Performance-based inference of selection on stomatal length and specific leaf area varies with climate-of-origin of the forest tree, *Eucalyptus ovata*

João Costa e Silva [a,*], Brad M. Potts [b,c], Suzanne M. Prober [d]

[a] *Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349–017 Lisboa, Portugal*
[b] *School of Natural Sciences, University of Tasmania, Private Bag 55, Hobart, Tasmania 7001, Australia*
[c] *ARC Training Centre for Forest Value, University of Tasmania, Private Bag 55, Hobart, Tasmania 7001, Australia*
[d] *CSIRO, Environment, Acton, Australian Capital Territory, Australia*

ARTICLE INFO

ABSTRACT

Understanding how functional traits affect plant performance and fitness is a key step in unravelling the role of natural selection in shaping the evolutionary trajectory of populations. We examined early-age selection acting on leaf traits via their effects on growth performance and fitness, measured in *Eucalyptus ovata* trees planted in a common-garden field trial embedded in a reforestation planting in Tasmania, Australia. We focused on two important leaf traits - stomatal length and specific leaf area (SLA) - measured two years after planting, and compared interplanted *E. ovata* groups originating from dry and wet home-site climates, with the trial site having intermediate long-term mean annual rainfall. Two-year height growth was used as the performance attribute, and the time-averaged tree survival over the subsequent six years as the fitness component. There was evidence for performance-based selection on the leaf traits, with the strength and form of selection depending on the trait and climate group being considered. In this sense, selection in the dry group operated mainly on stomatal length where a combination of directional (favouring longer stomata) and stabilizing selection was detected, whereas selection in the wet group acted only on SLA and was purely stabilizing. Estimates of performance-based correlational selection were not statistically significant. For both climate groups, estimates of fitness-based selection gradients provided evidence for significant directional (but not quadratic) selection on height performance, favouring individuals with faster growth, but did not indicate statistical support for direct effects of the leaf traits on tree survival, conditional on measured performance. These results validated qualitative inferences of selection from the performance-based analysis, and suggested that selection on the leaf traits appeared to be mediated by their effects on early-age height performance, which in turn directly influenced later-age survival. We discuss the mechanisms by which the focal traits may have affected height performance, and likely factors contributing to the different patterns of phenotypic selection observed in the two groups experiencing the same environment. We also provide expressions of analytical derivatives that were developed for the estimation of selection gradients based on a logistic regression model relating a binary fitness response to linear and nonlinear covariate terms for the target regressor variables.

1. Introduction

Understanding how plant functional traits affect fitness is a key focus of studies aimed at unravelling the role of natural selection in the evolution of population and species differences (Geber and Griffen, 2003; Kingsolver and Huey, 2003). While such studies require integration across multiple fields (such as evolutionary biology, genetics, functional ecology and ecophysiology), the identification of the functional traits under natural selection is an important first step (Geber and Griffen, 2003). The pathways linking a functional trait to individual fitness may

be complex (Arnold, 1983; Franklin and Morrissey, 2017; Franklin et al., 2018), and often ambiguous as the direct effects of selection on a focal trait may be confounded with the indirect effects from selection on other (omitted) phenotypically correlated characters (Caruso et al., 2020; Geber and Griffen, 2003; Lande and Arnold, 1983; Mitchell-Olds and Shaw, 1987; Walsh and Lynch, 2018). The effects of natural selection may be linear (directional selection) and/or non-linear in form (Lande and Arnold, 1983). Nonlinear selection may involve quadratic selection (e.g. stabilizing or disruptive selection) on single traits and correlational selection on pairs of traits, the latter indicating that selection on a focal

trait is dependent on its interaction with another trait (Lande and Arnold, 1983; Svensson et al., 2021; Walsh and Lynch, 2018). Selection may vary across space, time and environment (e.g. Carvalho et al., 2022; Dudley, 1996; Kingsolver et al., 2012; Siepielski et al., 2009). It may also differ between populations subjected to the same selection pressures (Costa e Silva et al., 2021; Donovan et al., 2009; Dudley, 1996; Etterson, 2004), which may in part reflect how close the multivariate phenotypes in the studied populations are to optimal fitness for the environmental conditions experienced (Colautti and Lau, 2015; Hendry, 2017; Walsh and Lynch, 2018).

Natural selection on a multivariate phenotype is usually quantified using *fitness-based* selection gradients. These selection parameters are estimated by regressing relative fitness (or a fitness component, such as survivorship or fecundity at a given period of the life cycle) on a set of focal traits, using fitness and trait measurements undertaken on the same individuals in the wild (Lande and Arnold, 1983; Walsh and Lynch, 2018). Arnold (1983, 2003) addressed natural selection by recognizing the putative links involving phenotypic traits, attributes of organismal performance and fitness, whereby selection on multiple traits is mediated by their effects on performance measures, which in turn are expected to influence fitness or a fitness component. In this context, for a causal pathway connecting the focal traits to fitness via a performance variable, the fitness consequences of trait variation are quantified by *performance-mediated* selection gradients, calculated as the product of regression coefficients obtained from the separate modelling of performance-trait and fitness-performance relationships. The performance-mediated analysis of selection can be applied to estimate directional selection (Arnold, 1983), as well as quadratic and correlational selection (Arnold, 2003; Franklin and Morrissey, 2017), on multiple traits. Arnold's framework (1983, 2003) enables exploring functional mechanisms that underlie natural selection on phenotypic traits (Franklin et al., 2018; Kingsolver and Huey, 2003; Opedal, 2021), while also facilitating the use of data from different cohorts of individuals to model the performance-trait and fitness-performance relationships. Both fitness-based and performance-mediated selection gradients are justified by evolutionary quantitative genetic theory (Arnold, 2003; Franklin and Morrissey, 2017; Lande and Arnold, 1983).

Performance attributes (e.g. growth measures) have been commonly used as fitness proxies in selection studies where field assessments of fitness cannot be attained (for a review, see Franklin and Morrissey, 2017). This leads to *performance-based* estimators of selection, which entail the regression coefficients obtained only from the performance-trait relationship segment in Arnold's framework (1983, 2003), and termed "performance gradients" by the author (hereafter, this terminology will also be used to refer to selection estimators from performance-based analysis). Although fitness proxies may be positively related to fitness or fitness components, they cannot be assumed to represent the demographic contribution of individuals to future generations. Consequently, by modelling a performance attribute in place of fitness, performance-based analysis may not result in accurate estimates of selection gradients, hence limiting their use for quantifying selection (Franklin and Morrissey, 2017).

Using data on the same individuals for focal traits, performance and fitness variables, Franklin and Morrissey (2017) investigated the sources of error that may appear when using performance as a substitute for fitness in quantitative inferences about selection. In this sense, the authors addressed two major sources of error: an error associated with the nature of the fitness-performance relationship; and an error due to alternative causal pathways through which the focal traits have direct effects on fitness. The performance gradients are affected equally by the former source of error, and thus can be *qualitatively* indicative of selection gradients; this is also applicable to the case of a nonlinear fitness-performance relationship, provided that fitness is monotonically related to the performance measure (Franklin and Morrissey, 2017; Franklin et al., 2018). However, as the latter source of error may influence the traits differentially, independence of traits and fitness,

conditional on measured performance, needs to be verified for valid qualitative within-study comparisons of selection based on performance gradient estimates (Franklin and Morrissey, 2017). As emphasized by Franklin and Morrissey (2017), whenever possible, a fitness proxy should be used in combination with information (including from other individuals or other studies within the studied species) on the fitness-performance relationship to assess whether a performance-based analysis is adequate for making inferences about natural selection. This will also be important to gain a broad view of selection and adaptation (Arnold, 1983).

Within the above framework, we evaluated performance-trait, fitness-performance, and fitness-trait relationships from the same cohort of trees tested in a common-garden population trial of the tree *Eucalyptus ovata* Labill. subsp. *ovata* (hereafter denoted as *E. ovata*), embedded in a reforestation planting towards the hotter-drier end of the species distribution in Tasmania (Prober et al., 2022). We classified the *E. ovata* trees into "wet" and "dry" groups based on their climate of origin in the native distribution, and compared selection estimates on two focal traits: one a leaf hydraulic trait - stomatal length - the other a leaf economic trait - specific leaf area (SLA).

Stomatal length is used as a measure of stomatal size (Jordan et al., 2015), a trait which can affect leaf gas exchange, water conservation, transpiration cooling and photosynthetic rates (e.g. Bertolino et al., 2019; Brodribb et al., 2020; Doheny-Adams et al., 2012; Franks et al., 2009; Harrison et al., 2020; Jordan et al., 2020; Schymanski et al., 2013). At the species-level, stomatal size is strongly positively associated with genome size, through its positive association with cell size, but over evolutionary time the genome-stomatal size relationship has been strongly modified by environmental adaptation (Veselý et al., 2020). For example, selection for small stomata is thought to have been a crucial adaptation for maintaining the gas-exchange capacity of land plants when atmospheric CO_2 levels were low during the Permo-Carboniferous and Cenozoic glaciations (Franks and Beerling, 2009). While stomatal size is often negatively associated with stomatal density, stomatal size has adaptive implications which are independent of stomatal density, as it affects stomatal construction costs (Franks et al., 2009) and the speed of the opening and closing of the individual stomata (Drake et al., 2013; Lawson and Vialet-Chabrand, 2019). However, despite the adaptive importance of stomatal size, there is a paucity of studies of contemporary phenotypic selection on this trait (Kosová et al., 2022).

In contrast to stomatal length, there are numerous studies of phenotypic selection involving SLA. SLA (and inverse indices such as leaf mass per area - LMA) reflects the leaf construction costs associated with light interception and carbon acquisition (John et al., 2017; Poorter et al., 2009). At the species level, SLA is one of the primary traits characterizing the Leaf Economic Spectrum between competitive ("fast species") and conservative ("slow species") growth strategies (Reich, 2014). The conservative growth strategy is often associated with a reduction in SLA, with leaves more sclerotic and having longer lifespans and higher survival in the face of abiotic and biotic stress (Díaz et al., 2016). While formulated at a global species scale, the evolution of these different growth strategies argues that the leaf economic traits should also be under contemporary selection within species (Ramírez-Valiente et al., 2022; Sartori et al., 2019). As expected, contemporary phenotypic selection on SLA has been shown to be environmentally dependent (e.g. Kimball et al., 2013; Ramírez-Valiente et al., 2014; Steinger et al., 2003). For example, high SLA has been shown to be favoured in light-limited environments (Lopez-Gallego and O'Neil, 2014), although such selection can be masked in the presence of herbivores (Salgado-Luarte and Gianoli, 2012). While most phenotypic selection studies involving SLA have been only focused on directional selection (e.g. Alexandre et al., 2020; Carlson et al., 2016; Gianoli and Saldaña, 2013; Ramírez-Valiente et al., 2014; Walter et al., 2023), some studies do report nonlinear selection (Blanco-Sánchez et al., 2022; Lopez-Gallego and O'Neil, 2014; Magnoli and Lau, 2020; Ramírez-Valiente et al., 2011; Steinger et al., 2003).

With a general hypothesis that the pattern of selection for the focal leaf traits in the same environment will differ between the *E. ovata* groups originating from wet and dry home-site climates, we explored the following questions:

1) Is there evidence for performance-based selection on the focal leaf hydraulic and economic traits within either climate group? If so: (a) do the wet and dry groups differ significantly in the strength, form and direction of performance-based estimates of linear and nonlinear selection acting on the leaf traits; and (b) given the dry climate of the common-garden site, is performance-based directional selection stronger in the wet than the dry group?
2) Is size-dependent mortality based on the relationship of early-age height performance with later-age tree survival valid? If so, is the expected probability of tree survival monotonically related to height performance in both groups?
3) Do the leaf traits have direct fitness consequences independent of variation in height performance, when considering both linear and nonlinear trait effects in the modelling of the fitness-trait relationship in either group?

In particular, answers to the latter two questions will be required to validate qualitative inferences of selection on the focal leaf traits from the performance-based analysis.

2. Materials and methods

2.1. Plant material and measurements

The common-garden field trial of *E. ovata* was established in August 2014 in a fenced reforestation area on an ex-pasture site adjacent to remnant native forest at Connorville in northern Tasmania, Australia (41.828S, 147.138E, altitude of 185 m) (for further details, see Bailey et al., 2021; Prober et al., 2022). *E. ovata* would have been a component of the forest originally cleared for pasture establishment. The trial included families derived from open-pollinated seed collections from mother trees sampled in native populations from across the geographic and ecological range of *E. ovata* in Tasmania. Within the trial, the families were arranged in a resolvable row-column experimental design comprising 8 blocks. Each block was arranged as a rectangular grid with an inter-tree spacing of 3 m (between rows) by 2.5 m (within rows), and each family was represented as a single-tree plot. Overall, the trial included 360 families from 51 populations (Bailey et al., 2021). Our study was based on a subsample of one plant from each of 34 populations and each of 5 blocks. This subset of populations was representative of the more-or-less continuous, eastern range of *E. ovata* in Tasmania (Harrison, 2017), and was subsequently divided into two main climate groups - denoted as "dry" and "wet" - of 17 populations each, based on the populations' home-site climate (see Methods S1 and Table S1; Appendix A). These groups differed markedly in home-site rainfall, with the dry group predicted to receive 32% less annual rainfall on average than the wet group (Methods S1). A unique family per population was sampled in each block, resulting in a population being represented by five, well-separated mother trees in their native population. Therefore, the cohort of individuals originally sampled for the current study comprised 170 plants in total (34 populations x 5 plants per population), and 85 plants per climate group. The long-term mean annual rainfall of the common-garden site (608 mm) was intermediate between that of the home-site means of the dry (mean 561 ± 13, range 451–663 mm) and wet (mean 824 ± 33, range 659–1077 mm) climate groups, but at the wetter end of the range of the dry group populations (Methods S1).

The assessment of growth performance and sampling of leaves for measuring the leaf traits were undertaken at age 2 years from planting, when the sampled cohort of trees were in the nonreproductive stage. Growth performance was assessed using maximum tree height (m), measured with height poles. The measurement details on the two focal leaf hydraulic and economic traits - stomatal length (μm) and specific leaf area (SLA, mm^2/mg) - are provided in Prober et al. (2022). Tree-level data for a given leaf trait was based on the average of measurements from mature, fully-expanded leaves (excluding the petiole) collected from three branches distributed around the mid-outer, sun-exposed part of the canopy of each tree. SLA was derived by measuring the leaf area (digitally) and dry mass of ten of these leaves. Stomatal length was measured as the stomatal pore length between guard cells. It was assessed from both surfaces of three leaves (10 measurements per surface) with a microscope at 1000x magnification, using the nail polish impression method described in Franks et al. (2009). At age 8 years from planting, the survival of the trees in the sampled cohort was scored as a binary outcome whereby a tree was classified as dead (scored as 0) when no live plant tissue was evident above ground or the plant was in very poor vegetative health, otherwise it was classified as alive (scored as 1). Thus, the fitness component studied at an early life-history stage was the time-averaged tree survival to 8 years (i.e. over a six-year period, starting at 2 years from planting in the sampled cohort). Most trees were non-reproductive during this phase (i.e. 5% of survivors were observed to be reproductive at age 4 years, but in the 8-year assessment no reproduction was observed). Table S2 (Appendix A) provides the estimated means and standard deviations for the leaf traits and the height performance measure. Of the sampled trees, 71% and 81% survived to the final assessment in the dry and wet climate groups, respectively.

2.2. Data analysis

Preliminary data analysis indicated that estimates of population variances were small and not statistically significant ($p > 0.20$) for height performance and the focal leaf traits within a climate group, resulting in low intraclass correlation coefficients (see Table S2 and its footnotes). A low level of between-population variation was also found for tree survival in either group (see Methods S2 in Appendix A for details on the frequentist and Bayesian estimation of the population variance). These results indicated that the putative dependence among observations due to clustering into populations was weak for the data used in the current study. Consequently, to simplify the models described below (in particular to reduce the number of parameters to be estimated in the logistic models used for the fitness-performance and fitness-trait relationships), we have not included a model term for populations (either under a fixed- or a random-effects specification), and thus we have assumed independence of the observations within populations. The data analyses were undertaken with the SAS 9.4 software (SAS, 2017).

2.2.1. Performance-based selection: estimation of performance gradients and their group differences

For each climate group, the following general linear model was used to evaluate the performance consequences of variation in the two studied leaf traits:

$$y_i = \beta_0 + r_m + \sum_{j=1}^{2} \beta_j x_j + \sum_{j=1}^{2} \frac{1}{2}\gamma_j x_j^2 + \gamma_{jk} x_j x_k + \varepsilon_i \qquad (1)$$

where y_i is the observation on the height performance of the ith individual; β_0 is an intercept term; r_m is the fixed effect of the mth block; β_j and γ_j are the partial regression coefficients for the linear and quadratic effects, respectively, of leaf trait x_j on performance; γ_{jk} is the partial regression coefficient for the interactive effect of the two leaf traits x_j and x_k on performance; and ε_i is a random residual. Together, γ_j and γ_{jk} characterize the nonlinear effects of the traits on performance. Eq. (1) included block effects, as initial analysis indicated that they were statistically significant ($p < 0.05$) for the height response variable at least in one of the groups.

Prior to analysis, the data on a leaf trait were mean-centered, and

then mean-standardized and expressed on a (dimensionless) percentage scale, within each group (see Methods S3, Appendix A, for details on the choice of within-group standardization). Subsequently, squared (x_j^2) and cross-product ($x_j x_k$) covariates were obtained for the estimation of nonlinear effects of the traits on performance. Standardizing in relation to a grand (group) mean is both permissible and meaningful for the focal traits (Houle et al., 2011). Thus, least-squares estimates of the regression parameters for the leaf traits in Eq. (1) refer to the expected change in average performance from increasing the values of the trait covariate terms expressed on a percentage scale.

The y_i height values were also relativized within each group (via the division by the group mean) to derive a proxy for relative fitness, so that the regression coefficients on the leaf traits in Eq. (1) can be interpreted as (mean-standardized) performance-based estimators of selection (Franklin and Morrissey, 2017). Thus, β_j and γ_j are the directional and quadratic performance gradients on trait x_j, respectively, and γ_{jk} is the correlational performance gradient on traits x_j and x_k (Arnold, 1983, 2003). The parameters β_j, γ_j and γ_{jk} can be geometrically interpreted as descriptors of the average slope, average curvature and orientation, respectively, of the individual performance surface acting on the observed distribution of the focal traits, under the assumption that a quadratic function is an accurate approximation of the true response surface (Phillips and Arnold, 1989; Walsh and Lynch, 2018). For performance-based quadratic selection on trait x_j, a negative (positive) $\widehat{\gamma}_j$ estimate reflects a downwardly (upwardly) curved performance surface along the trait axis, and may suggest stabilizing (disruptive) selection if a performance maximum (minimum) occurs within the observed trait range (Lande and Arnold, 1983; Mitchell-Olds and Shaw, 1987; Walsh and Lynch, 2018). In this context, note also that the values of the x_j^2 covariate terms in Eq. (1) were halved before modelling, so that the γ_j coefficients can adequately describe the average curvature of the individual performance surface (and thus properly quantify the strength of quadratic selection) (Stinchcombe et al., 2008). For performance-based correlational selection on a pair of traits, γ_{jk} indicates whether higher performance would favour either a positive ($\widehat{\gamma}_{jk} > 0$) or a negative ($\widehat{\gamma}_{jk} < 0$) correlation between traits x_j and x_k (Arnold, 2003; Lande and Arnold, 1983).

Departures from multivariate normality for the joint phenotypic distribution of a set of focal traits may cause a covariance between linear and nonlinear covariate terms in a regression model, which can lead to $\widehat{\beta}_j$ estimates measuring incorrectly the effects of directional selection on the phenotypic means of the traits (Lande and Arnold, 1983; Walsh and Lynch, 2018). To avoid this, we followed the two-step approach suggested by Lande and Arnold (1983): valid estimates of β_j were firstly obtained from including only linear covariate terms in Eq. (1); subsequently, γ_j and γ_{jk} were estimated by using the full model with both linear and nonlinear covariate terms.

The data pre-processed within each group as described above were combined to test whether performance gradient estimates differed significantly between the two groups. In this context, the full linear model fitted to the height response variable had: classification predictors for groups, and blocks within groups; linear and nonlinear covariate terms pertaining to mean effects of the traits across groups; and interaction effects of the groups with linear or nonlinear covariate terms, which relate to the differences between groups in directional or nonlinear performance gradients. Akin to the analyses within groups, we applied a two-step approach where group differences in directional performance gradients were firstly tested by modelling only linear covariate terms, and then group differences in nonlinear performance gradients were tested by using the full model. Heterogeneous residual variation between groups was incorporated by modelling a diagonal structure in which residual effects were estimated separately by group.

Following least-squares estimation of model parameters and their standard errors, statistical inferences about the performance gradient coefficients within groups and their difference between groups were provided by conducting *t*-tests and computing 95% confidence intervals. In particular, for the combined data analysis, the statistical inferences about the group differences in performance gradients used the effective degrees of freedom based on the Satterthwaite (1946) approximation, which is appropriate under the modelling of heterogeneous residual variation across groups (Steel et al., 1997). Residual plots and other diagnostic statistics indicated that model assumptions were reasonably met (i.e. normality of residual distributions, no residual heteroskedasticity, absence of statistically significant outliers, and variance inflation factors ≤ 1.5).

2.2.2. Visualization of the individual performance surface in each group

Within each group, a thin-plate spline (TPS) was used to visualize the individual performance surface defined by the relationship between height growth and the two leaf trait axes. Under this approach, a smooth multivariate response surface is produced from a non-parametric regression model fitted by a penalized least-squares method (Wahba, 1990). By not assuming any particular parametric form for the model [as opposed to a quadratic function, such as defined in Eq. (1)], a TPS fit allows for greater flexibility (e.g. unrevealing local features) in representing the shape of the response surface when compared to a best quadratic fit (Blows et al., 2003; Walsh and Lynch, 2018). Determining the smoothing parameter for a final TPS fit was based on the generalized cross validation (GCV) function described by Wahba (1990): the optimization process examined a range of smoothing values, and the value that minimized the GCV criterion was chosen as the optimal smoothing parameter. A two-dimensional data grid was generated to represent pairwise combinations of the modelled traits, and predicted height values from the final TPS fit were obtained for all points in the grid in order to plot the performance surface.

2.2.3. Fitness-performance relationship: size-dependent mortality

We evaluated whether the tree mortality observed in the data from each climate group was size-dependent. In this context, we assessed the strength and form of selection acting on height performance within each group, as well as the difference between groups. A logistic regression model was used to relate tree survival with height measures (that were mean-centered, and then expressed on a percentage scale following mean-standardization) within a group. The linear predictor of this model comprised a linear term, or both linear and quadratic terms, for the height variable. These model definitions did not include blocks as fixed effects, as previous generalized linear models fitted for tree survival in either group did not detect statistically significant ($p > 0.37$) block effects for the binary response.

In logistic regression models, the estimation of logit coefficients by maximum likelihood (ML) may suffer from finite sample bias (Firth, 1993; Nemes et al., 2009). This bias may not entirely disappear even in large samples, although it is expected to be greater in small to moderated sized data sets (van Smeden et al., 2016). In these situations, the Firth's (1993) bias-reducing method based on penalized maximum likelihood estimation can be applied to improve the accuracy of logistic regression coefficients, and also to provide reliable (finite) parameter estimates when separation occurs in the dataset (Heinze and Schemper, 2002; van Smeden et al., 2016). The problem of separation (which refers to a perfect prediction of the binary outcome by the modelled explanatory variables) was not detected in our logistic regression analyses, and this applies to both fitness-performance and fitness-trait relationships. Yet, given our sample sizes, the Firth's method was used in combination with penalized likelihood-ratio (LR) tests and 95% profile penalized-likelihood confidence intervals (Heinze and Schemper, 2002) to increase the efficiency of parameter estimation and statistical inference for the regression coefficients on the logit scale [penalized LR tests were pursued with the SAS macro developed by Heinze and Ploner (2004), available at github.com/georgheinze/flicflac].

The fitness consequences of variation in growth performance were

assessed via the estimation of (fitness-based) selection gradients for the height variable, using derivative-based methods (e.g. Costa e Silva et al., 2021; Franklin and Morrissey, 2017; Janzen and Stern, 1998; Morrissey and Sakrejda, 2013). An average directional selection gradient was estimated from a logistic model where the linear predictor included only a linear term for the height variable, whereas an average quadratic selection gradient was obtained under a linear predictor with both linear and quadratic terms for the height variable. In the estimation of selection gradients, the individual predictions of survival on the expected probability scale based on the Firth-type penalization were improved by using the "intercept correction" approach described by Puhr et al. (2017). Details on the derivative-based methods (i.e. derivatives calculated either by numerical approximation or analytically) used for estimation of selection gradients are given in Methods S3 (Appendix A).

Non-parametric bootstrapping was applied to provide the standard error, as well as the 95% confidence interval, for an estimated average selection gradient within a group and its difference between groups (for further details, see Methods S4 in Appendix A). Statistical support against a null hypothesis being true (i.e. against the absence of an effect for a selection gradient within a group or its difference between groups) was given by a 95% confidence interval not overlapping with zero. In particular, statistical support for the effect associated with the average directional selection gradient on height performance will be indicative of size-dependent mortality. Further, a lack of statistical support for the effect associated with the average quadratic selection gradient will indicate that the relationship between tree survival on the expected probability scale and growth performance is mainly monotonic (as characteristic of a logistic sigmoid function) over the observed range of the height variable. Under these circumstances, the performance gradients estimated for the leaf traits under Eq. (1) can be qualitatively indicative of selection gradients, provided that independence of traits and fitness, conditional on measured performance, also holds (Franklin and Morrissey, 2017; see also the Introduction).

2.2.4. Independence of traits and fitness, conditional on measured performance

Akin to selection gradients obtained by performance-mediated analysis, performance gradients also assume that the effects of the focal traits on fitness (or a fitness component) are completely mediated by the aspect(s) of performance measured, implying that: (i) traits do not influence unmeasured performance features; and (ii) traits do not affect fitness directly (Arnold, 1983; Franklin and Morrissey, 2017). Condition (i) pertains to unobserved aspects of organismal performance through which the traits could exert their effects on fitness, so that either analytical approach may provide a partial description of selection acting on the traits [i.e. referring only to the part of selection mediated by the specific performance attribute(s) measured]. Nevertheless, we assumed that the focal leaf traits would influence mainly height performance at the observed early stage of the life cycle of *E. ovata,* and thus their effects on other (unmeasured) features of tree performance would be less important. Therefore, we focused on condition (ii) to assess whether the two studied leaf traits had direct fitness consequences independent of variation in height performance. Using the time-averaged (i.e. over the observed period) survival as a fitness component, the following full logistic regression model was defined for each climate group:

$$g(E(w_i \mid \mathbf{x}_i)) = \eta_i = a + \sum_{j=1}^{2} b_j x_j + \sum_{j=1}^{2} c_j x_j^2 + c_{jk} x_j x_k + b_h x_h \tag{2}$$

where w_i is the observation on the survival of the ith individual (i.e. a binary outcome of 0 or 1); $g(.)$ is the logit link function, relating the conditional expectation $[E(w_i \mid \mathbf{x}_i)]$ of the binary response to the linear predictor η_i; $\mathbf{x}_i$ is a row vector for the ith individual from the design matrix $\mathbf{X}$ of all the explanatory variables included in η_i; a is an intercept term; b_j, c_j and c_{jk} are regression coefficients for the linear, quadratic and interactive effects, respectively, of a leaf trait x_j or the pair of leaf traits

x_j and x_k on the log-odds (logit scale) of survival; and b_h is a regression coefficient for the linear effect of the height variable x_h on the log-odds of survival. The binary responses were assumed to be Bernoulli distributed with a parameter μ_i, which is the conditional probability of tree survival in our case [i.e. $P(w_i = 1 \mid \mathbf{x}_i)$] and estimated by $P(w_i = 1 \mid \mathbf{x}_i) = g^{-1}(\eta_i) = \frac{\exp(\mathbf{x}_i \boldsymbol{\beta})}{1 + \exp(\mathbf{x}_i \boldsymbol{\beta})}$, where $g^{-1}(.)$ is the inverse of $g(.)$, exp denotes the exponential function, and $\boldsymbol{\beta}$ is a column vector of coefficients comprising all fitted parameters. The linear predictor η_i included covariate terms for the leaf traits alone, or the leaf traits plus height performance [i.e. as in Eq. (2)], with the trait and height data being previously mean-centered, and then mean-standardized and expressed on a percentage scale, within each group. In either of these model specifications, and to keep consistency with the performance-based analysis, the trait coefficients b_j were estimated separately from the trait coefficients c_j and c_{jk}, according to the two-step approach described above. Modelling the height variable x_h in η_i is intended to allow for the estimation of (direct) trait effects on survival that are not mediated by growth performance, hence enabling evaluation of whether the fitness component is independent of the focal traits, conditional on measured performance.

The finite sample bias in ML estimators of logistic regression coefficients is expected to be more likely with a decreasing ratio known as "events per variable" (EPV, defined by the number of observations on the rarer of the two outcomes - e.g. number of dead trees, in our case - relative to the number of estimated model parameters, excluding the intercept) (Courvoisier et al., 2011; van Calster et al., 2020; van Smeden et al., 2016). In this context, simulation studies have shown that the Firth's correction can reduce the finite sample bias of logit coefficients close to zero, while also substantially decreasing their mean squared error, in low EPV settings (e.g. EPV values from 3 to 6; Puhr et al., 2017; van Calster et al., 2020; van Smeden et al., 2016). We thus applied Firth-type estimation and statistical inference procedures for the logit coefficients. In addition, although nonlinear effects involving height performance could have been included in Eq. (2) (Franklin and Morrissey, 2017), this model extension was not pursued in our within-group analyses as we have limited the lowest EPV value at ≈ 3 [e.g. using Eq. (2) in the wet group]. However, independence of fitness and traits, given measured performance, was also assessed at the species level by combining the data across groups. Besides the terms defined in Eq. (2), the full logistic model for the combined data analysis comprised a group effect as a classification predictor, as well as covariate terms for the quadratic effect of height and the interaction effects between height and each of the leaf traits (EPV = 4). Estimates of variance inflation factors were < 1.5 in either the within- or across-group analysis, indicating a weak degree of multicollinearity.

Derivative-based estimates of selection gradients were obtained for the leaf traits and height performance under a given fitted logistic model. Average directional selection gradients were separately estimated from average quadratic and correlational selection gradients, following the two-step modelling approach described previously. In particular, averaging second partial derivatives over the range of the observations of the jth variable may measure adequately the average curvature of the individual selection surface along the axis of the focal variable (Walsh and Lynch, 2018). Therefore, the x_j^2 values were not halved in Eq. (2), in contrast to the performance-based analysis (where the x_j^2 values were halved to avoid underestimation of the strength of quadratic selection based on the γ_j coefficients; Stinchcombe et al., 2008). The details on the methods used to calculate partial derivatives and to estimate (fitness-based) average selection gradients are given in Methods S3 (Appendix A). In this context, analytical expressions of first, second and mixed partial derivatives are provided for the estimation of average directional, quadratic and correlational selection gradients, respectively, when a logistic regression model includes both linear and nonlinear covariate terms for the explanatory variables [in which case

the approach described by Janzen and Stern (1998) does not apply]. Non-parametric bootstrapping was used to obtain the standard error and the 95% confidence interval for an estimated average selection gradient or its difference between groups (see Methods S4, Appendix A, for further details).

3. Results

3.1. Performance-based selection: estimation of performance gradients and their group differences

The least-squares estimates of the performance gradients on the focal leaf traits in each climate group are presented in Table 1. For stomatal length in the dry group, the estimated directional performance gradient was statistically significant ($p < 0.05$) and positive, indicating performance-based linear selection favouring individuals with longer stomata. This tendency was opposite to that detected in the wet group, although the group's negative estimate of the directional performance gradient on stomatal length was not statistically significant. There was also a statistically significant estimate for the quadratic performance gradient on stomatal length in the dry group, with its negative sign reflecting a downwardly curved form of performance-based nonlinear selection (indicative of stabilizing selection) acting on the trait. In the wet group, the estimated quadratic performance gradient on stomatal length did not differ significantly from zero.

There was no evidence of statistically significant directional selection via growth performance on SLA in either group. Performance-based selection experienced by SLA was mainly quadratic, with the negative estimate of the corresponding performance gradient revealing a downward curvature along the axis of the trait in both groups. However, the estimated quadratic performance gradient on SLA was highly significant ($p < 0.001$) in the wet group, but was marginally non-significant ($p < 0.10$) in the dry group.

For either group, the estimated correlational performance gradient between stomatal length and SLA was not significantly different from zero, and thus there was no evidence for the presence of performance-based correlational selection influencing the covariance between the two leaf traits. When compared to a model specified with only block effects, extending the model to include both linear and nonlinear covariate terms [i.e. as in Eq. (1)] added 13.5% and 18.4% to the percentage of total variation explained in height performance for the dry and wet groups, respectively (not shown).

Statistical tests for differences in the magnitude of performance gradient estimates between the climate groups revealed a statistically significant difference in performance-based directional selection on stomatal length, and marginally non-significant differences in quadratic selection on stomatal length and SLA (Table 1). However, for the latter component of nonlinear selection, these results are likely to reflect the lack of sufficient statistical power for testing the difference between groups, as the visualization of the individual performance surfaces (see below) suggested that the dry and wet groups contrasted in the magnitude of performance-based quadratic selection operating on the leaf traits.

There was no evidence to suggest that performance-based directional selection in the wet group was stronger than in the dry group, rather contrary to our expectations based on differences between the population home-sites and the common-garden site in long-term mean annual rainfall. This was particularly the case of stomatal length on which statistically significant stronger performance-based directional selection was observed in the dry group. Further, the dry group had smaller stomata than the wet group ($p = 0.10$; Table S2) and, while not statistically significant in the wet group, the opposite signs of the estimated directional performance gradients for stomatal length (Table 1) suggested a performance peak at intermediate phenotypic values of the trait.

Table 1

Performance-based estimates of linear (directional) and nonlinear (quadratic and correlational) selection, and their difference between the dry and wet climate population groups, for stomatal length and specific leaf area assessed in a common-garden field trial of *Eucalyptus ovata*.

Trait covariate terms	Dry group	Wet group	Group difference
Directional selection			
Stomatal length	0.646 ± 0.293 $(0.062, 1.231)$ $p = 0.031$	-0.406 ± 0.321 $(-1.046, 0.233)$ $p = 0.209$	1.052 ± 0.435 $(0.193, 1.912)$ $p = 0.017$
Specific leaf area	0.237 ± 0.246 $(-0.253, 0.726)$ $p = 0.339$	-0.061 ± 0.275 $(-0.608, 0.486)$ $p = 0.824$	0.298 ± 0.369 $(-0.431, 1.026)$ $p = 0.421$
Quadratic selection			
Stomatal length^2	-0.095 ± 0.041 $(-0.178, -0.013)$ $p = 0.023$	0.030 ± 0.059 $(-0.087, 0.147)$ $p = 0.612$	-0.125 ± 0.072 $(-0.268, 0.017)$ $p = 0.083$
Specific leaf area^2	-0.048 ± 0.026 $(-0.099, 0.003)$ $p = 0.065$	-0.119 ± 0.030 $(-0.180, -0.059)$ $p < 0.001$	0.071 ± 0.040 $(-0.007, 0.150)$ $p = 0.074$
Correlational selection			
Stomatal length x Specific leaf area	-0.023 ± 0.029 $(-0.082, 0.035)$ $p = 0.430$	0.012 ± 0.030 $(-0.048, 0.071)$ $p = 0.697$	-0.035 ± 0.042 $(-0.118, 0.048)$ $p = 0.405$

The growth performance measure (tree height) and the leaf traits (stomatal length and specific leaf area) were assessed at age 2 years from planting. Evaluating the growth performance consequences of variation in the leaf traits within a climate (dry or wet) group was based on a general linear model with (relative) height performance as the response variable and covariate terms for the focal traits. The estimated performance gradients for the leaf traits within each group refer to the percentage change in relative performance (i.e. height change in % of its group mean) expected from increasing the values of the trait covariate terms expressed on a percentage scale. The performance-based estimates of directional selection were firstly obtained from the modelling of linear covariate terms for the leaf traits, and then the estimates of quadratic and correlational selection were provided from the modelling of both linear and nonlinear (i.e. squared and cross-product) covariate terms for the leaf traits (Lande and Arnold, 1983). This procedure was undertaken to estimate the performance gradients for the leaf traits within each climate group, as well as to estimate their differences between groups (i.e. in an analysis combining data across groups, and modelling the interaction effects between groups and each of the trait covariate terms). Standard errors, 95% confidence intervals (in parentheses) and *p*-values of statistical *t*-tests are given for the estimated performance gradients within a group and their difference between groups.
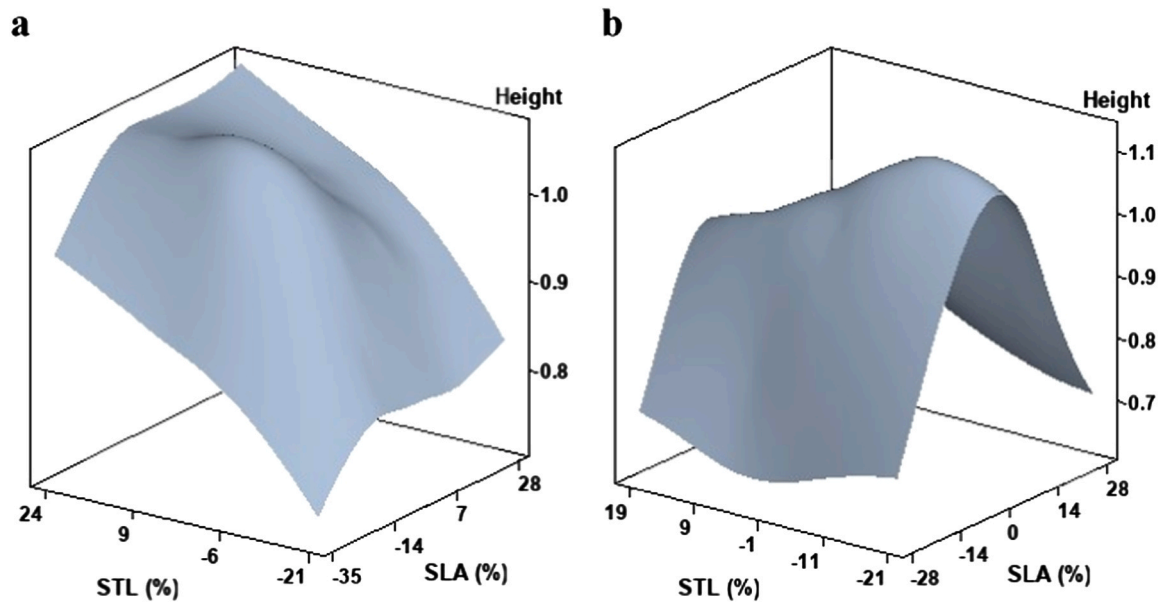
Fig. 1. Thin-plate spline visualization of the individual, performance-based selection surface on two leaf traits - stomatal length (STL) and specific leaf area (SLA) - assessed in a common-garden field trial of *Eucalyptus ovata* at age 2 years from planting. A performance surface was estimated for each of two climate population groups - (**a**) dry and (**b**) wet - by fitting a thin-plate spline model to relative height performance (i.e. height growth in proportion of its group mean - denoted as "Height" in the figure) as a function of the two focal leaf traits. The thin-plate spline modelling used observations of the leaf traits that were initially mean-centered with the grand means of their climate group, and then expressed on a percentage scale following mean-standardization within a group.

3.2. Visualization of the individual performance surface in each group

Fig. 1 shows the three-dimensional visualization of the individual performance surfaces that were fitted by using a thin-plate spline within each climate group. The thin-plate representation of the individual performance surfaces confirmed the patterns of performance-based selection inferred from the least-squares estimates of the performance gradients.

The individual performance surface in the dry group corroborated the presence of a combination of significant performance-based directional and quadratic selection on stomatal length, and the weaker (marginally non-significant) quadratic selection on SLA (Table 1). Reflecting a combination of linear and quadratic effects, the surface's downward curvature across stomatal length was asymmetric within the sampled space, being dominated by the effects of performance-based directional selection favouring individuals with longer stomata until the mid-upper range of the trait. A local performance maximum then appeared to occur within the observed range of stomatal length slightly above its group mean, suggesting that the quadratic form of

performance-based selection acting on the trait in the dry group corresponds to stabilizing selection.

In the wet group, the individual performance surface reflected the highly significant performance-based quadratic selection experienced by SLA, and the absence of other significant components of selection operating on this trait or stomatal length (Table 1). In this sense, the performance surface had a general downward curvature across the axis of SLA, with the highest relative height performance values occurring along a ridge (which rose slightly and peaked towards lower levels of stomatal length, consistent with the trait's non-significant negative estimate of the directional performance gradient) at an intermediate value close to the SLA mean. This suggests the presence of performance-based stabilizing selection acting on SLA within the range of the trait observations in the wet group.

3.3. Fitness-performance relationship: size-dependent mortality

Table 2 presents the (fitness-based) average directional and quadratic selection gradients on height performance, estimated from the

Table 2
Estimates of fitness-based average directional and quadratic selection gradients, and their difference between the dry and wet climate population groups, for growth performance assessed in a common-garden field trial of *Eucalyptus ovata*.

	Dry group	Wet group	Group difference
Average directional selection gradient	1.068 ± 0.248	0.807 ± 0.190	0.261 ± 0.311
	(0.625, 1.609)	*(0.485, 1.247)*	*(−0.328, 0.901)*
Average quadratic selection gradient	-0.016 ± 0.012	-0.019 ± 0.010	0.003 ± 0.015
	(−0.033, 0.009)	*(−0.032, 0.001)*	*(−0.031, 0.029)*

The growth performance measure (tree height) and tree survival (used as a fitness component) were assessed at ages 2 and 8 years from planting, respectively. Evaluating the fitness consequences of variation in growth performance within a climate (dry or wet) group was based on a logistic regression model with tree survival as the response variable and the linear predictor comprising a linear covariate term, or both linear and quadratic covariate terms, for the height performance measure. Estimates of (fitness-based) average directional and quadratic selection gradients, and their standard errors, are presented for height performance within each group, and refer to the percentage change (on average) in expected relative survival caused by increasing the values of the performance covariate terms expressed on a percentage scale (see Methods S3; Appendix A). The average directional selection gradient was firstly derived from the modelling of a linear covariate term for the performance measure, and then the average quadratic selection gradient was obtained from the modelling of both linear and quadratic covariate terms for the performance measure. Non-parametric bootstrapping was applied to provide the standard error, as well as the 95% confidence interval (in parentheses), for each estimated average selection gradient within a climate group and its difference between groups (see Methods S4; Appendix A). Statistical support against a null hypothesis being true (i.e. against the absence of an effect for either an average selection gradient within a group or its difference between groups) is given by a 95% confidence interval not overlapping with zero (indicated in Italics).
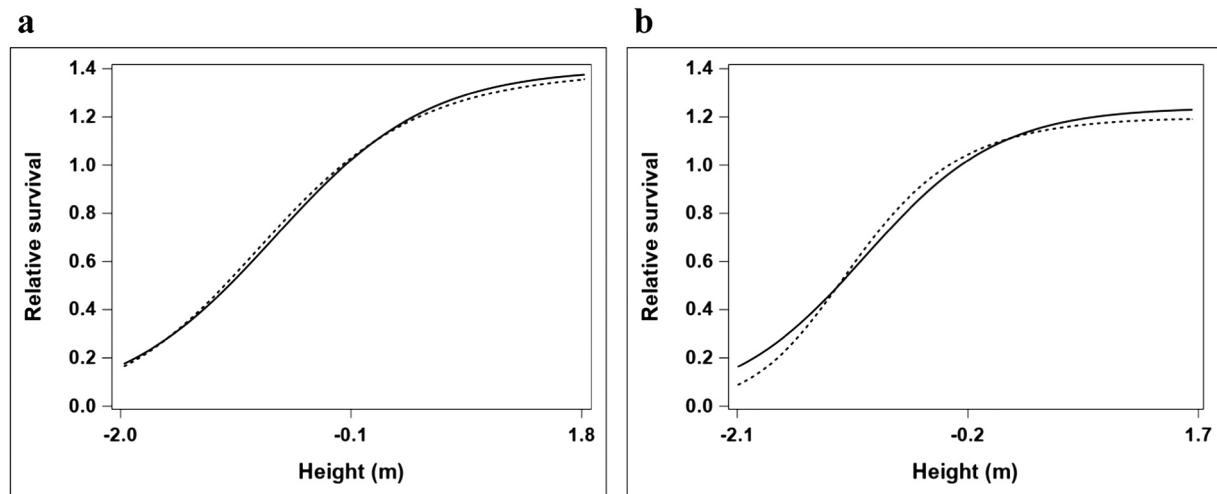
Fig. 2. Relationship of tree survival (used as a fitness component) with height growth (used as a performance measure), assessed in a common-garden field trial of *Eucalyptus ovata* at ages 8 and 2 years from planting, respectively. Individual fitness surfaces are represented for the height performance measure for two climate population groups - **(a)** dry and **(b)** wet. The fitness-performance relationship was estimated from a logistic regression model with tree survival as the (binary) response variable, and the fitted curves depicted for each group pertain to a model where the linear predictor included either a linear covariate term (solid line), or both linear and quadratic covariate terms (dashed line), for height performance. The relative survival in the vertical axis refers to individual predicted probabilities divided by the average predicted probability within a group. The bias-reducing, penalized maximum likelihood method described by Firth (1993) was used for the estimation of the model parameters on the logit scale. In this context, the "intercept-correction" procedure proposed by Puhr et al. (2017) was applied to improve the individual predicted probabilities of survival. The horizontal axis refers to mean-centered observations for the growth performance variable within a group.

modelling of the fitness-performance relationship in each climate group. Statistical support for the effect of the average directional selection gradient was provided by a bootstrapped 95% confidence interval (CI) not overlapping with zero in both groups. This is consistent with size-dependent mortality, and the positive estimate of the average directional selection gradient in both groups indicated selection favouring individuals with faster height growth. These results on the expected fitness scale were also consistent with the highly statistically significant ($p < 0.001$) estimates found on the logit scale for the regression coefficient associated with the linear term of the height variable (Table S3; Appendix A).

There was not statistical support for the effect of the average quadratic selection gradient on height performance in both groups (Table 2), nor was the logit quadratic regression coefficient estimated for the height variable found to be significant ($p > 0.10$; Table S3). Fitted functions relating survival on the expected probability scale to height performance showed a predominant monotonically increasing trend over the range spanned by the observed performance values in either group, regardless of whether or not a quadratic term was modelled for the height variable (Fig. 2). Overall, these results suggested that modelling a quadratic term for the height variable did not significantly contribute to a more accurate estimation of the individual selection surface acting on growth performance over the observed phenotypic range.

Differences between groups in the magnitude of the estimated average directional or quadratic selection gradient on height performance were not statistically supported (Table 2). We thus undertook an analysis combining the data across groups, in order to obtain a pooled, species-level estimate of the strength of directional selection on height performance. This was pursued by modelling only a linear term for the height variable (besides also a term for a group effect), as the results mentioned above suggested that selection on height performance was mainly directional given the observed phenotypic range. The estimate of the average directional selection gradient combined across groups was 0.943% (bootstrapped 95% CI: 0.665, 1.264), effectively indicating that a 1% increase in height contributed to a similar change (on average) in expected relative survival.

3.4. Independence of traits and fitness, conditional on measured performance

The estimated (fitness-based) average directional and nonlinear selection gradients on the leaf traits are shown for each climate group in Table 3. When growth performance was not modelled, the tendencies revealed by the estimated fitness-based selection gradients were in general compatible with those observed for the performance gradients. This was particularly the case of stomatal length, which indicated statistical support in the dry group for the positive and negative estimates of the average directional and quadratic selection gradients, respectively, as well as for the difference between groups in the magnitude of the average directional selection gradient. In addition, although both groups did not provide statistical support for the effect of the average quadratic selection gradient on SLA when growth performance was not modelled, its estimate was more negative in the wet group (Table 3). This also corresponded to an estimate of the logit quadratic regression coefficient for SLA that was found to be marginally non-significant ($p < 0.10$) in the wet group, but not in the dry group ($p > 0.10$) (Table S4; Appendix A).

When the logistic regression analysis within groups included also a linear term for the height variable, there was no evidence for directional or nonlinear selection to be acting on the leaf traits via their direct effects on tree survival. In this sense, the bootstrapped 95% CIs did not indicate statistical support for the effects of any of the average selection gradients estimated within groups or their differences between groups (Table 3). The latter results on group differences warranted an analysis combining the data across groups, and modelling both linear and nonlinear terms for the height variable. Such species-level analysis also did not provide statistical support for any of the estimates of average selection gradients on the leaf traits combined across groups (Table S5; Appendix A). Overall, these results suggested that the studied fitness component was independent of the focal leaf traits, conditional on the height performance measure. In contrast, statistical support for a direct effect of height performance on tree survival was always detected in either the within- or across-group analyses (e.g. the estimate of the average directional selection gradient combined across groups was 0.922%; bootstrapped 95% CI: 0.638, 1.273).

Table 3

Estimates of fitness-based average linear (directional) and nonlinear (quadratic and correlational) selection gradients, and their difference between the dry and wet climate population groups, for stomatal length and specific leaf area assessed in a common-garden field trial of *Eucalyptus ovata*.

Trait covariate terms	Excluding height performance as a predictor variable			Including height performance as a predictor variable		
	Dry group	Wet group	Group difference	Dry group	Wet group	Group difference
Directional selection						
Stomatal length	1.809 ± 0.812	-0.432 ± 0.671	2.241 ± 1.061	1.055 ± 0.773	-0.294 ± 0.498	1.349 ± 0.923
	(0.063, 3.264)	(−1.757, 0.894)	*(0.143, 4.335)*	(−0.478, 2.516)	(−1.243, 0.727)	(−0.379, 3.234)
Specific leaf area	0.622 ± 0.541	0.229 ± 0.616	0.393 ± 0.821	0.483 ± 0.458	0.340 ± 0.387	0.143 ± 0.600
	(−0.513, 1.609)	(−1.040, 1.358)	(−1.181, 2.034)	(−0.479, 1.355)	(−0.309, 1.242)	(−1.081, 1.272)
Quadratic selection						
Stomatal length^2	-0.209 ± 0.092	-0.119 ± 0.126	-0.090 ± 0.155	-0.154 ± 0.110	-0.079 ± 0.107	-0.075 ± 0.152
	(−0.437, −0.047)	(−0.408, 0.092)	(−0.384, 0.212)	(−0.419, 0.036)	(−0.227, 0.136)	(−0.385, 0.157)
Specific leaf area^2	-0.075 ± 0.064	-0.104 ± 0.082	0.029 ± 0.104	-0.038 ± 0.055	0.022 ± 0.062	-0.060 ± 0.082
	(−0.219, 0.053)	(−0.261, 0.066)	(−0.187, 0.225)	(−0.160, 0.066)	(−0.034, 0.344)	(−0.303, 0.043)
Correlational selection						
Stomatal length x Specific leaf area	0.024 ± 0.071	-0.039 ± 0.075	0.063 ± 0.104	0.021 ± 0.064	-0.023 ± 0.061	0.044 ± 0.089
	(−0.109, 0.181)	(−0.224, 0.088)	(−0.127, 0.286)	(−0.094, 0.168)	(−0.116, 0.148)	(−0.153, 0.207)

The leaf traits (stomatal length and specific leaf area) and the growth performance measure (tree height) were assessed at age 2 years from planting, and tree survival (used as a fitness component) was assessed at age 8 years from planting. Evaluating the fitness consequences of variation in the leaf traits within a climate (dry or wet) group was based on a logistic regression model with tree survival as the response variable and the linear predictor comprising covariate terms for the focal leaf traits, while also excluding or including a linear covariate term for the height performance measure. Modelling the height variable in the linear predictor aims at estimating the (direct) effects of the focal leaf traits on tree survival that are not mediated by (i.e. are independent of) variation in growth performance (e.g. Franklin and Morrissey, 2017). Estimates of (fitness-based) average linear and nonlinear selection gradients, and their standard errors, are presented for the leaf traits within each group, and refer to the percentage change (on average) in expected relative survival caused by increasing the values of the trait covariate terms expressed on a percentage scale (see Methods S3; Appendix A). The average directional selection gradients were firstly derived from the modelling of linear covariate terms for the leaf traits, and then the average quadratic and correlational selection gradients were obtained from the modelling of both linear and nonlinear (i.e. squared and cross-product) covariate terms for the leaf traits. Non-parametric bootstrapping was applied to provide the standard error, as well as the 95% confidence interval (in parentheses), for each estimated average selection gradient within a climate group and its difference between groups (see Methods S4; Appendix A). Statistical support against a null hypothesis being true (i.e. against the absence of an effect for either an average selection gradient within a group or its difference between groups) is given by a 95% confidence interval not overlapping with zero (indicated in Italics). In all cases where a linear covariate term was modelled for height performance, there was always statistical support for a significant average directional selection gradient for the growth performance measure (results not shown).

4. Discussion

4.1. Evidence for performance-based selection on the leaf traits

Performance-based phenotypic selection in our common-garden was suggested in both climate groups. The studied leaf traits explained a similar amount of variation in height performance for the dry (14%) and wet (18%) groups, when using the full model described in Eq. (1). However, the strength and form of performance-based selection depended on the trait and group being considered, with selection mainly detected on stomatal length in the dry group and on SLA in the wet group.

4.1.1. Selection on stomatal length

Selection on leaf hydraulic traits has been previously demonstrated (Carlson et al., 2016; Carvalho et al., 2022; Donovan et al., 2007, 2009; Dudley, 1996). However, few studies have included stomatal size proxies such as stomatal length and, even then, only directional selection has been studied (Costa e Silva et al., 2022; Kosová et al., 2022). In the clonal grass *Festuca rubra*, for example, Kosová et al. (2022) reported environment-specific directional selection on stomatal length acting through one of the fitness proxies studied. Our current study in *E. ovata* provided evidence for contemporary performance-based selection on stomatal length in the dry climate group. This comprised a combination of directional and stabilizing forms of selection, with the former favouring individuals with longer stomata. These two forms of selection can be concurrent (Brooks et al., 2005; Lewis et al., 2011; Mitchell-Olds and Shaw, 1987; Phillips and Arnold, 1989), and for the dry group appeared to indicate an optimal performance value within the data range but somewhat above the observed mean for stomatal length (Fig. 1a) (e.g. see also Walsh and Lynch, 2018, Figure 29.8 D).

The most obvious mechanism by which stomatal size affects plant performance is through stomatal gas conductance, which functionally connects the water transport system of the plant with the rate of photosynthetic assimilation (Brodribb et al., 2020; Drake et al., 2013). At the individual stoma-level, large stomata have greater maximum gas conductance than small stomata (Parlange and Waggoner, 1970). The maximum stomatal conductance determines the maximum rate of leaf carbon gain (a major determinant of the maximum rate of photosynthesis) and water vapour loss (which affects water-use efficiency and transpiration cooling) (Bertolino et al., 2019; Brodribb et al., 2020; Jordan et al., 2020; Lin et al., 2017; Liu et al., 2021; Schymanski et al., 2013). However, at the leaf-level, the greater maximum gas conductance of large stomata is often countered by a reduction in stomatal density (Bertolino et al., 2019; Doheny-Adams et al., 2012; Jordan et al., 2020; Kardiman and Ræbild, 2018; Lawson and Vialet-Chabrand, 2019; Liu et al., 2021). This reduction in stomatal density may occur to the extent that the maximum stomatal conductance at the leaf-level is diminished in large-stoma species (Jordan et al., 2020) or plants (Franks et al., 2009). In such cases, gains may be made in water-use efficiency, but at the expense of the maximum rate of photosynthesis (Bertolino et al., 2019; Drake et al., 2013) and potential for transpiration cooling of the leaf (Lin et al., 2017; Schymanski et al., 2013).

The performance consequences of variation in stomatal length will in part depend upon the relationship between stomatal length and maximum stomatal conductance at the leaf-level. In the present case, using total stomatal length per unit leaf area as a proxy for stomatal conductance at the leaf-level (Prober et al., 2022), we found a weak positive correlation with stomatal length in the dry group ($r = 0.28$; $p = 0.011$) but not in the wet group ($r = -0.04$; $p = 0.741$) (not shown), suggesting that the trait-performance pathway is mediated more through maximum leaf-level stomatal conductance in the dry group.

Increasing rates of photosynthesis arising from greater maximum gas conductance at the leaf-level could explain the positive directional component of performance-based selection acting on stomatal length in the dry group for moderate phenotypic values of the trait. The observed stabilizing component of performance-based selection at higher values of stomatal length may reflect a drop-off in the growth performance of the large-stoma plants due to an overall leaf-level, trade-off between CO_2 uptake and water loss with increasing stomatal conductance (Bertolino et al., 2019; Drake et al., 2013). However, other factors may also have contributed to the observed performance-based stabilizing selection in the dry group, such as the greater construction costs of large stomata (Franks et al., 2009), as well as their possible slower opening and closing response to short-term environment fluctuations (Drake et al., 2013; Kardiman and Ræbild, 2018; Lawson and Vialet-Chabrand, 2019; Raven, 2014), such as variable light (e.g. sunflecks; Raven, 2014; Schymanski et al., 2013) and water deficit (Bertolino et al., 2019; Drake et al., 2013; Fanourakis et al., 2015).

4.1.2. Selection on specific leaf area

Performance-based selection was also suggested for SLA, but of quite different form to that on stomatal length. It was most evident in the wet group where only stabilizing selection favouring an average SLA phenotype was detected (Fig. 1b). Evidence for quadratic selection on SLA has been reported in other studies, but is usually environment specific (Blanco-Sánchez et al., 2022; Lopez-Gallego and O'Neil, 2014; Ramírez-Valiente et al., 2011), and is rarely reported in the absence of a directional component of selection (Blanco-Sánchez et al., 2022; Steinger et al., 2003).

SLA (the inverse of LMA) is a complex trait integrating different components of leaf anatomy and chemistry (John et al., 2017; Niinemets, 2001; Poorter et al., 2009). SLA affects growth performance through the trade-off between light capture, photosynthetic capacity, biomass investment in leaf construction, leaf longevity and relative growth rate (Díaz et al., 2016; Kimball et al., 2013; Poorter et al., 2009; Sartori et al., 2019). High SLA is beneficial in low light environments as leaf area, and thus light interception, is increased per unit leaf biomass, whereas low SLA is beneficial under high light due to greater photosynthetic biomass per unit area (Steinger et al., 2003). Per unit area, low-SLA leaves have greater construction costs than high-SLA leaves, but this is countered by better defence against herbivores and physical hazards leading to greater leaf longevity (Díaz et al., 2016; Poorter et al., 2009). High SLA tends to be favoured in productive, high-resource, habitats, whereas low SLA is favoured in more stressful environments, including those subjected to drought and freezing stress (Anderegg et al., 2021; Ball et al., 2002; McLean et al., 2014; Poorter et al., 2009; Ramírez-Valiente et al., 2022; Ramírez-Valiente et al., 2011; Sartori et al., 2019). However, our dry group of populations had higher SLA than the wet group (Table S2), which suggests that factors other than drought adaptation (e.g. biotic stressors, Barton and Koricheva, 2010; genetic constraints, Costa e Silva et al., 2020) may have driven the evolution of group differences. The pure performance-based stabilizing selection we observed in the wet group suggested that the SLA phenotype which optimised growth performance of the group in the test environment was close to the observed SLA mean (i.e. selection does not lead to a change in the SLA mean; e.g. see Walsh and Lynch, 2018, Figure 29.8 C). Reduced growth performance of extreme SLA phenotypes could be due to multiple genetic factors (Hendry, 2017), but may well signal the trade-offs involved in over- and under-investment of resources for leaf construction at the expense of whole plant growth and leaf defence, respectively (Poorter et al., 2009; Salgado-Luarte and Gianoli, 2012, 2017).

4.1.3. Correlational selection

Correlational selection has been documented in natural populations for many organisms in both plant and animal species (Geber and Griffen, 2003; Svensson et al., 2021). Correlational selection can signal trait synergies and trade-offs, and can shape the genetic and phenotypic architecture of traits over evolutionary time (Roff and Fairbairn, 2007, 2012; Svensson et al., 2021). Significant correlational selection involving SLA (or LMA) has been detected in other studies (e.g. Etterson, 2004; Magnoli and Lau, 2020; Sartori et al., 2019). Despite the potential functional roles of stomatal and leaf economic traits in photosynthesis and water-use efficiency, in our case, the performance-based correlational selection estimate between SLA and stomatal length was close to zero and not statistically significant in either group.

Using fitness components or measures of plant growth as fitness proxies, the absence of evidence for correlational selection involving SLA has been reported not only for pairwise combinations where SLA is functionally distally related to the other interacting trait (Lopez-Gallego and O'Neil, 2014; Steinger et al., 2003), but also for pairwise combinations where SLA interacts with another leaf economic trait (Ramírez-Valiente et al., 2011). Such results may reflect the need of large sample sizes for an accurate estimation and detection of correlational selection (Simon et al., 2022; Walsh and Lynch, 2018). However, they could signal true independence of the physiological mechanisms by which variation in the focal leaf traits affects growth performance. In the present case, such independence could arise from stomatal length impacting external gas exchange, but SLA affecting internal characteristics such as gas diffusion within the leaf, the amount of chlorophyll per unit leaf area, as well as biotic and abiotic defence (Niinemets, 2001; Poorter et al., 2009).

4.1.4. Differences between the dry and wet climate groups in performance-based selection

The modelling of the fitness-performance and fitness-trait relationships indicated that qualitative comparisons of performance-based selection of the two groups were justified (see below the discussion on the "Relationship of fitness with growth performance and/or leaf traits"). However, estimates of both linear and nonlinear selection gradients are influenced by numerous factors, including the study environment and the phenotypic range of the focal traits (Geber and Griffen, 2003; Hendry, 2017; Walsh and Lynch, 2018). Indeed, group comparisons of selection gradients will rely upon the assumption that the phenotypic (co)variances of the focal traits do not differ substantially among groups (Chenoweth et al., 2013). This was the case for the phenotypic variances of the studied leaf traits (Table S2), and also as indicated from the Box's *M*-test (Box, 1949) which could not reject (i.e. $p > 0.05$) the null hypothesis of homogeneity of phenotypic variance-covariance matrices between the climate groups for the two leaf traits (not shown). In addition, environmental covariance between fitness and traits may induce bias in estimates of phenotypic selection gradients (Stinchcombe et al., 2002). Nevertheless, the modelling of experimental block effects in Eq. (1) may have reduced the influence of the environmental covariance between height performance and the leaf traits on our performance gradient estimates (McGoey and Stinchcombe, 2009). Regardless, any estimation bias would equally affect the two groups as the individual genotypes representing the groups were randomly distributed within blocks.

The differences in performance-based selection observed between the dry and wet groups for both SLA and stomatal length may partly reflect hidden interactions arising from selection acting on unmeasured traits which affect growth performance (Delph et al., 2011; Ludwig et al., 2004). Group differences in trait plasticity could also contribute to differences in performance-based selection (Kosová et al., 2022; Steinger et al., 2003). Genetic differences in trait plasticity may have altered the effects of a trait on the growth performance measure and/or the observed fitness component, as suggested by the study of Walter et al. (2023). Further, trait plasticity itself may incur a fitness cost (Steinger et al., 2003) and even be constrained by unmeasured correlated traits (Franks et al., 2009).

In the present study, plasticity may have changed the performance-trait relationship, diminishing the effects of SLA on height

performance in our dry but not the wet climate group. SLA is a highly plastic trait (Niinemets, 2001; Poorter et al., 2009; Steinger et al., 2003), and this plasticity may vary dependent on genotype. For example, a study of *Eucalyptus tricarpa* populations grown in common-gardens in regions of low (470 mm) and high (840 mm) annual rainfall revealed that SLA plasticity was greater in populations originating from a dry climate compared with those originating from wetter regions (McLean et al., 2014). This finding is consistent with past selection for greater trait plasticity in populations from more variable dry home-site environments (Lázaro-Nogal et al., 2015), which may mask the effects of the focal traits on performance or fitness (Hendry, 2015).

For stomatal length, the observed group difference may be due to a disconnect of stomatal size from variation in leaf-level maximum stomatal conductance in the wet group, leaving only a (weak) signal of the performance-trait pathway which may have favoured smaller stomata in the wet group due to their lower construction costs or faster response to environmental fluctuations (as discussed above). These different ways by which stomatal size may affect gas exchange, and ultimately growth performance and fitness, could explain the opposing sign of the dry and wet groups for performance-based (Table 1) and fitness-based (Table 3) estimates of directional selection on stomatal length. However, while these directional selection estimates statistically differed between the two studied groups, as for functional traits in other studies (e.g. Alexandre et al., 2020; Costa e Silva et al., 2021), the groups were not replicated to allow a broader inference (Chenoweth et al., 2013).

4.2. Relationship of fitness with growth performance and/or leaf traits

The current work provided statistical support for a positive direct effect of height growth on tree survival in both *E. ovata* climate groups. Given such a positive fitness-performance relationship, the modelling of the fitness-trait relationship further indicated that, in our studied system, the influence of the focal leaf traits on the observed fitness component could occur via (or mediated by) their effects on height performance, under the environmental conditions of the trial site and time period covered in the current study.

A linear fitness-performance relationship is assumed in both performance-based and Arnold's (1983, 2003) performance-mediated analyses of selection (Franklin and Morrissey, 2017). Under the logistic regression model applied here, the relationship of survival on the expected probability scale with height performance is intrinsically nonlinear, reflecting the cumulative distribution function of the logistic distribution. Yet, when allowing for a nonlinear, monotonically increasing fitness-performance function in a performance-mediated analysis, Franklin and Morrissey (2017) found selection gradients that agreed quite closely with corresponding estimates obtained from a performance-mediated analysis that assumed a linear fitness-performance relationship, with both approaches leading to similar biological interpretations. Thus, as long as the fitness-performance relationship is monotonic, it may also be expected that qualitative inferences of selection from performance-based estimators of selection [e.g. as provided by Eq. (1)] will not be largely affected by the error associated with the fitness-performance relationship not meeting the assumption of linearity. In this sense, performance gradients can be representative of major qualitative features (such as the presence of minima or maxima) of the true selection surface describing the relationship between relative fitness and the focal traits (Franklin and Morrissey, 2017; Franklin et al., 2018). This may apply to our *E. ovata* system, as the estimated average selection gradients on the height variable in either climate group (Table 2) indicated prevalent monotonically increasing (sigmoid) functions relating survival on the expected probability scale to growth performance over the observed phenotypic range (Fig. 2). In addition, the modelling of the fitness-trait relationship revealed that the leaf traits studied were independent of tree survival, conditional on the height performance measure (Table 3 and Table S5). That is, there was no evidence indicating that the leaf traits had direct fitness consequences independent of variation in

growth performance, hence suggesting that the focal traits appeared to affect tree survival indirectly through their effects on height performance. Overall, these results suggested that the performance gradients estimated for the leaf traits under Eq. (1) could be qualitatively indicative of selection gradients, hence validating qualitative inferences of selection from the performance-based analysis (Franklin and Morrissey, 2017).

Over the two- to eight-year period studied in *E. ovata*, faster growth clearly had fitness benefits, with mortality being size-dependent, consistent with the fitness-performance relationship reported for *E. pauciflora* planted at the same time and site (Costa e Silva et al., 2022). While the positive size-dependent mortality is commonly reported in forest trees, including eucalypts (see discussion in Costa e Silva et al., 2022), context is important as in any relationship involving fitness (or a fitness component) and the focal characteristics (Walsh and Lynch, 2018). Our results on the modelling of relationships involving tree survival as a response variable pertained to a post-establishment, early-age fitness component in the life-cycle of a long-lived forest tree species. They also encompassed a growth period where no extreme experiment-wide abiotic stress events occurred (e.g. drought, frost or fire), with mortality gradually accumulating through time consistent with biotic causes, including inter-tree competition (Costa e Silva et al., 2017).

Given the growth period studied, we found no evidence to suggest that the overall survival and mean height performance, as well as the results on selection parameters, were predictable from the long-term climate differences between the home-sites of the source populations and the common-garden site. The long-term climate differences suggest that the wet group could have been more maladapted to the common-garden site than the dry group. However, over the growth period studied, the dry group had lower survival and exhibited greater directional selection than the wet group, as evidenced by both the fitness-based (Table 3) and performance-based (Table 1) estimators of selection on stomatal length. Greater mean annual rainfall during the studied growth period than the long-term average is unlikely to explain these discrepancies, as the region has experienced a hotter-drier climate since the common-garden was planted, potentially reflecting anthropogenic climate change (Harrison, 2021). These discrepancies may have been caused by multiple factors such as group differences in survival being due to inbreeding depression arising from maternal outcrossing variation (Hardner et al., 1996), and selection estimates being possibly influenced by changed site conditions due to agricultural use or cultivation for planting (Davidson et al., 2021), which may increase water availability in the soil (Mendham et al., 2011). The latter may initially reduce water-deficit stress at the common-garden site and contribute to the stronger and positive early-age directional selection on stomatal length (favouring large stomata) in the dry compared with the wet group.

Variation in fitness-based selection gradients on growth performance and/or on the focal leaf traits may also occur at later ages in our *E. ovata* restoration planting, although evidence for frequent temporal changes in the strength and sign of directional selection in natural populations may be limited (Kingsolver et al., 2012; Morrissey and Hadfield, 2012). Changes in the directional selection gradient on height growth, for example, may result from early-age trade-offs in resource allocation being altered between above- as opposed to below-ground growth, as well as between growth and biotic/abiotic defences (Costa e Silva et al., 2021; Ferrenberg et al., 2023; Orians et al., 2010). The changing endogenous stress landscape as trees grow in height (e.g. biotic defence, Barton and Koricheva, 2010; water transport, McDowell et al., 2011) may also impact on the variation in selection gradient estimates on growth performance. In addition, temporal changes in natural selection may occur with short-term events (e.g. climate fluctuations, Blanco-Sánchez et al., 2022; Siepielski et al., 2009). For example, in contrast to the present finding, higher mortality of well-established larger trees compared with smaller trees is often reported in response to major selection events such as high winds (Peterson, 2000), droughts (Stovall

et al., 2019) and insect outbreaks (Pfeifer et al., 2011), although in the case of eucalypts the reverse is more likely for wildfires (Hernández et al., 2022). With predictions of increases in the frequency of such extreme selection events with climate change (Australian Bureau of Meteorology, 2020), combined with increasing depletion of soil water reserves as the stand ages (Mendham et al., 2011), these more-extreme selection events may well have a greater impact on the relationships of later-age fitness components with functional traits or growth performance.

5. Conclusions

Our study in a common-garden field trial of the forest tree *E. ovata* provided evidence for early-age, performance-based phenotypic selection on two leaf functional traits, stomatal length and specific leaf area (SLA), the former being a proxy of stomatal size and rarely included in selection studies despite its importance in plant evolution. The strength and form of performance-based selection depended on the trait and climate group (dry or wet) being considered, with selection mainly detected on stomatal length in the dry group and on SLA in the wet group. In this context, the combination of directional and stabilizing selection on stomatal length in the dry group suggested an optimal performance value slightly above the observed trait mean, whereas a performance maximum at the mean of SLA was indicated by a pure stabilizing selection on the trait in the wet group. Under the environmental conditions of the trial site and over the time period studied, fitness-based selection gradients on height performance indicated statistical support for directional selection but not for quadratic selection in the two population groups of different climate origin. The modelling of these fitness-performance relationships suggested that mortality was size-dependent, with early-age height growth having a consistent positive effect on the later-age tree survival. Further, fitness-based selection gradients did not provide statistical support for direct effects of the focal traits on the observed fitness component, conditional on measured performance. These results validated qualitative inferences of selection from the performance-based analysis, and indicated that selection on the focal leaf traits appeared to be mediated by their effects on early-age height performance, which in turn directly influenced later-age tree survival. Nevertheless, temporal factors are likely to have an impact on inferences about natural selection, and longer-term monitoring will be required to determine the stability of both performance-based and fitness-based estimators of phenotypic selection across the life cycle of these long-lived forest trees.

CRediT authorship contribution statement

João Costa e Silva (JCS) conceived and conceptualised the research work undertaken in this study. Brad M. Potts (BMP) and Suzanne M. Prober (SMP) were responsible for planning and coordinating the common-garden field trial, sampling design and trait measurements. JCS performed the data analyses and prepared the associated document provided in Appendix A, except for the case of the analysis on the allocation of the *Eucalyptus ovata* populations into two climate groups (Methods S1) that was undertaken by BMP. JCS and BMP wrote the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data can be requested for consideration from the corresponding author.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ppees.2023.125765.

References

Alexandre, H., Truffaut, L., Klein, E., Ducousso, A., Chancerel, E., Lesur, I., Dencausse, B., Louvet, J.-M., Nepveu, G., Torres-Ruiz, J.M., Lagane, F., Musch, B., Delzon, S., Kremer, A., 2020. How does contemporary selection shape oak phenotypes? Evol. Appl. 13, 2772–2790.

Anderegg, L.D., Loy, X., Markham, I.P., Elmer, C.M., Hovenden, M.J., HilleRisLambers, J., Mayfield, M.M., 2021. Aridity drives coordinated trait shifts but not decreased trait variance across the geographic range of eight Australian trees. N. Phytol. 229, 1375–1387.

Arnold, S.J., 1983. Morphology, performance and fitness. Am. Zool. 23, 347–361.

Arnold, S.J., 2003. Performance surfaces and adaptive landscapes. Integr. Comp. Biol. 43, 367–375.

Australian Bureau of Meteorology, 2020. State of the Climate 2020. Commonwealth Scientific and Industrial Research Organisation [CSIRO].

Bailey, T.G., Harrison, P.A., Davidson, N.J., Weller-Wong, A., Tilyard, P., Steane, D.A., Vaillancourt, R.E., Potts, B.M., 2021. Embedding genetics experiments in restoration to guide plant choice for a degraded landscape with a changing climate. Ecol. Manag. Restor. 22, 92–105.

Ball, M.C., Wolfe, J., Canny, M., Hofmann, M., Nicotra, A.B., Hughes, D., 2002. Space and time dependence of temperature and freezing in evergreen leaves. Funct. Plant Biol. 29, 1259–1272.

Barton, K.E., Koricheva, J., 2010. The ontogeny of plant defense and herbivory: characterizing general patterns using meta-analysis. Am. Nat. 175, 481–493.

Bertolino, L.T., Caine, R.S., Gray, J.E., 2019. Impact of stomatal density and morphology on water-use efficiency in a changing world. Front. Plant Sci. 10, 225.

Blanco-Sánchez, M., Ramos-Muñoz, M., Pías, B., Ramírez-Valiente, J.A., Díaz-Guerra, L., Escudero, A., Matesanz, S., 2022. Natural selection favours drought escape and an acquisitive resource-use strategy in semi-arid Mediterranean shrubs. Funct. Ecol. 36, 2289–2302.

Blows, M.W., Brooks, R., Kraft, P.G., 2003. Exploring complex fitness surfaces: multiple ornamentation and polymorphism in male guppies. Evolution 57, 1622–1630.

Box, G.E.P., 1949. A general distribution theory for a class of likelihood criteria. Biometrika 36, 317–346.

Brodribb, T.J., Sussmilch, F., McAdam, S.A.M., 2020. From reproduction to production, stomata are the master regulators. Plant J. 101, 756–767.

Brooks, R., Hunt, J., Blows, M.W., Smith, M.J., Bussiére, L.F., Jennions, M.D., 2005. Experimental evidence for multivariate stabilizing sexual selection. Evolution 59, 871–880.

Carlson, J.E., Adams, C.A., Holsinger, K.E., 2016. Intraspecific variation in stomatal traits, leaf traits and physiology reflects adaptation along aridity gradients in a South African shrub. Ann. Bot. 117, 195–207.

Caruso, C.M., Maherali, H., Martin, R.A., 2020. A meta-analysis of natural selection on plant functional traits. Int. J. Plant Sci. 181, 44–55.

Carvalho, C., Davis, R., Connallon, T., Gleadow, R.M., Moore, J.L., Uesugi, A., 2022. Multivariate selection mediated by aridity predicts divergence of drought-resistant traits along natural aridity gradients of an invasive weed. N. Phytol. 234, 1088–1100.

Chenoweth, S.F., Hunt, J., Rundle, H.D., 2013. Analyzing and comparing the geometry of individual fitness surfaces. In: Svensson, E., Calsbeek, R. (Eds.), The Adaptive Landscape in Evolutionary Biology (online edition). Oxford Academic, Oxford, pp. 126–149.

Colautti, R.I., Lau, J.A., 2015. Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. Mol. Ecol. 24, 1999–2017.

Costa e Silva, J., Potts, B.M., Gilmour, A.R., Kerr, R.J., 2017. Genetic-based interactions among tree neighbors: identification of the most influential neighbors, and estimation of correlations among direct and indirect genetic effects for leaf disease and growth in *Eucalyptus globulus*. Heredity 119, 125–135.

Costa e Silva, J., Jordan, R., Potts, B.M., Pinkard, E., Prober, S.M., 2021. Directional selection on tree seedling traits driven by experimental drought differs between mesic and dry populations. Front. Ecol. Evol. 9, 722964.

Costa e Silva, J., Potts, B.M., Harrison, P.A., 2020. Population divergence along a genetic line of least resistance in the tree species *Eucalyptus globulus*. Genes 11, 1095.

Costa e Silva, J., Potts, B.M., Wiehl, G., Prober, S.M., 2022. Linking leaf economic and hydraulic traits with early-age growth performance and survival of *Eucalyptus pauciflora*. Front. Plant Sci. 13, 973087.

Courvoisier, D.S., Combescure, C., Agoritsas, T., Gayet-Ageron, A., Perneger, T.V., 2011. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure. J. Clin. Epidemiol. 64, 993–1000.

Davidson, N.J., Bailey, T.G., Burgess, S., Potts, B.M., 2021. New approaches for revegetating agricultural landscapes to provide connectivity for wildlife: The example of the Tasmanian Midlands, Australia. Ecol. Manag. Restor. 22, 47–60.

Delph, L.F., Andicoechea, J., Steven, J.C., Herlihy, C.R., Scarpino, S.V., Bell, D.L., 2011. Environment-dependent intralocus sexual conflict in a dioecious plant. N. Phytol. 192, 542–552.

Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Colin Prentice, I., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P.B., Moles, A.T., Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Joseph Wright, S., Sheremet'ev, S.N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M.D., Gorné, L.D., 2016. The global spectrum of plant form and function. Nature 529, 167–171.

Doheny-Adams, T., Hunt, L., Franks, P.J., Beerling, D.J., Gray, J.E., 2012. Genetic manipulation of stomatal density and size, plant growth and tolerance to restricted water supply across a growth carbon dioxide gradient. Philos. Trans. R. Soc. B: Biol. Sci. 367, 547–555.

Donovan, L.A., Dudley, S.A., Rosenthal, D.M., Ludwig, F., 2007. Phenotypic selection on leaf water use efficiency and related ecophysiological traits for natural populations of desert sunflowers. Oecologia 152, 13–25.

Donovan, L.A., Ludwig, F., Rosenthal, D.M., Rieseberg, L.H., Dudley, S.A., 2009. Phenotypic selection on leaf ecophysiological traits in *Helianthus*. N. Phytol. 183, 868–879.

Drake, P.L., Froend, R.H., Franks, P.J., 2013. Smaller, faster stomata: scaling of stomatal size, rate of response, and stomatal conductance. J. Exp. Bot. 64, 495–505.

Dudley, S.A., 1996. Differing selection on plant physiological traits in response to environmental water availability: a test of adaptive hypotheses. Evolution 50, 92–102.

Etterson, J.R., 2004. Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. I. Clinal patterns of selection along an environmental gradient in the Great Plains. Evolution 58, 1446–1458.

Fanourakis, D., Giday, H., Milla, R., Pieruschka, R., Kjaer, K.H., Bolger, M., Vasilevski, A., Nunes-Nesi, A., Fiorani, F., Ottosen, C.-O., 2015. Pore size regulates operating stomatal conductance, while stomatal densities drive the partitioning of conductance between leaf sides. Ann. Bot. 115, 555–565.

Ferrenberg, S., Vázquez-González, C., Lee, S.R., Kristupaitis, M., 2023. Divergent growth-differentiation balance strategies and resource competition shape mortality patterns in ponderosa pine. Ecosphere 14, e4349.

Firth, D., 1993. Bias reduction of maximum likelihood estimates. Biometrika 80, 27–38.

Franklin, O.D., Morrissey, M.B., 2017. Inference of selection gradients using performance measures as fitness proxies. Methods Ecol. Evol. 8, 663–677.

Franklin, O.D., Skúlason, S., Morrissey, M.B., Ferguson, M.M., 2018. Natural selection for body shape in resource polymorphic Icelandic Arctic charr. J. Evol. Biol. 31, 1498–1512.

Franks, P.J., Beerling, D.J., 2009. Maximum leaf conductance driven by CO_2 effects on stomatal size and density over geologic time. Proc. Natl. Acad. Sci. 106, 10343–10347.

Franks, P.J., Drake, P.L., Beerling, D.J., 2009. Plasticity in maximum stomatal conductance constrained by negative correlation between stomatal size and density: an analysis using *Eucalyptus globulus*. Plant Cell Environ. 32, 1737–1748.

Geber, M.A., Griffen, L.R., 2003. Inheritance and natural selection on functional traits. Int. J. Plant Sci. 164, S21–S42.

Gianoli, E., Saldaña, A., 2013. Phenotypic selection on leaf functional traits of two congeneric species in a temperate rainforest is consistent with their shade tolerance. Oecologia 173, 13–21.

Hardner, C.M., Vaillancourt, R.E., Potts, B.M., 1996. Stand density influences outcrossing rate and growth of open-pollinated families of *Eucalyptus globulus*. Silvae Genet 45, 226–228.

Harrison, E.L., Cubas, Arce, Gray, L., Hepworth, C, J.E., 2020. The influence of stomatal morphology and distribution on photosynthetic gas exchange. Plant J. 101, 768–779.

Harrison, P., 2017. Integrating climate change into conservation and restoration strategies: the case of the Tasmanian eucalypts. University of Tasmania. ⟨https://eprints.utas.edu.au/27322/⟩.

Harrison, P.A., 2021. Climate change and the suitability of local and non-local species for ecosystem restoration. Ecol. Manag. Restor. 22, 75–91.

Heinze, G., Ploner, M., 2004. A SAS macro, S-PLUS library and R package to perform logistic regression without convergence problems. Section of Clinical Biometrics, Department of Medical Computer Sciences, Medical University of Vienna, Vienna. Technical Report 2/2004.

Heinze, G., Schemper, M., 2002. A solution to the problem of separation in logistic regression. Stat. Med. 21, 2409–2419.

Hendry, A.P., 2015. Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. J. Hered. 107, 25–41.

Hendry, A.P., 2017. Eco-evolutionary Dynamics. Princeton University Press, Oxford, U.K.

Hernández, M.A., Butler, J.B., Ammitzboll, H., Freeman, J.S., O'Reilly-Wapstra, J., Vaillancourt, R.E., Potts, B.M., 2022. Genetic variation in fire recovery and other fire-related traits in a global eucalypt species. Tree Genet. Genom. 18, 42.

Houle, D., Pélabon, C., Wagner, G.P., Hansen, T.F., 2011. Measurement and meaning in biology. Q Rev. Biol. 86, 3–34.

Janzen, F.J., Stern, H.S., 1998. Logistic regression for empirical studies of multivariate selection. Evolution 52, 1564–1571.

John, G.P., Scoffoni, C., Buckley, T.N., Villar, R., Poorter, H., Sack, L., 2017. The anatomical and compositional basis of leaf mass per area. Ecol. Lett. 20, 412–425.

Jordan, G.J., Carpenter, R.J., Koutoulis, A., Price, A., Brodribb, T.J., 2015. Environmental adaptation in stomatal size independent of the effects of genome size. N. Phytol. 205, 608–617.

Jordan, G.J., Carpenter, R.J., Holland, B.R., Beeton, N.J., Woodhams, M.D., Brodribb, T. J., 2020. Links between environment and stomatal size through evolutionary time in Proteaceae. Proc. R. Soc. B 287, 20192876.

Kardiman, R., Ræbild, A., 2018. Relationship between stomatal density, size and speed of opening in Sumatran rainforest species. Tree Physiol. 38, 696–705.

Kimball, S., Gremer, J.R., Huxman, T.E., Venable, D.L., Angert, A.L., 2013. Phenotypic selection favors missing trait combinations in coexisting annual plants. Am. Nat. 182, 191–207.

Kingsolver, J.G., Huey, R.B., 2003. Introduction: The evolution of morphology, performance, and fitness. Integr. Comp. Biol. 43, 361–366.

Kingsolver, J.G., Diamond, S.E., Siepielski, A.M., Carlson, S.M., 2012. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. Evol. Ecol. 26, 1101–1118.

Kosová, V., Hájek, T., Hadincová, V., Münzbergová, Z., 2022. The importance of ecophysiological traits in response of *Festuca rubra* to changing climate. Physiol. Plant. 174, e13608.

Lande, R., Arnold, S.J., 1983. The measurement of selection on correlated characters. Evolution 37, 1210–1226.

Lawson, T., Vialet-Chabrand, S., 2019. Speedy stomata, photosynthesis and plant water use efficiency. N. Phytol. 221, 93–98.

Lázaro-Nogal, A., Matesanz, S., Godoy, A., Pérez-Trautman, F., Gianoli, E., Valladares, F., 2015. Environmental heterogeneity leads to higher plasticity in dry-edge populations of a semi-arid Chilean shrub: insights into climate change responses. J. Ecol. 103, 338–350.

Lewis, Z., Wedell, N., Hunt, J., 2011. Evidence for strong intralocus sexual conflict in the Indian meal moth, *Plodia interpunctella*. Evolution 65, 2085–2097.

Lin, H., Chen, Y., Zhang, H., Fu, P., Fan, Z., 2017. Stronger cooling effects of transpiration and leaf physical traits of plants from a hot dry habitat than from a hot wet habitat. Funct. Ecol. 31, 2202–2211.

Liu, C., Muir, C.D., Li, Y., Xu, L., Li, M., Zhang, J., de Boer, H.J., Sack, L., Han, X., Yu, G., He, N., 2021. Scaling between stomatal size and density in forest plants. bioRxiv, 2021.2004.2025.441252.

Lopez-Gallego, C., O'Neil, P., 2014. Genetic variation and the potential response to selection on leaf traits after habitat degradation in a long-lived cycad. Evol. Ecol. 28, 775–791.

Ludwig, F., Rosenthal, D.M., Johnston, J.A., Kane, N., Gross, B.L., Lexer, C., Dudley, S.A., Rieseberg, L.H., Donovan, L.A., 2004. Selection on leaf ecophysiological traits in a desert hybrid *Helianthus* species and early-generation hybrids. Evolution 58, 2682–2692.

Magnoli, S.M., Lau, J.A., 2020. Evolution in novel environments: Do restored prairie populations experience strong selection? Ecology 101, e03120.

McDowell, N.G., Bond, B.J., Dickman, L.T., Ryan, M.G., Whitehead, D., 2011. Relationships between tree height and carbon isotope discrimination. In: Meinzer, F. C., Lachenbruch, B., Dawson, T.E. (Eds.), Size- and Age-Related Changes in Tree Structure and Function. Springer, New York, NY, USA, pp. 255–286.

McGoey, B.V., Stinchcombe, J.R., 2009. Interspecific competition alters natural selection on shade avoidance phenotypes in *Impatiens capensis*. N. Phytol. 183, 880–891.

McLean, E.H., Prober, S.M., Stock, W.D., Steane, D.A., Potts, B.M., Vaillancourt, R.E., Byrne, M., 2014. Plasticity of functional traits varies clinally along a rainfall gradient in *Eucalyptus tricarpa*. Plant Cell Environ. 37, 1440–1451.

Mendham, D.S., White, D.A., Battaglia, M., McGrath, J.F., Short, T.M., Ogden, G.N., Kinal, J., 2011. Soil water depletion and replenishment during first- and early second-rotation *Eucalyptus globulus* plantations with deep soil profiles. Agricultural and Forest Meteorology 151, 1568–1579.

Mitchell-Olds, T., Shaw, R.G., 1987. Regression analysis of natural selection: statistical inference and biological interpretation. Evolution 41, 1149–1161.

Morrissey, M.B., Hadfield, J.D., 2012. Directional selection in temporally replicated studies is remarkably consistent. Evolution 66, 435–442.

Morrissey, M.B., Sakrejda, K., 2013. Unification of regression-based methods for the analysis of natural selection. Evolution 67, 2094–2100.

Nemes, S., Jonasson, J.M., Genell, A., Steineck, G., 2009. Bias in odds ratios by logistic regression modelling and sample size. BMC Med. Res. Methodol. 9, 56.

Niinemets, Ü., 2001. Global-scale climatic controls of leaf dry mass per area, density, and thickness in trees and shrubs. Ecology 82, 453–469.

Opedal, Ø., 2021. A functional view reveals substantial predictability of pollinator-mediated selection. J. Pollinat. Ecol. 30, 273–288.

Orians, C.M., Hochwender, C.G., Fritz, R.S., Snäll, T., 2010. Growth and chemical defense in willow seedlings: trade-offs are transient. Oecologia 163, 283–290.

Parlange, J.-Y., Waggoner, P.E., 1970. Stomatal dimensions and resistance to diffusion. Plant Physiol. 46, 337–342.

Peterson, C.J., 2000. Catastrophic wind damage to North American forests and the potential impact of climate change. Sci. Total Environ. 262, 287–311.

Pfeifer, E.M., Hicke, J.A., Meddens, A.J.H., 2011. Observations and modeling of aboveground tree carbon stocks and fluxes following a bark beetle outbreak in the western United States. Glob. Change Biol. 17, 339–350.

Phillips, P.C., Arnold, S.J., 1989. Visualizing multivariate selection. Evolution 43, 1209–1222.

Poorter, H., Niinemets, Ü., Poorter, L., Wright, I.J., Villar, R., 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. N. Phytol. 182, 565–588.

Prober, S.M., Potts, B.M., Harrison, P.A., Wiehl, G., Bailey, T.G., Costa e Silva, J., Price, M.R., Speijers, J., Steane, D.A., Vaillancourt, R.E., 2022. Leaf economic and hydraulic traits signal disparate climate adaptation patterns in two co-occurring woodland eucalypts. Plants 11, 1846.

Puhr, R., Heinze, G., Nold, M., Lusa, L., Geroldinger, A., 2017. Firth's logistic regression with rare events: accurate effect estimates and predictions? Stat. Med. 36, 2302–2317.

Ramírez-Valiente, J.A., Valladares, F., Huertas, A.D., Granados, S., Aranda, I., 2011. Factors affecting cork oak growth under dry conditions: local adaptation and contrasting additive genetic variance within populations. Tree Genet. Genom. 7, 285–295.

Ramírez-Valiente, J.A., Valladares, F., Sanchéz-Goméz, D., Delgado, A., Aranda, I., 2014. Population variation and natural selection on leaf traits in cork oak throughout its distribution range. Acta Oecol 58, 49–56.

Ramírez-Valiente, J.A., Santos del Blanco, L., Alía, R., Robledo-Arnuncio, J.J., Climent, J., 2022. Adaptation of Mediterranean forest species to climate: Lessons from common garden experiments. J. Ecol. 110, 1022–1042.

Raven, J.A., 2014. Speedy small stomata? J. Exp. Bot. 65, 1415–1424.

Reich, P.B., 2014. The world-wide 'fast-slow' plant economics spectrum: a traits manifesto. J. Ecol. 102, 275–301.

Roff, D.A., Fairbairn, D.J., 2007. The evolution of trade-offs: where are we? J. Evol. Biol. 20, 433–447.

Roff, D.A., Fairbairn, D.J., 2012. A test of the hypothesis that correlational selection generates genetic correlations. Evolution 66, 2953–2960.

Salgado-Luarte, C., Gianoli, E., 2012. Herbivores modify selection on plant functional traits in a temperate rainforest understory. Am. Nat. 180, E42–E53.

Salgado-Luarte, C., Gianoli, E., 2017. Shade tolerance and herbivory are associated with RGR of tree species via different functional traits. Plant Biol. 19, 413–419.

Sartori, K., Vasseur, F., Violle, C., Baron, E., Gerard, M., Rowe, N., Ayala-Garay, O., Christophe, A., Garcia de Jalón, L., Masclef, D., Harscouet, E., Granado, Md.R., Chassagneux, A., Kazakou, E., Vile, D., 2019. Leaf economics and slow-fast adaptation across the geographic range of *Arabidopsis thaliana*. Sci. Rep. 9, 10758.

SAS, 2017. SAS/STAT® 14.3 User's Guide. SAS Institute Inc, Cary, NC.

Satterthwaite, F.E., 1946. An approximate distribution of estimates of variance components. Biom. Bull. 2, 110–114.

Schymanski, S.J., Or, D., Zwieniecki, M., 2013. Stomatal control and leaf thermal and hydraulic capacitances under rapid environmental fuctuations. PLoS One 8, e54231.

Siepielski, A.M., DiBattista, J.D., Carlson, S.M., 2009. It's about time: the temporal dynamics of phenotypic selection in the wild. Ecol. Lett. 12, 1261–1276.

Simon, M.N., Marroig, G., Arnold, S.J., 2022. Detecting patterns of correlational selection with sampling error: a simulation study. Evolution 76, 207–224.

Steel, R., Torrie, J., Dicky, D., 1997. Principles and Procedures of Statistics: A Biometrical Approach, 3rd edition. McGraw Hill, Inc. Book Co., New York, NY.

Steinger, T., Roy, B.A., Stanton, M.L., 2003. Evolution in stressful environments II: adaptive value and costs of plasticity in response to low light in *Sinapis arvensis*. J. Evol. Biol. 16, 313–323.

Stinchcombe, J.R., Rutter, M.T., Burdick, D.S., Tiffin, P., Rausher, M.D., Mauricio, R., 2002. Testing for environmentally induced bias in phenotypic estimates of natural selection: theory and practice. Am. Nat. 160, 511–523.

Stinchcombe, J.R., Agrawal, A.F., Hohenlohe, P.A., Arnold, S.J., Blows, M.W., 2008. Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? Evolution 62, 2435–2440.

Stovall, A.E., Shugart, H., Yang, X., 2019. Tree height explains mortality risk during an intense drought. Nat. Commun. 10, 4385.

Svensson, E.I., Arnold, S.J., Bürger, R., Csilléry, K., Draghi, J., Henshaw, J.M., Jones, A.G., De Lisle, S., Marques, D.A., McGuigan, K., Simon, M.N., Runemark, A., 2021. Correlational selection in the age of genomics. Nat. Ecol. Evol. 5, 562–573.

van Calster, B., van Smeden, M., De Cock, B., Steyerberg, E.W., 2020. Regression shrinkage methods for clinical prediction models do not guarantee improved performance: simulation study. Stat. Methods Med. Res. 29, 3166–3178.

van Smeden, M., de Groot, J.A.H., Moons, K.G.M., Collins, G.S., Altman, D.G., Eijkemans, M.J.C., Reitsma, J.B., 2016. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. BMC Med. Res. Methodol. 16, 163.

Veselý, P., Šmarda, P., Bureš, P., Stirton, C., Muasya, A.M., Mucina, L., Horová, L., Veselá, K., Šilerová, A., Šmerda, J., Knápek, O., 2020. Environmental pressures on stomatal size may drive plant genome size evolution: evidence from a natural experiment with Cape geophytes. Ann. Bot. 126, 323–330.

Wahba, G., 1990. Spline Models for Observational Data, CBMS-NSF regional conference series in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia.

Walsh, B., Lynch, M., 2018. Evolution and Selection of Quantitative Traits. Oxford University Press, Oxford, UK.

Walter, G.M., Clark, J., Terranova, D., Cozzolino, S., Cristaudo, A., Hiscock, S.J., Bridle, J., 2023. Hidden genetic variation in plasticity provides the potential for rapid adaptation to novel environments. N. Phytol. 239, 374–387.