New Phytologist Supporting Information

Article title: Bounds on stomatal size can explain scaling with stomatal density in forest plants

Authors: Congcong Liu, Christopher D. Muir, Lawren Sack, Ying Li, Li Xu, Mingxu Li, Jiahui Zhang, Hugo Jan de Boer, Xingguo Han, Guirui Yu, Nianpeng He

Article acceptance date: 11 September 2025

Note S1 Covariance between f_S and $g_{s,max}$

Here we prove that the covariance between f_S and $g_{s,max}$ as defined in Eqn. 1 and Eqn. 2 must almost certainly be positive on a log-log scale. The problem can be formulated as follows:

$$\mathrm{Cov}\left(\log f_S, \log g_{s,max}\right) > 0. \tag{S1}$$

Substituting the right-hand sides of Eqn. 1 and 2 into Eqn S1 and removing constants, we obtain:

$$\mathrm{Cov}\left(d_S + a_S, d_S + \beta a_S\right) > 0. \tag{S2}$$

Applying the addition rule and bilinearity of covariance property:

$$\mathrm{Cov}\left(d_S + a_S, d_S + \beta a_S\right) = \mathrm{Var}\left(d_S\right) + \beta \mathrm{Var}\left(a_S\right) + (1+\beta)\,\mathrm{Cov}\left(d_S, a_S\right). \tag{S3}$$

Based on Eqn S3, $\mathrm{Cov}\left(\log f_S, \log g_{s,max}\right)$ must be greater than 0 if $\mathrm{Cov}\left(d_S, a_S\right) > 0$, so we restrict the analysis to the domain where $\mathrm{Cov}\left(d_S, a_S\right) < 0$. The covariance will not be positive unless:

$$\mathrm{Cov}\left(d_S, a_S\right) < -\frac{\mathrm{Var}\left(d_S\right) + \beta \mathrm{Var}\left(a_S\right)}{1+\beta}. \tag{S4}$$

The minimum value of the covariance for a given correlation ρ between d_S and a_S is:

$$\mathrm{Cov}\left(d_S, a_S\right) \geq -\rho\sqrt{\mathrm{Var}\left(d_S\right)\mathrm{Var}\left(a_S\right)}. \tag{S5}$$

From the minimum and maximum values of the covariance in Eqn S4 and Eqn S5, it is only possible to for f_S and $g_{s,max}$ to negatively covary if $\beta < 1$ and the ratio of their variances is between:

$$\beta^2 < \frac{\mathrm{Var}\left(d_S\right)}{\mathrm{Var}\left(a_S\right)} < 1 \tag{S6}$$

Within this range, the least negative possible correlation between d_S and a_S that would result is positive covariance between f_S and $g_{s,max}$ is where $\mathrm{Var}\left(d_S\right)/\mathrm{Var}\left(a_S\right) = \beta$. Substituting this into Eqn S4 and calculating the correlation, we find that the correlation d_S and a_S must be:

$$-1 \geq \rho \geq \frac{2\sqrt{\beta}}{1+\beta} \tag{S7}$$

For $\beta = 0.5$, the relationship f_S and $g_{s,max}$ must be positive unless the correlation between d_S and a_S is less than -0.943. Hence, there is extremely limited scope for the covariance between f_S and $g_{s,max}$ to be negative.

Note S2 Background on evolutionary quantitative genetics and estimating trait relationships

Scaling up microevolutionary processes to macroevolutionary patterns requires assumptions about how the fitness optima change through time in different species. Macroevolutionary landscape theory (Hansen, 1997; Pennell & Harmon, 2013; Uyeda & Harmon, 2014; Boucher *et al.*, 2018; Rolland *et al.*, 2023) provides a framework to model the dynamics and bounds on fitness optima. One key result is that the long-term dynamics of traits are probably not governed primarily by evolution on a static or unbounded landscape, because such models predict patterns of trait variance (Estes & Arnold, 2007) and rates of trait evolution (Hansen, 2012) that are inconsistent with observations. What this means for stomatal density and size in vascular plants is that covariance among species may be primarily determined by the dynamics of the macroevolutionary adaptive landscape, regardless of genetic (co)variance and response to selection within species. Specifically, we applied the multivariate Ornstein-Uhlenbeck (OU) model originally derived from quantitative genetics for intraspecific (population) trait microevolution by Lande (1976), and extended by Hansen (1997) and others (Pennell & Harmon, 2013) to macroevolutionary interspecific trait variation. In the macroevolutionary OU model, interspecific trait variation expands through time until it reaches a stationary distribution around a long-term average (Hansen, 1997). Within each species, microevolutionary forces (selection, genetic drift, mutation, and migration) drive genetic and plastic trait variation such that species' trait means should be near their current adaptive optimum. The among-species stationary distribution in the OU model arises from independent shifts in species' optimum trait values. At stationarity, an OU process leads to stable trait means and (co)variances among species. Fitness tradeoffs likely limit the breadth of values for adaptive trait optima, given that extreme trait values will rarely optimize competing functions (Sack & Buckley, 2020). The OU and other stochastic models of trait evolution do not specify what factors constrain movement in the adaptive optimum, but empirically the OU model best describes empirical patterns of functional trait evolution in angiosperms (Pennell *et al.*, 2015).

We use the trait covariance at stationarity to evaluate whether standardized major axis (SMA), ordinary least squares (OLS), or neither can reliably distinguish among competing hypotheses. SMA is commonly used in comparative ecology for estimating trait-trait relationships (Warton *et al.*, 2006), but this approach has been criticized in the context of phylogenetic comparative analyses because its estimates are biased when there is biological variation in the relationship between traits (Hansen & Bartoszek, 2012; Kilmer & Rodríguez, 2017). Phylogenetic information can be incorporated into both OLS (often referred to as generalized least-squares; Grafen (1989)) and SMA methods (e.g. Ives *et al.* (2007)) to account for statistical nonindependence in the residuals.

Table S1: Glossary of symbols.

Parameter	Description
$A_{\mathrm{S}},\ a_{\mathrm{S}}$	Stomatal size (area of stomatal complex). Lowercase is log-transformed.
$\overline{a}_{\mathrm{S}}$	Mean stomatal size within species
$a_{\mathrm{S,opt}}$	Optimal stomatal size
α	Return rate toward θ
$\mathbf{A}$	Return rate toward $\vec{\theta}$
b	Biophysical constant (Eqn 8)
Γ	Interspecific covariance matrix between $a_{\mathrm{S,opt}}$ and $g_{\mathrm{s,opt}}$
β	Scaling exponent between stomatal size and density in Eqns 3-4
$D_{\mathrm{S}},\ d_{\mathrm{S}}$	Stomatal density. Lowercase is log-transformed.
$\overline{d}_{\mathrm{S}}$	Mean stomatal density within species
f_{S}	Proportion of epidermal area allocated to stomata
$\overline{f}_{\mathrm{S}}$	Average proportion of epidermal area allocated to stomata within species
$f_{\mathrm{S,min}}$	Minimum proportion of epidermal area allocated to stomata
$\mathbf{G}$	Additive genetic covariance matrix
$\mathbf{G}^*$	Equilibrium additive genetic covariance matrix
$G_{a_{\mathrm{S}}}$	Genetic variance in stomatal size
$G^*_{a_{\mathrm{S}}}$	Equilibrium genetic variance in stomatal size
$G_{d_{\mathrm{S}}}$	Genetic variance in stomatal density
$G^*_{d_{\mathrm{S}}}$	Equilibrium genetic variance in stomatal density
$G_{d_{\mathrm{S}},a_{\mathrm{S}}}$	Genetic covariance in stomatal density and size
$G^*_{d_{\mathrm{S}},a_{\mathrm{S}}}$	Equilibrium genetic covariance in stomatal density and size
$g_{\mathrm{s,max}}$	Anatomical maximum stomatal conductance
$\overline{g}_{\mathrm{s,max}}$	Average anatomical maximum stomatal conductance within species
$g_{\mathrm{s,opt}}$	Optimal $g_{\mathrm{s,max}}$
$\mathbf{H}$	Curvature and orientation of the log-fitness landscape with respect to the component traits
λ	Scalar constant in Eqn S34
m	Morphological constant (Eqn 9)
$\mathbf{M}$	Mutation covariance matrix
$M_{a_{\mathrm{S}}}$	Mutational variance in stomatal size
$M_{d_{\mathrm{S}}}$	Mutational variance in stomatal density
$M_{d_{\mathrm{S}},a_{\mathrm{S}}}$	Mutational covariance in stomatal density and size
$\vec{\mu}$	Vector of trait means among species $[\mu_{d_{\mathrm{S}}}, \mu_{a_{\mathrm{S}}}]$
$\vec{\mu}^*$	Equilibrium vector of trait means among species $[\mu^*_{d_{\mathrm{S}}}, \mu^*_{a_{\mathrm{S}}}]$
ϕ_a	Strength of selection on stomatal size relative to ω
ϕ_f	Strength of selection on f_{S} relative to ω
ρ_{M}	Mutational correlation between in stomatal size and density
σ	Rate of random movement in $\log\left(g_{\mathrm{s,opt}}\right)$ each generation

Parameter	Description
Σ	Matrix of rate of random movement in $\vec{\theta}$
θ	Long-run average value of $\log\left(g_{\mathrm{s,opt}}\right)$
$\vec{\theta}$	Vector of long-run averages in $a_{\mathrm{S,opt}}$ and $\log\left(g_{\mathrm{s,opt}}\right)$
$\mathbf{V}$	Among-species trait covariance matrix
$\mathbf{V}^*$	Equilibrium among-species trait covariance matrix
$V_{a_{\mathrm{S}}}$	Among-species variance in stomatal size
$V^*_{a_{\mathrm{S}}}$	Equilibrium among-species variance in stomatal size
$V_{d_{\mathrm{S}}}$	Among-species variance in stomatal density
$V^*_{d_{\mathrm{S}}}$	Equilibrium among-species variance in stomatal density
$V_{d_{\mathrm{S}},a_{\mathrm{S}}}$	Among-species covariance in stomatal density and size
$V^*_{d_{\mathrm{S}},a_{\mathrm{S}}}$	Equilibrium among-species covariance in stomatal density and size
W	Individual absolute fitness
$\overline{W}$	Average absolute fitness
$W_{\max}$	Maximum individual absolute fitness
$\vec{\mathbf{x}}$	Vector of trait means within species $[\overline{d}_{\mathrm{S}}, \overline{a}_{\mathrm{S}}]$
$\vec{\mathbf{x}}^*$	Equilibrium vector of trait means within species $[\overline{d}^*_{\mathrm{S}}, \overline{a}^*_{\mathrm{S}}]$
z	Composite stomatal trait (log-transformed)
z_0	Scalar constant in Eqn 4
ω	Strength of stabilizing selection around $g_{\mathrm{s,opt}}$

Note S3 Evolutionary quantitative genetic model: Theory

Here we describe a evolutionary quantitative genetic (EQG) modeling framework and derive predictions about stomatal size-density scaling under competing hypotheses. A glossary of symbols is provided in Table S1. The model integrates microevolutionary (within species) and macroevolutionary (among species) processes. Within species, we use quantitative genetic theory to simulate the evolution of continuous phenotypes in response to selection and mutation. We use macroevolutionary theory to simulate how the adaptive optima fluctuate through time.

As mentioned in the Introduction, both $\log(g_{\mathrm{s,max}})$ and $\log(f_{\mathrm{S}})$ are derived from constituent traits stomatal density (d_{S}) and size (a_{S}), which we treat as random variables with different constants, z_0 and β:

$$z = z_0 + d_{\mathrm{S}} + \beta a_{\mathrm{S}} \tag{S8}$$

Recall that d_{S} and a_{S} are on a log-transformed scale. The composite trait z is equivalent to f_S when $z_0 = 0$ and $\beta = 1$; z is equivalent to $g_{\mathrm{s,max}}$ when $z_0 = \log(bm)$ and $\beta = 0.5$. We assume

that traits evolve due to selection and mutation in a large ensemble of independent "species". Every species has infinite population size, random mating, and there is no gene flow between species, so they evolve as independent lineages. In the next section, we describe short-term change in trait means and variances due to selection and mutation within species. We then describe the long-term macroevolution of trait optima. After describing micro- and macroevolutionary assumptions, we describe the general process we used to solve for the equilibria trait means and additive genetic (co)variance within species and the analogous stationary distribution of trait means and (co)variance among species. Finally, we apply these methods to derive predictions for each competing hypothesis.

S3.1 Microevolutionary quantitative genetic model

S3.1.1 Selection

We modeled the short term change in component trait means at time t in species j using the multivariate Lande equation (Lande, 1979):

$$\Delta\vec{\bar{\mathbf{x}}}_{j,t} = \mathbf{G}_{j,t}\nabla \log \overline{W}_{j,t}. \tag{S9}$$

In this equation, $\Delta\vec{\bar{\mathbf{x}}}_{j,t}$ is the vector of changes in the component trait means, $\left[\Delta\bar{d}_{\mathrm{S},j,t}, \Delta\bar{a}_{\mathrm{S},j,t}\right]$, the log-transformed stomatal density and size. The additive genetic (co)variance matrix between size and density is $\mathbf{G}_{j,t}$. The elements of $\mathbf{G}_{j,t}$ are:

$$\mathbf{G}_{j,t}[1,1] = G_{d_{\mathrm{S}},j,t}, \tag{S10}$$

$$\mathbf{G}_{j,t}[1,2] = \mathbf{G}_{j,t}[2,1] = G_{d_{\mathrm{S}},a_{\mathrm{S}},j,t}, \text{ and} \tag{S11}$$

$$\mathbf{G}_{j,t}[2,2] = G_{a_{\mathrm{S}},j,t}. \tag{S12}$$

These are the genetic variances in stomatal density (Eqn S10) and size (Eqn S12) and their genetic covariance (Eqn S11) in species j at time t.

The final term in Eqn S9, $\nabla \log \overline{W}_{j,t}$ is the gradient of the log mean relative fitness with respect to the component traits:

$$\nabla \log \overline{W}_{j,t} = \begin{pmatrix} \frac{\partial \log \overline{W}_{j,t}}{\partial \bar{d}_{\mathrm{S},j,t}} \\ \frac{\partial \log \overline{W}_{j,t}}{\partial \bar{a}_{\mathrm{S},j,t}} \end{pmatrix} \tag{S13}$$

The short term change in the genetic covariance between component traits at time t in species j is given by another multivariate Lande equation (Lande, 1980):

$$\Delta \mathbf{G}_{j,t} = \mathbf{G}_{j,t}\mathbf{H}\mathbf{G}_{j,t}. \tag{S14}$$

The matrix $\mathbf{H}$ is the curvature and orientation of the log-fitness landscape with respect to the component traits:

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 \log \overline{W}}{\partial \bar{d}_{\mathrm{S}}^2} & \frac{\partial^2 \log \overline{W}}{\partial \bar{d}_{\mathrm{S}} \partial \bar{a}_{\mathrm{S}}} \\ \frac{\partial^2 \log \overline{W}}{\partial \bar{d}_{\mathrm{S}} \partial \bar{a}_{\mathrm{S}}} & \frac{\partial^2 \log \overline{W}}{\partial \bar{a}_{\mathrm{S}}^2} \end{pmatrix}. \tag{S15}$$

Individual fitness W is a function of component and composite traits. As described below, the relationship between traits and fitness differs among the competing hypotheses.

S3.1.2 Mutation

Mutation occurs after selection. We assume that mutation is bivariate normal ($\mathcal{N}_2$) with a mean vector $\vec{0}$ and mutational covariance matrix $\mathbf{M}$. We assume that mutation is constant through time and the same in all species. $\mathbf{M}$ is a 2×2 covariance matrix with the diagonals equal to the mutational variance of stomatal size and density and the covariance on the off-diagonals:

$$\mathbf{M} = \begin{bmatrix} M_{d_{\mathrm{S}}} & M_{d_{\mathrm{S}},a_{\mathrm{S}}} \\ M_{d_{\mathrm{S}},a_{\mathrm{S}}} & M_{a_{\mathrm{S}}} \end{bmatrix}. \tag{S16}$$

Let $\rho_{\mathrm{M}} = M_{d_{\mathrm{S}},a_{\mathrm{S}}} / \sqrt{M_{d_{\mathrm{S}}} M_{a_{\mathrm{S}}}}$ be the mutational correlation between in stomatal size and density.

S3.1.3 Within species equilibrium

In each species, we analyzed the trait means and genetic variance to solve for their equilibria at mutation-selection balance. Strictly speaking, there is no equilibrium for trait means because the fitness optima randomly fluctuates through time. The process determining these fluctuations are described below in the section on macroevolutionary dynamics. However, we derive approximate equilibria by assuming each species has sufficient genetic variation to track the changing optimum. The genetic covariance $\mathbf{G}$ may reach an equilibrium distribution where the effect of selection and mutation balance. We assume sufficiently large N to ignore the effect of genetic drift.

S3.1.3.1 Trait means, $\vec{\bar{\mathbf{x}}}^*$

Let $\vec{\bar{\mathbf{x}}}^*$ be the vector of mean trait values ($\bar{d}_S^*$ and $\bar{a}_S^*$) at equilibrium:

$$0 = \mathbf{G}^* \nabla \log \overline{W}^*. \tag{S17}$$

In this equation $\mathbf{G}^*$ is the equilibrium genetic covariance matrix (described in the next subsection) and $\nabla \log \overline{W}^*$ is Eqn S13 evaluated at $\vec{\bar{\mathbf{x}}}^*$. Note that for clarity we have omitted the subscript for each species.

S3.1.3.2 Genetic covariance, $\mathbf{G}^*$

Let $\mathbf{G}^*$ be the additive genetic covariance matrix at equilibrium, such that:

$$\mathbf{0} = \mathbf{G}^* \mathbf{H} \mathbf{G}^* + \mathbf{M} \tag{S18}$$

In this equation, the reduction in genetic variance due to selection (Eqn S14) is balanced by the increase due to mutation (Eqn S16). We solved for mutation-balance numerically and confirmed solutions through recursions.

S3.2 Macroevolutionary peak-controller model

Once fitness is defined as a function of stomatal traits, the above EQG model describes how traits evolve within a species. We need a second model layered on top of that to describe how the fitness peak itself changes through time in each species. We refer to this as a "macroevolutionary adaptive landscape" *sensu* Pennell & Jiang (2024) to emphasize that the fitness optimum varies among species and through time (Arnold (2023) uses the term "dynamic adaptive landscape"). We assume that all species experience the same macroevolutionary adaptive landscape defined by a univariate or multivariate Ornstein-Uhlenbeck (OU) peak-controller process (Arnold, 2023). In other words, the fitness optimum fluctuates following an OU process. We assume that the OU process is homogeneous through time in all lineages, meaning parameters are constant. The macroevolutionary OU process has been described extensively elsewhere (Hansen, 1997; Hansen *et al.*, 2008; Hansen & Bartoszek, 2012; Pennell & Harmon, 2013). The process is defined by a long-run average trait value, or "primary adaptive optimum" *sensu* Hansen (1997), a rate of random movement away from the long-run average, and a strength of pull back toward the long-run average. We describe the univariate or multivariate parameterization of the OU process for specific competing hypotheses below. Although simplistic, compared to other peak-controller processes, the OU most adequately describes the dynamics of many trait optima (Pennell *et al.*, 2015).

S3.2.1 Among species stationarity

Analogous to the equilibrium genetic variance at mutation-selection balance within species, the macroevolutionary adaptive landscape may reach a stationary distribution of trait means among species. Let the among species vector of trait means be $\vec{\mu}_t$ at time t. The elements of this vector are μ_{d_S} (the mean of $\bar{d}^*_S$ among species) and μ_{a_S} (the mean of $\bar{a}^*_S$ among species). Let the covariance matrix between density and size be $\mathbf{V}_t$ at time t. The elements of $\mathbf{V}_t$ are:

$$\mathbf{V}_t[1,1] = V_{d_S,t}, \tag{S19}$$

$$\mathbf{V}_t[1,2] = \mathbf{V}_t[2,1] = V_{d_S,a_S,t}, \text{ and} \tag{S20}$$

$$\mathbf{V}_t[2,2] = V_{a_S,t}. \tag{S21}$$

These are the among species variances in stomatal density (Eqn S19) and size (Eqn S21) and their covariance (Eqn S20) at time t. A stationary distribution occurs when the mean trait values $\vec{\mu}_t$ and the among species covariance matrix $\mathbf{V}_t$ do not change through time due to selection. The effect of mutation is incorporated through its effect on $\mathbf{G}^*$. We ignore random fluctuations caused by genetic drift. For each competing hypothesis, we solved for the stationary trait means $\vec{\mu}^*$ and covariance $\mathbf{V}^*$ analytically or numerically and confirmed solutions through recursions.

S3.2.2 Predicted SMA and OLS scaling exponents

Rearranging Eqn S8 demonstrates why linear regression might seem like a method to estimate the scaling exponent β by estimating the slope of stomatal density on size:

$$a_S = \frac{z - z_0 - d_S}{\beta}, \tag{S22}$$

or stomatal size on density:

$$d_S = z - z_0 - \beta a_S. \tag{S23}$$

Using Eqn S22, the estimated scaling exponent would be $\hat{\beta} = -\text{slope}^{-1}$; using Eqn S23, the estimated scaling exponent would be $\hat{\beta} = -\text{slope}$. Since there is no cause-and-effect relationship stomatal traits, the choice of which variable to regress on the other should be arbitrary and result in the same estimate. However, using regression to estimate β implicitly assumes that z is a constant. Regression will not reliably estimate β when z is a variable, rather than a constant, that covaries with its constituent traits. Nonetheless, we can use the stationary covariance matrix $\mathbf{V}^*$ to predict the scaling exponents we would estimate with either

standardized major axis (SMA) or ordinary least squares (OLS) regression. The expected value of the estimated slopes treating d_{S} as the explanatory (aka independent) variable would be:

$$\mathrm{slope}_{\mathrm{sma},d_{\mathrm{S}}} = \mathrm{sign}\left(V^*_{d_{\mathrm{S}},a_{\mathrm{S}}}\right)\sqrt{\frac{V^*_{a_{\mathrm{S}}}}{V^*_{d_{\mathrm{S}}}}},\ \text{and} \tag{S24}$$

$$\mathrm{slope}_{\mathrm{ols},d_{\mathrm{S}}} = \frac{V^*_{d_{\mathrm{S}},a_{\mathrm{S}}}}{V^*_{d_{\mathrm{S}}}}. \tag{S25}$$

The expected value of the estimated slopes treating a_{S} as the explanatory (aka independent) variable would be:

$$\mathrm{slope}_{\mathrm{sma},a_{\mathrm{S}}} = \mathrm{sign}\left(V^*_{d_{\mathrm{S}},a_{\mathrm{S}}}\right)\sqrt{\frac{V^*_{d_{\mathrm{S}}}}{V^*_{a_{\mathrm{S}}}}},\ \text{and} \tag{S26}$$

$$\mathrm{slope}_{\mathrm{ols},a_{\mathrm{S}}} = \frac{V^*_{d_{\mathrm{S}},a_{\mathrm{S}}}}{V^*_{a_{\mathrm{S}}}}. \tag{S27}$$

With SMA, the estimated β is unaffected by the choice of which variable is treated as the explanatory variable because $\mathrm{slope}^{-1}_{\mathrm{sma},d_{\mathrm{S}}} = \mathrm{slope}_{\mathrm{sma},a_{\mathrm{S}}}$, but this is not necessarily true with OLS. Note that for predicting the estimate we can use the standard nonphylogenetic equations because we are treating all the species in the model as independent lineages. In practice one needs to account for phylogenetic nonindependence to obtain the most reliable estimates using either regression approach.

S3.3 Competing hypotheses

We considered three competing hypotheses for selection in stomatal size and density described in the main text:

- H_1: Stomatal-area adaptation hypothesis
- H_2: Stomatal-area minimization hypothesis
- H_3: stomatal adaptation + bounded size hypothesis

Note that these hypotheses are nested in that H_3 add assumptions to H_2, which adds assumptions to H_1. For each hypothesis, we define a fitness function and a peak controller process. Then we derive the equilibrium trait means ($\vec{\mathbf{x}}^*$) and genetic covariance matrix ($\mathbf{G}^*$) within species and the stationary distribution of traits means among species ($\mathbf{V}^*$). It is principally the among species covariance structure that informs our understanding of what processes cause inverse size-density scaling in forest plants globally.

S3.3.1 H_1: Stomatal-area adaptation hypothesis

S3.3.1.1 Overview

In this hypothesis, we assume univariate stabilizing selection on a fluctuating optimal $g_{\mathrm{s,max}}$. There is no fitness cost of increasing f_{S} or bounds on stomatal size. The optimal $g_{\mathrm{s,max}}$ fluctuates according to a univariate OU process.

S3.3.1.2 Fitness function

As described in the main text, we assume an individual fitness Gaussian fitness function with stabilizing selection around a fluctuating optimal $g_{\mathrm{s,max}}$, which we denote as $g_{\mathrm{s,opt}}$. The fitness function is:

$$W = W_{\max} e^{-\frac{(\log(g_{\mathrm{s,max}}) - \log(g_{\mathrm{s,opt}}))^2}{2\omega}}, \tag{S28}$$

where W is absolute fitness, $W_{\max}$ is the maximum absolute fitness when $g_{\mathrm{s,max}} = g_{\mathrm{s,opt}}$, and ω is the strength of stabilizing selection around $g_{\mathrm{s,opt}}$. Higher values of ω indicate weaker stabilizing selection. An individual's $g_{\mathrm{s,max}}$ is determined by its stomatal density and size following Eqn. 2. However, there is no direct selection on size or density in this model. For simplicity, we assume that $W_{\max} = 1$ for all time and species. This has no affect on the result since we assume soft selection and hence, only relative fitness determines the response to selection. For clarity, we do not show implied subscripts denoting each individual trait and fitness value or subscripts for different species and times.

We assume that species are generally close to their adaptive optima and that selection is weak. With these assumptions (Lande, 1979) showed that an approximation of the log mean-fitness can be derived as:

$$\log \overline{W} \approx -\frac{(\log(\bar{g}_{\mathrm{s,max}}) - \log(g_{\mathrm{s,opt}}))^2}{2\omega}. \tag{S29}$$

In this equation, $\log(\bar{g}_{\mathrm{s,max}})$ is the mean $\log(g_{\mathrm{s,max}})$ in the species, derived from $\vec{\vec{\mathbf{x}}}$.

S3.3.1.3 Peak controller process

Within each species, the $\log(g_{\mathrm{s,opt}})$ changes through time, independently in each species, following a univariate OU process:

$$\log(g_{\mathrm{s,opt}})[t+1] = \alpha(\log(g_{\mathrm{s,opt}})[t] - \theta) + \sigma W t, \tag{S30}$$

where θ is the long-run average $\log(g_{s,opt})$, σ is the rate of random movement in the optimum each generation, and α is the rate of return to θ. The W indicates this is a Wiener process. Note that in this model, there are no independent adaptive optima for stomatal size or density.

S3.3.1.4 Within species trait means at equilibrium $\vec{\bar{\mathbf{x}}}^*$

In each species, the trait mean evolves according to Eqn S9 where:

$$\nabla \log \overline{W} = \begin{pmatrix} -\frac{z_0+\bar{d}_S+\beta\bar{a}_S-\log(g_{s,opt})}{\omega} \\ -\beta\frac{z_0+\bar{d}_S+\beta\bar{a}_S-\log(g_{s,opt})}{\omega} \end{pmatrix}. \tag{S31}$$

We derived Eqn S31 by applying Eqn S29 to Eqn S13. There is a neutral plane of equilibria wherever stomatal density and size combine to equal $g_{s,opt}$:

$$\log(g_{s,opt}) = z_0 + \bar{d}_S^* + \beta\bar{a}_S^*. \tag{S32}$$

This equilibrium is fragile because any perturbation causes the system to move to a new equilibrium along the neutral plane.

S3.3.1.5 Equilibrium genetic covariance matrix $\mathbf{G}^*$

In this model, the genetic variance in stomatal density and size never reaches an equilibrium, but rather grows to to ∞ as long as long as the mutational variance is positive. Using the fitness function in Eqn S29, we determined the curvature and orientation of the fitness surface (Eqn S15) to be:

$$\mathbf{H} = \begin{pmatrix} \frac{-1}{\omega} & \frac{-\beta}{\omega} \\ \frac{-\beta}{\omega} & \frac{-\beta^2}{\omega} \end{pmatrix}. \tag{S33}$$

With Eqn S33, there is no solution to Eqn S18 where $M_{a_S} > 0$ and $M_{d_S} > 0$. A formal proof is beyond the scope of our study, but recursions (not shown) revealed that the genetic variance in stomatal density and size increases without bound in this model. However, the structure of $\mathbf{G}^*$ converges onto the form:

$$\mathbf{G}^* = \lambda \begin{pmatrix} 1 & -\beta \\ -\beta & -\beta^2 \end{pmatrix}, \tag{S34}$$

where λ is a scalar.

S3.3.1.6 Stationary distribution ($\vec{\mu}^*$ and $\mathbf{V}^*$)

The stationary distribution of $\log(g_{\mathrm{s,max}})$ is determined by the parameters of the OU peak controller process (Section S3.3.1.3). At stationarity, the among species mean will be θ and the variance will be $\sigma^2/2\alpha$. However, there is no stationary distribution for stomatal density and size. Since the genetic variance within species increases without bound, it follows that there no stationary distribution of traits among species. Since infinite variances are impossible, we asked what the stationary distribution would be if there was a finite G-matrix.

To find the stationary distribution, we treat $\vec{\bar{\mathbf{x}}}_t$ as a vector of random variables with mean vector $\vec{\mu}_t$ and covariance $\mathbf{V}_t$. By treating the trait means as random variables, $\Delta\vec{\bar{\mathbf{x}}}$ in Eqn S9 becomes a function of random variables with $\nabla \log \overline{W}$ given by Eqn S31. We applied rules from random variable algebra, specifically the linearity of expectations and bilinearity of covariance, to derive the vector of means and covariance matrix of $\Delta\vec{\bar{\mathbf{x}}}$. We also used the fact that at stationarity, there is no covariance between change in $\log(g_{\mathrm{s,opt}})$ and change in stomatal traits, though there can be transitory nonindependence (results not shown). For the interspecific trait means, we obtain:

$$\begin{aligned}\mu_{d_{\mathrm{S}},t+1} = \; & \mu_{d_{\mathrm{S}},t} - \\ & \beta\mathbf{G}[1,2]\frac{z_0 + \mu_{d_{\mathrm{S}},t} + \beta\mu_{a_{\mathrm{S}},t} - \log(\bar{g}_{\mathrm{s,opt}})}{\omega} - \\ & \mathbf{G}[1,1]\frac{z_0 + \mu_{d_{\mathrm{S}},t} + \beta\mu_{a_{\mathrm{S}},t} - \log(\bar{g}_{\mathrm{s,opt}})}{\omega}, \text{ and}\end{aligned} \tag{S35}$$

$$\begin{aligned}\mu_{a_{\mathrm{S}},t+1} = \; & \mu_{a_{\mathrm{S}},t} - \\ & \beta\mathbf{G}[2,2]\frac{z_0 + \mu_{d_{\mathrm{S}},t} + \beta\mu_{a_{\mathrm{S}},t} - \log(\bar{g}_{\mathrm{s,opt}})}{\omega} - \\ & \mathbf{G}[1,1]\frac{z_0 + \mu_{d_{\mathrm{S}},t} + \beta\mu_{a_{\mathrm{S}},t} - \log(\bar{g}_{\mathrm{s,opt}})}{\omega}.\end{aligned} \tag{S36}$$

Analogous to the result for within species equilibria, Eqns S35 and S36 show there a neutral plane of solutions wherever:

$$z_0 + \mu^*_{d_{\mathrm{S}}} + \beta\mu^*_{a_{\mathrm{S}}} = \log(\bar{g}_{\mathrm{s,opt}}) = \theta. \tag{S37}$$

In other words, any combination of stomatal density and size that equals θ is an equilibrium because $\log\left(\bar{g}_{\mathrm{s,opt}}\right)$ converges to θ in the long run.

To derive the stationary among species trait covariance $\mathbf{V}^*$, we asked what the among species covariance matrix would be after one generation of selection in a large ensemble of species. To

do this, we again applied the rules of random variable algebra. Then we numerically solved for when $\mathbf{V}_{t+1} = \mathbf{V}_t$. The among-species variance in stomatal density at time $t+1$ is:

$$\begin{aligned} V_{d_{\mathrm{S}},t+1} = &\ a_1^2 V_{d_{\mathrm{S}},t} + a_2^2 V_{a_{\mathrm{S}},t} + a_3^2 \mathrm{Var}\left(\log\left(g_{\mathrm{s,opt}}\right)\right) + 2a_1 a_2 V_{d_{\mathrm{S}},a_{\mathrm{S}},t}, \text{ where} \\ a_1 = &\ 1 - \frac{\beta \mathbf{G}[1,2] + \mathbf{G}[1,1]}{\omega}, \\ a_2 = &\ -\frac{\beta^2 \mathbf{G}[1,2] + \beta \mathbf{G}[1,1]}{\omega}, \text{ and} \\ a_3 = &\ \frac{\beta \mathbf{G}[1,2] + \mathbf{G}[1,1]}{\omega}. \end{aligned} \tag{S38}$$

The among-species variance in stomatal size at time $t+1$ is:

$$\begin{aligned} V_{a_{\mathrm{S}},t+1} = &\ a_1^2 V_{d_{\mathrm{S}},t} + a_2^2 V_{a_{\mathrm{S}},t} + a_3^2 \mathrm{Var}\left(\log(g_{\mathrm{s,opt}})\right) + 2a_1 a_2 V_{d_{\mathrm{S}},a_{\mathrm{S}},t}, \text{ where} \\ a_1 = &\ -\frac{\beta \mathbf{G}[2,2] + \mathbf{G}[1,2]}{\omega}, \\ a_2 = &\ 1 - \frac{\beta^2 \mathbf{G}[2,2] + \beta \mathbf{G}[1,2]}{\omega}, \text{ and} \\ a_3 = &\ \frac{\beta \mathbf{G}[2,2] + \mathbf{G}[1,2]}{\omega}. \end{aligned} \tag{S39}$$

The among-species covariance between density and size at time $t+1$ is:

$$\begin{aligned} V_{d_{\mathrm{S}},a_{\mathrm{S}},t+1} = &\ a_{11} a_{12} V_{d_{\mathrm{S}},t} + (a_{11} a_{22} + a_{21} a_{12}) V_{d_{\mathrm{S}},a_{\mathrm{S}},t} + a_{21} a_{22} V_{a_{\mathrm{S}},t} + a_{31} a_{32} \mathrm{Var}\left(\log(g_{\mathrm{s,opt}})\right), \text{ where} \\ a_{11} = &\ 1 - \frac{\beta \mathbf{G}[1,2] + \mathbf{G}[1,1]}{\omega}, \\ a_{21} = &\ -\frac{\beta^2 \mathbf{G}[1,2] + \beta \mathbf{G}[1,1]}{\omega}, \\ a_{31} = &\ \frac{\beta \mathbf{G}[1,2] + \mathbf{G}[1,1]}{\omega}, \\ a_{12} = &\ -\frac{\beta \mathbf{G}[2,2] + \mathbf{G}[1,2]}{\omega}, \\ a_{22} = &\ 1 - \frac{\beta^2 \mathbf{G}[2,2] + \beta \mathbf{G}[1,2]}{\omega}, \text{ and} \\ a_{32} = &\ \frac{\beta \mathbf{G}[2,2] + \mathbf{G}[1,2]}{\omega}. \end{aligned} \tag{S40}$$

The variance in $\log(g_{\mathrm{s,opt}})$ follows the univariate OU process and at stationarity will be:

$$\mathrm{Var}\left(\log\left(g_{\mathrm{s,opt}}\right)\right) = \frac{\sigma^2}{2\alpha} \tag{S41}$$

We numerically solved for $\mathbf{V}^*$ by iterating Eqn S38, Eqn S39, and Eqn S40 until $\mathbf{V}_{t+1} = \mathbf{V}_t$. All solutions converged on a neutral plane where:

$$\mathbf{V}^* = \begin{pmatrix} V^*_{d_{\mathrm{S}}} & -V^*_{d_{\mathrm{S}}}/\beta \\ -V^*_{d_{\mathrm{S}}}/\beta & V^*_{d_{\mathrm{S}}}/\beta^2 \end{pmatrix}. \tag{S42}$$

In other words, there is no single stationary distribution, but any stationary distribution will have a fixed covariance structure proportional to the overall trait variance.

S3.3.1.7 Predicted scaling exponent estimation

We derived the predicted scaling exponents for this hypothesis by substituting the elements of Eqn S42 into Eqn S24, Eqn S25, and Eqn S27:

$$\hat{\beta}_{\mathrm{sma},d_{\mathrm{S}}} = -\mathrm{slope}^{-1}_{\mathrm{sma},d_{\mathrm{S}}} = -\mathrm{sign}\left(-V^*_{d_{\mathrm{S}}}/\beta\right)\sqrt{\frac{V^*_{d_{\mathrm{S}}}}{V^*_{d_{\mathrm{S}}}/\beta^2}} = \beta \tag{S43}$$

$$\hat{\beta}_{\mathrm{ols},d_{\mathrm{S}}} = -\mathrm{slope}^{-1}_{\mathrm{ols},,d_{\mathrm{S}}} = -\frac{V_{d^*_{\mathrm{S}}}}{-V^*_{d_{\mathrm{S}}}/\beta} = \beta \tag{S44}$$

$$\hat{\beta}_{\mathrm{ols},a_{\mathrm{S}}} = -\mathrm{slope}_{\mathrm{ols},a_{\mathrm{S}}} = -\frac{V_{d^*_{\mathrm{S}}}/\beta}{-V^*_{d_{\mathrm{S}}}/\beta^2} = \beta \tag{S45}$$

Hence, this model predicts that both SMA and OLS should accurately estimate the scaling exponent to be $\hat{\beta} = 0.5$. The result is not sensitive to the method (OLS and SMA regression) or which variable is the explanatory variable.

S3.3.2 H_2: Stomatal-area minimization hypothesis

S3.3.2.1 Overview

This model adds to H_1 (Section S3.3.1) by assuming directional selection for lower f_{S}. The strength of selection on f_{S} relative to $g_{\mathrm{s,opt}}$ is determined by a new parameter, ϕ_f.

S3.3.2.2 Fitness function

With directional selection for lower f_{S} toward a minimum $f_{\mathrm{S,min}}$, the fitness function becomes:

$$W = W_{\max} e^{-\frac{(\log(f_{\mathrm{S}})-\log(f_{\mathrm{S,min}}))^2}{2\phi_f \omega} - \frac{(\log(g_{\mathrm{s,max}})-\log(g_{\mathrm{s,opt}}))^2}{2\omega}}. \tag{S46}$$

Following the same assumptions as Section S3.3.1.2, the log mean fitness is approximately:

$$\log \overline{W} \approx -\frac{(\log(\bar{f}_{\mathrm{S}}) - \log(f_{\mathrm{S,min}}))^2}{2\phi_f \omega} - \frac{(\log(\bar{g}_{\mathrm{s,max}}) - \log(g_{\mathrm{s,opt}}))^2}{2\omega}. \tag{S47}$$

S3.3.2.3 Peak controller process

The peak controller process is the same univariate OU process as in H_1 (Section S3.3.1.3).

S3.3.2.4 Within species trait means at equilibrium $\vec{\bar{\mathbf{x}}}^*$

In each species, the trait mean evolves according to Eqn S9 where:

$$\nabla \log \overline{W} = \begin{pmatrix} -\frac{\bar{d}_{\mathrm{S}} + \bar{a}_{\mathrm{S}} - \log(f_{\mathrm{S,min}})}{\phi_f \omega} - \frac{z_0 + \bar{d}_{\mathrm{S}} + \beta \bar{a}_{\mathrm{S}} - \log(g_{\mathrm{s,opt}})}{\omega} \\ -\beta \left(\frac{\bar{d}_{\mathrm{S}} + \bar{a}_{\mathrm{S}} - \log(f_{\mathrm{S,min}})}{\phi_f \omega} - \frac{z_0 + \bar{d}_{\mathrm{S}} + \beta \bar{a}_{\mathrm{S}} - \log(g_{\mathrm{s,opt}})}{\omega} \right) \end{pmatrix}. \tag{S48}$$

We derived Eqn S48 by applying Eqn S47 to Eqn S13. At equilibrium, d_{S}^* and a_{S}^* will satisfy:

$$\log(f_{\mathrm{S,min}}) = d_{\mathrm{S}}^* + a_{\mathrm{S}}^* \text{ and } \log(g_{\mathrm{s,opt}}) = z_0 + d_{\mathrm{S}}^* + \beta a_{\mathrm{S}}^*.$$

These conditions are satisfied when:

$$d_{\mathrm{S}}^* = \frac{\log(g_{\mathrm{s,opt}}) - z_0 - \log(f_{\mathrm{S,min}})}{1-\beta} \text{ and } a_{\mathrm{S}}^* = \frac{\log(f_{\mathrm{S,min}}) + z_0 - \log(g_{\mathrm{s,opt}})}{1-\beta}. \tag{S49}$$

S3.3.2.5 Equilibrium genetic covariance matrix $\mathbf{G}^*$

Using the fitness function in Eqn S47, we determined the curvature and orientation of the fitness surface (Eqn S15) to be:

$$\mathbf{H} = \begin{pmatrix} -\frac{1+\phi_f}{\phi_f \omega} & -\frac{1+\beta\phi_f}{\phi_f \omega} \\ -\frac{1+\beta\phi_f}{\phi_f \omega} & -\frac{1+\beta^2\phi_f}{\phi_f \omega} \end{pmatrix}. \tag{S50}$$

With Eqn S50, we solved for $\mathbf{G}^*$ numerically using equation Eqn S18. The range of parameter values for which we calculated $\mathbf{G}^*$ was is given in Table S2.

Table S2: Range of parameter values for numerically solving $\mathbf{G}^*$ for H_2.

Parameter	Values
M_{d_S}	$[3.77 \times 10^{-7}, 3.77 \times 10^{-5}]$
M_{a_S}	$[3.77 \times 10^{-7}, 3.77 \times 10^{-5}]$
ρ_M	$[-0.7, 0.7]$
ω	$\{0.0189, 0.0377, 0.0755\}$
ϕ_f	$[0.5, 2]$

S3.3.2.6 Stationary distribution ($\vec{\mu}^*$ and $\mathbf{V}^*$)

We took a different approach to finding the stationary distribution than in H_1 because unlike H_1, simulations indicated a finite, stable stationary distribution could be approximated by making the simplifying assumption that all species perfectly track the fluctuating $g_{s,opt}$. The ensemble of species converge on the same stationary distribution regardless of differences in genetic (co)variance. With this assumption, we can derive $\vec{\mu}^*$ and $\mathbf{V}^*$ by treating $\log(g_{s,opt})$ as a random variable with mean θ and variance $\sigma^2/2\alpha$ at stationarity (see Section S3.3.1.3 for description of macroevolutionary parameters).

The interspecific trait means $\mu^*_{d_S}$ and $\mu^*_{a_S}$ are the expected values of Eqn S49:

$$\mu^*_{d_S} = \frac{\theta - z_0 - \log(f_{S,min})}{1-\beta} \text{ and } \mu^*_{a_S} = \frac{\log(f_{S,min}) + z_0 - \theta}{1-\beta}. \tag{S51}$$

The interspecific trait variances and covariance can be derived from rules of random variable algebra as:

$$V^*_{d_S} = \frac{\sigma^2}{2\alpha\beta^2},\ V^*_{a_S} = \frac{\sigma^2}{2\alpha\beta^2},\ \text{and } V^*_{d_S,a_S} = -\frac{\sigma^2}{\alpha\beta}. \tag{S52}$$

S3.3.2.7 Predicted scaling exponent estimation

We derived the predicted scaling exponents for this hypothesis by substituting Eqn S52 into Eqn S24, Eqn S25, and Eqn S27:

$$\hat{\beta}_{\text{sma},d_S} = -\text{slope}^{-1}_{\text{sma},d_S} = -\text{sign}\left(-\frac{\sigma^2}{\alpha\beta}\right)\sqrt{\frac{\sigma^2/(2\alpha\beta^2)}{\sigma^2/(2\alpha\beta^2)}} = 1 \tag{S53}$$

$$\hat{\beta}_{\text{ols},d_S} = -\text{slope}^{-1}_{\text{ols},d_S} = -\frac{\sigma^2/(\alpha\beta)}{\sigma^2/(2\alpha\beta^2)} = 2\beta. \tag{S54}$$

$$\hat{\beta}_{\text{ols},a_S} = -\text{slope}_{\text{ols},a_S} = -\frac{\sigma^2/(\alpha\beta)}{\sigma^2/(2\alpha\beta^2)} = 2\beta. \tag{S55}$$

Note that the sign of the covariance in Eqn S53 must be negative because $\sigma > 0$, $\alpha > 0$, and $\beta > 0$ by assumption. Hence, this model predicts that both SMA and OLS should estimate the scaling exponent to be $\hat{\beta} = 1.0$ when the actual value is $\beta = 0.5$. This occurs the same when either d_S or a_S is the explanatory variable.

S3.3.3 H_3: stomatal adaptation + bounded size hypothesis

S3.3.3.1 Overview

This model builds on H_2 (Section S3.3.2) by adding stabilizing selection around a fluctuating optimal stomatal size. The strength of selection on stomatal size relative to $g_{s,\max}$ is determined by a new parameter, ϕ_a. As in H_2, directional selection for lower f_S that is inversely proportional to $\phi_f\omega$. As in both H_1 and H_2, there is stabilizing selection around a fluctuating optimal $g_{s,\max}$.

S3.3.3.2 Fitness function

With directional selection for lower f_S and stabilizing selection on optimal stomatal size $a_{S,\text{opt}}$, the fitness function becomes:

$$W = W_{\max} e^{-\frac{\log(f_S)}{2\phi_f\omega} - \frac{(a_S - a_{S,\text{opt}})^2}{2\phi_a\omega} - \frac{(\log(g_{s,\max}) - \log(g_{s,\text{opt}}))^2}{2\omega}}. \tag{S56}$$

Following the same assumptions as Section S3.3.1.2, the log mean fitness is approximately:

$$\log \overline{W} \approx -\frac{\log(f_S)}{2\phi_f\omega} - \frac{(a_S - a_{S,\text{opt}})^2}{2\phi_a\omega} - \frac{(\log(g_{s,\max}) - \log(g_{s,\text{opt}}))^2}{2\omega}. \tag{S57}$$

Note that while optimal stomatal density is not explicitly assumed, it necessarily emerges because of the mathematical relationship between density, size, and $g_{s,max}$:

$$g_{s,opt} = z_0 + d_{S,opt} + \beta a_{S,opt}. \tag{S58}$$

S3.3.3.3 Peak controller process

Selection on both stomatal size and $g_{s,max}$ requires a multivariate OU process where $\vec{\theta}$ is the vector of long-run average stomatal size and $\log(g_{s,max})$ among species, Σ is a matrix specifying random movement in trait optima, and $\mathbf{A}$ is a matrix specifying the rate of return to $\vec{\theta}$. The first element of $\vec{\theta}$, denoted θ_1, is the long-run average $a_{S,opt}$; θ_2 is the long-run average $\log(g_{s,max})$. Elements in the matrices Σ and $\mathbf{A}$ are ordered the same way. We solved the equations using numerical integration in *R*. We denote the stationary 2×2 interspecific covariance matrix of $a_{S,opt}$ and $\log(g_{s,opt})$ as Γ, where:

$$\Gamma_{11} = \mathrm{Var}(a_{S,opt}),\ \Gamma_{12} = \Gamma_{21} = \mathrm{Cov}(a_{S,opt}, \log(g_{s,opt})),\ \text{and } \Gamma_{22} = \mathrm{Var}(\log(g_{s,opt})). \tag{S59}$$

S3.3.3.4 Within species trait means at equilibrium $\vec{\mathbf{x}}^*$

In each species, the trait mean evolves according to Eqn S9 where:

$$\nabla \log \overline{W} = \begin{pmatrix} -\frac{2\phi_a\phi_f(z_0+d_S+\beta a_S-\log(g_{s,opt}))-\phi_a}{2\phi_f\omega} \\ -\frac{2\beta\phi_a\phi_f(z_0+d_S+\beta a_S-\log(g_{s,opt}))+\phi_a+2\phi_f(a_S-a_{S,opt})}{2\phi_a\phi_f\omega} \end{pmatrix}. \tag{S60}$$

We derived Eqn S60 by applying Eqn S57 to Eqn S13. At equilibrium, d_S^* and a_S^* are:

$$d_S^* = \log(g_{s,opt}) - z_0 - \beta a_{S,opt} - \frac{1-\beta\phi_a(1-\beta)}{2\phi_f} \text{ and } a_S^* = a_{S,opt} - \frac{\phi_a(1-\beta)}{2\phi_f}. \tag{S61}$$

S3.3.3.5 Equilibrium genetic covariance matrix $\mathbf{G}^*$

Using the fitness function in Eqn S57, we determined the curvature and orientation of the fitness surface (Eqn S15) to be:

$$\mathbf{H} = \begin{pmatrix} -\frac{1}{\omega} & -\frac{\beta}{\omega} \\ -\frac{\beta}{\omega} & -\frac{\beta^2\phi_a+1}{\phi_a\omega} \end{pmatrix}. \tag{S62}$$

With Eqn S62, we solved for $\mathbf{G}^*$ numerically using equation Eqn S18. The range of parameter values for which we calculated $\mathbf{G}^*$ was is given in Table Table S3.

Table S3: Range of parameter values for numerically solving $\mathbf{G}^*$ for H_3.

Parameter	Values
M_{d_S}	$[3.77 \times 10^{-7}, 3.77 \times 10^{-5}]$
M_{a_S}	$[3.77 \times 10^{-7}, 3.77 \times 10^{-5}]$
ρ_M	$[-0.7, 0.7]$
ω	$\{0.0189, 0.0377, 0.0755\}$
ϕ_a	$[0.5, 2]$

S3.3.3.6 Stationary distribution ($\vec{\mu}^*$ and $\mathbf{V}^*$)

We used the same general approach to analyzing the stationary distribution as in H_2 (Section S3.3.2.6). The main difference is that both $g_{s,opt}$ and $a_{S,opt}$ are random variables. The interspecific (co)variance of these parameters are determined by the multivariate OU process (Section S3.3.3.3).

The interspecific trait means $\mu^*_{d_S}$ and $\mu^*_{a_S}$ are the expected values of Eqn S61:

$$\mu^*_{d_S} = \theta_2 - z_0 - \beta\theta_1 - \frac{1 - \beta\phi_a(1-\beta)}{2\phi_f} \text{ and } \mu^*_{a_S} = \theta_1 - \frac{\phi_a(1-\beta)}{2\phi_f}. \tag{S63}$$

Recall that θ_1 and θ_2 are the long-run average values of $a_{S,opt}$ and $\log(g_{s,opt})$, respectively (Section S3.3.3.3).

The interspecific trait variances and covariance can be derived from rules of random variable algebra as:

$$\begin{aligned} V^*_{d_S} &= \beta^2 \text{Var}(a_{S,opt}) + \text{Var}(\log(g_{s,opt})) - 2\beta \text{Cov}(a_{S,opt}, \log(g_{s,opt})), \\ V^*_{a_S} &= \text{Var}(a_{S,opt}), \text{ and} \\ V^*_{d_S,a_S} &= \text{Cov}(a_{S,opt}, \log(g_{s,opt})) - \beta \text{Var}(a_{S,opt}). \end{aligned} \tag{S64}$$

The value of macroevolutionary parameters $\text{Var}(\log(g_{s,opt}))$, $\text{Var}(a_{S,opt})$, and $\text{Cov}(a_{S,opt}, \log(g_{s,opt}))$ are not known *a priori* and must be estimated from the elements of $\mathbf{V}^*$. However, since these macroevolutionary parameters are meant to *explain* $\mathbf{V}^*$ the apparent correspondence between macroevolutionary parameter values and $\mathbf{V}^*$ is circular. We therefore ask the less restrictive question: what do the estimated values of $\mathbf{V}^*$ imply about the macroevolutionary parameters if our model were correct?

Let $\hat{V}^*_{a_S}$, $\hat{V}^*_{d_S}$, and $\hat{V}^*_{a_S,d_S}$ denote the estimates of $V^*_{a_S}$, $V^*_{d_S}$, and $V^*_{a_S,d_S}$, respectively. The implied values of $\text{Var}(\log(g_{s,opt}))$, $\text{Var}(a_{S,opt})$, and $\text{Cov}(a_{S,opt}, \log(g_{s,opt}))$ are:

$$\begin{aligned}
\mathrm{Var}\left(\log\left(g_{\mathrm{s,opt}}\right)\right) &= \beta^2 \hat{V}^*_{a_\mathrm{S}} + \hat{V}^*_{d_\mathrm{S}} + 2\beta \hat{V}^*_{a_\mathrm{S},d_\mathrm{S}}, \\
\mathrm{Var}\left(a_{\mathrm{S,opt}}\right) &= \hat{V}^*_{a_\mathrm{S}}, \text{ and} \\
\mathrm{Cov}\left(a_{\mathrm{S,opt}}, \log\left(g_{\mathrm{s,opt}}\right)\right) &= \beta \hat{V}^*_{a_\mathrm{S}} + \hat{V}^*_{a_\mathrm{S},d_\mathrm{S}}.
\end{aligned} \tag{S65}$$

S3.3.3.7 Predicted scaling exponent estimation

We derived the predicted scaling exponents for this hypothesis by substituting Eqn S64 into Eqn S24 and Eqn S25:

$$\hat{\beta}_{\mathrm{sma},d_\mathrm{S}} = -\,\mathrm{slope}^{-1}_{\mathrm{sma},d_\mathrm{S}} = -\mathrm{sign}\left(\mathrm{Cov}\left(a_{\mathrm{S,opt}}, \log\left(g_{\mathrm{s,opt}}\right)\right) - \beta\mathrm{Var}\left(a_{\mathrm{S,opt}}\right)\right) \times \sqrt{\frac{\mathrm{Var}\left(a_{\mathrm{S,opt}}\right)}{\beta^2\mathrm{Var}\left(a_{\mathrm{S,opt}}\right) + \mathrm{Var}\left(\log\left(g_{\mathrm{s,opt}}\right)\right) - 2\beta\mathrm{Cov}\left(a_{\mathrm{S,opt}}, \log\left(g_{\mathrm{s,opt}}\right)\right)}} \tag{S66}$$

$$\hat{\beta}_{\mathrm{ols},d_\mathrm{S}} = -\mathrm{slope}^{-1}_{\mathrm{ols},d_\mathrm{S}} = -\frac{\mathrm{Cov}\left(a_{\mathrm{S,opt}}, \log\left(g_{\mathrm{s,opt}}\right)\right) - \beta\mathrm{Var}\left(a_{\mathrm{S,opt}}\right)}{\beta^2\mathrm{Var}\left(a_{\mathrm{S,opt}}\right) + \mathrm{Var}\left(\log\left(g_{\mathrm{s,opt}}\right)\right) - 2\beta\mathrm{Cov}\left(a_{\mathrm{S,opt}}, \log\left(g_{\mathrm{s,opt}}\right)\right)} \tag{S67}$$

The predicted scaling exponents for this hypothesis substituting Eqn S64 into Eqn S27:

$$\hat{\beta}_{\mathrm{ols},a_\mathrm{S}} = -\mathrm{slope}_{\mathrm{ols},a_\mathrm{S}} = -\frac{\mathrm{Cov}\left(a_{\mathrm{S,opt}}, \log\left(g_{\mathrm{s,opt}}\right)\right) - \beta\mathrm{Var}\left(a_{\mathrm{S,opt}}\right)}{\mathrm{Var}\left(a_{\mathrm{S,opt}}\right)} \tag{S68}$$

Both SMA and OLS slopes can take on a range of values. For example, if $\mathrm{Cov}\left(a_{\mathrm{S,opt}}, \log\left(g_{\mathrm{s,opt}}\right)\right) > \beta\mathrm{Var}\left(a_{\mathrm{S,opt}}\right)$, then the covariance between stomatal density and size would be *positive* rather than negative. In this domain of parameter space, there would also be very little interspecific variance in stomatal density (Eqn S64). Other than the sign of the slope, little generalization can be made from the equations alone without empirical constraints on parameter values. The main result is that we should not expect SMA and OLS to be the same or reliable estimates of β except under very special circumstances. The OLS regression will also estimate β differently depending on whether a_S or d_S is treated as the explanatory variables. These discrepancies among methods occur because the combination of selection on $g_{\mathrm{s,max}}$ and f_S with bounds on stomatal size shapes the covariance between density and size such it is not predictable from a single selective target in isolation.

Note S4 Evolutionary quantitative genetic model: Simulations

We simulated large ensembles of species under plausible areas of parameter space for H_2 and H_3 to evaluate our assumption that among-species trait covariance can be approximated using the stationary distribution of $g_{\mathrm{s,opt}}$ (H_2) or $a_{\mathrm{S,opt}}$ and $g_{\mathrm{s,opt}}$ (H_3). We first introduce baseline parameter values and then discuss parameters that we varied over select ranges.

S4.1 Baseline parameter values

We restrict our analysis to the domain where selection and mutation are the dominant evolutionary forces. We also assume that changes in fitness optima are small enough that species can generally track the fluctuating optima. A limitation of our model is that it does not account for cases where genetic drift is important and/or the rate of environmental change outstrips the rate of adaptation.

In order to simulate within a parameter space that would recapitulate realistic patterns, we estimated the stationary variance of $\log\left(g_{\mathrm{s,max}}\right)$ across all species in our data set using the *R* package **phylolm** version 2.6.5 (Ho & Ané, 2014) with the `OUrandomRoot` model. The model removed affects of major group (Angiosperm, Pteridophyte, Gymnosperm) and growth form (grass, herb, shrub, tree). We did the same for a_{S} and d_{S} separately.

The per-generation mutational variance is small (10^{-5}) relative to the stationary variance of $\log\left(g_{\mathrm{s,max}}\right)$. Since we estimated the stationary variance to be 0.377, we set $M_{d_{\mathrm{S}}} = M_{a_{\mathrm{S}}} = 3.77 \times 10^{-6}$. We assumed no mutational correlation, $\rho_{\mathrm{M}} = 0$. We set the strength of selection on $g_{\mathrm{s,max}}$ to 10^{-1} the value of the stationary variance in $g_{\mathrm{s,max}}$, $\omega = 0.0377$. With the rate of peak movement (see below), this value ensures that selection is moderate or weak. The Lande equation approximations described above are most accurate in this domain.

We used $\beta = 0.5$ for calculating $g_{\mathrm{s,max}}$ and $\beta = 1$ for calculating f_{S}. The relative strengths of selection of f_{S} and a_{S} were $\phi_f = 1$ and $\phi_a = 1$, respectively. For H_2, we set $f_{\mathrm{S,min}} = 10^{-3}$, which is far below observed values of f_{S} in our data set.

For macroevolutionary parameters, we set the long-run average values of $\log\left(g_{\mathrm{s,opt}}\right)$ and $a_{\mathrm{S,opt}}$ to the mean values of these traits estimated from the phylogenetic regression described above. The α and σ parameters were set from regression estimates as well. For $g_{\mathrm{s,max}}$, these values were: $\theta = 0.091$, $\alpha = 1.56 \times 10^{-6}$, and $\sigma = 0.00109$; for a_{S}, these values were: $\theta = 5.81$, $\alpha = 5.79 \times 10^{-7}$, and $\sigma = 6.85 \times 10^{-4}$. For H_3, we assumed the off-diagonals of the $\mathbf{A}$ and Σ matrices were 0.

S4.2 Variable parameter values

We evaluated the sensitivity of among-species trait (co)variance to the microevolutionary details by varying the mutational variance, mutational correlation, and strength of selection. The

variable parameters values are given in Table S2 and Table S3 above. We performed simulations for H_2 and H_3, but not H_1 because we derived the stationary distribution analytically.

S4.3 Simulation procedure

For simulations, we discretized the OU process in time by randomly sampling fluctuating optima each generation following a one-dimensional (H_2) or two-dimensional (H_3) OU process for an ensemble of 10^3 independently evolving species using the *Python* package **thermox** version 0.0.3 (Duffield *et al.*, 2024). We started all species at the long-run average trait values, but ran simulations for 5×10^6 generations, which was sufficient to reach stationarity, meaning the starting values did not affect the results. In each generation, we approximated the deterministic change in trait means and (co)variance using Eqn S9 and Eqn S14 with $\nabla \log\left(\overline{W}\right)$ and $\mathbf{H}$ from Eqn S48 and Eqn S50 for H_2 and Eqn S60 and Eqn S62 for H_3. We added mutational (co)variance to the additive genetic variance based on $\mathbf{M}$. To compare simulation results to theoretical expectations, we estimated the sample covariance matrix among species in the final generation and calculated 95% confidence intervals using 10^3 bootstrap replicates.

Supporting Figures

Figure S1: The fraction of epidermal area allocated to stomata (f_S) is strongly positively correlated with the maximum stomatal conductance ($g_{s,max}$) across species of Angiosperms (left), Gymnosperms (center), and Pteridophytes (right). Dark points are species' means within focal group; grey points are data from all species for reference.

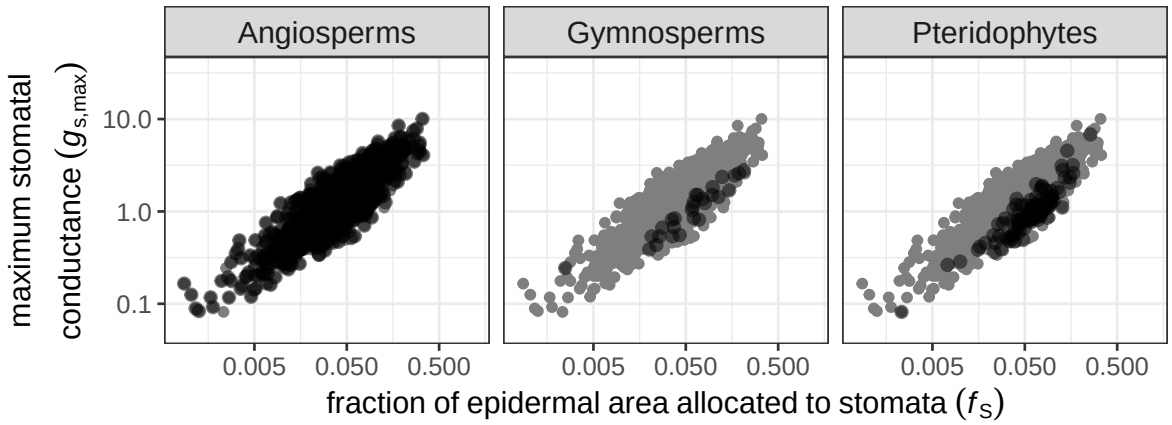


Figure S2: The effect of mutational variance and strength of selection on $g_{\mathrm{s,max}}$ on the additive genetic variance within species under H_2: stomatal-area minimization hypothesis. The additive genetic variance within species at mutation-selection balance in stomatal density (left panel) and size (right panel) depends on mutational variance ($M_{d_{\mathrm{S}}}$ and $M_{a_{\mathrm{S}}}$) and the strength of selection on $g_{\mathrm{s,max}}$ (ω). In the left panel, $M_{a_{\mathrm{S}}} = 3.77 \times 10^{-7}$ while $M_{d_{\mathrm{S}}}$ varied; in the right panel, $M_{d_{\mathrm{S}}} = 3.77 \times 10^{-7}$ while $M_{a_{\mathrm{S}}}$ varied. The strength of selection on $g_{\mathrm{s,max}}$ varied between $\omega = 7.54 \times 10^{-2}$ (weak, solid line), $\omega = 3.77 \times 10^{-2}$ (medium, dotted line), and $\omega = 1.89 \times 10^{-2}$ (strong, dashed line). We also assumed no mutational correlation and $\phi_f = 2$. See text for explanation of parameter values. Note that both stomatal density and size are expressed on a log-scale.

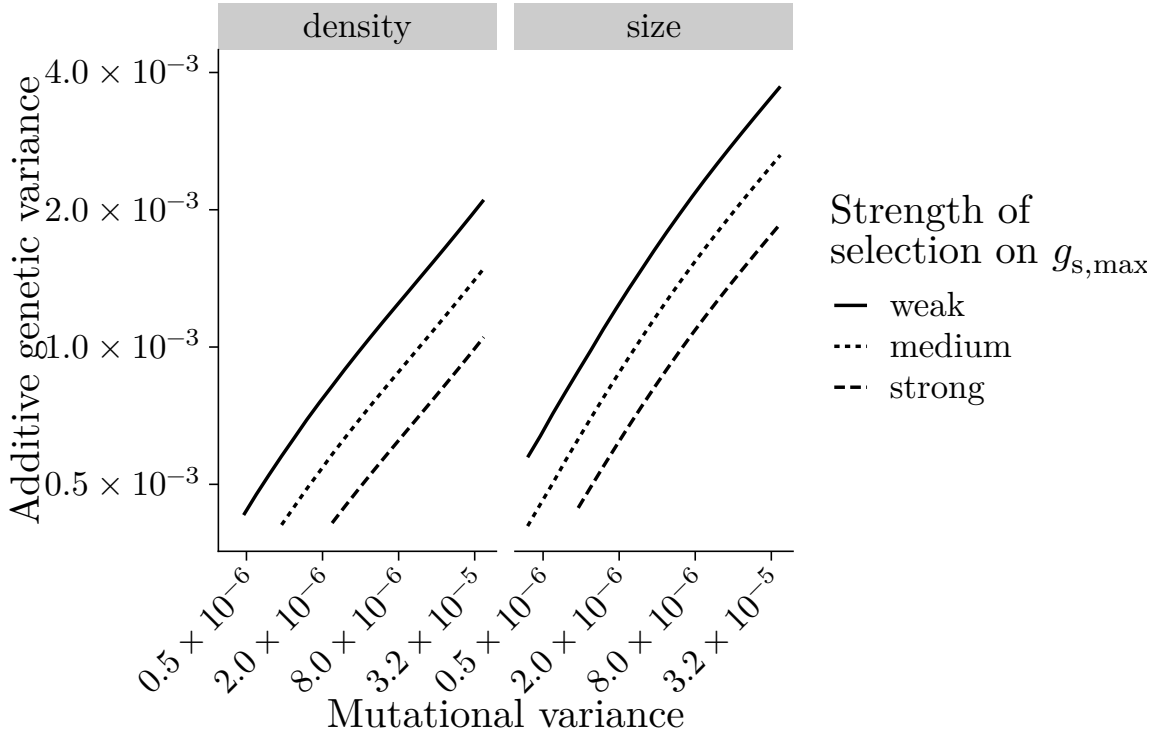


Figure S3: The effect of mutational variance and strength of selection on $g_{\mathrm{s,max}}$ on the additive genetic variance within species under H_3: stomatal adaptation + bounded size hypothesis. The additive genetic variance within species at mutation-selection balance in stomatal density (left panel) and size (right panel) depends on mutational variance ($M_{d_{\mathrm{S}}}$ and $M_{a_{\mathrm{S}}}$) and the strength of selection on $g_{\mathrm{s,max}}$ (ω). In the left panel, $M_{a_{\mathrm{S}}} = 3.77 \times 10^{-7}$ while $M_{d_{\mathrm{S}}}$ varied; in the right panel, $M_{d_{\mathrm{S}}} = 3.77 \times 10^{-7}$ while $M_{a_{\mathrm{S}}}$ varied. The strength of selection on $g_{\mathrm{s,max}}$ varied between $\omega = 7.54 \times 10^{-2}$ (weak, solid line), $\omega = 3.77 \times 10^{-2}$ (medium, dotted line), and $\omega = 1.89 \times 10^{-2}$ (strong, dashed line). We also assumed no mutational correlation and $\phi_a = 2$. See text for explanation of parameter values. Note that both stomatal density and size are expressed on a log-scale.

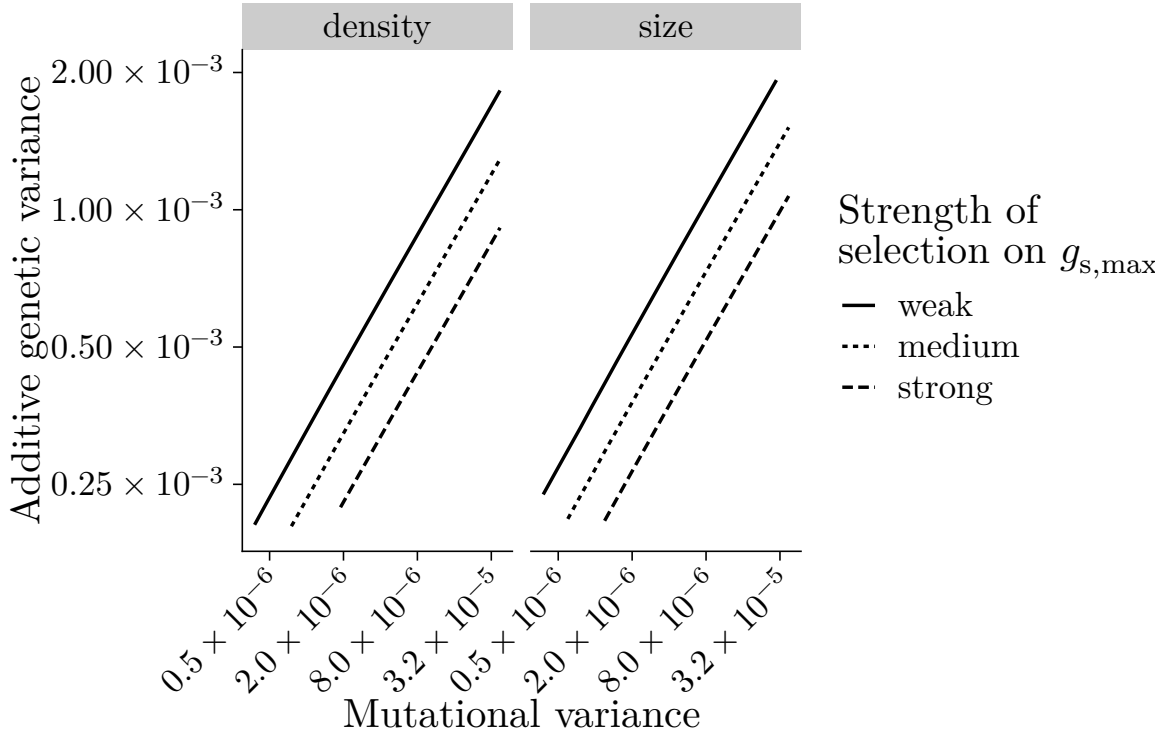


Figure S4: The effects of selection on minimizing f_{S} and mutational variance in stomatal size on the additive genetic variance under H_2: stomatal-area minimization hypothesis. The ratio of additive genetic variance within species at mutation-selection balance in stomatal density and size depends on the relative mutational variance ($M_{d_{\mathrm{S}}}/M_{a_{\mathrm{S}}}$) and the strength of selection on f_{S} ($\phi_f\omega$). Results are shown for $M_{d_{\mathrm{S}}} = 3.77 \times 10^{-7}$ while $M_{a_{\mathrm{S}}}$ varied from $0.1M_{d_{\mathrm{S}}}$ (low, solid line), $M_{d_{\mathrm{S}}}$ (medium, dotted line), and $10M_{d_{\mathrm{S}}}$ (high, dashed line). The relative strength of selection on f_{S} varied between $\phi_f = 0.5$ and 2. We also assumed no mutational correlation and $\omega = 3.77 \times 10^{-2}$. See text for explanation of parameter values.

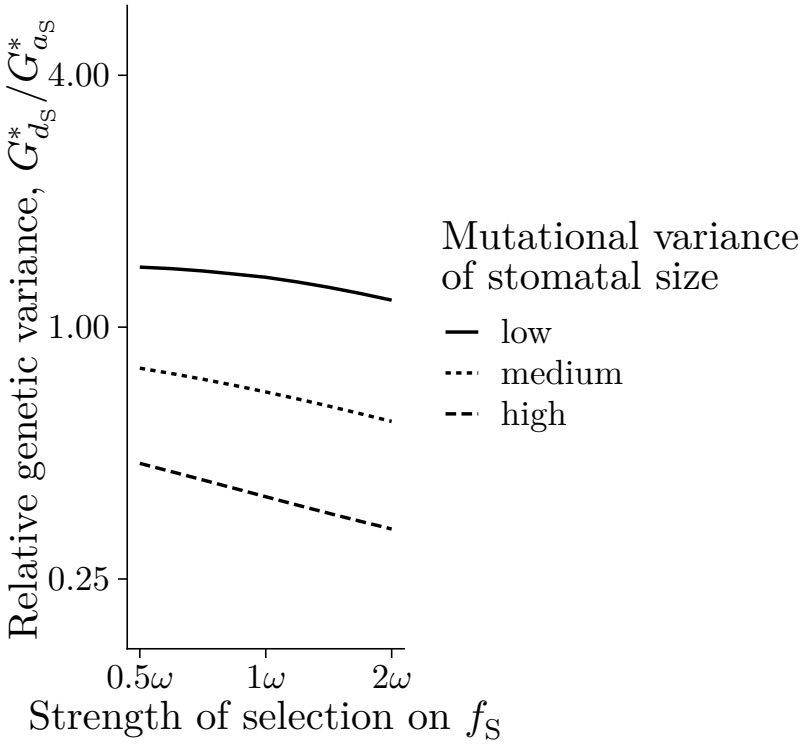


Figure S5: The effects of selection on and mutational variance in stomatal size on the additive genetic variance under H_3: stomatal adaptation + bounded size hypothesis. The ratio of additive genetic variance within species at mutation-selection balance in stomatal density and size depends on the relative mutational variance (M_{d_S}/M_{a_S}) and the strength of selection on stomatal size ($\phi_a\omega$). Results are shown for $M_{d_S} = 3.77 \times 10^{-7}$ while M_{a_S} varied from $0.1M_{d_S}$ (low, solid line), M_{d_S} (medium, dotted line), and $10M_{d_S}$ (high, dashed line). The relative strength of selection on stomatal size varied between $\phi_a = 0.5$ and 2. We also assumed no mutational correlation and $\omega = 3.77 \times 10^{-2}$. See text for explanation of parameter values.

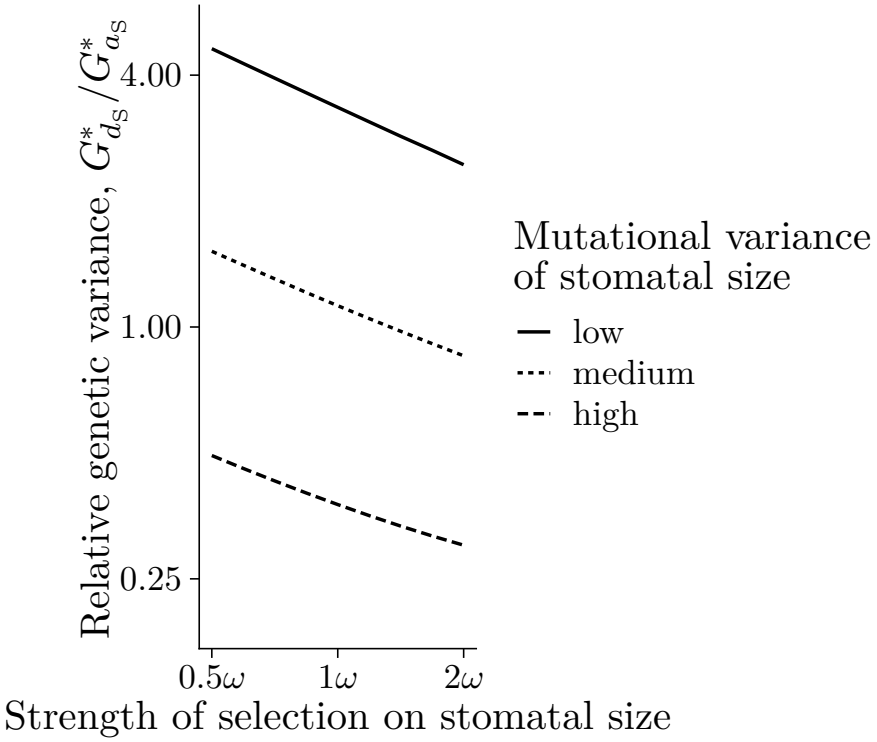


Figure S6: The effect of mutational correlation between stomatal density and size (ρ_{M}) on the additive genetic correlation under H_2: stomatal-area minimization hypothesis. The additive genetic correlation between stomatal density and size within species at mutation-selection balance depends on mutational correlation and the strength of selection on f_{S} (ϕ_f). The strength of selection on stomatal size varied between $\phi_f = 0.5$ (strong, solid line), $\phi_f = 1$ (medium, dotted line), and $\phi_f = 2$ (weak, dashed line). Other parameters were set to $M_{d_{\mathrm{S}}} = M_{a_{\mathrm{S}}} = 3.77 \times 10^{-7}$ and $\omega = 3.77 \times 10^{-2}$. See text for explanation of parameter values.

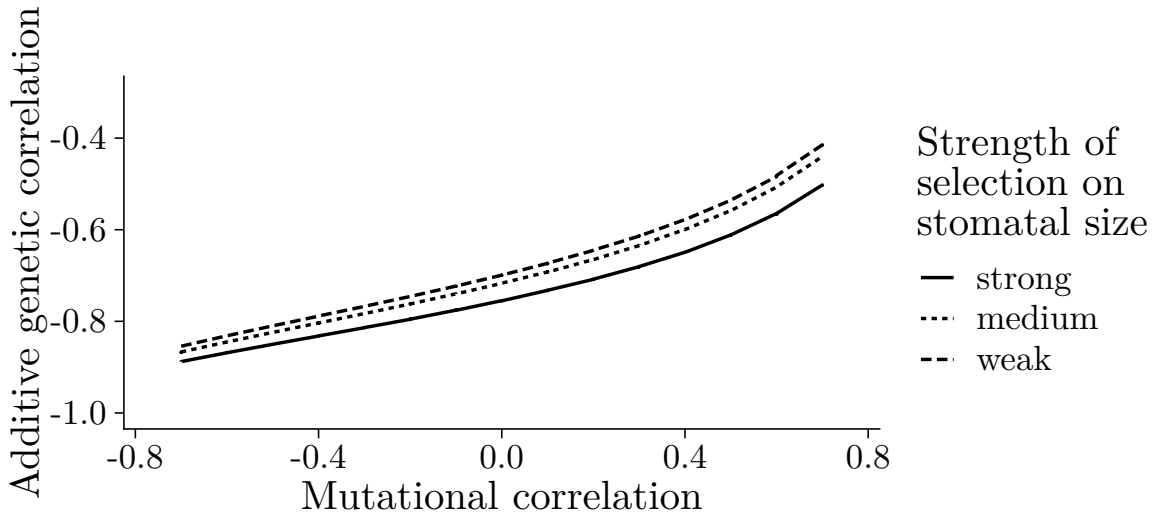


Figure S7: The effect of mutational correlation between stomatal density and size (ρ_{M}) on the additive genetic correlation under H_3: stomatal adaptation + bounded size hypothesis. The additive genetic correlation between stomatal density and size within species at mutation-selection balance depends on mutational correlation and the strength of selection on stomatal size (ϕ_a). The strength of selection on stomatal size varied between $\phi_a = 0.5$ (strong, solid line), $\phi_a = 1$ (medium, dotted line), and $\phi_a = 2$ (weak, dashed line). Other parameters were set to $M_{d_{\mathrm{S}}} = M_{a_{\mathrm{S}}} = 3.77 \times 10^{-7}$ and $\omega = 3.77 \times 10^{-2}$. See text for explanation of parameter values.

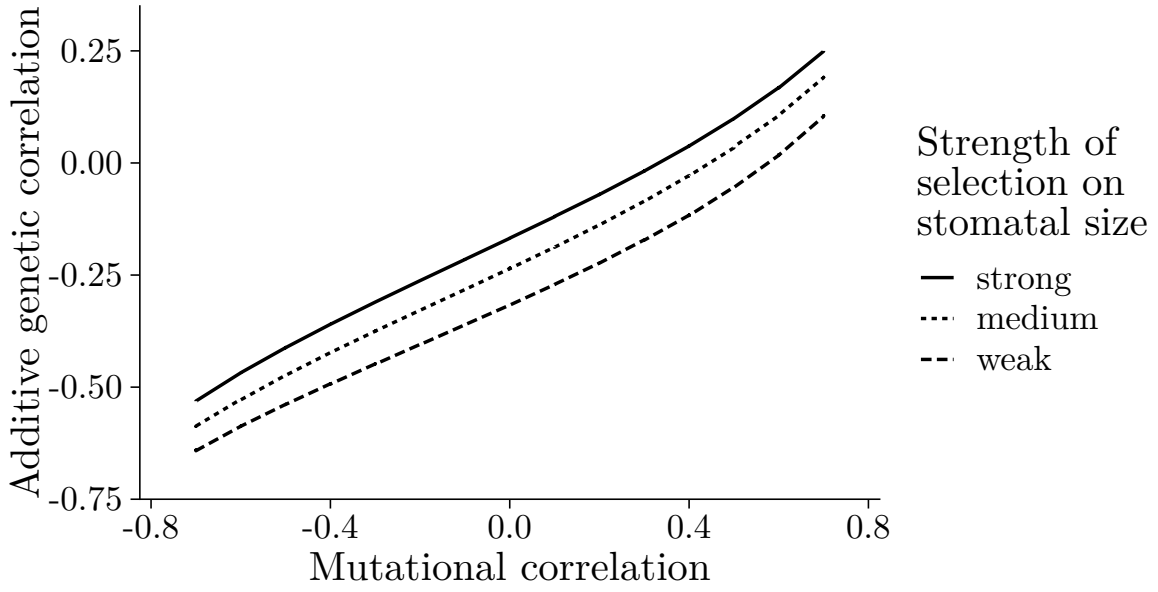


Figure S8: The predicted response to selection on greater $g_{s,max}$ under three hypotheses: H_2: stomatal-area adaptation (blue); H_1: stomatal-area minimization (green); and H_3: stomatal adaptation + bounded size (red). Each vector shows the relative response of stomatal density (log-scale, x-axis) and stomatal size (log-scale, y-axis). The predicted response assumes the additive genetic variance is at equilibrium and there is no other change in parameters from the baseline scenario described in the Supporting Information. The left and right panels show how mutational correlation between stomatal density and size affect the response vectors.

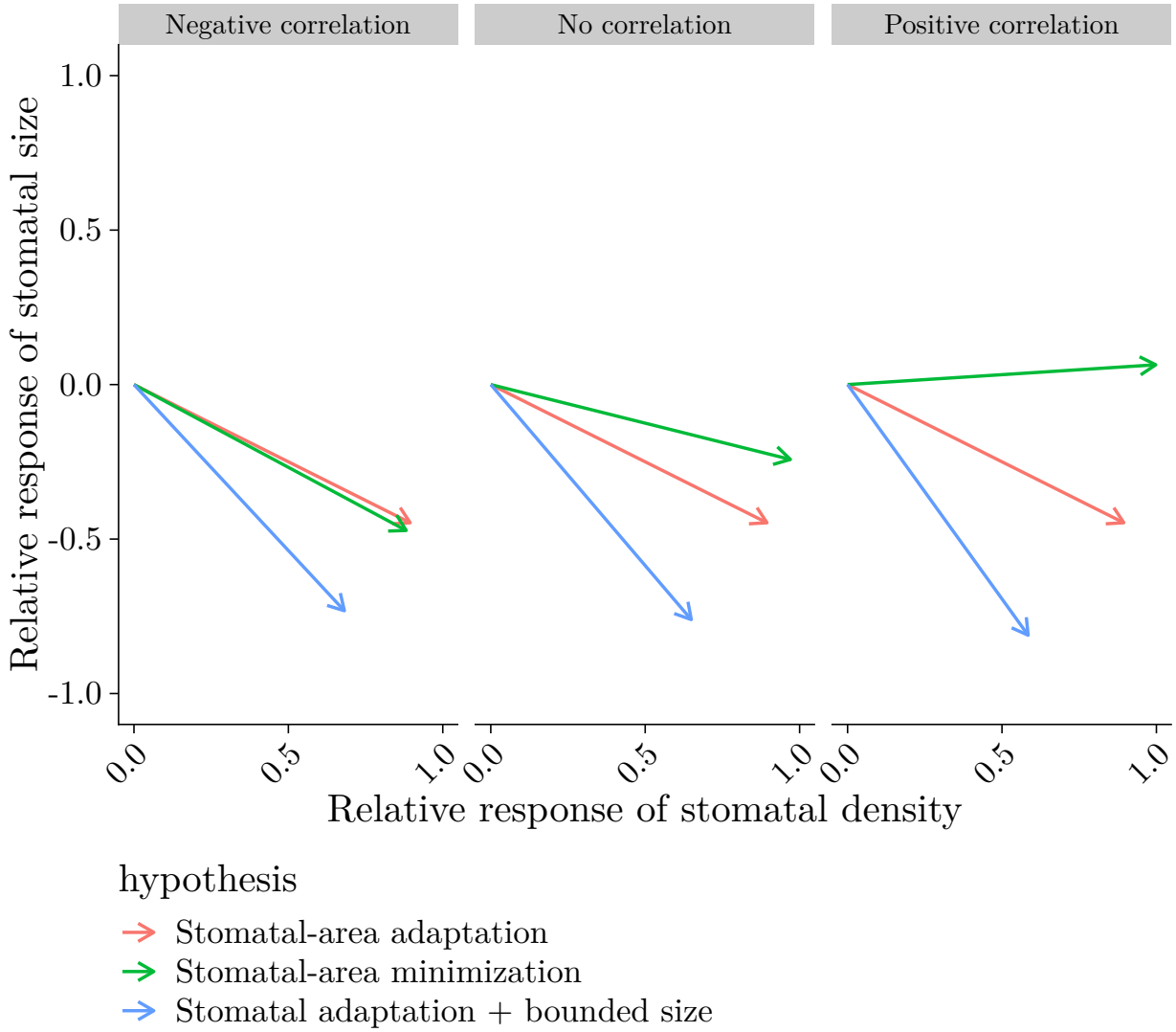


Figure S9: (Caption on next page.)

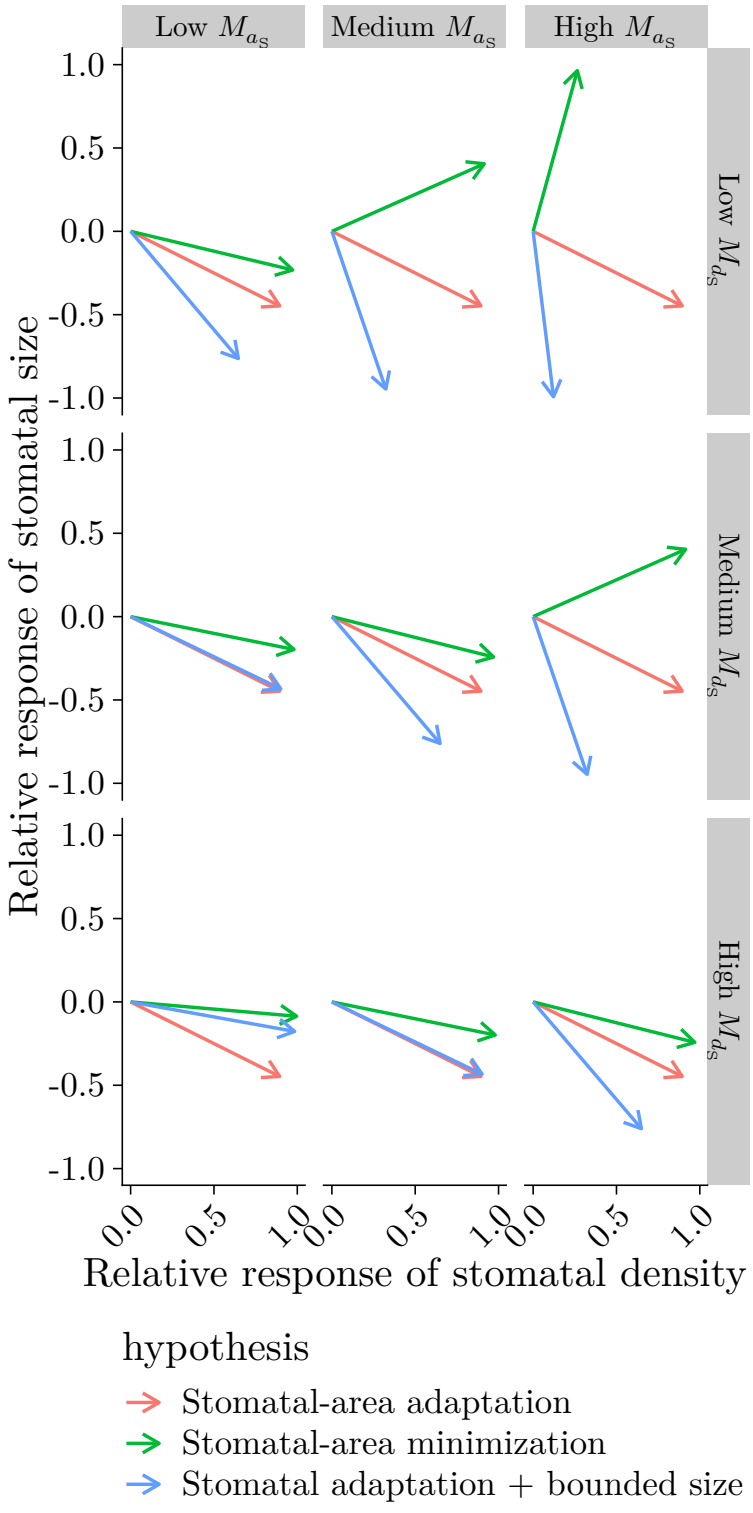


Figure S9: (previous page) The predicted response to selection on greater $g_{\mathrm{s,max}}$ under three hypotheses: H_2: stomatal-area adaptation (blue); H_1: stomatal-area minimization (green); and H_3: stomatal adaptation + bounded size (red). Each vector shows the relative response of stomatal density (log-scale, x-axis) and stomatal size (log-scale, y-axis). The predicted response assumes the additive genetic variance is at equilibrium and there is no other change in parameters from the baseline scenario described in the Supporting Information. The left and right columns of panels show how mutational variance in stomatal size ($M_{a_{\mathrm{S}}}$) affects the response vectors; the bottom and top rows of panels show how mutational variance in stomatal density ($M_{d_{\mathrm{S}}}$) affects the response vectors.

Figure S10: The predicted response to selection on greater $g_{s,max}$ under three hypotheses: H_2: stomatal-area adaptation (blue); H_1: stomatal-area minimization (green); and H_3: stomatal adaptation + bounded size (red). Each vector shows the relative response of stomatal density (log-scale, x-axis) and stomatal size (log-scale, y-axis). The predicted response assumes the additive genetic variance is at equilibrium and there is no other change in parameters from the baseline scenario described in the Supporting Information. The left and right panels show how the strength of selection to minimize f_S affects the response vectors.

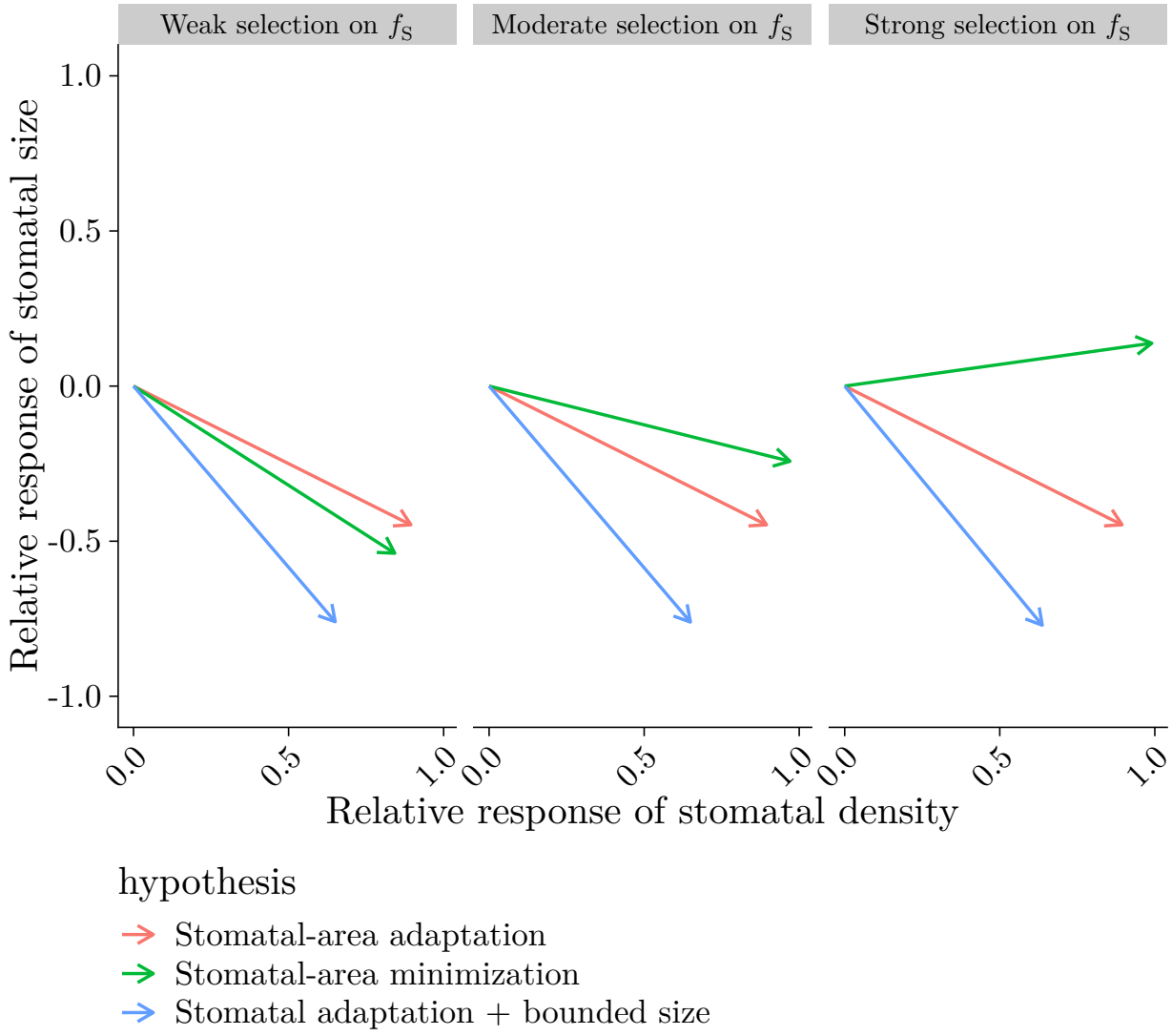


Figure S11: Among-species trait (co)variances are not sensitive to mutational ratio at stationarity. Dashed lines are the theoretical expectations for the interspecific variance in stomatal density ($V^*_{d_S}$), stomatal size ($V^*_{a_S}$), anatomical maximum stomatal conductance ($V^*_{g_{s,max}}$), and covariance between density and size ($V^*_{d_S,a_S}$) at stationarity. The theoretical expectations differ between the hypotheses, H_2: stomatal-area minimization (upper panel) and H_3: stomatal adaptation + bounded size (lower panel). The points are estimated from simulations of an ensemble of 1000 over a range of mutational ratios (M_{d_S}/M_{a_S}), where a value of 1 corresponds to the baseline scenario. The solid line behind each points indicates the 95% bootstrap confidence interval. All other parameters are set to their baseline value as described in the "Baseline parameter values" subsection.

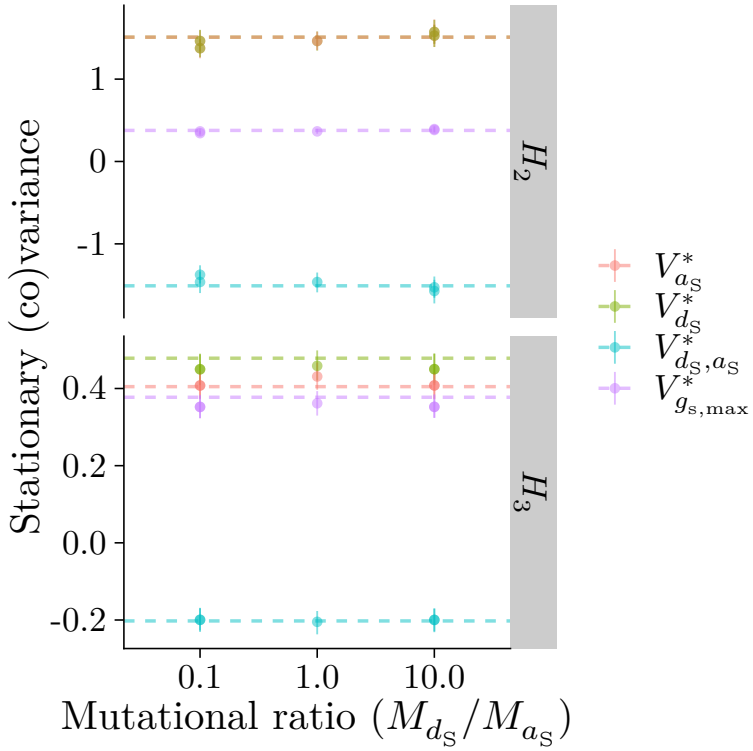


Figure S12: Among-species trait (co)variances are not sensitive to mutational correlation (ρ_{M}) at stationarity. Dashed lines are the theoretical expectations for the interspecific variance in stomatal density ($V^*_{d_{\mathrm{S}}}$), stomatal size ($V^*_{a_{\mathrm{S}}}$), anatomical maximum stomatal conductance ($V^*_{g_{\mathrm{s,max}}}$), and covariance between density and size ($V^*_{d_{\mathrm{S}},a_{\mathrm{S}}}$) at stationarity. The theoretical expectations differ between the hypotheses, H_2: stomatal-area minimization (upper panel) and H_3: stomatal adaptation + bounded size (lower panel). The points are estimated from simulations of an ensemble of 1000 over a range of mutational correlations, where a value of 0 corresponds to the baseline scenario. The solid line behind each points indicates the 95% bootstrap confidence interval. All other parameters are set to their baseline value as described in the "Baseline parameter values" subsection.

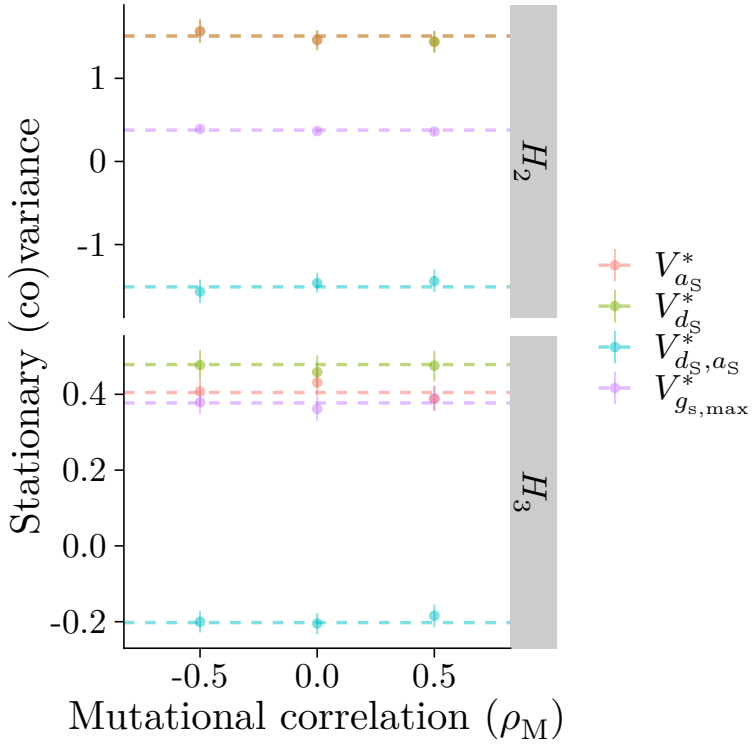


Figure S13: Among-species trait (co)variances are not sensitive to the strength of selection on $g_{\mathrm{s,max}}$ at stationarity. Dashed lines are the theoretical expectations for the interspecific variance in stomatal density ($V^*_{d_{\mathrm{S}}}$), stomatal size ($V^*_{a_{\mathrm{S}}}$), anatomical maximum stomatal conductance ($V^*_{g_{\mathrm{s,max}}}$), and covariance between density and size ($V^*_{d_{\mathrm{S}},a_{\mathrm{S}}}$) at stationarity. The theoretical expectations differ between the hypotheses, H_2: stomatal-area minimization (upper panel) and H_3: stomatal adaptation + bounded size (lower panel). The points are estimated from simulations of an ensemble of 1000 over a range of selections strengths (ω), where a value of 0.0377 corresponds to the baseline scenario. The solid line behind each points indicates the 95% bootstrap confidence interval. All other parameters are set to their baseline value as described in the "Baseline parameter values" subsection.

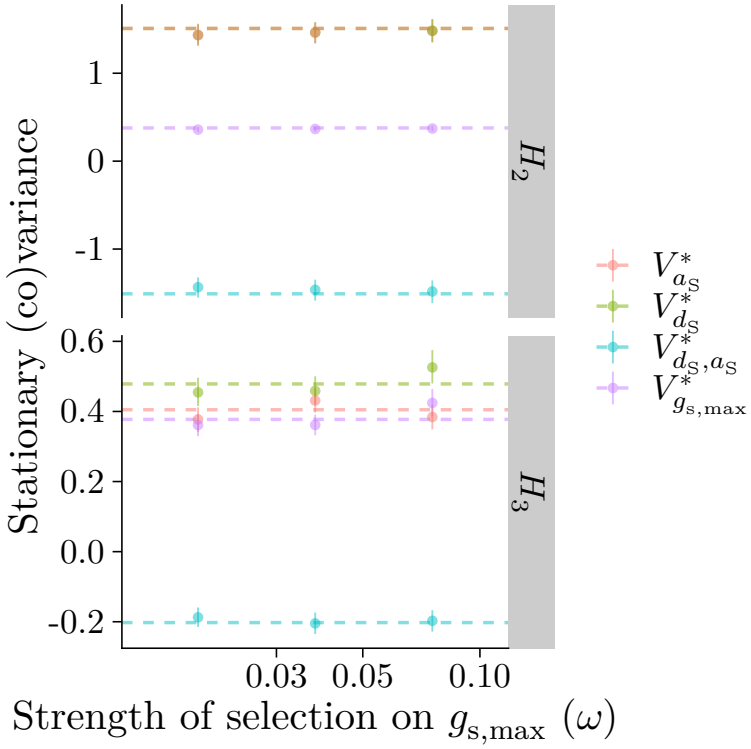


Figure S14: Among-species trait (co)variances are not sensitive to the strength of selection to minimize f_{S} (ϕ_f) at stationarity. Dashed lines are the theoretical expectations for the interspecific variance in stomatal density ($V^*_{d_{\mathrm{S}}}$), stomatal size ($V^*_{a_{\mathrm{S}}}$), anatomical maximum stomatal conductance ($V^*_{g_{\mathrm{s,max}}}$), and covariance between density and size ($V^*_{d_{\mathrm{S}},a_{\mathrm{S}}}$) at stationarity. The theoretical expectations differ between the hypotheses, H_2: stomatal-area minimization (upper panel) and H_3: stomatal adaptation + bounded size (lower panel). The points are estimated from simulations of an ensemble of 1000 over a range of selection strengths (ϕ_f), where a value of 1 corresponds to the baseline scenario. The solid line behind each points indicates the 95% bootstrap confidence interval. All other parameters are set to their baseline value as described in the "Baseline parameter values" subsection.

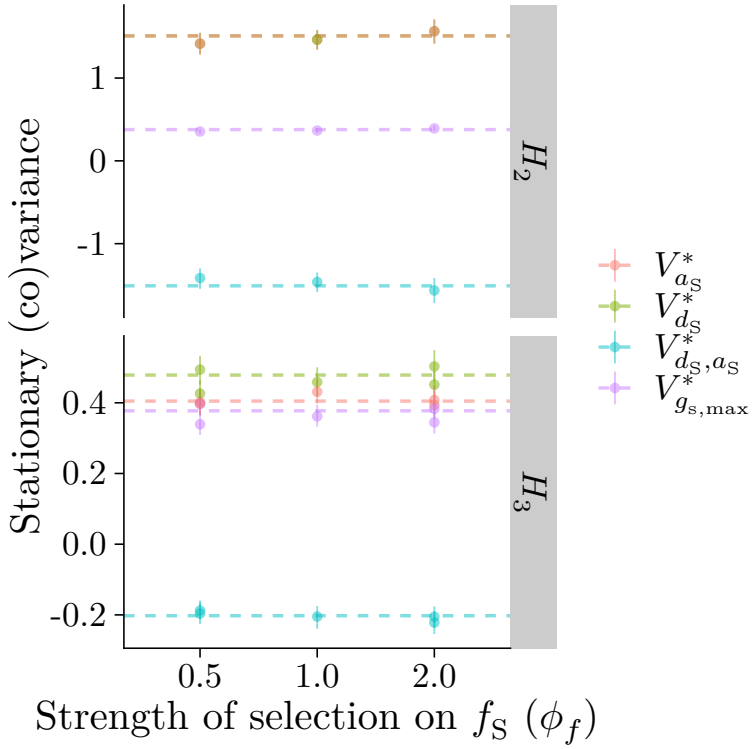
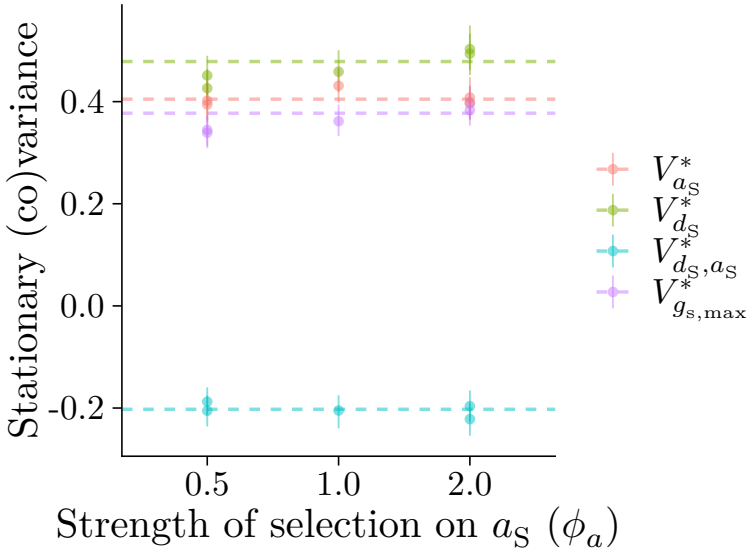


Figure S15: Among-species trait (co)variances are not sensitive to the strength of stabilizing selection on stomatal size (ϕ_a) at stationarity. Dashed lines are the theoretical expectations for the interspecific variance in stomatal density ($V^*_{d_S}$), stomatal size ($V^*_{a_S}$), anatomical maximum stomatal conductance ($V^*_{g_{s,max}}$), and covariance between density and size ($V^*_{d_S,a_S}$) at stationarity. The theoretical expectations differ only apply to H_3: stomatal adaptation + bounded size hypothesis. The points are estimated from simulations of an ensemble of 1000 over a range of selection strengths (ϕ_a), where a value of 1 corresponds to the baseline scenario. The solid line behind each points indicates the 95% bootstrap confidence interval. All other parameters are set to their baseline value as described in the "Baseline parameter values" subsection.



References

Arnold SJ. 2023. *Evolutionary Quantitative Genetics*. Oxford, United Kingdom: Oxford University Press.

Boucher FC, Démery V, Conti E, Harmon LJ, Uyeda J. 2018. A General Model for Estimating Macroevolutionary Landscapes. *Systematic Biology* **67**: 304–319.

Duffield S, Donatella K, Melanson D. 2024. *Thermox: Exact OU processes with JAX*.

Estes S, Arnold SJ. 2007. Resolving the Paradox of Stasis: Models with Stabilizing Selection Explain Evolutionary Divergence on All Timescales. *The American Naturalist* **169**: 227–244.

Grafen A. 1989. The phylogenetic regression. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* **326**: 119–157.

Hansen TF. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51**: 1341.

Hansen TF. 2012. Adaptive landscapes and the comparative analysis of adaptation. In: Svensson E, Calsbeek R, eds. The Adaptive Landscape in Evolutionary Biology. Oxford, UK: Oxford University Press, 205–226.

Hansen TF, Bartoszek K. 2012. Interpreting the evolutionary regression: The interplay between observational and biological errors in phylogenetic comparative studies. *Systematic Biology* **61**: 413–425.

Hansen TF, Pienaar J, Orzack SH. 2008. A comparative method for studying adaptation to a randomly evolving environment. *Evolution*: 1965–1977.

Ho LST, Ané C. 2014. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Systematic Biology* **63**: 397–408.

Ives AR, Midford PE, Garland T. 2007. Within-Species Variation and Measurement Error in Phylogenetic Comparative Methods (T Oakley, Ed.). *Systematic Biology* **56**: 252–270.

Kilmer JT, Rodríguez RL. 2017. Ordinary least squares regression is indicated for studies of allometry. *Journal of Evolutionary Biology* **30**: 4–12.

Lande R. 1976. Natural Selection and Random Genetic Drift in Phenotypic Evolution.

Evolution **30**: 314.

Lande R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: Body size allometry. *Evolution* **33**: 402–416.

Lande R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* **94**: 203–215.

Pennell MW, FitzJohn RG, Cornwell WK, Harmon LJ. 2015. Model adequacy and the macroevolution of Angiosperm functional traits. *The American Naturalist* **186**: E33–E50.

Pennell MW, Harmon LJ. 2013. An integrative view of phylogenetic comparative methods: Connections to population genetics, community ecology, and paleobiology: Integrative comparative methods. *Annals of the New York Academy of Sciences* **1289**: 90–105.

Pennell MW, Jiang D. 2024. The macroevolutionary adaptive landscape: More than a metaphor? *Evolution* **78**: 792–795.

Rolland J, Henao-Diaz LF, Doebeli M, Germain R, Harmon LJ, Knowles LL, Liow LH, Mank JE, Machac A, Otto SP, *et al.* 2023. Conceptual and empirical bridges between micro- and macroevolution. *Nature Ecology & Evolution* **7**: 1181–1193.

Sack L, Buckley TN. 2020. Trait Multi-Functionality in Plant Stress Response. *Integrative and Comparative Biology* **60**: 98–112.

Uyeda JC, Harmon LJ. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Systematic Biology* **64**: 902–918.

Warton DI, Wright IJ, Falster DS, Westoby M. 2006. Bivariate line-fitting methods for allometry. *Biological Reviews* **81**: 259.