## Methods

## Model

For each of n species, we assumed that data were generated from the following sampling distribution:

$$y_j \sim \mathcal{N}(\mu_j, \sigma_e^2)$$
 (1)

where

$$\mu_i = \alpha_i + \beta_{1,i} X_2 + \beta_{2,i} X_2 + \beta_{3,i} X_3 \tag{2}$$

Predictors  $X_1, X_2, X_3$  are standardized forcing, chilling, and photoperiod, and their effects on the phenology of species j are determined by parameters  $\beta_{1,j}, \beta_{2,j}, \beta_{3,j}$  representing traits. These traits, including the species-specific intercept  $\alpha_j$ , are elements of the following normal random vectors:

$$\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_n\}^T \text{ such that } \boldsymbol{\alpha} \sim \mathcal{N}(\mu_{\alpha}, \boldsymbol{\Sigma}_{\alpha})$$
(3)  
$$\boldsymbol{\beta_1} = \{\beta_{1,1}, \dots, \beta_{1,n}\}^T \text{ such that } \boldsymbol{\beta_1} \sim \mathcal{N}(\mu_{\beta_1}, \boldsymbol{\Sigma}_{\beta_1})$$
  
$$\boldsymbol{\beta_2} = \{\beta_{2,1}, \dots, \beta_{2,n}\}^T \text{ such that } \boldsymbol{\beta_2} \sim \mathcal{N}(\mu_{\beta_2}, \boldsymbol{\Sigma}_{\beta_2})$$
  
$$\boldsymbol{\beta_3} = \{\beta_{3,1}, \dots, \beta_{3,n}\}^T \text{ such that } \boldsymbol{\beta_3} \sim \mathcal{N}(\mu_{\beta_3}, \boldsymbol{\Sigma}_{\beta_3})$$

where the means of the multivariate normal distributions are root trait values (i.e., trait values prior to evolving across a phylogenetic tree) and  $\Sigma_i$  are  $n \times n$  phylogenetic variance-covariance matrices of the form:

$$\begin{bmatrix} \sigma_i^2 & \lambda_i \times \sigma_i \times \rho_{12} & \dots & \lambda_i \times \sigma_i \times \rho_{1n} \\ \lambda_i \times \sigma_i \times \rho_{21} & \sigma_i^2 & \dots & \lambda_i \times \sigma_i \times \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_i \times \sigma_i \times \rho_{n1} & \lambda_i \times \sigma_i \times \rho_{n2} & \dots & \sigma_i^2 \end{bmatrix}$$

$$(4)$$

where  $\sigma_i^2$  is the rate of evolution across a tree for trait i (here assumed to be constant along all branches),  $\lambda_i$  scales branch lengths and therefore is a measure of the "phylogenetic signal" within a species trait, and  $\rho_{xy}$  is the phylogenetic correlation between species x and y, or the fraction of the tree shared by the two species.

The above specification is exactly equivalent to writing equation 2 in terms of root trait values and residuals, such that:

$$\mu_j = \mu_\alpha + \mu_{\beta_1} X_1 + \mu_{\beta_2} X_2 + \mu_{\beta_3} X_3 + e_{\alpha_j} + e_{\beta_{1,j}} + e_{\beta_{2,j}} + e_{\beta_{3,j}}$$
 (5)

where the residual error terms (e.g.,  $e_{\alpha_j}$ ) are elements of normal random vectors from multivariate normal distributions centered on 0 with the same phylogenetic variance-covariance matrices as in equation 4.