American Journal of Botany

Botanical Society of America
leading scientists and educators
since 1893

RESEARCH ARTICLE

BRIEF COMMUNICATION

Macroevolutionary history predicts flowering time but not phenological sensitivity to temperature in grasses

Barbara M. Neto-Bradley[1,2,5], Jeannette Whitton[1,2], Linda P. J. Lipsen[1,3], and Matthew W. Pennell[2,4]

Manuscript received 8 July 2020; revision accepted 23 December 2020.

[1] Department of Botany, University of British Columbia, Vancouver V6T 1Z4, Canada

[2] Biodiversity Research Centre, University of British Columbia, Vancouver V6T 1Z4, Canada

[3] UBC Herbarium, University of British Columbia, Vancouver V6T 1Z4, Canada

[4] Department of Zoology, University of British Columbia, Vancouver V6T 1Z4, Canada

[5] Author for correspondence (e-mail: barbaranetobradley@gmail.com)

Citation: Neto-Bradley, B. M., J. Whitton, L. P. J. Lipsen, and M. W. Pennell. 2021. Macroevolutionary history predicts flowering time but not phenological sensitivity to temperature in grasses. *American Journal of Botany* 108(5): 1–10.

doi:10.1002/ajb2.1647

PREMISE: Long-term observations show that flowering phenology has shifted in many lineages in response to climate change. However, it remains unclear whether these results can be generalized to predict the presence, direction, or magnitude of responses in lineages for which we lack long time-series data. If phenological responses are phylogenetically conserved, we can extrapolate from species for which we have data to predict the responses of close relatives. While several studies have found that closely related species flower at similar times, fewer have evaluated whether phylogenetically proximal species respond to environmental change similarly.

METHODS: We paired flowering time data from 3161 manually scored herbarium specimens of 72 species of grasses (Poaceae) with historical climate data and analyzed the phylogenetic signal and phylogenetic half-life of phenological sensitivity. We also ran these analyses on a subset of species showing statistically significant sensitivities, in order to assess the role of sampling bias on phylogenetic signal.

RESULTS: Closely related grass species tend to flower at similar times, but flowering times respond to temperature changes in species-specific ways. We also show that only including species for which there is strong evidence of phenological shifts results in overestimating phylogenetic signal.

CONCLUSIONS: In agreement with other recent studies, our results suggest caution in extrapolating from evidence of phylogenetic similarity to predicting shared responses in this ecologically relevant trait. Future work is needed to better understand the discrepancy between the phylogenetic signal in observed phenological shifts and absence of such signal in sensitivity.

KEY WORDS Herbarium records; macroevolution; phenological sensitivity; phenology; phylogenetic half-life; phylogenetic signal; Poaceae; temperature change.

Phenology is changing in response to anthropogenic climate change (Parmesan and Yohe, 2003) and understanding exactly how phenological shifts track changing environmental conditions can help prioritize mitigation as well as management actions (Morellato et al., 2016). For example, if we can predict the phenology of recently arrived invasive species based on their closest relatives, we can narrow down treatment windows for manual removal or herbicide application.

Phenological studies of flowering plants use combinations of historical observations and herbarium specimens to track changes in leaf, floral, and fruiting phenology (Primack et al., 2004; Munson and Long, 2017; Willis et al., 2017; Meineke et al., 2019). The multiple lifetimes of collection efforts housed in herbaria span space and time, and provide irreplaceable insight into long-term trends and broadscale patterns of changes in phenology (Primack et al., 2004; Munson and Long, 2017; Willis et al., 2017; Meineke et al., 2019).

Despite these strengths, herbarium records suffer from some key limitations, including biases in taxonomic representation, in latitudinal/elevational coverage, and targeted collection based on the particular interests of specific botanists at specific times (Graham et al., 2004). In addition, while collection of a plant in flower on a specific date indicates that at least one individual was in flower at that time, this individual represents a single draw from the distribution of flowering dates in a population (Pearse et al., 2017). Finally, while there has been a considerable effort to digitize herbarium specimens (which may itself introduce some bias; very old and/or very new specimens are often prioritized for digitization (Thiers et al., 2016)), there is simply not enough data in digital repositories to examine phenological shifts of all species. Given these limitations, and the urgency of identifying and managing species most vulnerable to climate change, it is important to define the conditions under which we can generalize results from phenological studies of focal taxa to the rest of the Tree of Angiosperms.

Incorporating phylogeny provides one way forward, because species' shared ancestry leads us to hypothesize that close relatives may share both phenological traits (e.g., similar flowering times) and phenological responses to environmental change (e.g., similar changes in flowering time in response to similar climatic changes) as a result of shared constraints on their life histories (reviewed in Elzinga et al., 2007). If measured effects can be predicted and generalized from knowledge of relationships among taxa, phylogenies will provide an opportunity to interpret a subset of findings more broadly and say something about the probability that a species for which there is currently little data will experience (or has already experienced) a particular phenological response to climate change (Davis et al., 2010; Davies et al., 2013; Rafferty and Nabity, 2017).

While several studies have already assessed the level of phylogenetic signal of overall phenological change over time (e.g., change in flowering date over the past decade; Lessard-Therrien et al., 2014; Panchen et al., 2014; CaraDonna and Inouye, 2015; Rafferty and Nabity, 2017; Li et al., 2019), fewer studies have tested for phylogenetic signal in the relationships between changes in climate variables and phenological shifts (i.e., phenological sensitivity; but see (CaraDonna and Inouye, 2015; Daru et al., 2019; Li et al., 2019)). This distinction is important, because in fact there is some evidence for both lack of phylogenetic signal in phenological sensitivity (CaraDonna and Inouye, 2015; Li et al., 2019) and for different phenological cues in closely related species (Zohner and Renner, 2014). Studies that quantify the phylogenetic signal of phenological sensitivity are crucial if we hope to extend our understanding of whether and how phenology will track ongoing climate change.

Furthermore, in attempting to generalize the results of few species to many, the role of reporting biases becomes particularly important. Unreported phenological stasis may skew estimates of global rate-of-change (Kopp et al., 2020) and in the case of phylogenetic signal, this effect could possibly be compounded. If a dataset is not representative of a set of taxa due to the omission of taxa showing phenological stasis, we may either: (1) incorrectly detect a signal; or (2) not detect a signal when there is one. Given that incorporating phenological stasis extends the range of traits values evaluated (to include very small effect sizes of phenological change), which in turn introduces greater trait variation to evolutionary models, we hypothesize that including phenological stasis will decrease the chance of detecting phylogenetic signal.

Here, we present a case study where we use a selection of grasses (Poaceae) as a focal group to understand how the ability to respond to environmental change of this set of organisms has been shaped by evolution over deep time. We chose to work with grasses for several reasons. Grass phenology is of broad economic importance, but phenological responses of grasses to climate shifts have not been well-studied. In addition, because grasses generally do not rely on biotic pollination, we might expect a more direct link between flowering time and climatic variables. Furthermore, a well-resolved phylogeny is available (Spriggs et al., 2014), and lastly, although grasses are often considered difficult to identify, the presence of a relatively large and well-curated collection of grasses at the University of British Columbia herbarium (UBC) reduced the likelihood that errors in identification would confound analysis of phenological data. Focusing our study on the collections from a single, large, regional herbarium means that most of the specimens that we include occur within limited geographic boundaries (85% of records are between northern latitudes of 49 and 70 degrees, and western longitudes of –110 and –140 degrees), a region that has both already undergone significant climate change, and that is predicted to change further in the coming decades (Environment and Climate Change Canada, 2020).

We model the macroevolutionary dynamics of the magnitude and direction of phenological sensitivity using a Bayesian approach for fitting Ornstein-Uhlenbeck (OU) models (Uyeda and Harmon, 2014) that allows for different parts of the tree to have different evolutionary optima and simultaneously identifies branches on the tree where there is evidence that optima shifted. This allows us to ask three complementary questions in the context of Poaceae: first, are different clades responding to climate in different ways? Second, what is the overall level of phylogenetic signal in the data (accounting for clade-specific differences)? And third, how might the detection of phylogenetic signals be influenced by unreported phenological stasis? Determining the answers to all three of these questions is essential for determining whether we can make predictions about how data-deficient species will respond to changing climate from information on their close relatives.

MATERIALS AND METHODS

Specimen data

We focused on species of grasses included in the family-wide time-calibrated phylogeny of Poaceae (Spriggs et al., 2014) and represented in the collections of the University of British Columbia (UBC) herbarium (Vancouver, Canada). We included all available accessioned specimens, totaling 8674 specimens across 418 species. To address the potential for non-independent records, we removed duplicate specimens, which we define as those collected at the same site on the same day. Each specimen was manually inspected with the aid of a dissecting microscope and scored for a single trait, i.e., the presence or absence of flowers with anthers (Appendix S1). This trait was scored as binary, with anthers present on a specimen corresponding to flowering and no anthers present corresponding to not flowering. This 'reproductiveCondition' data was entered into the UBC database, following DarwinCore standards (Wieczorek, 2012). We focused on anthesis, as this is an identifiable phenological stage that occurs for a limited duration of time (Abel and Boelt, 2018)—*Lolium perenne* L. for example, is typically in anthesis for

less than three weeks. Note that for simplicity, we refer to anthesis as flowering in the remainder of the manuscript. The taxonomic identity of these specimens was recently confirmed by specialists in grass taxonomy working with the collection, reassuring us of a low likelihood of misidentified specimens. After removing species represented by fewer than 10 records, our dataset comprised 3161 records from 72 species (see Appendix S2 for latitudinal range and number of records per species). We chose this threshold number of records to capture a reasonable amount of variation in collection dates and to ensure that we could fit a linear regression with the data for each species.

Historical climate data

For each specimen, we used the latitude and longitude of the collection location and downloaded two climate variables, average temperature, and average precipitation—for each month during the year of collection, using the desktop version of climateNA (Wang et al., 2016). This historical climate model provides climate data at 4×4 km resolution. Latitude and longitude were sourced from the UBC database; these coordinates were either originally provided by collectors or estimated using a UBC algorithm that matches known place names both in British Columbia and surrounding provinces/states to latitude and longitude coordinates.

Sensitivity to climate

For each species, we interpreted the mean collection date of specimens containing anthers as the mean flowering date and used those dates to assign temperature and precipitation data to each observation for the months prior to flowering (Prevéy et al., 2017). To select the variable that best predicted mean flowering date, we evaluated the effects of temperature and precipitation in the months prior to flowering using our entire dataset, with species identity set as a random effect. We compared AIC scores for the set of linear models considered to identify a single variable to assess sensitivity of phenology of grasses (Burnham et al., 2011), and therefore allowing us to evaluate whether this sensitivity can be predicted for data-deficient species on the tree. Average temperature one and two months prior to mean flowering date showed the lowest AIC scores (Appendix S3). To assess the sensitivity of our phylogenetic analysis to our climate variable selection, we ran these in turn using each of the two temperature variables, as well as precipitation two months prior to flowering (the precipitation variable with the lowest AIC score). We note that in an earlier preliminary analysis, we evaluated the effect of latitude on flowering date and found that this explained less than 1% of the variation. As a result, this variable is omitted from analyses presented here.

For each species we estimated their phenological sensitivity to each climate variable as the slope and standard error of a linear model between each climate variable and date of flowering. While this regression slope is not a direct measurement of phenological sensitivity to temperature of any given population, we followed the approach used in previous studies of other emergent traits (Voje et al., 2014; Bolstad et al., 2015; Muir and Thomas-Huebner, 2015; Uyeda et al., 2017) and assumed that the slope, an estimate of the species' phenological sensitivity to climate, as assessed by our sampling and within our study region, is an underlying quantitative trait of each species.

Analysis of phylogenetic signal in phenological traits

We determined the shape of the macroevolutionary adaptive landscape (Simpson, 1944; Hansen, 2012) of the phenological sensitivity to climate in our sample of grass species by using a published phylogeny of Poaceae (Spriggs et al., 2014). We used the R package phytools (Revell, 2012) to estimate phylogenetic signal (Blomberg's K; Blomberg et al., 2003) in flowering date as well as in phenological sensitivity. Values of $K < 1$ indicate that closely related taxa are less similar than expected under a model of Brownian motion, and $K > 1$ indicates greater similarity than expected under Brownian motion (Blomberg et al., 2003). We also modeled phylogenetic shifts using another phylogenetic comparative method, bayou (Uyeda and Harmon, 2014) to retrieve an alternate metric of phylogenetic signal, i.e., phylogenetic half-life. This method is based on an OU model of trait evolution (Hansen, 1997); (Butler and King, 2004), which assumes that lineages are pulled towards a macroevolutionary optimum. Bayou uses a reversible-jump Markov chain Monte Carlo (MCMC) process to identify branches on the phylogeny where this optimum has shifted (Green, 1995). In effect, bayou explores the entire spectrum of possible OU models, ranging from the simplest model, in which all lineages share a single macroevolutionary optimum, to the most complex, in which each lineage has its own evolutionary dynamics. This approach serves the dual purpose of allowing us to investigate the macroevolutionary dynamics of our traits and provides us with a measure of phylogenetic signal—the phylogenetic half-life (Hansen, 1997; Uyeda and Harmon, 2014; Boucher et al., 2018)—that is not based on Brownian motion or overly simplistic single-optimum OU models, which often poorly describe the patterns in real comparative data (Harmon et al., 2010; Pennell et al., 2015). Phylogenetic half-life quantifies the extent to which the trait displays evolutionary inertia (Hansen and Orzack, 2005); if the half-life is long relative to the depth of the phylogeny, then the macroevolutionary history of a trait is a good predictor of its current value and if it is short, it is not. These models have been used to describe the macroevolution of function-based traits (Muir and Thomas-Huebner, 2015; Uyeda et al., 2017), as we have done here. We think that phylogenetic half-life is more meaningful to interpret than other measures of phylogenetic signal because it is in units of time rather than a 0 to 1 scale. As such, it provides a measure of evolutionary constraint that is better linked to the evolutionary process we are interested in (e.g., how long it takes for a lineage to move towards a new adaptive optimum).

As bayou is a Bayesian approach, we needed to set priors for the model. We set equal probability of a shift on any given branch, with a one-shift maximum per branch, and the priors for the following parameters, as given: $P(\alpha) \sim$ half-Cauchy(scale = 0.1); $P(\sigma^2) \sim$ half-Cauchy(scale = 0.1); $P(\theta) \sim$ LogNormal($\ln \mu$ = mean(trait values), $\ln \sigma$ = 1.5(sd(trait values))); $P(K) \sim$ ConditionalPoisson(λ=10, Kmax = 50). These standard, non-informative priors are taken from Uyeda and Harmon (2014). For each trait (mean flowering date and our three phenological sensitivity variables), we ran two MCMC chains for 1 million generations each. We assessed convergence of parameter estimates using Gelman's R statistic, where values substantially greater than 1 indicate the chains have not converged. We also ran two MCMC chains as above using only the subset of sensitivity estimates from species for which we had evidence of significant sensitivity to climate. The aim of this last analysis was to explore the potential effect of sampling bias in increasing the likelihood of

inferring macroevolutionary signals, i.e., the effect of only including species for which there are evident shifts in phenology.

The code used for these analyses is available at website https://github.com/bnetobradley/grassphenology.

RESULTS

We detect significant relationships between temperature two months prior to mean flowering and date of flowering for 22 of the 72 species (Appendix S4). Across all 72 taxa, our temperature variable explains between <1% and 84% of the variation in flowering dates (Appendix S4). The slopes for the linear models across all species in our dataset vary from −12.6 to 6.0, representing a range of sensitivities from 12.6-day advances to 6.0-day delays in flowering for every increase of 1°C (Fig. 1). All the 22 significant relationships have negative slopes indicating advances in flowering dates with increases in mean temperature, with sensitivities ranging from −12.6 to −1.0.

Across our dataset, we find that closely related species share similar mean flowering dates (Fig. 2). This result emerged in both phylogenetic analyses, i.e., close relatives are almost exactly as similar as expected under a Brownian motion model (n = 72, K = 0.94, $p = 0.001$, Table 1), and the half-life of mean flowering time is estimated as 29.80 myr (Table 1), also suggesting phylogenetic signal persists over moderately deep time. Despite this, closely related species do not share similar sensitivity to temperature; across all 72 estimates of phenological sensitivity and accounting for the standard error of those estimates, the sensitivity of flowering date to temperature had significant phylogenetic signal (compared to the null hypothesis of no phylogenetic signal at all) but substantially less phylogenetic structure than predicted under a Brownian motion model (n = 72, K = 0.65, $P = 0.002$, Table 1). But in this case, we find support for a model of many (14) small evolutionary shifts in phenological sensitivity, distributed across the tree, with a phylogenetic half-life of sensitivity estimated as 5.1 myr (Table 1). This half-life estimate represents only approximately 10% of the tree depth.

Our results differed when we only used the 22 species with significant sensitivity estimates. With only this subset included, we find that closely related species are very similar to each other (even more so than expected under a Brownian motion model; n = 22, K = 1.51, $p = 0.003$) but that this is only true of recent divergences; phylogenetic signal falls off precipitously as one looks further back in time; the inferred phylogenetic half-life is only 3.9 myr, with even more inferred switches in trait optima (number of supported shifts = 36, Table 1). Overall, this pattern of results is indicative of a non-random sample of possible sensitivity values across the phylogeny for this set of species.

To assess the sensitivity of our results to our chosen climate variables, we also ran our analyses with two other climate variables: (1) temperature one month prior; and (2) precipitation two months prior to mean flowering). In both cases, the magnitude of phylogenetic half-life and the strength of phylogenetic signal are comparable to those reported with temperature two months prior to mean flowering (Table 1). Further, as seen with our main climate variable, both phylogenetic signal and the number of supported shifts increase as a result of subsetting to only include species with significant slope estimates.

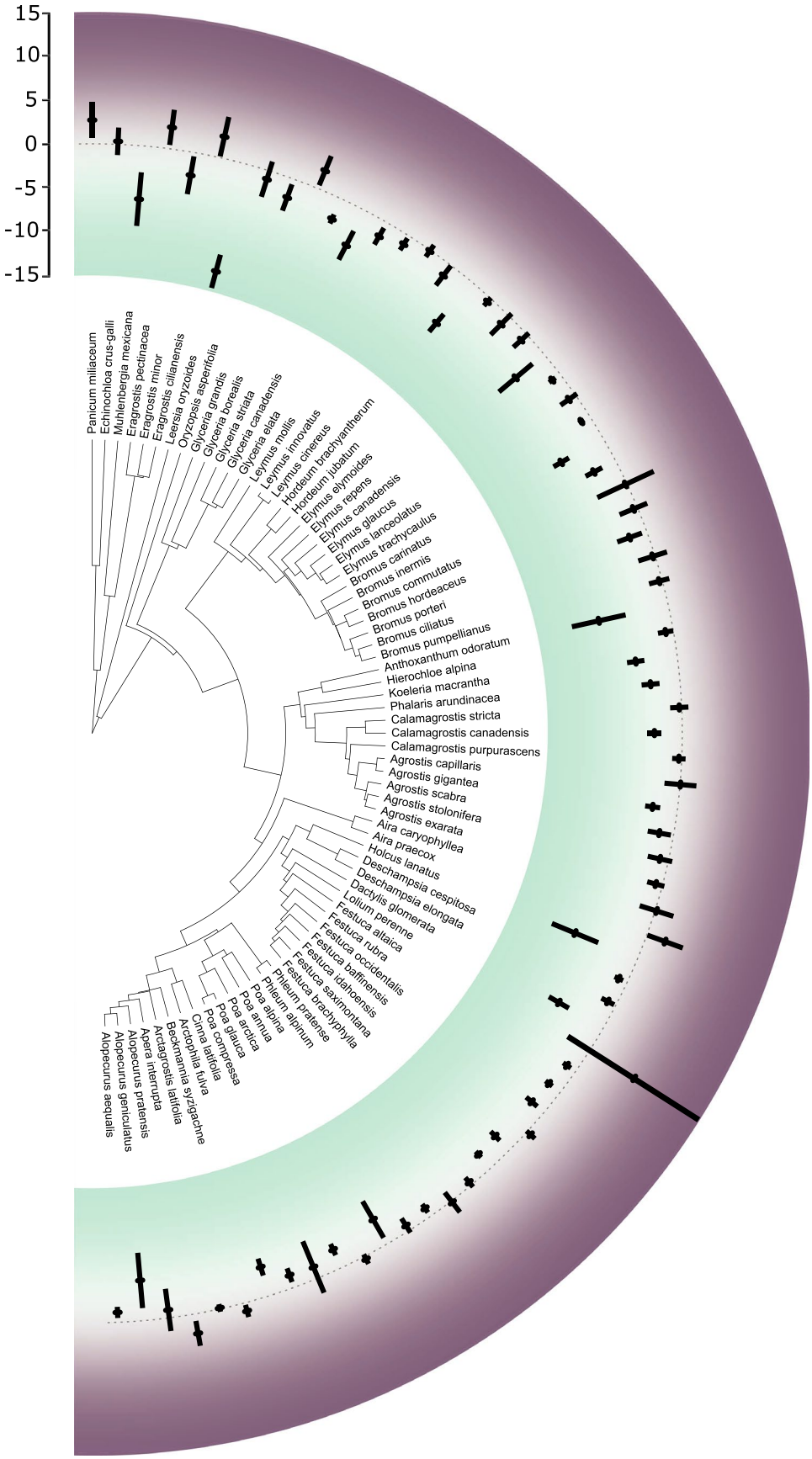
DISCUSSION

Using a novel dataset for grass species focused on the Pacific Northwest, we find a complex macroevolutionary relationship between phenology and our environmental variables. We detect significant phylogenetic signal in the estimates of species' mean flowering dates, with close relatives being almost as similar as expected under a model of Brownian motion (i.e., $K > 1$). Despite phylogenetic signal in flowering time, we find mixed evidence for phylogenetic patterns in phenological sensitivity to temperature. There is no strong phylogenetic signal to indicate that close relatives respond similarly to variation in temperature when using data from all 72 species, but a signal emerges when we include only species with statistically significant estimates of sensitivity. As outlined below, we interpret this difference in the inferred signal in phenological sensitivity as an underappreciated potential source of error in previously reported phylogenetic comparisons (CaraDonna and Inouye, 2015; Rafferty and Nabity, 2017). In contrast to temperature, across phylogenetic methods and sampling of significant effects, the phylogenetic patterns of sensitivity to precipitation are more consistent, with K > 1 and similar half-life estimates regardless of the inclusion of all or only significant species-level estimates of sensitivity (Table 1).

Consistent with the findings of Davies et al. (2013), who detected a broad scale phylogenetic signal in the pattern of first leaf-out dates and first flowering dates across angiosperms, we reach a similar conclusion using slightly different metrics of phylogenetic signal for mean flowering date in our sample of grass species (Fig. 2). While these findings are not necessarily indicative of predictive power, it is noteworthy that mean flowering appears to be reasonably well-conserved, especially within genera. For example, most species of *Glyceria* in our survey had mean flowering dates in early June, whereas species of *Festuca* typically flowered a month or more later (Fig. 2).

We recognize that there are sources of error and/or uncertainty that can impact the detection and significance of relationships between flowering date and temperature. For example, large effects are more likely to be statistically significant in comparison to estimates of stasis. To better estimate and assess impacts related to the magnitude of effects, we incorporate the standard error of our estimate of sensitivity into our phylogenetic analyses. In doing so, we find that species' sensitivity to temperature appears not to be phylogenetically conserved, because while some taxa display large advances

FIGURE 1. Phylogenetic distribution of phenological sensitivity to temperature (mean temperature 2 months prior to species' mean flowering date). Estimates in the green portion of the outer arc (to the left of the zero line) indicate advances in flowering dates with increases in temperature; estimates in the purple area (to the right of the zero line) show delays in flowering dates with increases in temperature. For each species, the central ellipse on each bar shows the estimate of phenological sensitivity (measured as the change in flowering date per degree Celsius), and the bar depicts one standard error deviation either side of this estimate. (For *Lolium perenne*, the upper limit of this deviation has been clipped for visual ease.) Phylogeny and data were plotted using iTOL (Letunic and Bork, 2016).


15
10
5
0
-5
-10
-15
Panicum miliaceum
Echinochloa crus-galli
Muhlenbergia mexicana
Eragrostis pectinacea
Eragrostis minor
Eragrostis ciliatensis
Leersia oryzoides
Oryzopsis asperifolia
Glyceria grandis
Glyceria borealis
Glyceria striata
Glyceria canadensis
Leymus elata
Leymus mollis
Leymus innovatus
Leymus cinereus
Hordeum brachyantherum
Hordeum jubatum
Elymus elymoides
Elymus repens
Elymus canadensis
Elymus glaucus
Elymus lanceolatus
Elymus trachycaulus
Elymus carinatus
Bromus inermis
Bromus commutatus
Bromus hordeaceus
Bromus porteri
Bromus ciliatus
Bromus pumpellianus
Anthoxanthum odoratum
Hierochloe alpina
Koeleria macrantha
Phalaris arundinacea
Calamagrostis stricta
Calamagrostis canadensis
Calamagrostis purpurascens
Agrostis capillaris
Agrostis gigantea
Agrostis scabra
Agrostis stolonifera
Agrostis exarata
Aira caryophyllea
Aira praecox
Holcus lanatus
Deschampsia cespitosa
Deschampsia elongata
Dactylis glomerata
Lolium perenne
Festuca altaica
Festuca rubra
Festuca occidentalis
Festuca baffinensis
Festuca idahoensis
Festuca saximontana
Festuca brachyphylla
Phleum pratense
Phleum alpinum
Poa alpina
Poa annua
Poa arctica
Poa glauca
Poa compressa
Cinna latifolia
Arctophila fulva
Beckmannia syzigachne
Arctagrostis latifolia
Apera interrupta
Alopecurus pratensis
Alopecurus geniculatus
Alopecurus aequalis

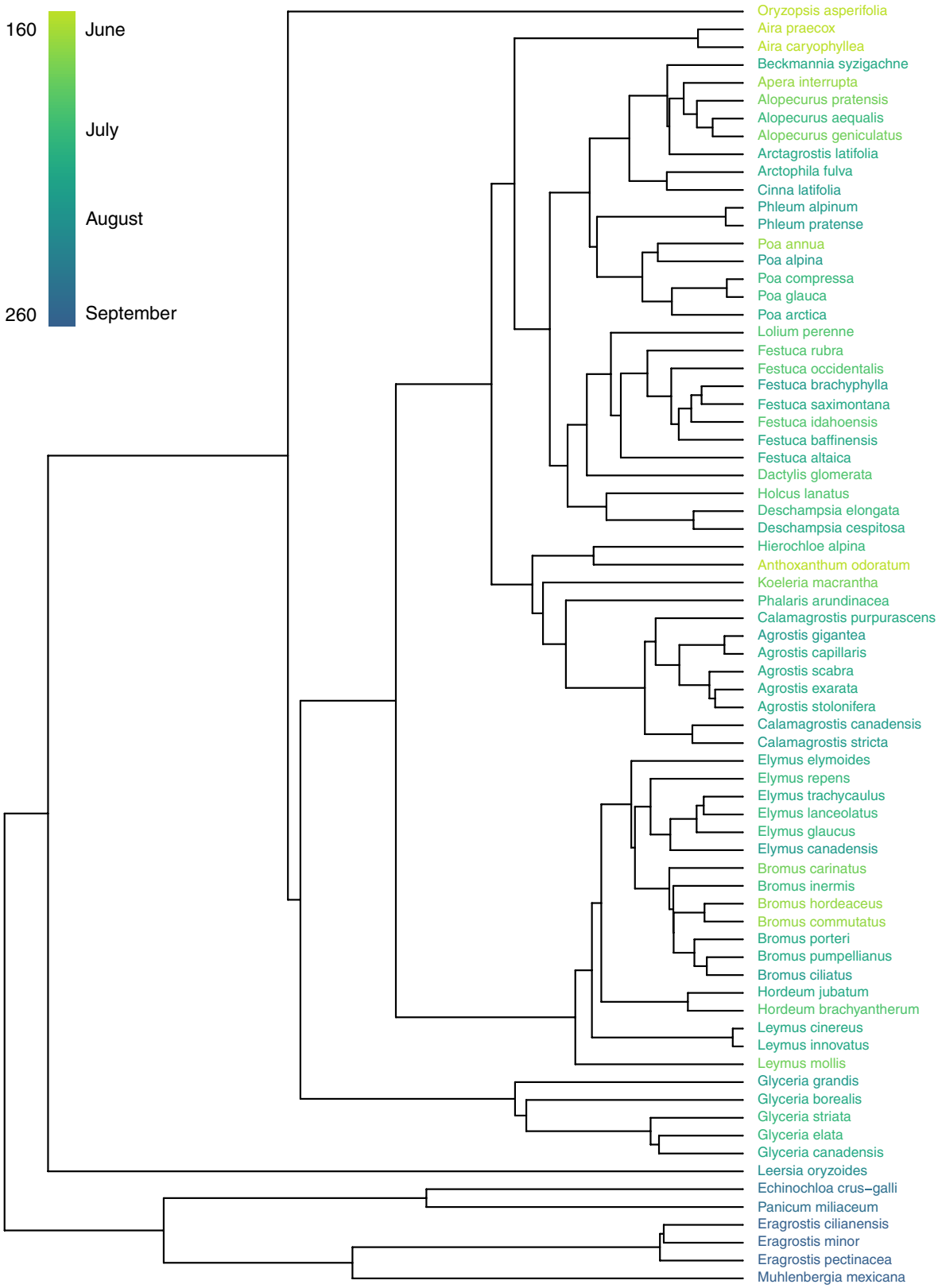


FIGURE 2. Trimmed phylogenetic tree from Spriggs et al. (2014) showing the distribution of mean flowering date. For each species, the mean day-of-year of herbarium specimen collection is indicated via the color of the tip label, from light green (early mean flowering date) to blue (later mean flowering dates), as indicated on the scale bar.

TABLE 1. Results of phylogenetic half-life and number of shifts.

		Bayou Analysis		Phylogenetic signal	
	N	Half-life (Mya)	Supported shifts (posterior probabilities > 0.1)	Blomberg's K	*p*-value
Flowering time					
Only anthers (mean)	72	29.80	10	0.94	0.001
Flowering sensitivity					
Temperature month prior	72	1.55	13	0.79	0.002
Temperature month prior, only significant	20	2.49	30	1.50	0.06
Temperature 2 months prior	72	5.07	14	0.65	0.002
Temperature 2 months prior, only significant	22	3.90	36	1.51	0.003
Precipitation 2 months prior	72	4.20	8	1.39	0.001
Precipitation 2 months prior, only significant	9	5.70	16	1.50	0.001

in flowering dates relative to temperature, many taxa show little or no sensitivity and these differential effects are not phylogenetically structured. Perhaps more tellingly, the very small half-life of sensitivity on the phylogeny suggests that there is sparse retention of information on phenological sensitivity throughout the phylogeny. Below, we present two ways of resolving this apparent discordance in phylogenetic signals between flowering dates and flowering sensitivity.

(1) Our finding of phylogenetic signal in flowering date and lack thereof in flowering sensitivity suggests the possibility that flowering time may be decoupled from geographic location. These results are consistent with our understanding of the contribution of genetic factors in determining flowering times (Andrés and Coupland, 2012; Du et al., 2015), and of the phenotypic responses of flowering time to changes in environmental conditions (CaraDonna and Inouye, 2015). Closely related species that inhabit similar conditions are likely to flower at similar times both because of similarities in developmental pathways, and because they are likely to experience similar climates. However, as species diverge in experienced environmental conditions, it becomes more apparent that the sensitivity of flowering times may actually be a highly plastic trait where the extent of the plasticity displayed may be linked to other, possibly non-phylogenetically related traits, such as the propensity for a species' dispersal to new environments (Lande, 2015; Murren et al., 2015).

Even though life-history theory predicts stabilizing selection on flowering times (reviewed in (Elzinga et al., 2007)), it is inappropriate to directly interpret the parameters estimated from bayou as reflecting the strength of stabilizing selection (Lande, 1976; Hansen and Martins, 1996). It is more justifiable to view the OU model's fit to comparative data as being a phenomenological description of the macroevolutionary adaptive landscape (Hansen, 1997; Pennell and Harmon, 2013). Furthermore, because we are describing the dynamics of a function- based trait which is not literally a property of the organism but rather is a proxy for an unknown trait or combination of traits, we cannot use this information to infer genetic or developmental constraints that might explain the macroevolutionary distribution of the observed floral sensitivities.

(2) An alternative explanation for our finding of low phylogenetic half-life in climate sensitivity is that the slope estimates we used are not an accurate proxy for the true sensitivity of grass species to their environmental cues for flowering. This could be the case for example, if different populations within a species have different sensitivities to climate variables. In this study, we implicitly assume that phenological sensitivity is a species-level trait, accurately estimated by the samples included in our analysis. Our regional emphasis, with geographic sampling mainly focused on British Columbia (from which 85% of specimens were collected), reduced the range of flowering dates and the range of climate variables used to infer sensitivities, especially for species with wide ranges outside this area. This approach may have introduced error into our phenological trait estimates, but our focus here is on comparing the relative phylogenetic signal in flowering date with the signal in phenological sensitivity, and we see no simple reason to expect that these relationships should be biased in a particular direction by sampling approach. On balance, we favored an approach that emphasized taxonomic consistency and narrower scoring of flowering status (i.e., using the presence of anthers to indicate flowering), which likely reduced other sources of error that may have been introduced with different data sets. Ultimately, if a consensus about the macroevolutionary signal in phenological traits emerges, it will be on the strength of multiple studies in diverse taxa over varying spatial, environmental, and phylogenetic scales.

Another drawback in our use of herbarium specimens is that for any given collection, it is impossible to know how the phenology of the taxa being studied differs from that of its collector and their reasons for collecting. If a collector aims to collect flower instead of seed, there is no post-hoc way of knowing how representative this record is. While we acknowledge that working with herbarium specimens as a proxy for measuring flowering date under controlled conditions increases the uncertainty in our estimates of sensitivity of flowering to temperature, our finding of phylogenetic signal in mean flowering dates in contrast to sensitivity indicates that this is unlikely a pivotal element to our findings. Since uncertainty around collection and flowering dates should then apply to both traits (mean flowering date and sensitivity), it seems more plausible that our results are explained by species having evolved unique floral cues as a result of their dispersal to and colonization of different regions (Colasanti and Coneva, 2009). In contrast to the simplifying assumption we make—that species share a common climatic driver—if the evolution of species specific climatic floral cues is the rule and not the exception, a broad approach such as the one taken here (as have others; see van de Pol and Cockburn, 2011) would be inadequate in tackling how and whether species are responding —let alone for use in attempting to estimate the presence and magnitude of a phylogenetic pattern.

We want to underscore that had we included only those species that display significant sensitivity to temperature, we would have

biased our results towards species with large effect sizes. This exclusion constrains the variation in our estimates of phylogenetic signal, such that the range of trait values remaining appears narrower than the true range of trait values in the dataset. When using Blomberg's K to measure phylogenetic signal, this creates an inflated estimate of similarity among taxa (Table 1). However, when using phylogenetic half-life, the bias introduced by including only species with significant sensitivities tends to reduce the estimate of half-life (for two of our three traits), while increasing the number of inferred shifts in the rate of evolution of the trait. Estimating half-life elucidates some of the evolutionary dynamics in ways that Blomberg's K cannot, and in this case, both paints a picture of a very disjointed evolution of these traits and provides an example of how biased inclusion of data may severely limit our ability to uncover underlying evolutionary dynamics of a system. This is especially important when agglomerating published data to search for phylogenetic patterns, and we hope our analysis can help reduce the likelihood that this type of confirmation bias leads to overestimates of phylogenetic signal.

CONCLUSIONS

While many studies quantify temporal trends in phenology (reviewed in ((Willis et al., 2017)), they may hold little predictive value for estimating how species will respond to future variable climate. As new phenological studies lean toward measuring sensitivity to climate variables in comparative frameworks (CaraDonna and Inouye, 2015; Daru et al., 2019; Li et al., 2019), we suggest this is a more generalizable metric, and likely to relay more information regarding the underlying effects of climate on mechanisms that underlie flowering phenology. Contrary to our expectation, we observe both a lack of clear shifts in evolutionary dynamics as well as poor retention of information across the tree. This work provides evidence that there may be little predictive value across species in this single-environmental cue framework. There is a need to continue assessing other genus- and family-level macroevolutionary dynamics of sensitivity to understand whether the lack of conserved sensitivity is a widespread phenomenon. Future work should also attempt to consider phenological responses in a macroevolutionary framework as variable in both their environmental cues as well their sensitivity.

ACKNOWLEDGMENTS

We thank the many individuals who have contributed to the collection, identification, and curation of specimens used here—notably the efforts of J. Saarela, F. Lomer, and P. Zika that have provided a taxonomically reliable dataset. We thank D. Rowswell, K. Van Roost, and J. Oshiro for help with databasing herbarium specimens, as well as Z. Panchen, J. Rolland, K. Kaur, L.F. Henao Diaz, M. Sugawara, O. Li, and E. Ho for early feedback on the manuscript. We also thank the two anonymous reviewers whose insightful comments helped improve and clarify this manuscript. This work was supported by a NSERC Discovery Grant to M.W.P.

AUTHOR CONTRIBUTIONS

B.N.B., J.W., L.P.L.J., and M.W.P. designed the study. B.N.B. collected and analyzed the data. B.N.B., J.W., L.P.L.J., and M.W.P. wrote the paper.

DATA AVAILABILITY STATEMENT

The herbarium data used in this study are available in the online supplement (Appendix S5). For the most up-to-date versions of these records (which may include, for example, new additional data, metadata, or taxonomic updates) we recommend directly accessing the original records via the UBC herbarium database at website https://herbweb.botany.ubc.ca/herbarium/.

The code used for these analyses is available at website https://github.com/bnetobradley/grassphenology.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

APPENDIX S1. A visual description of the number and proportion of herbarium records for each species that held anthers included in our analyses. The total number of herbarium specimens housed at UBC is colored in grey. The subset of these specimens that also had anthers present on them is shaded in dark blue. Each specimen was inspected for anthers with the aid of a dissecting microscope.

APPENDIX S2. Map of herbarium records, i.e., a depiction of the geographic spread of all herbarium records (part 1) and of each of the 72 species (parts 2–7) included in our analyses that held anthers (i.e., flowering specimens). The color of the point representing a herbarium record indicates the day of year (doy) on which it was collected and for which it has subsequently been recorded as being in flower. Flowering was determined by inspecting each specimen for the presence of anthers with the aid of a dissecting microscope. A 2D-density plot is shaded in navy on each individual species map, this shading depicts the density of North American records available for that species on GBIF. Note that this shading scheme does not always overlap with the locations we have records from (see for example, *Agrostis gigantea* Roth and *Bromus carinatus* Hook. & Arn.). Since shading is based on the density of GBIF records, this lack of overlap is driven by the fact that there are higher densities of records in other locations. These records were accessed through the R package "spocc" (Chamberlain, 2020) on 20 November 2020.

APPENDIX S3. Summary table of the comparison between models using temperature or precipitation in the months prior to mean flowering, as an explanatory variable for the day of year in which flowering occurs. Average temperature one and two months prior to mean flowering show the lowest AIC scores. We use two months prior as our main climate variable and use temperature in the month prior as well as precipitation two months prior to mean flowering to assess the sensitivity of our choice of climate variables.

APPENDIX S4. Summary table reporting for each species in our analysis, i.e., the slope, standard error, *p*-value, and r squared for the linear model fit using average temperature two months prior to flowering as an explanatory variable for flowering time. (Uploaded as a .csv file, see: Neto-Bradley_et_al_AJB_2020_Appendix_S4.csv)

APPENDIX S5. Voucher information for the UBC herbarium records used in this study. (Uploaded as a .csv file, see: Neto-Bradley_et_al_AJB_2020_Appendix_S5.csv)

LITERATURE CITED

Abel, S., and B. Boelt. 2018. Precipitation during anthesis reduces seed set in perennial ryegrass. *Grass and Forage Science* 73: 239–246.

Andrés, F., and G. Coupland. 2012. The genetic basis of flowering responses to seasonal cues. *Nature Reviews Genetics* 13: 627–639.

Blomberg, S. P., T. Garland Jr, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution* 57: 717–745.

Bolstad, G. H., J. A. Cassara, E. Márquez, T. F. Hansen, K. van der Linde, D. Houle, and C. Pélabon. 2015. Complex constraints on allometry revealed by artificial selection on the wing of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, USA* 112: 13284–13289.

Boucher, F. C., V. Démery, E. Conti, L. J. Harmon, and J. Uyeda. 2018. A general model for estimating macroevolutionary landscapes. *Systematic Biology* 67: 304–319.

Burnham, K. P., D. R. Anderson, and K. P. Huyvaert. 2011. AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. *Behavioral Ecology and Sociobiology* 65: 23–35.

Butler, M. A., and A. A. King. 2004. Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *American Naturalist* 164: 683–695.

CaraDonna, P. J., and D. W. Inouye. 2015. Phenological responses to climate change do not exhibit phylogenetic signal in a subalpine plant community. *Ecology* 96: 355–361.

Chamberlain, S. 2020. spocc: Interface to Species Occurrence Data Sources. R package version 1.1.0. Website: https://CRAN.R-project.org/package=spocc.

Colasanti, J., and V. Coneva. 2009. Mechanisms of floral induction in grasses: Something borrowed, something new. *Plant Physiology* 149: 56–62.

Daru, B. H., M. M. Kling, E. K. Meineke, and A. E. van Wyk. 2019. Temperature controls phenology in continuously flowering *Protea* species of subtropical Africa. *Applications in Slant Sciences* 7: e01232.

Davies, T. J., E. M. Wolkovich, N. J. B. Kraft, N. Salamin, J. M. Allen, T. R. Ault, J. L. Betancourt, et al. 2013. Phylogenetic conservatism in plant phenology. *Journal of Ecology* 101: 1520–1530.

Davis, C. C., C. G. Willis, R. B. Primack, and A. J. Miller-Rushing. 2010. The importance of phylogeny to the study of phenological response to global climate change. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 365: 3201–3213.

Du, Y., L. Mao, S. A. Queenborough, R. P. Freckleton, B. Chen, and K. Ma. 2015. Phylogenetic constraints and trait correlates of flowering phenology in the angiosperm flora of China: Angiosperm flowering phenology in China. *Global Ecology and Biogeography* 24: 928–938.

Elzinga, J. A., A. Atlan, A. Biere, L. Gigord, A. E. Weis, and G. Bernasconi. 2007. Time after time: Flowering phenology and biotic interactions. *Trends in Ecology & Evolution* 22: 432–439.

Environment and Climate Change Canada. 2020. Canadian Centre for Climate Services, Gatineau, Quebec, Canada. Website: https://www.canada.ca/en/environment-climate-change/services/climate-change/canadian-centre-climate-services/display-download.html [accessed 24 November 2020].

Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution* 19: 497–503.

Green, P. J. 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82: 711–732.

Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51: 1341–1351.

Hansen, T. F. 2012. Adaptive landscapes and macroevolutionary dynamics. *In* E. I. Svensson and Calsbeek, R. [eds.], The adaptive landscape in evolutionary biology, 203–226.Oxford University Press, Oxford, UK.

Hansen, T. F., and E. P. Martins. 1996. Translating between microevolutionary process and macroevolutionary patterns: The correlation structure of interspecific data. *Evolution* 50: 1404.

Hansen, T. F., and S. H. Orzack. 2005. Assessing current adaptation and phylogenetic inertia as explanations of trait evolution: The need for controlled comparisons. *Evolution* 59: 2063–2072.

Harmon, L. J., J. B. Losos, T. J. Davies, R. G. Gillespie, J. L. Gittleman, W. Bryan Jennings, K. H. Kozak, et al. 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64: 2385–2396.

Kopp, C. W., B. M. Neto-Bradley, L. P. J. Lipsen, J. Sandhar, and S. Smith. 2020. Herbarium records indicate variation in bloom-time sensitivity to temperature across a geographically diverse region. *International Journal of Biometeorology* 64: 873–880.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30: 314–334.

Lande, R. 2015. Evolution of phenotypic plasticity in colonizing species. *Molecular Ecology* 24: 2038–2045.

Lessard-Therrien, M., T. J. Davies, and K. Bolmgren. 2014. A phylogenetic comparative study of flowering phenology along an elevational gradient in the Canadian subarctic. *International Journal of Biometeorology* 58: 455–462.

Letunic, I., and P. Bork. 2016. Interactive tree of life (iTOL) version 3: An online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research* 44: W242–5.

Li, D., B. J. Stucky, J. Deck, B. Baiser, and R. P. Guralnick. 2019. The effect of urbanization on plant phenology depends on regional temperature. *Nature Ecology & Evolution* 3: 1661–1667.

Meineke, E. K., T. J. Davies, B. H. Daru, and C. C. Davis. 2019. Biological collections for understanding biodiversity in the Anthropocene. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 374: 20170386.

Morellato, L. P. C., B. Alberton, S. T. Alvarado, B. Borges, E. Buisson, M. G. G. Camargo, L. F. Cancian, et al. 2016. Linking plant phenology to conservation biology. *Biological Conservation* 195: 60–72.

Muir, C. D., and M. Thomas-Huebner. 2015. Constraint around quarter-power allometric scaling in wild tomatoes (*Solanum* sect. *Lycopersicon*; Solanaceae). *American Naturalist* 186: 421–433.

Munson, S. M., and A. L. Long. 2017. Climate drives shifts in grass reproductive phenology across the western USA. *New Phytologist* 213: 1945–1955.

Murren, C. J., J. R. Auld, H. Callahan, C. K. Ghalambor, C. A. Handelsman, M. A. Heskel, J. G. Kingsolver, et al. 2015. Constraints on the evolution of phenotypic plasticity: Limits and costs of phenotype and plasticity. *Heredity* 115: 293–301.

Panchen, Z. A., R. B. Primack, B. Nordt, E. R. Ellwood, A.-D. Stevens, S. S. Renner, C. G. Willis, et al. 2014. Leaf out times of temperate woody plants are related to phylogeny, deciduousness, growth habit and wood anatomy. *New Phytologist* 203: 1208–1219.

Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.

Pearse, W. D., C. C. Davis, D. W. Inouye, R. B. Primack, and T. J. Davies. 2017. A statistical estimator for determining the limits of contemporary and historic phenology. *Nature Ecology & Evolution* 1: 1876–1882.

Pennell, M. W., R. G. FitzJohn, W. K. Cornwell, and L. J. Harmon. 2015. Model adequacy and the macroevolution of angiosperm functional traits. *American Naturalist* 186: E33–50.

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

Prevéy, J., M. Vellend, N. Rüger, R. D. Hollister, A. D. Bjorkman, I. H. Myers-Smith, S. C. Elmendorf, et al. 2017. Greater temperature sensitivity of plant phenology at colder sites: Implications for convergence across northern latitudes. *Global Change Biology* 23: 2660–2671.

Primack, D., C. Imbres, R. B. Primack, A. J. Miller-Rushing, and P. Del Tredici. 2004. Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. *American Journal of Botany* 91: 1260–1264.

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. 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.

Simpson, G. G. 1944. Tempo and mode in evolution. Columbia University Press, New York, New York, USA.

Spriggs, E. L., P.-A. Christin, and E. J. Edwards. 2014. C_4 photosynthesis promoted species diversification during the Miocene grassland expansion. *PLoS One* 9: e97722.

Thiers, B. M., M. C. Tulig, and K. A. Watson. 2016. Digitization of The New York Botanical Garden herbarium. *Brittonia* 68: 324–333.

Uyeda, J. C., and L. J. Harmon. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Systematic Biology* 63: 902–918.

Uyeda, J. C., M. W. Pennell, E. T. Miller, R. Maia, and C. R. McClain. 2017. The evolution of energetic scaling across the vertebrate tree of life. *American Naturalist* 190: 185–199.

van de Pol, M., and A. Cockburn. 2011. Identifying the critical climatic time window that affects trait expression. *American Naturalist* 177: 698–707.

Voje, K. L., T. F. Hansen, C. K. Egset, G. H. Bolstad, and C. Pélabon. 2014. Allometric constraints and the evolution of allometry. *Evolution* 68: 866–885.

Wang, T., A. Hamann, D. Spittlehouse, and C. Carroll. 2016. Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS One* 11: e0156720.

Wieczorek, J., D. Bloom, R. Guralnick, S. Blum, M. Döring, et al. 2012. Darwin Core: an evolving community-developed biodiversity data standard. *PLoS One* 7: e29715.

Willis, C. G., E. R. Ellwood, R. B. Primack, C. C. Davis, K. D. Pearson, A. S. Gallinat, J. M. Yost, et al. 2017. Old plants, new tricks: Phenological research using herbarium specimens. *Trends in Ecology & Evolution* 32: 531–546.

Zohner, C. M., and S. S. Renner. 2014. Common garden comparison of the leaf-out phenology of woody species from different native climates, combined with herbarium records, forecasts long-term change. *Ecology Letters* 17: 1016–1025.