Unravelling the phenology-phylogeny tangle.

March 22, 2022

Authors:

The Wolkovich Lab in 2019 & collaborators [1,2,3,4]

Author affiliations:

[1]Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4;

[2]Arnold Arboretum of Harvard University, 1300 Centre Street, Boston, Massachusetts, USA;

[3]Organismic & Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts, USA;

[4]Edificio Ciencias, Campus Universitario 28805 Alcalá de Henares, Madrid, Spain

*Corresponding author: ignacio.moralesc@uah.es

Abstract

Plants have evolved responses to environmental cues able to inform them about the temporal distribution of key resources—i.e. energy and light. The responses to individual cues such as forcing (or spring warming) have shown to be subjected to some degree of evolutionary conservatism. Yet, plants do not respond to isolated cues but to a combination of interacting cues, which difficults accurate predictions of phenology in the face of environmental change. Whether and how evolution has constrained phenological responses to combinations of interacting cues is not yet understood even when this knowledge could enhance model predictions and inform how different plant lineages have adapted to environmental change along their evolutionary histories. Here we use Bayesian hierarchical models and the most complete dataset on tree species phenological responses measured in experimental conditions to: (a) test if phenological responses to three major interacting cues are conserved phylogenetically when considered jointly, (b) compare the phylogenetic signal in the responses to different cues and, (c) test whether coefficient estimates differ between models assuming phylogenetic independence among species and models that explicitly incorporate phylogeny. Results show non-random phylogenetic structuring of phenological responses, highly variable across species and cues. More interestingly, regression coefficients shift when models control for phylogenetic effects, particularly so for forcing, which becomes the most important cue. Taken together, our results suggest that phylogeny should be incorporated into studies modelling multi-species phenological responses, as such responses have been jointly constrained through evolution and thus are not independent.

Introduction

Forecasting remains a major challenge in ecology, especially as anthropogenic climate change drives demand for more accurate forecasts across species. As global temperatures warm, plant species are shifting the timing of their life cycles (Cleland et al., 2007) and their geographical distributions towards higher latitudes and altitudes (Chen et al., 2011), but responses are highly variable across species (Menzel et al.). Some of this variability is due to the complexity of climate change itself—the regional and seasonal variation in warming underlying average trends and shifts in other climate axes (e.g. precipitation)—however, much of it is driven by species-specific variation in response to environmental change, which we can only predict for a few well-studied species. More accurate forecasts will require efforts to scale our understanding across more plant species. Understanding how different plant lineages have evolved their phenotypic responses to the combined effects of multiple dimensions of environmental change would greatly aid prediction.

Decades of research show that plants use environmental cues to time their phenotypic responses with the

temporal distribution of key resources and to avoid periods of high abiotic or biotic stress (Larcher, 1980; Chuine, 2000). Commonly, responses to environmental cues, and their evolution, are studied individually, assuming that a given phenotypic response is predominantly linked to a single cue: for example, that time of leafout is driven by summed heat during early spring (Wolkovich et al., 2012; Davies et al., 2013). Such efforts ignore a more likely scenario for most phenotypic traits where multiple cues interacting along evolutionary history have shaped plant responses (). For example, species-level growth rates or plant height may be determined by the interaction among several cues—e.g. soil nutrients, water availability and light (Larcher, 1980). Similarly, the timing of recurring life cycle events (phenology) is determined by a combination of temperature and light (Chuine and Regniere, 2017).

Phenology may provide an ideal case study to gain insights on how species responses to interacting environmental cues have evolved because the basic cue system is well established (Chuine and Regniere, 2017). In temperate plants, phenology is generally determined by forcing—warm temperatures during the growing season, chilling—cool temperatures during dormancy period over winter, and photoperiod (Chuine and Regniere, 2017). Further, phenology is one of few phenotypic traits where decades of research provide multi-species experimental data on plant responses to these three major cues. Recent multi-species analyses considering forcing, chilling and photoperiod have shown that chilling and forcing together determine complex non-linear responses to warming (Flynn and Wolkovich, 2018; Ettinger et al., 2020), complicating forecasting. In addition, studies found remarkable variation across species in their responses to cues, raising the question of whether such variation is phylogenetically structured. If phylogenetically close species have evolved similar responses to cues it could facilitate forecasting to unmeasured species, and provide insights into how cues have evolved in the past.

Accummulating literature on the evolutionary constraints of phenological responses to cues suggests that phenology is phylogenetically conserved, at least to some extent (Davies et al., 2013). For example, dates of budburst, leafout or first flowering, and shifts in these dates in response to warming are significantly conserved (Davies et al., 2013; Joly et al., 2019), as are traits correlated with phenology, such as seed size (Bolmgren and D. Cowan, 2008; Willis et al., 2008). Almost all of this literature has focused on the phenotype, which may be more strongly determined by the local environment (e.g., the climate where phenology was measured), rather than species' intrinsic responses to the environment through time (but see). And out of the few studies looking at phylogenetic structuring of responses to cues, the typical focus is on one single cue (), even if the evolution of responses to one environmental cue would have not occurred in isolation from the influence of others. Tackling the evolutionary constraints of phenological responses to multiple cues simultaneously would allow answering questions such as: have specific lineages adapted more strongly to some of the cues or

to any combination of cues? Is there any cue that is particularly labile? These questions are highly relevant because answering them would (i) inform about the need to account for phylogeny in phenological models and predictions, and (ii) expand our knowledge on how phenological responses have been constrained so far, which would be relevant in a context where species' sensitivities to warming temperatures seem to decline.

While current methods have advanced much of our understanding of how specific lineages adapted phenotypic responses to the environment, they were not originally designed to capture the complexity of phenotypes evolved in response to multiple interacting cues. For example, typical phylogenetic regression accounts for phylogenetic relationships as a grouping factor either explicitly (Phylogenetic Mixed Model; Housworth et al. (2004)) or implicitly (Phylogenetic Generalized Linear Models; Revell (2010)), assuming all phylogenetic structuring to only affect model intercepts (or residuals). This assumption is known to be little reallistic, particularly in presence of type II error hidden by markedly different clade-structured responses (or phylogenetic non-stationarity) (Davies et al., 2019). Here we present one possible approach that accounts for more complex interactions among predictors, which would be reflected in the species-level slopes being allowed to vary as a function of the phylogeny, rather than keeping slopes constant and only allowing the intercepts (or residuals) to vary. Beyond answering the above questions, our approach has the potential to provide further insights as to whether or not accounting for phylogeny is needed in multi-species phenological studies.

Based on previous research on phylogenetic signal of phenological responses, we expect non-random phylogenetic structuring of the responses to environmental cues (Davies et al., 2013; Rafferty and Nabity, 2017; Joly et al., 2019) and expect that temperature-related cues display higher phylogenetic signal than photoperiod because the latter has remained more constant through evoutionary time. Yet, rather than specific hypotheses for different lineage-level responses, our work aims at exploring and discussing the following questions:

1. Do we need to account for phylogeny in multi-species, multi-cue modelling of the magnitude (strength) and variation of phenological responses to cues? This is, we worry about what are the biggest cues, and we think we may know which are those but if we have the wrong model, we may make the wrong inference or get estimates wrong.
2. If so, can accounting for phylogeny shed light on the ongoing debate on declining sensitivities? For example, if particular lineages have very different evolutionary constraints on their responses to the cues, they may also display very differt declines in their sensitivities to the cues.
3. How can we interpret lambdas and sigmas for each cue, and for the intercept?
4. What are the implications for phenological predictions and forecasts?
5. Is this approach transferable to different taxa or biological responses?

Methods

Phenological and Phylogenetic Data

Phenological data: To estimate phenological responses to chilling, forcing and photoperiod we used data from phenological experiments of temperate woody species conducted in controlled environments, brought together in the Observed Spring Phenology Responses in Experimental Environments (OSPREE) database. In July 2019, we updated an earlier version of this database (Wolkovich et al., 2019) by reviewing all papers found through searching ISI Web of Science and Google Scholar with the following terms:

1. TOPIC = (budburst OR leaf-out) AND (photoperiod OR daylength) AND temperature*, which yielded 623 publications
2. TOPIC = (budburst OR leaf-out) AND dorman*, which yielded 270 publications

We scraped data from all papers of woody species that tested for photoperiod and/or temperature effects on budburst, leafout, or flowering, resulting in 56 papers. **?** used a portion (72 experiments across 49 papers) of the earlier OSPREE database and provides extensive methods on the database creation and cleaning. For our analysis here, we included all budburst experiments where we could quantify chilling, forcing and photoperiod levels, resulting in 44 studies from 33 papers. Across experiments chilling treatments were often fully or partially applied in the field, thus we estimated field chilling ourselves using daily temperature data from ... [Cat and Nacho – add here: Be sure to include updated info on our datasets and which chilling metric we used]. **?** provides additional details on these calculations (however, to have climate data through all our study years, we used a different climate dataset here for North America).
We analyze 2 different subsets of species in the OSPREE database to explore differences across two major groups of taxa, angiosperms and gymnosperms, for which there are markedly different number of species (194 angiosperms vs. 19 gymnosperms), and whose deep evolutionary divergence advises for separate analyses ().

We used the phylogenetic megatree for seed plant from Smith and Brown (2018) to extract a subset phylogenetic tree containing only the species in the OSPREE dataset (Wolkovich et al., 2019). We pruned the megatree to generate to sub-trees containing only the species in each subset of data. The species that were not present in the megatree were added as polytomies at the generic level (using the function *congeneric.merge*; (Pearse et al., 2015)) with a branch length of zero. Polytomies represent 26.8% of the full angiosperm dataset. To test for the ability of polytomies to bias our results we run sensitivity analyses excluding these species from models (which lead to 142 angiosperms; see Supporting Information).

Bayesian hierarchical phylogenetic model

Commonly used phylogenetic regression methods today (e.g., PGLS and PMM) were originally conceived as statistical corrections for phylogenetic non-independence across observations—generally species—thus allowing multi-species studies to meet the assumptions linear regression (Freckleton et al., 2002). These corrections incorporated phylogenetic structure in the regression by modifying the residual variance-covariance matrix to substitute off-diagonal elements of zero (the value given the assumption of independence across observations) for shared phylogenetic branch lengths representing pairwise covariances (under phylogenetic non-independece among observations). Off-diagonals were also allowed to include a multiplying parameter—generally referred to as lambda—which is a transformation indicating the amount of phylogenetic relatedness among species (see below). Because the original aim of these methods was to correct for statistical nuance, the underlying assumption of phylogenetic regressions is that phylogenetic relatedness would only affect either model residuals (in PGLS approaches, Freckleton et al., 2002), or the model intercepts (e.g., in many PMM approaches, Housworth et al., 2004).

Because our aim is to understand how evolution may have imprinted biological responses to multiple interactive cues, our approach expands the above methods by explicitly incorporating phylogenetic structure accross model intercepts and slopes. Doing so allows explicitly estimating the amount of phylogenetic relatedness in species' sensitivities to each cue, when these sensitivities are modelled in a multi-predictor regression setting.

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

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

where

$$\mu_j = \alpha_j + \beta_{1,j} X_2 + \beta_{2,j} X_2 + \beta_{3,j} X_3 \tag{2}$$

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 species' responses (or sensitivities) to each of the cues. These responses, including the species-specific intercept α_j, are elements of the following

normal random vectors:

$$\begin{aligned}
\boldsymbol{\alpha} &= \{\alpha_1, \ldots, \alpha_n\}^T \text{ such that } \boldsymbol{\alpha} \sim \mathcal{N}(\mu_\alpha, \boldsymbol{\Sigma_\alpha}) \\
\boldsymbol{\beta_1} &= \{\beta_{1,1}, \ldots, \beta_{1,n}\}^T \text{ such that } \boldsymbol{\beta_1} \sim \mathcal{N}(\mu_{\beta_1}, \boldsymbol{\Sigma_{\beta_1}}) \\
\boldsymbol{\beta_2} &= \{\beta_{2,1}, \ldots, \beta_{2,n}\}^T \text{ such that } \boldsymbol{\beta_2} \sim \mathcal{N}(\mu_{\beta_2}, \boldsymbol{\Sigma_{\beta_2}}) \\
\boldsymbol{\beta_3} &= \{\beta_{3,1}, \ldots, \beta_{3,n}\}^T \text{ such that } \boldsymbol{\beta_3} \sim \mathcal{N}(\mu_{\beta_3}, \boldsymbol{\Sigma_{\beta_3}})
\end{aligned} \tag{3}$$

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

$$\begin{bmatrix}
\sigma_i^2 & \lambda_i \times \sigma_i \times \rho_{12} & \ldots & \lambda_i \times \sigma_i \times \rho_{1n} \\
\lambda_i \times \sigma_i \times \rho_{21} & \sigma_i^2 & \ldots & \lambda_i \times \sigma_i \times \rho_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_i \times \sigma_i \times \rho_{n1} & \lambda_i \times \sigma_i \times \rho_{n2} & \ldots & \sigma_i^2
\end{bmatrix} \tag{4}$$

where σ_i^2 is the rate of evolution across a tree for trait i (here assumed to be constant along all branches), λ_i scales branch lengths and therefore is a measure of the "phylogenetic signal" or extent of phylogenetic relatedness on each model parameter (i.e., α_j, $\beta_{1,j}$, $\beta_{2,j}$, $\beta_{3,j}$), and ρ_{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 equivalent to writing equation 2 in terms of root trait values and residuals, such that:

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

where the residual error terms (e.g., e_{α_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.

For the sake of comparison, the PGLS version of equation 5 would be written:

$$\mu_j = \mu_\alpha + \mu_{\beta_1} X_1 + \mu_{\beta_2} X_2 + \mu_{\beta_3} X_3 + e_{\alpha_j} \tag{6}$$

and the PMM version:

$$\mu_j = \mu_\alpha + \mu_{\beta_1} X_1 + \mu_{\beta_2} X_2 + \mu_{\beta_3} X_3 + e_{\alpha_j} + e_{\epsilon_j} \quad (7)$$

where eϵ is the error term independent of phylogenetic effects. Thus, the model specification used here 5 can be seen as an extended version of the above, where the error term is decomposed one step further, allowing to partition residual variance into predictor-specific components.

Interpretation of λ_j and σ_j^2 on slopes and intercepts

In classic phylogenetic regression approaches aimed at controlling for phylogenetic non-independence of analysis units (i.e. species), see (Revell, 2010), the λ scaling parameter is assumed constant across the specific set of predictors that enter the model. λ is estimated as a single parameter based on one single residual term VCV matrix, which means it does not allow inferring differences across predictors (i.e. environmental cues in our example, cues hereafter) in how species responses to those cues are phylogenetically structured. For example, in a PGLS model with three cues, it would be possible to have a high (i.e., close to 1) value of λ, due to either a strong phylogenetic signal in the response but no phylogenetic structuring in the cues, or one or multiple predictors being strongly phylogenetically structured. In the latter case, phylogenetic structuring of responses to cues could be correlated (i.e., responses to cues evolving in a correlated fashion) or uncorrelated (i.e., independent evolution of responses to cues). Discerning these different situations is not trivial as they would inform whether responses to predictors configure in a structured fashion along the evolutionary process. However, classic approaches act as a black box regarding this information, and merely inform whether or not model residuals are phylogenetically structured (i.e. in PGLS) or the amount of model variance attributable to the phylogeny and independent from other sources of variation (i.e., in PMM, see Housworth et al. (2004)).

Our approach goes one step beyond in explicitly partitioning model residual variance into specific components relative to the model intercept and predictor slopes (see equation 5). The multivariate normal distributions of intercept and slopes include each a variance term (see equation 3), modelled with a λ scaling parameter. The interpretation of λs in our models are analogous to Pagel's Pagel (1999) λ parameter (Housworth et al., 2004), constrained to range from 0 to 1, with values of 0 indicating absence of phylogenetic relatedness, and values of 1 indicating *Brownian Motion* evolution (BM). Estimated λs are not fully equivalent to computing phylogenetic signal of the slopes of each cue separately (i.e., fitting a multilevel regression model with species as a grouping factor on intercepts, and subsequently estimating phylogenetic signal for model slopes). Instead, they are a relative metric of phylogenetic relatedness allowing to compare among responses known to interact with each other and estimated simultaneously. While classic approaches compute only one λ, our

approach computes four, one independent of the predictors, and one for each predictor.

A traditional interpretation of σ^2s under Brownian Motion evolution, is an 'evolutionary rate' or phenotypic accummulation over time (Revell et al., 2008). In PGLS, σ^2_ϵ is estimated for the model error term, which is distributed as a multivariate normal with variance–covariance matrix given by $\sigma^2_\epsilon \mathbf{\Sigma_i}$. Here, we estimate four σ^2scorresponding to each model parameter. In our particular case (i.e., modelling a phenological response to three intearcting environmental cues), σ^2_α for the intercept could be interpreted as the phenological variation across species accummulated along evolution independently from the cues. The $\sigma_{\beta}{}^2_1$, $\sigma_{\beta}{}^2_2$, and $\sigma_{\beta}{}^2_3$, corresponding to model slopes would represent the phylogenetic variance linked to species responses to each of the modelled cues (i.e., forcing, chilling, and photoperiod, respectively). This is, the variability in how species shift their phenology responding to temperature and light, accummulated along the evolutionary process and considered in concert.

Results

Most analyzed angiosperm species are sensitive to all three environmental cues—i.e., forcing, chilling, and photoperiod (Figs. **??**, Supporting Table **??**). Cue sensitivity led to average phenological advances of 7.2 days per unit of standardized chilling, of 5.8 days per unit of forcing, and 1.4 days/standard unit of photoperiod (see Table **??**). Direction of the effects and ranking of cue sentivities were qualitatively similar for gymnosperms—chilling ¿ forcing ¿ photoperiod—but with marked quantitative shifts and larger variability around parameters, which led 97.5% Credible Intervals for estimates of sensitivity to forcing and photoperiod to overlap with zero (see Table **??**). Gymnosperm based results must be interpreted with caution as they are based on a very limited set of observations.

The above patterns were strongly phylogenetically clustered for forcing (λ $\bar{0}$.68), moderately so for chilling (λ $\bar{0}$.56) and weakly for photoperiod (λ $\bar{0}$.24) (see Fig. 3, Table **??**). This indicates that species responses to interacting cues, when modelled jointly, widely differ in how they are affected by evolutionary relatedness. Sensitivigy to photoperiod treatments does not vary across clades while responses to forcing tend to be more similar among closely related species (Fig. 2).

If we want to make the point of with and without lambda results, we could just include a scatterplot? With

Cue sensitivities: are there major shifts when phylogeny is accounted for?

1. Perhaps we can compare the new results against those in Ailene's paper more closely **??**, and **??**.
2. A first glance comparing results here with those in the NCC paper suggest that after taking phylogeny into account, the associations with photoperiod may decrease and the variance around estimations of sensitivity to chilling gets larger.

Phylogenetic signal in phenological responses

1. Phenological responses to the three studied cues are overall phylogenetically conserved but estimates of phylogenetic signal differ strongly across species subsets (angio vs. gymno).
2. When angiosperm species (from main model) are considered, responses to forcing are more conserved ($\lambda = 0.64$) than responses to chilling ($\lambda = 0.66$) or to photoperiod ($\lambda = 0.35$) (see Figure 6).
3. When gymnosperm species are considered, all responses to cues are similarly low (yet different from zero): forcing ($\lambda = 0.36$), chilling ($\lambda = 0.32$) and photoperiod ($\lambda = 0.37$) and show almost overlapping posterior distributions, which may be driven by a low number of species (19) **??**).

Budburst models, phylogenetic vs. non-phylogenetic

1. Here goes text comparing results with lambda $\bar{0}$ against results with estimated lambda.

Discussion

To be fleshed out.

1. Random discussion points with no home, yet ...
   (a) This is a case where phylogeny makes a big difference! Changes overall forcing cues?
   (b) Reduced uncertainty in species estimates (I think?) with including phylogeny (goes with above point perhaps also)
   (c) Even with phylogeny added FagSyl is still freakish for photoperiod cue ... suggesting we've been studying an extreme species as one of our focal species (maybe?)

References

Bolmgren, K., and P. D. Cowan. 2008. Time–size tradeoffs: A phylogenetic comparative study of flowering time, plant height and seed mass in a north-temperate flora. Oikos 117:424–429.

Chen, I.-C., J. K. Hill, R. Ohlemüller, D. B. Roy, and C. D. Thomas. 2011. Rapid range shifts of species associated with high levels of climate warming. Science 333:1024–1026.

Chuine, I. 2000. A unified model for budburst of trees. Journal of Theoretical Biology 207:337 – 347.

Chuine, I., and J. Regniere. 2017. Process-based models of phenology for plants and animals. Annual Review of Ecology, Evolution, and Systematics 48:159–182.

Cleland, E. E., I. Chuine, A. Menzel, H. A. Mooney, and M. D. Schwartz. 2007. Shifting plant phenology in response to global change. Trends in Ecology & Evolution 22:357–365.

Davies, T., E. Wolkovich, N. Kraft, N. Salamin, and S. E. Travers. 2013. Phylogenetic conservatism in plant phenology. Journal of Ecology 101:1520–1530.

Davies, T. J., J. Regetz, E. M. Wolkovich, and B. J. McGill. 2019. Phylogenetically weighted regression: A method for modelling non-stationarity on evolutionary trees. Global ecology and biogeography 28:275–285.

Ettinger, A., C. Chamberlain, I. Morales-Castilla, D. Buonaiuto, D. Flynn, T. Savas, J. Samaha, and E. Wolkovich. 2020. Winter temperatures predominate in spring phenological responses to warming. Nature Climate Change pages 1–6.

Flynn, D. F. B., and E. M. Wolkovich. 2018. Temperature and photoperiod drive spring phenology across all species in a temperate forest community. New Phytologist 219:1353–1362.

Freckleton, R. P., P. H. Harvey, and M. Pagel. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. The American Naturalist 160:712–726.

Housworth, E. A., E. P. Martins, and M. Lynch. 2004. The phylogenetic mixed model. The American Naturalist 163:84–96.

Joly, S., D. F. Flynn, and E. M. Wolkovich. 2019. On the importance of accounting for intraspecific genomic relatedness in multi-species studies. Methods in Ecology and Evolution .

Larcher, W. 1980. Plant Physiological Ecology. Springer-Verlag.

Menzel, A., Y. Yuan, M. Matiu, T. Sparks, H. Scheifinger, R. Gehrig, and N. Estrella. ???? Climate change fingerprints in recent european plant phenology. Global Change Biology .

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877–884.

Pearse, W. D., M. W. Cadotte, J. Cavender-Bares, A. R. Ives, C. M. Tucker, S. C. Walker, and M. R. Helmus. 2015. Pez: Phylogenetics for the environmental sciences. Bioinformatics 31:2888–2890.

Rafferty, N. E., and P. D. Nabity. 2017. A global test for phylogenetic signal in shifts in flowering time under climate change. Journal of Ecology 105:627–633.

Revell, L. J. 2010. Phylogenetic signal and linear regression on species data. Methods in Ecology and Evolution 1:319–329.

Revell, L. J., L. J. Harmon, and D. C. Collar. 2008. Phylogenetic signal, evolutionary process, and rate. Systematic biology 57:591–601.

Smith, S. A., and J. W. Brown. 2018. Constructing a broadly inclusive seed plant phylogeny. American journal of botany 105:302–314.

Willis, C. G., B. Ruhfel, R. B. Primack, A. J. Miller-Rushing, and C. C. Davis. 2008. Phylogenetic patterns of species loss in thoreau's woods are driven by climate change. Proceedings of the National Academy of Sciences 105:17029–17033.

Wolkovich, E. M., B. I. Cook, J. M. Allen, T. M. Crimmins, J. L. Betancourt, S. E. Travers, S. Pau, J. Regetz, T. J. Davies, N. J. B. Kraft, T. R. Ault, K. Bolmgren, S. J. Mazer, G. J. McCabe, B. J. McGill, C. Parmesan, N. Salamin, M. D. Schwartz, and E. E. Cleland. 2012. Warming experiments underpredict plant phenological responses to climate change. Nature 485:494–497.

Wolkovich, E. M., A. K. Ettinger, D. Flynn, T. Savas, C. Chamberlain, D. Buonaiuto, and J. Samaha. 2019. Observed Spring Phenology Responses in Experimental Environments (OSPREE). doi:10.5063/F1CZ35KB.

Tables and Figures

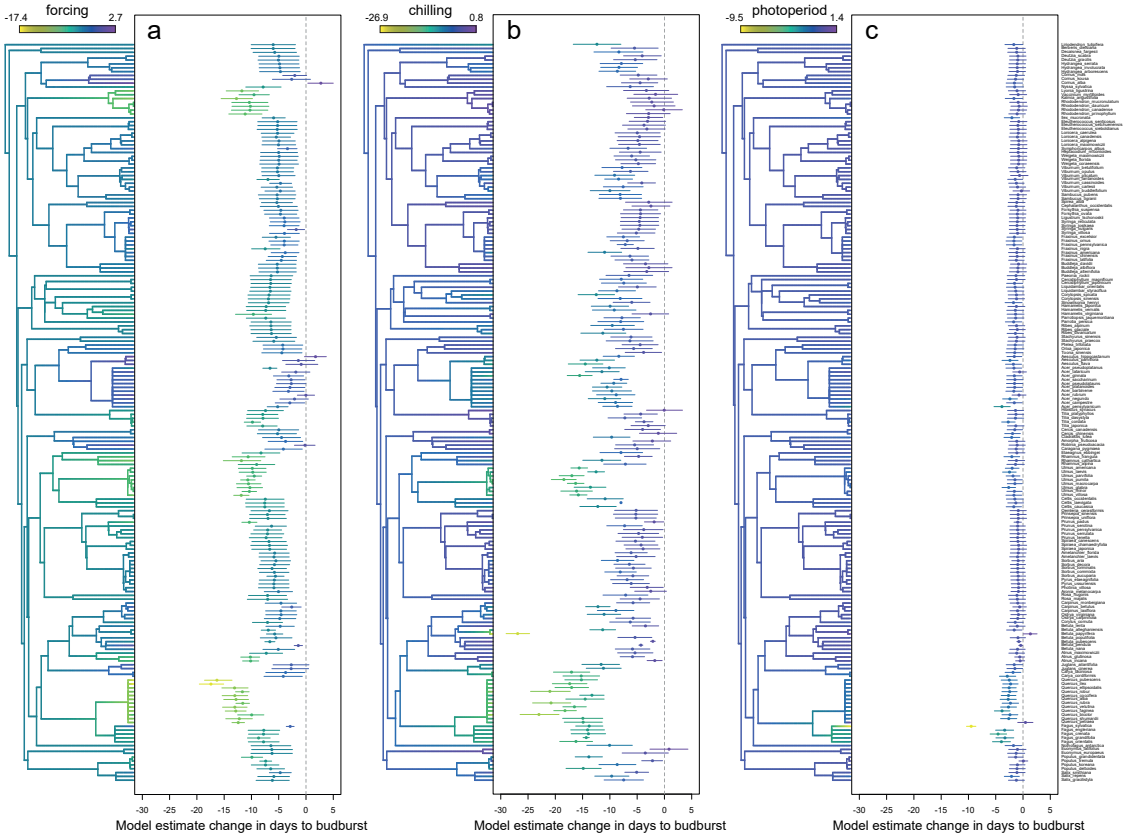

Figure 1: Phenological sensitivity to thee environmental cues, forcing (a), chilling (b) and photoperiod (c) measured in change in days to budburst per standardized unit (z-transformation) of the cues across 192 angiosperm species. The same phylogenetic tree is shown in each panel, colored acording to an estimation of ancestral character states, being the states at the tips the model slopes of our hierarchical phylogenetic model. Note that the color scale varies in each panel. Total tree depth is 81. My.

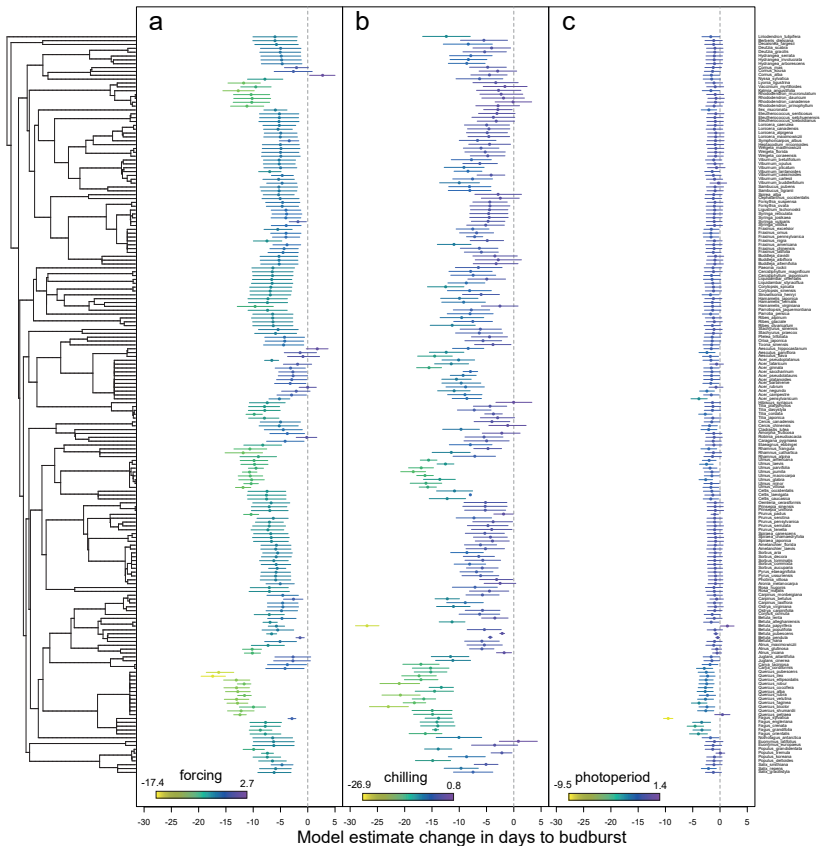

Figure 2: Phenological sensitivity to thee environmental cues, forcing (a), chilling (b) and photoperiod (c) measured in change in days to budburst per standardized unit (z-transformation) of the cues across 192 angiosperm species. The same phylogenetic tree is shown in each panel, colored acording to an estimation of ancestral character states, being the states at the tips the model slopes of our hierarchical phylogenetic model. Note that the color scale varies in each panel. Total tree depth is 81. My.

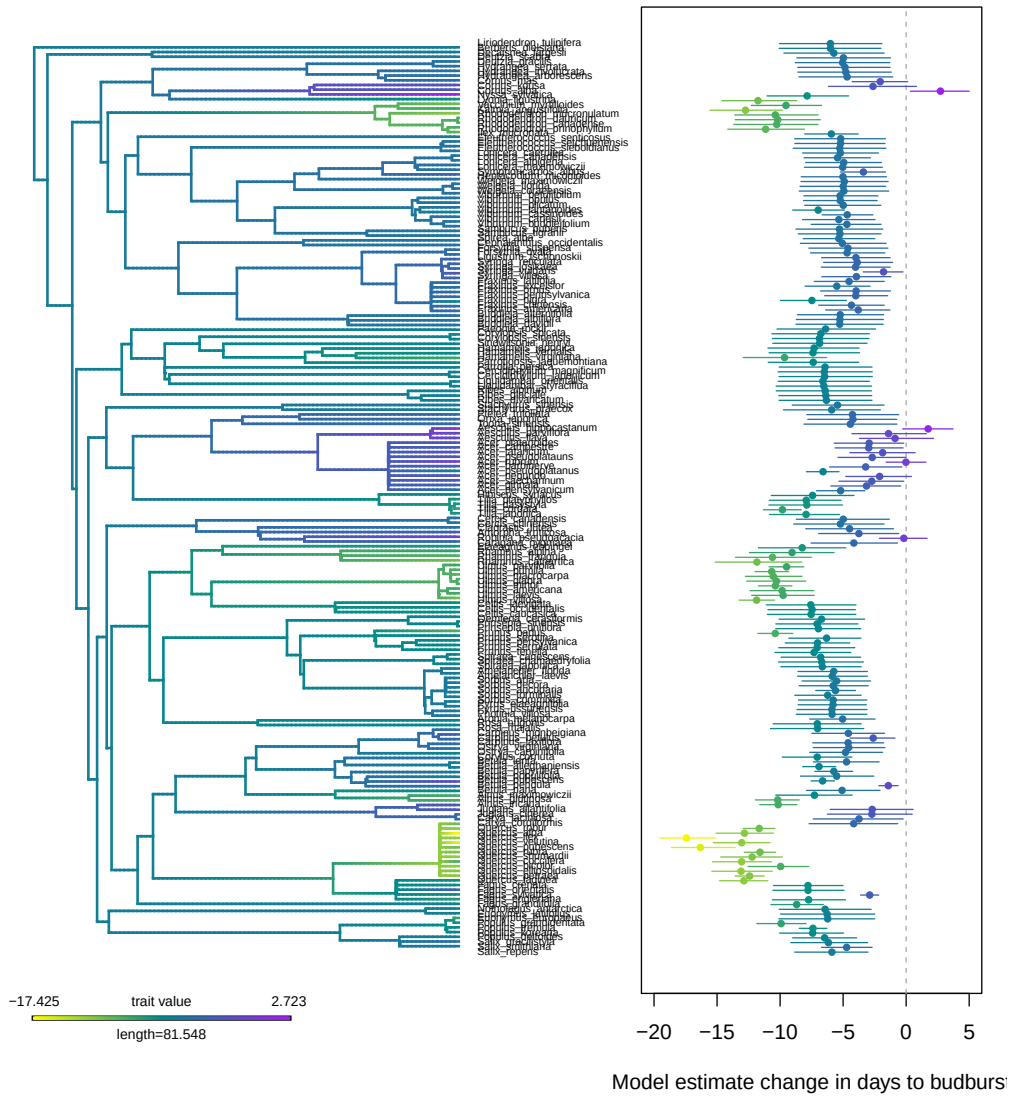

Figure 3: Cue sensitivity estimation by hierarchical phylogenetic model showing slopes for forcing, for 194 angiosperm species.

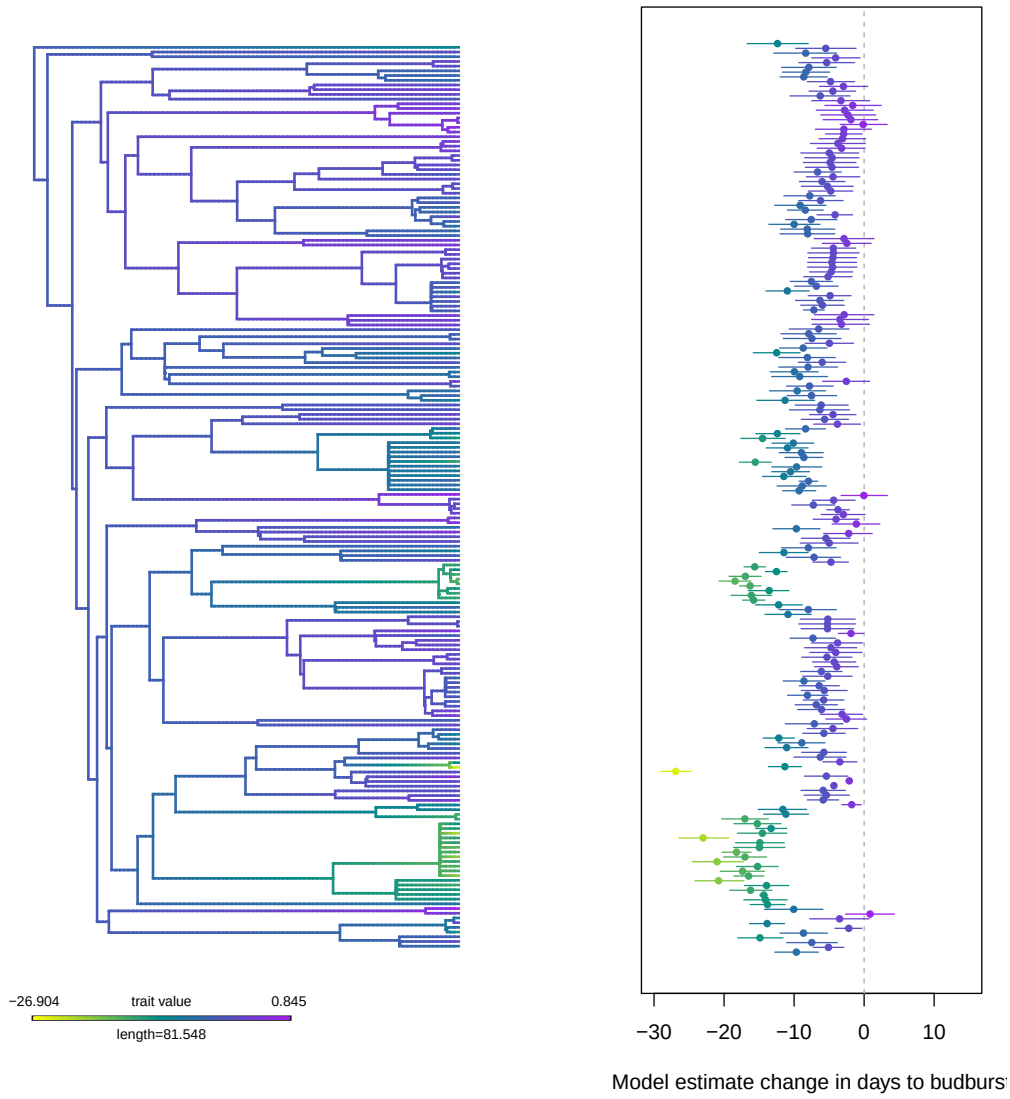

Figure 4: Cue sensitivity estimation by hierarchical phylogenetic model showing slopes for chilling, for 194 angiosperm species.

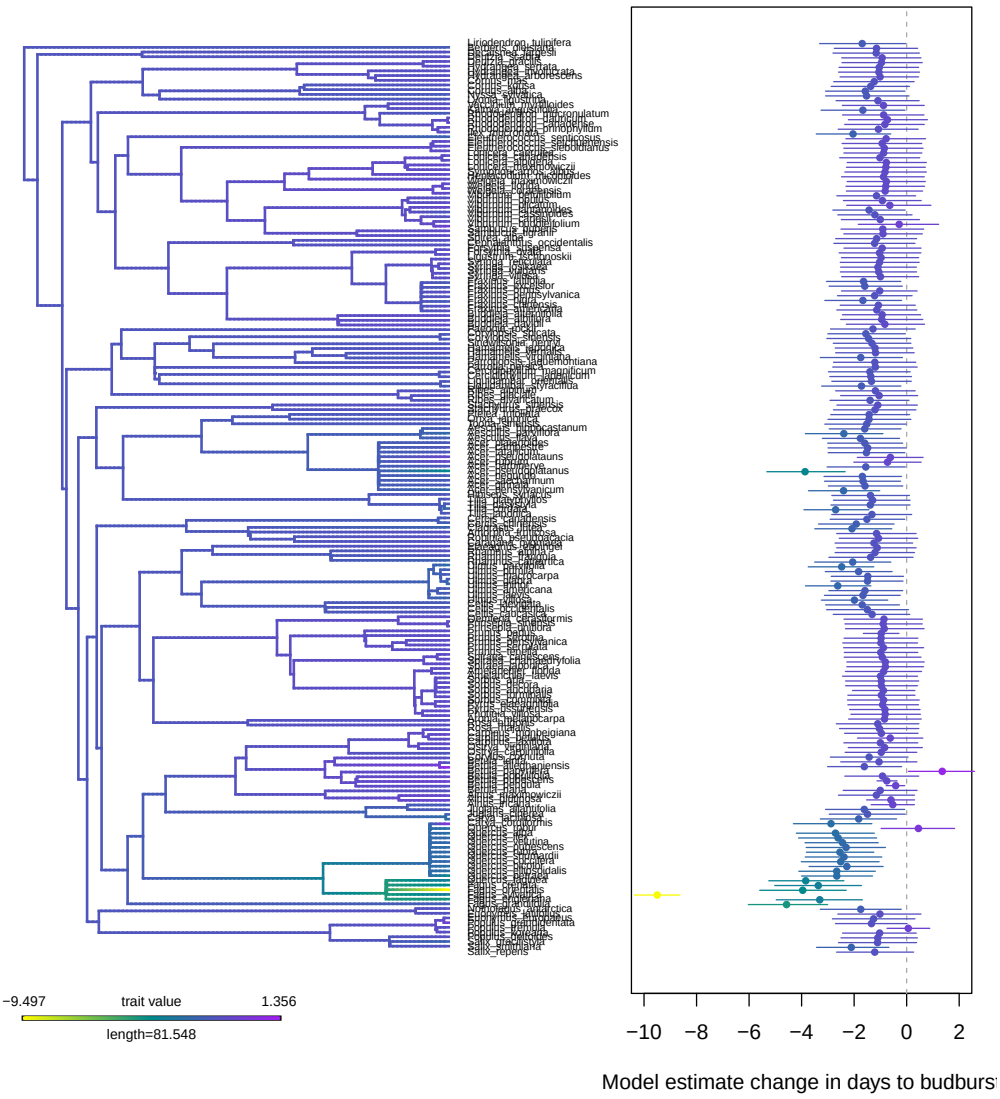

Figure 5: Cue sensitivity estimation by hierarchical phylogenetic model showing slopes for photoperiod, for 194 angiosperm species.

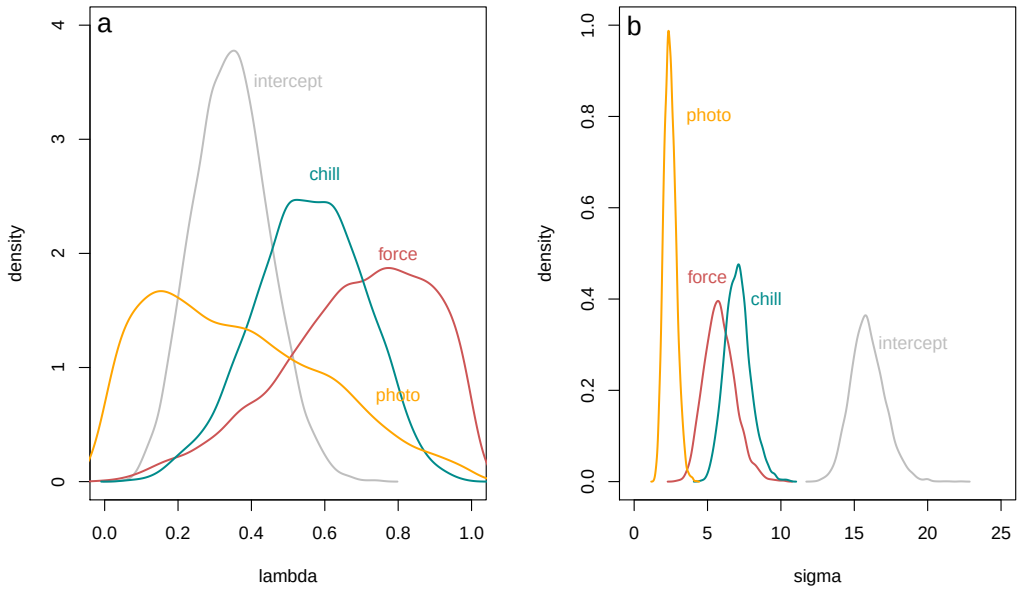

Figure 6: Posterior distribution of phylogenetic signal measured by lambda for each cue included as a predictor in the model for angiosperms: forcing (red), chilling (blue), photoperiod (orange) and for the model intercept (grey).

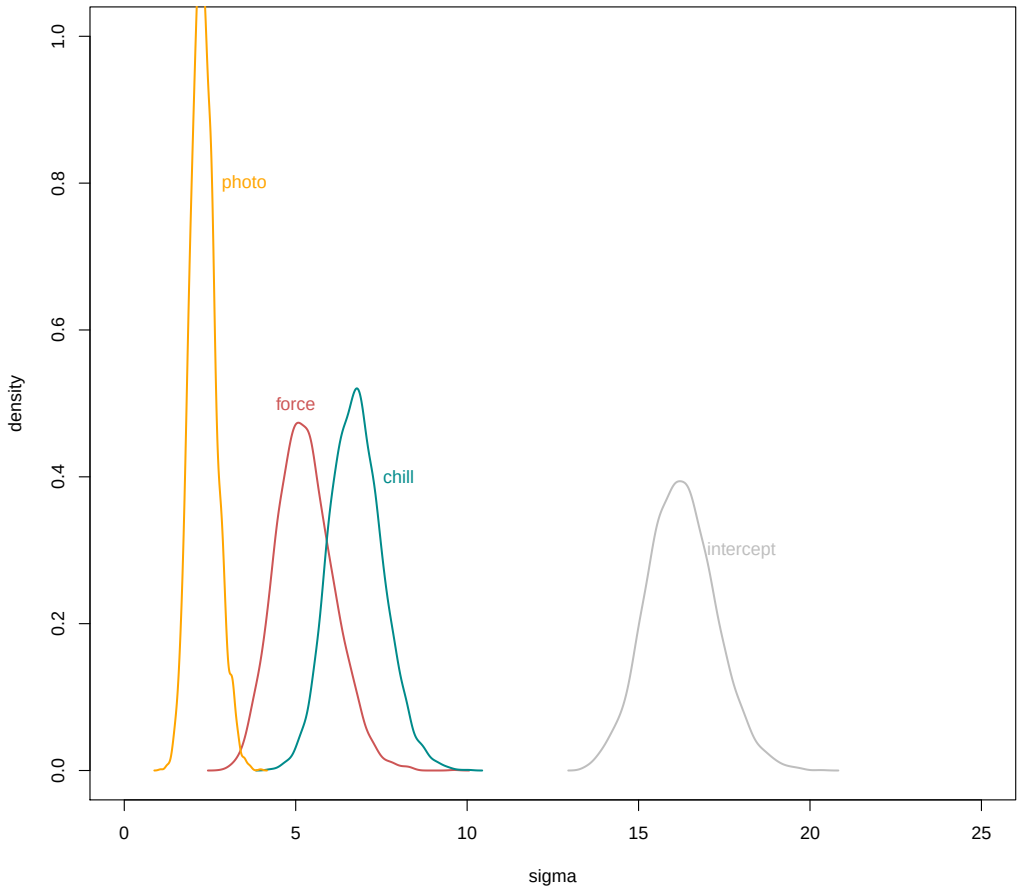

Figure 7: Posterior distribution of sigma for each cue included as a predictor in the model for angiosperms: forcing (red), chilling (blue), photoperiod (orange) and for the model intercept (grey).

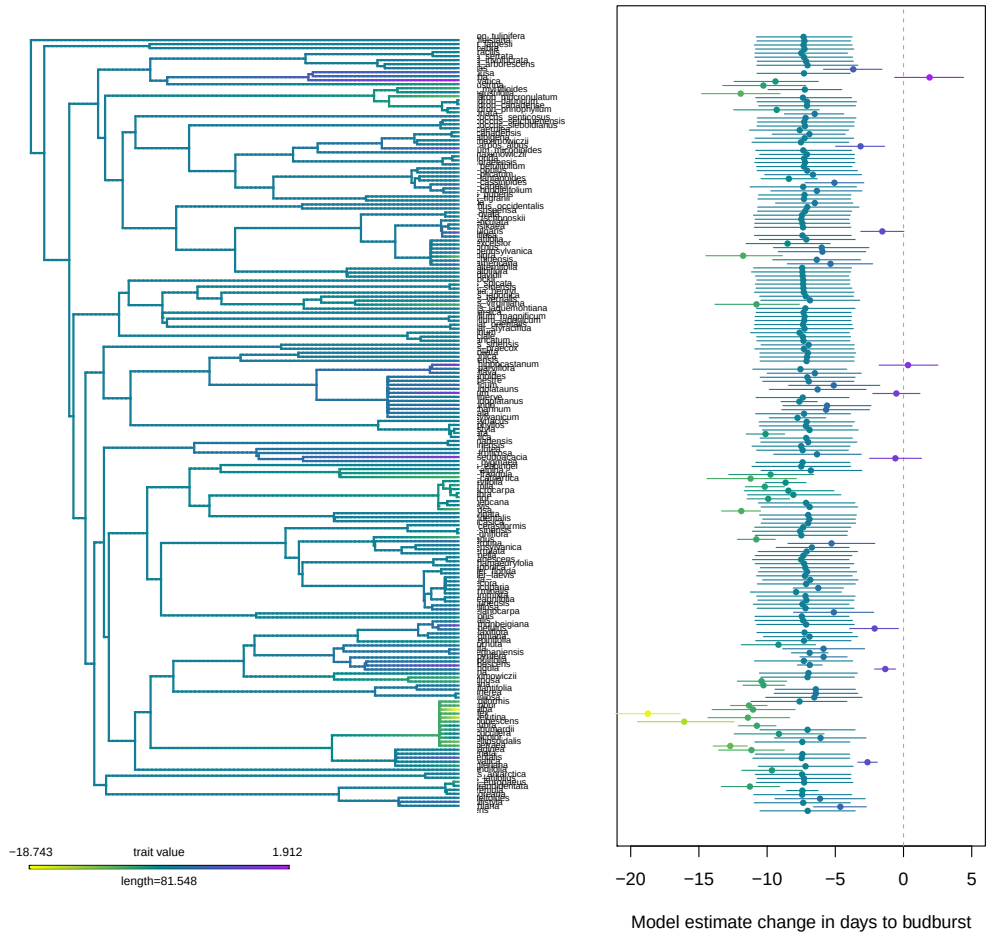

Figure 8: Cue sensitivity estimation by hierarchical phylogenetic model showing slopes for forcing making lambda $\bar{0}$, for 194 angiosperm species.

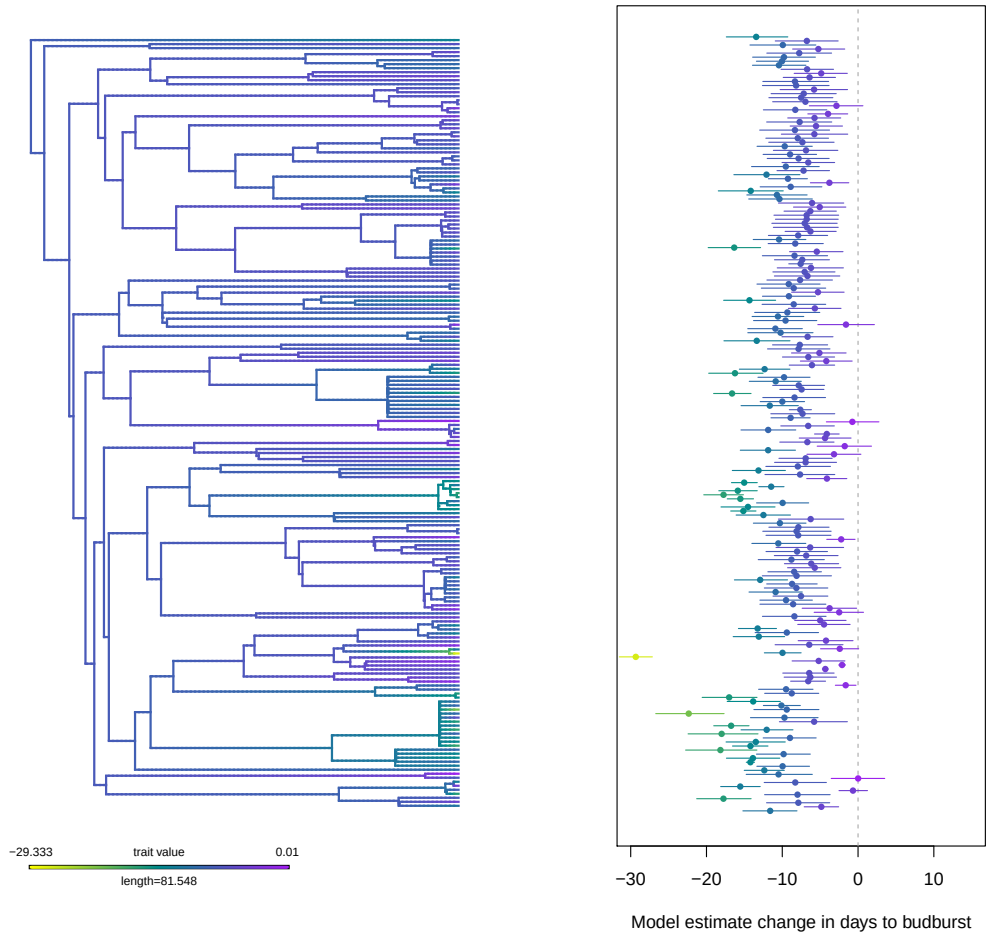

Figure 9: Cue sensitivity estimation by hierarchical phylogenetic model showing slopes for chilling making lambda $\bar{0}$, for 194 angiosperm species.

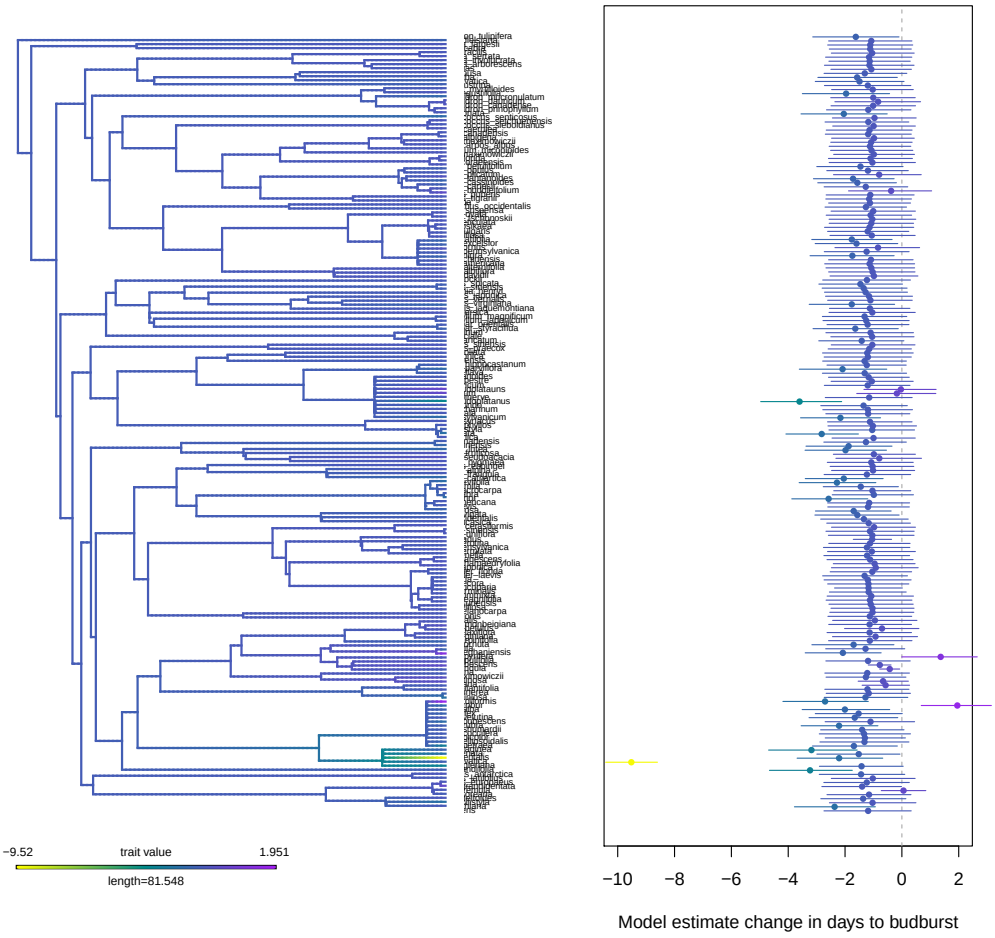

Figure 10: Cue sensitivity estimation by hierarchical phylogenetic model showing slopes for photoperiod making lambda $\bar{0}$, for 194 angiosperm species.

Table 1: Full model parameters estimated for 192 angiosperm species.

parameter	**mean**	**sd**	**2.50%**	**50%**	**97.50%**	**n_eff**
μ_α	30.57	3.41	23.68	30.59	37.14	5031.19
$\mu_{\beta forcing}$	-5.84	2.01	-9.72	-5.89	-1.79	2374.73
$\mu_{\beta chilling}$	-7.19	2.03	-11.15	-7.18	-3.18	3694.93
$\mu_{\beta photoperiod}$	-1.37	0.76	-2.92	-1.35	0.14	1565.41
λ_α	0.35	0.10	0.16	0.34	0.56	3416.51
$\lambda_{\beta forcing}$	0.68	0.20	0.23	0.71	0.98	185.35
$\lambda_{\beta chilling}$	0.56	0.15	0.25	0.56	0.83	738.57
$\lambda_{\beta photoperiod}$	0.36	0.24	0.02	0.33	0.88	296.51
σ^2_α	15.93	1.17	13.84	15.85	18.41	2988.37
$\sigma^2_{\beta forcing}$	5.84	1.04	4.03	5.78	8.15	502.74
$\sigma^2_{\beta chilling}$	7.05	0.87	5.48	7.02	8.92	1026.77
$\sigma^2_{\beta photoperiod}$	2.45	0.41	1.74	2.42	3.32	469.46
σ^2_y	12.81	0.18	12.47	12.80	13.17	4017.16

Table 2: Full model parameters estimated for 19 gymnosperm species.

parameter	**mean**	**sd**	**2.50%**	**50%**	**97.50%**	**n_eff**
μ_α	25.75	4.50	16.88	25.73	34.73	33151.86
$\mu_{\beta forcing}$	-5.92	3.80	-12.97	-6.05	1.90	16443.03
$\mu_{\beta chilling}$	-8.11	3.63	-15.31	-8.09	-0.94	21379.81
$\mu_{\beta photoperiod}$	-0.88	3.33	-8.01	-0.67	5.19	16301.93
λ_α	0.47	0.26	0.02	0.48	0.90	15934.03
$\lambda_{\beta forcing}$	0.36	0.23	0.02	0.33	0.84	14336.60
$\lambda_{\beta chilling}$	0.32	0.23	0.01	0.28	0.82	13230.88
$\lambda_{\beta photoperiod}$	0.37	0.24	0.02	0.34	0.88	11199.49
σ^2_α	23.47	6.20	13.87	22.59	37.81	18272.58
$\sigma^2_{\beta forcing}$	8.89	2.45	4.96	8.60	14.51	8126.51
$\sigma^2_{\beta chilling}$	10.47	2.66	5.78	10.30	16.17	8539.38
$\sigma^2_{\beta photoperiod}$	7.18	2.29	3.29	6.96	12.25	5625.69
σ^2_y	15.81	0.41	15.04	15.81	16.63	28640.16