kappa_j + (\delta J_j + \delta I_{Gj} + \delta I_{Vj})/2$. The κ' values in (S67) are the modified species average fitness differences that account for asymmetrical effects of stabilizing mechanisms (Chesson 2018). A species average fitness can be calculated directly via the equation $\kappa_j' = \bar{r}_j/\beta_j - A = \bar{r}_j/\beta_j - (\bar{r}_1/\beta_1 + \bar{r}_2/\beta_2)/2$, which only requires the calculation of invasion growth rates of both species. In addition, the species average fitness difference is $\kappa_1' - \kappa_2' = \bar{r}_1/\beta_1 - \bar{r}_2/\beta_2$.

Stable coexistence occurs when both species have positive $\bar{r}_i/\beta_i > 0$ for both species. From the last line of (S67), it follows that $\bar{r}_1/\beta_1 > 0$ when $\bar{A} > \kappa_2' - \kappa_1'$. The invasion condition for species 2 is $\bar{r}_2/\beta_2 > 0$, which occurs when $\bar{A} > \kappa_1' - \kappa_2'$. Putting these two conditions together yields the condition for stable coexistence,

$$\bar{A}>\left|\kappa_i'-\kappa_j'\right|. \tag{S68}$$

Calculating invader growth rates from simulations

The section directly above shows that total stabilizing mechanisms and species average fitness differences can be calculated directly from invader growth rates of species. To estimate these, we used the following process in simulations for 50,000 times steps in Matlab.

1. Randomly sample T independent draws from a four-dimensional multivariate distribution with mean zero and variance matrix

$$\Sigma = \begin{pmatrix} 1 & 0 & \rho & 0 \\ 0 & 1 & 0 & \rho \\ \rho & 0 & 1 & 0 \\ 0 & \rho & 0 & 1 \end{pmatrix}. \tag{S69}$$

   Σ satisfies the assumption in the main text for the vector of environmental variables (X_1, X_2, Z_1, Z_2).

2. Transform the traits, (θ_{G1}, θ_{G2}, θ_{V1}, θ_{V2}), using the relationship $e_{j1} = \sin(\theta_{Gj})$, $e_{j2} = \cos(\theta_{Gj})$, $v_{j1} = \sin(\theta_{Vj})$, and $v_{j2} = \cos(\theta_{Vj})$.

3. Using the transformed values in step 2, calculate environmental response, E_G and E_V, for each species using eqns (6) and (8) from the main text for the T values of the environmental factors sampled in step 1.

4. Initiate the population of each species, $N_j(0)$ with the density corresponding to the equilibrium population size in a constant environment,

$$N_j^* = \left[\frac{y_j \bar{G}^* \bar{V}^*}{1 - s(1 - \bar{G}^*)} - 1 \right] \frac{1}{\alpha \bar{G}^* \bar{V}^*}, \tag{S70}$$

   where $\bar{G}^* = \exp(\mu_{E_G})/(1 + \exp(\mu_{E_G}))$ and $\bar{V}^* = \exp(\mu_{E_V})$.

5. Using eqns (1) – (3) in the main text, simulate population growth of each species assuming only one species is present, i.e., $C' = 1 + \alpha G_j V_j N_j$ for focal species j while retaining the T values of C'.

6. For each species, calculate the invader growth rates after discarding the first 1000 time steps to get rid of effects on initial values. Using the values of C' for the corresponding competing species as resident from step 5. This means that species 1's invader growth rate is

$$\frac{\bar{r}_1}{\beta_1} = \frac{\frac{1}{(T-1000)} \sum_{t=1001}^{T} \log\left\{ s(1 - G_1(t)) + \frac{y_1 G_1(t) V_1(t)}{1 + \alpha G_2(t) V_2(t) N_2(t)} \right\}}{1 - s\left(1 - \frac{\exp\left(E_{G_1}^*\right)}{1 + \exp\left(E_{G_1}^*\right)} \right)} \tag{S71}$$

   and species 2's is

$$\frac{\bar{r}_2}{\beta_2} = \frac{\frac{1}{(T-1000)} \sum_{t=1001}^{T} \log\left\{ s(1 - G_2(t)) + \frac{y_2 G_2(t) V_2(t)}{1 + \alpha G_1(t) V_1(t) N_1(t)} \right\}}{1 - s\left(1 - \frac{\exp\left(E_{G_2}^*\right)}{1 + \exp\left(E_{G_2}^*\right)} \right)}, \tag{S72}$$

   where $E_{G_j}^*$ is given in equation (S30) above.

Calculating selection gradients from simulations

To calculate the selection gradient, we used the approximation

$$\left. \frac{d\bar{r}_j'}{d\theta_{G_j}'} \right|_{\theta_{G_j}' = \theta_{G_j}} \approx \frac{\bar{r}_j' - \bar{r}}{\theta_{G_j}' - \theta_{G_j}}, \tag{S73}$$

for small $\theta'_{G_j} - \theta_{G_j}$ ($\theta'_{G_j} > \theta_{G_j}$). The approximation is better for smaller differences between the resident and mutant phenotypes. We choose $\theta'_{G_j} - \theta_{G_j} = 0.05 \approx 0.016\pi$. Each mutant-resident comparison represents approximately 0.8% of phenotype space. The mutant invader growth rate is calculated using the formula

$$\bar{r}'_j = \frac{1}{(T-1000)} \sum_{t=1001}^{T} \log\left\{ s\left(1-G'_j(t)\right) + \frac{y_j G'_j(t) V_j(t)}{1+\alpha\left(G_1(t)V_1(t)N_1(t)+G_2(t)V_2(t)N_2(t)\right)} \right\}, \tag{S74}$$

and the resident growth rate is calculated using the formula

$$\bar{r}_j = \frac{1}{(T-1000)} \sum_{t=1001}^{T} \log\left\{ s\left(1-G_j(t)\right) + \frac{y_j G_j(t) V_j(t)}{1+\alpha\left(G_1(t)V_1(t)N_1(t)+G_2(t)V_2(t)N_2(t)\right)} \right\}. \tag{S75}$$

In both cases, we used values of $N_1(t)$ and $N_2(t)$ from stationary distributions of the two species when in competition with each other. That is to say that we simulated equations (1) – (3) in the main text for $N_1(0) > 0$ and $N_2(0) > 0$ for T time units. As in the calculations above, $T = 50,000$.

Calculating the evolutionarily stable $|\theta_{G1} - \theta_{G2}|$.
We numerically calculated the ESS $|\theta_{G1} - \theta_{G2}|$ in figures 4-6 of the main text by simulating the adaptive dynamics process for alternating mutations in each species. Initial trait values for each species were set to $\theta_{G1} = \pi/10$ and $\theta_{G2} = -\pi/10$. Starting with these trait values, mutations alternated between species, creating mutant phenotypes that could invade the resident population. Mutant phenotypes are a random sample from a uniform distribution, $\theta'_{Gj} \sim \text{Uniform}(\theta_{Gj} - 0.01, \theta_{Gj} + 0.01)$, meaning that resident and mutant phenotypes differ at most 0.01 units. When this mutation process led to $\theta'_{G_j} > \pi/2$ or $\theta'_{G_j} < -\pi/2$, meaning the phenotype lies outside our defined phenotype space, we set the mutant to the boundary phenotype θ'_{G_j} = $\pi/2$ in the case of $\theta'_{G_j} > \pi/2$ and $\theta'_{G_j} = -\pi/2$ when $\theta'_{G_j} < -\pi/2$.

For each resident-mutant comparison, we calculated the mutant advantage, $\bar{r}_j' - \bar{r}_j$, for species j subject to mutation, using equations (S74-S75) above. If $\bar{r}_j' > \bar{r}_j$, we set the mutant as the new resident phenotype value. If $\bar{r}_j' \le \bar{r}_j$, we assumed the mutant failed to invade and the resident phenotype was retained for the next mutation.

We repeated these steps for 3500 mutations, corresponding to 1750 mutations for each species, and recorded the end trait values θ_{G1} and θ_{G2}. We visually inspected that trait values had stabilized across a range of parameter values after the 3500 mutations. 3500 mutations were sufficient in all cases to reach evolutionary equilibrium.

S7. Explanation for increasing coexistence mechanism strength with competition

Factors affecting resource demand in the seed bank model
An important factor in the model is the demand for resources as it controls the strength of density-dependent selection that drives character displacement. We track the amount of resource demand using $C' = 1 + \alpha G_1 V_1 N_1 + \alpha G_2 V_2 N_2$, which combines *per-capita* competitive effects of species (α) and *biomass weighted population density* in a year (GVN). Together, these two factors together give the *total* competitive effect in a year. A more direct measure of total competitive effect is given by a simple translation of C' to $C' - 1 = \alpha G_1 V_1 N_1 + \alpha G_2 V_2 N_2$, which measures total community biomass weighted by competitive effect. It is on a scale of $[0,\infty)$, where $C' - 1 = 0$ means no plants and thus no demand for resources.

One question a reader might ask is why we have chosen to change the value of y to change the level of resource demand – which is conspicuously absent in C' – rather than the competition coefficient α, which is present in C''? The answer to this question lies in the relationship between population density, N, and the competition coefficient, α. The two parameters act together as a composite parameter that is effectively constant with respect to changes in α. This is most easily illustrated in the constant environment version of the model for a single species. With a single species, we drop subscripts and take $G(t) = g$ and $V(t) = v$ for all t. The finite rate of increase is now

$$\lambda(t) = s(1-g) + \frac{yvg}{1+\alpha vgN(t)}.$$
(S76)

At equilibrium, $\lambda = 1$, and we have the following equality

$$vg\alpha N_{eq} = \frac{yvg}{1-s(1-g)} - 1,$$
(S77)

where the left-hand side is the equilibrium demand for resources ($= C'_{eq} - 1$) and the right-hand side is a suite of parameters related to the demographic rates of the species (g, v, s, and y). As illustrated from this equation, the equilibrium demand for competition cannot change by changing only the parameter α.

The same principle applies in a fluctuating environment. Figure S1 show the average value of $C'(t)$ for different values of α and y for a single species and for two coexisting species.
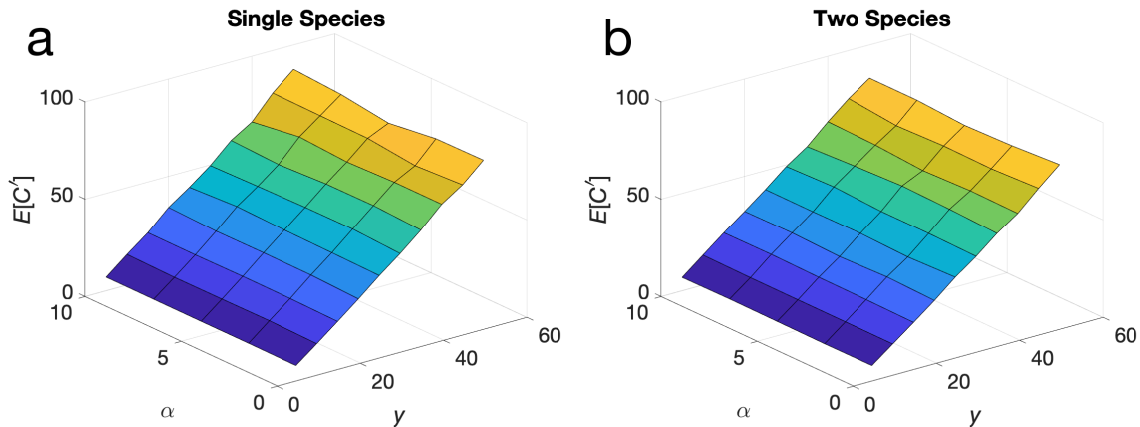

Figure S1. Values of average demand for resources, $E[C']$, for a single species (a) and two coexisting species (b) when G and V varies and $y_1 = y_2 = y$. Parameters: $s = 0.9$, $\mu_{E_G} = \mu_{E_V} = 0$, $\sigma_{E_G} = \sigma_{E_V} = 1$, $\rho = 0.5$, $\theta_{G1} = 0.4\pi$, $\theta_{G2} = -0.4\pi$, and $\theta_{V1} = \theta_{V2} = 0$.

Understanding the size of the coexistence region and the intensity of competition
In this model (and many others), the size of the coexistence region is larger for more intense competition. For example, the coexistence region is larger for larger values of y. The ecological justification for this result is as follows. Two species coexist when each species inhibits its own growth more than it inhibits the growth of a competitor (Chesson 2018). A common restatement is "*intraspecific competition is greater than interspecific competition.*" The strength of competition here plays an important role in the level of intraspecific limitation. When competition is strong, a species as resident is restricted by competition. If ecological differences reduce interspecific competition *proportionately* (as we will show they do here), species as invaders experience greater benefits at fixed ecological differences with greater intensity of competition (Chesson 1994, 2000, 2003, 2018).

To illustrate this effect from Figure 3, consider the case of identical vigor responses (i.e., $\theta_{V1} = \theta_{V2}$ and $V_1(t) = V_2(t) = V(t)$ for all t) and symmetric competitors ($y_1 = y_2 = y$). In this case, the only coexistence

mechanism is the storage effect from germination (see Table 3). Hence, the low-density growth rate can be written as

$$\frac{\overline{r_i}}{\beta_i} \approx \kappa_i - \kappa_j + \Delta I_{G_i} \ . \tag{S78}$$

From Table 2 of the main text, the expression for the storage effect from germination is

$$\Delta I_{G_i} = (1-\beta)\mathrm{Cov}\left(E_{G_j} - E_{G_i}, C^{\{-i\}}\right) \tag{S79}$$

Using the first order approximation of $C^{\{-i\}}$ (S43) from above in (S79) yields the following expression

$$\Delta I_{G_i} \approx (1-\beta)\mathrm{Cov}\left(E_{G_j} - E_{G_i}, \hat{C}^{\{-i\}} + p_j^*(1-G_j^*)\left(E_{G_j} - E_{G_j}^*\right) + p_j^*\left(E_{V_j} - E_{V_j}^*\right)\right)$$

$$= (1-\beta)\mathrm{Cov}\left(E_{G_j} - E_{G_i}, (1-G_j^*)E_{G_j} + E_{V_j}\right)p_j^* \ , \tag{S80}$$

$$= (1-\beta)\left[(1-G_j^*)\sigma_{E_G}^2\left(1-\rho_{G_i,G_j}\right) + \sigma_{E_G}\sigma_{E_V}\rho\left(\rho_{G_j,V} - \rho_{G_i,V}\right)\right]p_j^*$$

where we have set $p_i^* = 0$ since the invader species is at zero density and written the covariances in terms of their correlations, $\rho_{U,Z} = \mathrm{corr}(U,Z)$ for compactness in the final line.

To relate (S80) to the size of the coexistence region, we take the community average because the community average is proportional to the size of the coexistence region (Chesson 2018). Hence, larger values of the community average indicate larger regions of trait space that allow for coexistence of species. The community average storage effect writes as

$$\overline{\Delta I_G} \approx (1-\beta)\sigma_{E_G}^2\left(1-\overline{G^*}\right)\left(1-\rho_{G_1,G_2}\right)\overline{p^*} + (1-\beta)\sigma_{E_G}\sigma_{E_V}\rho\left(\rho_{G_2,V} - \rho_{G_1,V}\right)E\left[\frac{p_1^* - p_2^*}{2}\right], \tag{S81}$$

where an overbar indicates an average over species as residents. The intensity of competition is in part determined by $\overline{p^*}$, which can be seen by the explicit formula for p^*

$$p_j^* = E\left[\frac{\alpha G_j^* V^* N_j(t)}{1 + \alpha G_j^* V^* N_j(t)}\right]. \tag{S82}$$

The bounds on these functions are $0 < p_j^* < 1$ and they are larger for species with greater resident population sizes. Hence, more abundant species have greater competitive effect, which we can control with y, as illustrated in SI S7 "Factors affecting resource demand in the seed bank model."

We simplify the problem a bit by noting that $\rho_{G2,V} - \rho_{G1,V} \approx 0$ for species with similar absolute distances $|\theta_{Gi} - \theta_V|$, in which case (S81) simplifies greatly to

$$\overline{\Delta I_G} \approx \underbrace{(1-\beta)}_{\substack{\text{buffered population}\\\text{growth}}} \cdot \underbrace{\left(1-\overline{G^*}\right)\overline{p^*}}_{\substack{\text{intensity of}\\\text{competition}}} \cdot \underbrace{\sigma_{E_G}^2}_{\substack{\text{environmental}\\\text{variability}}} \cdot \underbrace{\left(1-\rho_{G_1,G_2}\right)}_{\substack{\text{species-specific}\\\text{germination}}}. \tag{S83}$$

Equation (S83) takes a form that is interpretable in terms of the intensity of competition. The first three terms together determine the effects of competition on resident species, on average. This amount is reduced for invaders by a factor ρ_{G_1,G_2}, determined by species-specific germination responses (i.e., $|\theta_{G1} - \theta_{G2}|$). Hence, for fixed ρ_{G_1,G_2}, $\overline{\Delta I_G}$ increases with increasing intensity of competition. Stated differently, the size of the coexistence region increases as the distribution of population densities of the resident is larger, all else equal. This is the situation in Figure 3 where y increases in each panel, increasing N for each point in each panel.

In the section where we consider species with different vigor responses, we find that the size of the coexistence region—as measured by $\overline{A}$—increases the most in sympatry when there is a strong cue. In this case, total stabilizing mechanisms for one species is the sum of storage effects from both germination and vigor:

$$A_i = \Delta I_{G_i} + \Delta I_{V_i}$$
$$= (1-\beta)\mathrm{Cov}\left(\tilde{E}_j - \tilde{E}_i, C^{\{-i\}}\right)' \tag{S84}$$

where $\tilde{E}_j = E_{Gj} + E_{Vj}$ is the *total environmental response* in a year and combines both germination and vigor. We can now understand the storage effect in terms of single environmental response that combines both germination and vigor. For symmetric competitors (which we have in Fig. 6), we expect (S84) to have the general form:

$$A_i = (1-\beta)\sigma_{\tilde{E}}^2\left(1 - \rho_{\tilde{E}}\right)B \ , \tag{S85}$$

where $\sigma_{\tilde{E}}^2 = \mathrm{Var}(\tilde{E})$, $\rho_{\tilde{E}} = \mathrm{Corr}(\tilde{E}_j, \tilde{E}_i)$, and B is related to the average density of the residents (Chesson 2003). The effect of the cue is present in the variance of the total environmental response because the expression for the variance of a sum is

$$\sigma_{\tilde{E}}^2 = \sigma_{E_G}^2 + \sigma_{E_V}^2 + 2\mathrm{Cov}(E_G, E_V)$$
$$= \sigma_{E_G}^2 + \sigma_{E_V}^2 + 2\sigma_{E_G}\sigma_{E_V}\rho\cos\left(\theta_G - \theta_V\right) \ . \tag{S86}$$

This makes clear that variances in total environmental responses in environments with stronger cues (ρ) provided the covariance between E_G and E_V is positive. The covariance is positive when germination is predictive, i.e., $\cos(\theta_G - \theta_V) > 0$. In terms of trait differences, $|\theta_G - \theta_V| < \pi/2$. It appears that the evolutionarily stable θ_G correspond to predictive germination for both species in Fig. 6b for large values of ρ. We suspect increased variance in total environmental response under strong cues increases the magnitude of competition and elevates total stabilizing mechanisms.

References

Abrams, P. A., Tucker, C. M. and Gilbert, B. 2013. Evolution of the storage effect. *Evolution* **67**: 315–327.

Chesson, P. 1994. Multispecies Competition in Variable Environments. Theoretical Population Biology **45**:227-276.

Chesson, P. 2000. Mechanisms of maintenance of species diversity. Annual Review of Ecology and Systematics **31**:343-366.

Chesson, P. 2003. Quantifying and testing coexistence mechanisms arising from recruitment fluctuations. Theor Popul Biol **64**:345-357.

Chesson, P. 2018. Updates on mechanisms of maintenance of species diversity. Journal of Ecology **106**:1773-1794.

Chesson, P.L. and S. Ellner. 1989. Invasibility and stochastic boundedness in monotonic competition models. *Journal of Mathematical Biology* **27**:117–138.

Holt, G., and P. Chesson. 2014. Variation in moisture duration as a driver of coexistence by the storage effect in desert annual plants. Theoretical Population Biology **92**:36-50.

Johnson, N. L., and S. Kotz. 1972. Continuous multivariate distributions. Wiley, New York.

Johnson, N. L. 1994. Continuous univariate distributions. 2nd ed. edition. New York : Wiley & Sons, New York.

Gilbert, G.T. 1991. Positive definite matrices and Sylvester's criterion. *The American Mathematical Monthly*. Mathematical Association of America, **98** (1): 44-46.

Schreiber, S. J., M. Benaim, and K. A. Atchade. 2011. Persistence in fluctuating environments. Journal of Mathematical Biology **62**:655-683.