Check for updates

New Phytologist

Review

Research review

Integrating experiments to predict interactive cue effects on spring phenology with warming

Author for correspondence:
E. M. Wolkovich
Email: e.wolkovich@ubc.ca

Received: *15 December 2021*
Accepted: *21 April 2022*

E. M. Wolkovich[1,2,3], **C. J. Chamberlain**[2,3], **D. M. Buonaiuto**[2,3], **A. K. Ettinger**[4] and **I. Morales-Castilla**[5]

[1]Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada; [2]Arnold Arboretum of Harvard University, 1300 Centre Street, Boston, MA 02131, USA; [3]Organismic & Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA; [4]The Nature Conservancy, 74 Wall Street, Seattle, WA 98121, USA; [5]Department of Life Sciences, Global Change Ecology and Evolution Group, Universidad de Alcalá, Alcalá de Henares 28805, Spain

New Phytologist (2022)
doi: 10.1111/nph.18269

Key words: budburst, chilling, climate change, forcing, leafout, nonlinear responses, photoperiod, spring warming.

Summary

Climate change has advanced plant phenology globally 4–6 d °C^{-1} on average. Such shifts are some of the most reported and predictable biological impacts of rising temperatures. Yet as climate change has marched on, phenological shifts have appeared muted over recent decades – failing to match simple predictions of an advancing spring with continued warming. The main hypothesis for these changing trends is that interactions between spring phenological cues – long-documented in laboratory environments – are playing a greater role in natural environments due to climate change. Here, we argue that accurately linking shifts observed in long-term data to underlying phenological cues is slowed by biases in observational studies and limited integration of insights from laboratory studies. We synthesize seven decades of laboratory experiments to quantify how phenological cue-space has been studied and how treatments compare with shifts caused by climate change. Most studies focus on one cue, limiting our ability to make accurate predictions, but some well-studied forest species offer opportunities to advance forecasting. We outline how greater integration of controlled-environment studies with long-term data could drive a new generation of laboratory experiments, built on physiological insights, that would transform our fundamental understanding of phenology and improve predictions.

Introduction

Shifts in spring plant phenology are one of the most reported and predictable changes with climate change. Decades of research have documented advancing budburst, leafout, and flowering with climate change (Delpierre *et al.*, 2009; Yu *et al.*, 2010; Ellwood *et al.*, 2013; Jochner *et al.*, 2013; Hereford *et al.*, 2017), especially in temperate systems where long-term records highlight how humans have altered the timing of spring (Schwartz, 1997; Menzel *et al.*, 2006). Recently, however, these advances have appeared to slow (Fu *et al.*, 2015) or even reverse in some regions (Yu *et al.*, 2010), failing to match simple predictions of an advancing spring with continued warming (Ellwood *et al.*, 2013). The main hypothesis for this failure is that spring warming – which most observational studies focus on – is no longer the only environmental cue that matters to predicting responses to warming (Chuine *et al.*, 2016; Gauzere *et al.*, 2019).

Despite the strong focus on spring temperatures, increasing evidence in climate change studies suggests a more complicated physiology (e.g. Zohner *et al.*, 2016; Gauzere *et al.*, 2019; Ettinger *et al.*, 2020). For decades, studies of agricultural and model species have found that three major cues underlie spring phenology (Maurya & Bhalerao, 2017; Satake *et al.*, 2022): chilling (cool temperatures, generally occurring in the fall and winter), forcing (warm temperatures, generally occurring in the late winter and early spring), and photoperiod (day length).

Together, chilling, forcing, and photoperiod may produce nonlinear responses. Predicting these nonlinearities is a common goal in plant phenology research today (Güsewell *et al.*, 2017; Chen *et al.*, 2019), but this has been slowed by data gaps and the complexity

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.

of spring phenological cues. Nonlinear responses can generally occur through two alternative pathways. First, though each cue alone may yield linear responses in the mid-range of cue levels, extremely high or low levels can produce threshold responses (Heide, 1993; Partanen *et al.*, 1998; Singh *et al.*, 2017; Rinne *et al.*, 2018). Such extreme cue levels are uncommon in natural environmental regimes and thus unlikely to explain recent observations of shifts. Second, interactions can produce nonlinearities.

Most hypotheses for the observed decline instead invoke the interaction between cues to produce nonlinear responses to warming. Yet estimating the effect of interactions from observational data, where cues themselves often covary, may produce spurious results (Ettinger *et al.*, 2020). By contrast, controlled-environment (e.g. growth-chamber) studies are designed to understand and estimate multiple cues and can tease out their interactive effects. Such studies have been critical both in our understanding of interactions from a more simple statistical perspective (where we estimate how the effect of one cue depends on another cue) and from a deeper understanding of the developmental pathways that produce observed interactions.

Here, we argue that controlled-environment studies to estimate cue interactions are critical to robust forecasts of phenological shifts with climate change. Such studies have been conducted for over 70 yr, but they have rarely been reviewed and are often poorly integrated into the current phenological literature on climate change (e.g. Fu *et al.*, 2015; Richardson *et al.*, 2018). To address this gap, we synthesize these studies (Supporting Information Fig. S1) to quantify how much of the cue-space (i.e. the possible range of each cue and interactions across cues) has been studied and how treatments (i.e. levels of cues) compare with shifts in cues caused by climate change. Based on this, we suggest a path forward where advances in physiology, greater integration of controlled environment studies into forecasting, and population-to-species modeling can yield more robust predictions.

Given our aim to improve understanding of current trends and forecasts, we focus on early vegetative phases (budburst and leafout), which are critical to plant growth and carbon storage, have shifted the most with climate change (Cleland *et al.*, 2007), and have been studied extensively in experiments (see Methods S1). Our review concentrates on woody species phenology, where most research on budburst and leafout has occurred. Most of our conclusions and suggested approaches, however, could be adapted to nonwoody species and other phenophases with similar underlying interactive cues.

Why observational studies are not enough to robustly estimate cues

The first step towards robust phenological predictions is accurate measurements of chilling, forcing, and photoperiod cues. Recently, efforts have focused on estimating these cues from long-term observational data (e.g. Luedeling *et al.*, 2013). Yet observational studies may struggle to accurately estimate any of the three major cues due to two types of statistical issues. First, in most observational data these cues are correlated: forcing increases alongside longer photoperiods (Elmendorf & Ettinger, 2020), and changes to chilling and photoperiod cues with warming predict the same response – later spring phenology (but see Notes S1 for how chilling may lead to advances). Approaches that attempt to disentangle some of these correlations, such as leveraging trends across elevation or latitude, may run afoul of other correlations (Tansey *et al.*, 2017). Second, most observational studies focus on linear models of each cue, often without estimating interactions between cues (Visser & Holleman, 2001; Polgar *et al.*, 2014). Even when interactions are invoked to explain declining responses to warming, they are often not estimated statistically (e.g. Fu *et al.*, 2015).

In contrast to observational studies, controlled-environment studies can manipulate each cue independently. By experimentally decoupling cues, interactions between them can be teased out, allowing both statistical estimates of interactions and insights into the developmental pathways that create them. Perhaps unsurprisingly then, most of our fundamental understanding of these cues comes from controlled-environment studies.

Controlled-environment studies have shown that chilling, forcing, and photoperiod cues together determine the transition from dormancy to growth (usually first observed as budburst) each year for most temperate woody species (van der Schoot *et al.*, 2014; Chuine & Regniere, 2017; Singh *et al.*, 2017; Chang *et al.*, 2021). Chilling and forcing cues are generally understood to be accumulation processes, where plants integrate chilling and forcing experienced over time to meet a threshold value at which they can budburst, leafout, or flower (Chuine & Regniere, 2017). By contrast, photoperiod is generally not considered an integral cue, but one evaluated daily. In practice, these two types of processes are often abstracted into the effects of average experienced values (mean temperature or day length), generally over some temporal window in long-term data (e.g. Wolkovich *et al.*, 2012; Fu *et al.*, 2015), or by holding temperatures and light regimes more constant in experiments (e.g. Worrall & Mergen, 1967; Heide, 1993; Skuterud & Dietrichson, 1994).

Combined with molecular approaches, controlled-environment studies have led to our current models of the pathways that control phenological events and create interactions. The most well developed of these, the flowering time pathway for *Arabidopsis thaliana* (reviewed in Satake *et al.*, 2022), links closely to chilling, photoperiod, and forcing: flowering time is mediated by cool temperatures though a vernalization pathway, day length through circadian clock control, and together these pathways limit or promote development driven by warm temperatures. Similar multicontrolled pathways appear across other species and events (Maurya & Bhalerao, 2017; Satake *et al.*, 2022), including in peach flowering and *Populus* budbreak (Azeez *et al.*, 2021). Across species, these pathways almost always have far greater complexity than generally considered in observational studies of spring events. Molecular studies show that plants may monitor temperatures (especially cool temperatures) over long time periods while also responding to short-term fluctuations, and new work on budbreak suggests epigenetic control, with multiple regulatory layers (Azeez *et al.*, 2021).

How climate change impacts cues

Climate change is likely altering each of the three major cues. Long-term data suggest that warming has already translated into shifts in

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.

forcing and chilling (Fu *et al.*, 2015; Piao *et al.*, 2017), but such estimates are complicated by our limited understanding of chilling (see Fig. S2; Notes S1). Additionally, the relevant photoperiod a plant experiences during important physiological transitions (e.g. when dormancy ends) may change dramatically with warming (Ettinger *et al.*, 2021).

Researchers hypothesize that these shifts – in forcing, chilling, and photoperiod experienced near the time of an event – produce nonlinearities in plant responses to warming. Though a growing number of observational studies have attempted to document nonlinear shifts in phenology due to interactive cues (Fu *et al.*, 2015; Gauzere *et al.*, 2019), understanding exactly when cues cause nonlinearities may be more complicated than often presented (Fig. 1), and almost always relies on knowledge gained from controlled-environment studies (e.g. Fu *et al.*, 2015; Richardson *et al.*, 2018; Gauzere *et al.*, 2019).

Interactions alone are unlikely to produce nonlinearities with warming

Interactions are well documented in controlled-environment studies, and many may be explained via molecular pathways where chilling and photoperiod modify a developmental rate driven by forcing (Chew *et al.*, 2012). Supporting this, multiple studies now show that the threshold of forcing needed for budburst depends on the sum of chilling over the fall and winter and also the photoperiod experienced in the spring (e.g. Zohner & Renner, 2014; Flynn & Wolkovich, 2018). Higher forcing is generally needed, given the lower chilling (Fig. 2) and shorter photoperiods (Basler & Körner, 2014; Fu *et al.*, 2019). Statistically, this yields a subadditive interaction of forcing × chilling and of forcing × photoperiod, where high values of both cues together produce a more muted response than would be expected from examining either cue alone.

These interactions may produce a smaller response, but they do not necessarily produce nonlinearities (Fig. 1). Instead, most hypotheses for nonlinear responses tacitly require both the interaction of cues and covarying shifts in cues with warming to produce nonlinearities. For example, if some species have a critical photoperiod for budburst and warming leads to forcing cues being met before the critical threshold – effectively shifting both the forcing and photoperiod cues – then we would expect incomplete or highly delayed budburst (Singh *et al.*, 2017; Rinne *et al.*, 2018). Alternatively, the threshold could be crossed in the opposite direction; in this case, pre-climate-change conditions would cause budburst to occur at the extreme values of some cues, but warming could push budburst into values where responses are more linear. This is the mechanism often suggested for declining responses to warming in some temperate trees (Fu *et al.*, 2015; Piao *et al.*, 2017; Gauzere *et al.*, 2019); specifically, that plants previously accumulated sufficient chilling such that forcing was the dominant cue, whereas warming has now reduced chilling such that more forcing is needed for budburst, producing an overall muted effect when estimated as change in days per degree Celsius; see Fu *et al.* (2015) for one example. As this example highlights, however, changes in a single cue are unlikely to occur without additional effects on other cues, complicating how well we can understand them in long-term data without robust understanding of the exact cue requirements from other methods.

Forecasting nonlinear responses: do we have the necessary data?

Controlled-environment studies can help predict nonlinear responses by allowing researchers to examine the effects of one cue with the others held constant and examine interactive effects – given the appropriate study design. Such experiments may be especially useful for forecasting if they contain enough variation in treatments to capture precisely where nonlinearities occur and are designed across a range of levels relevant to current vs future conditions (Shen *et al.*, 2015). Indeed, one of the major advantages of experiments is that they allow treatments outside of the historical range of a species' (or region's) climate – an option that observational data cannot provide.

To understand how valuable currently available data may be to forecasting, we reviewed controlled-environment studies over the last seven decades, quantifying the range of treatments and how

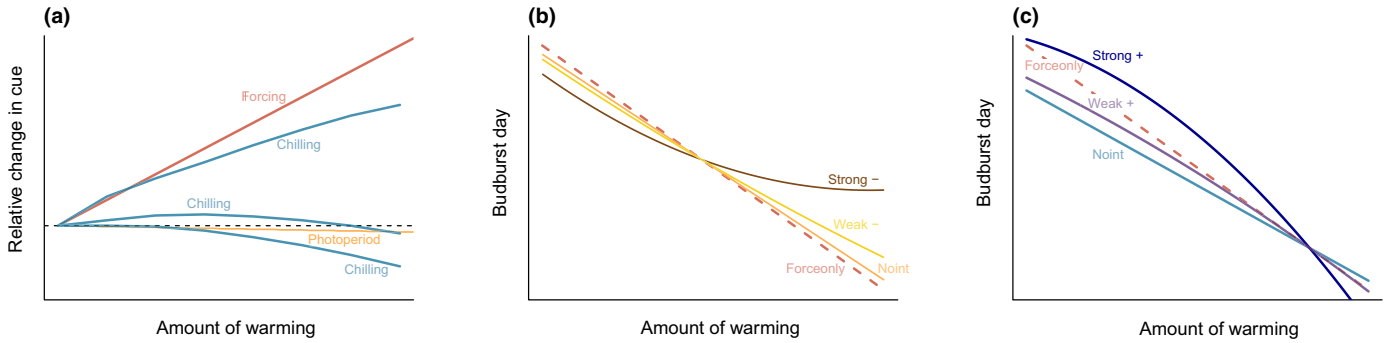

Fig. 1 Nonlinear phenological responses to warming can occur when multiple cues shift in concert with climate change. (a) Much research focuses on how warming increases forcing, but warming may also alter other cues, including photoperiod experienced near the time of the event (yellow line), which is expected to shorten, and chilling (blue lines), which may either increase or decrease. (b, c) Shifts in forcing alongside shifts in a second cue may produce nonlinearities due to the interaction between cues showing the effect of forcing only (dashed line), both cues without an interaction ('noint'), and both cues with an interaction (labeled as 'weak' and 'strong'). The overall change in budburst day predicted with warming depends on the sign of the second cue (negative in (b), positive in (c)), as well as its strength (weak/strong).

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.

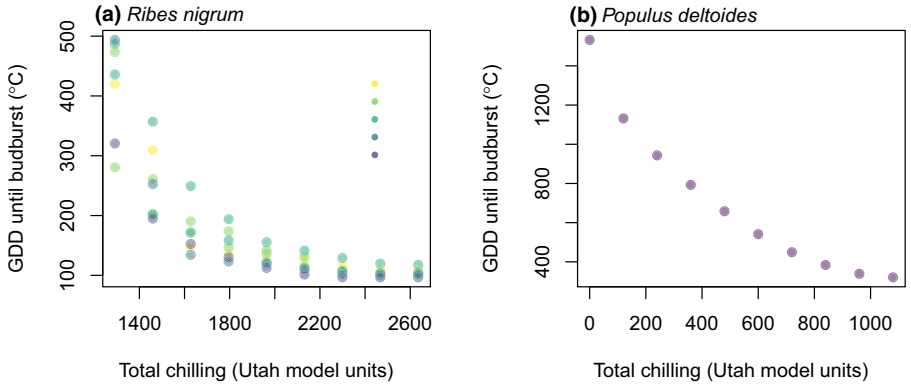

Fig. 2 A common example of how the level of one cue can modify the required level of another cue comes from experiments finding that the amount of chilling affects the amount of forcing needed for budburst. Here, we show this from experiments in (a) black currant (*Ribes nigrum*) from Sønsteby & Heide (2014) and (b) eastern cottonwood (*Populus deltoides*) from Thielges & Beck (1976). Forcing is estimated as growing degree days (GDD, sum of temperatures > 0°C), whereas chilling is estimated using the Utah model (Richardson, 1974).

they compare with current and future conditions. These studies were rarely conducted for climate change research, and most often done for reasons of fundamental or applied science (e.g. horticulture or forestry). Yet, they are some of the best available data for how plants respond to the environment and are thus a critical resource for climate change research today.

How studies and their experimental treatments vary globally

Controlled-environment studies have been conducted with 226 woody species across the globe, with the majority of papers reporting research occurring in Europe (54 of 84 papers; and 93 of 136 studies across papers; a study is a unique experiment within a paper), followed by North America (22 papers and 32 studies; Fig. S1; Table S1). Most studies manipulate one cue, though studies of two or three cues have occurred in almost every decade (Fig. S3). Across study designs, chilling was the most commonly studied cue (69% of 117 studies that manipulated at least one cue; with roughly half of these studies, 47%, using outdoor chilling through multiple field sampling dates), followed by forcing and photoperiod (43% and 40%, respectively).

Studies have covered only a small portion of cue-space, especially for experimentally applied chilling (Fig. 3), with certain levels of cues much more common than others. This suggests we have greater inference at certain cue levels (e.g. 12 h photoperiods and 20°C forcing), but also more limited understanding beyond them, which could limit forecasting. The levels of cues also varied across latitude, with a general trend towards examining more extreme values at higher latitudes (Fig. S4; Notes S2). This trend appears related to latitudinal clines in the environment – higher latitudes experience colder temperatures and longer photoperiods, and they see similar shifts in their controlled-environment study designs. But this trend also introduces a bias in the results, as any comparisons of studies from lower and higher latitudes are also comparing a different range of cues (see Notes S2).

How studies manipulate cues

Robust statistical tests for interactions come from experimental designs that include most or all combinations of all levels of cues studied (factorial design). Of the studies manipulating at least one cue, 37% (43 studies) additionally manipulated another cue in a factorial design, with studies almost evenly split across all possible two-way interactions. Studies examining three cues directly were rare: we identified only three studies examining all three cues at once. Two of these were on *Picea abies* (Worrall & Mergen, 1967; Søgaard *et al.*, 2008), and the other was on *Betula pendula* (Skuterud & Dietrichson, 1994). A slightly larger set of studies (five studies from four papers) examined three cues indirectly – manipulating photoperiod and forcing in controlled environments but equating chilling with sequential removal of tissue from the field – for 11 species (Schnabel & Wample, 1987; Heide, 1993; Partanen *et al.*, 1998; Basler & Körner, 2014).

Do we have enough data to estimate interactions?

The small number of studies examining multiple cues hinders forecasting by limiting our understanding of how, when combined, cues will determine future budburst with continued warming. Because the cues are all known to be interactive, estimates of any one cue are influenced by the level of each other cue. But, unfortunately, current studies often do not report the levels of each cue (see 'NA' in Fig. 3), and many studies use field conditions to apply treatments (see Fig. 3, and 'Paths forward' section).

Of the 136 studies we reviewed, only 72 reported enough information to estimate forcing, chilling, and photoperiod treatments (Ettinger *et al.*, 2020). This, combined with the uneven sampling of cue-space (Fig. 3), makes it impossible to estimate the effect of interactions between cues with current data (Ettinger *et al.*, 2020). Although a few studies estimated interaction terms (e.g. Zohner & Renner, 2014), most studies did not, likely because of the increased effort in study design. In addition to needing enough controlled environments to apply a full-factorial design, the statistical power needed to robustly estimate an interaction is much greater than a simple main effect (such as the effect of forcing or photoperiod alone) – requiring a 16 times greater sample size (Gelman *et al.*, 2020).

How relevant are treatments to current and future conditions?

The utility of controlled-environment studies to forecasting also depends on how relevant the treatments are to current and future conditions. Estimating such relevance is difficult as it depends on a

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.

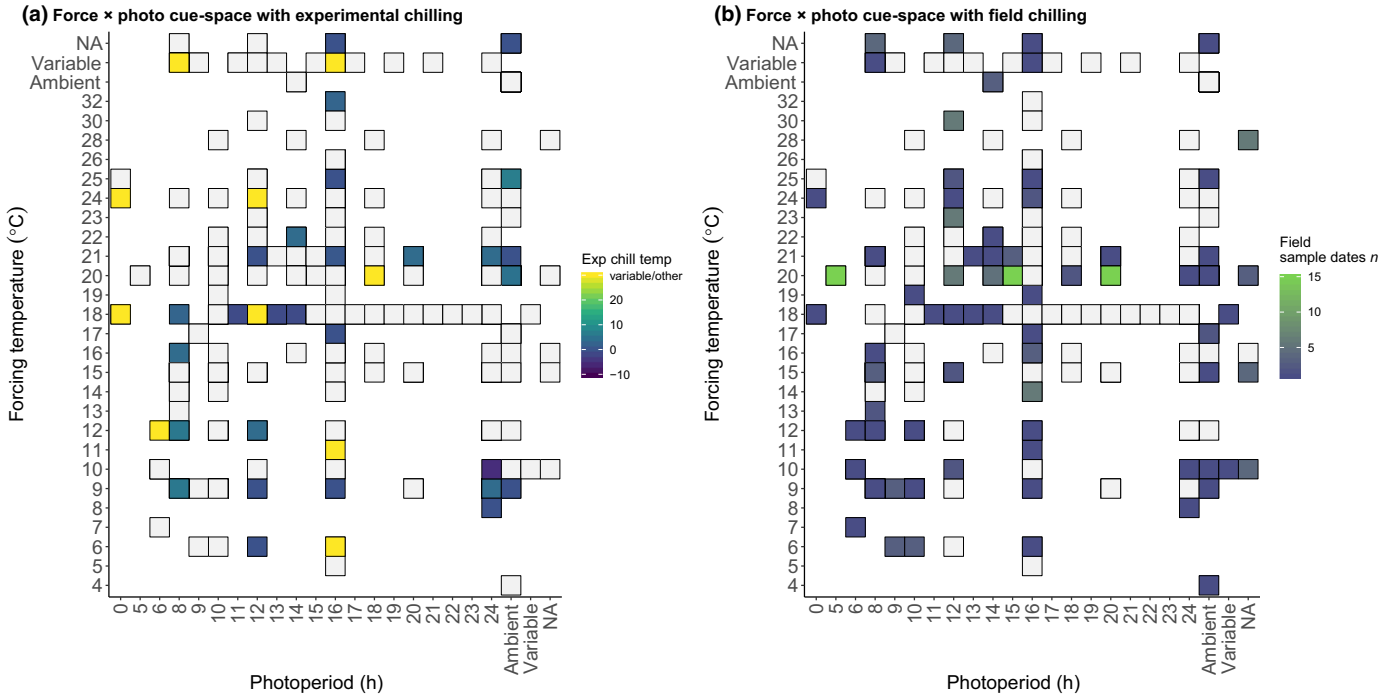

Fig. 3 Studies have examined a range of cue-space (forcing temperature × photoperiod across two methods to manipulate chilling: (a) experimentally or (b) using multiple field sampling dates). However, a focus on certain levels of cues and limited reporting of all cues makes using these data to estimate nonlinearities challenging (see the 'Forecasting nonlinear responses: do we have the necessary data?') section. Gray squares indicate a combination of forcing × photoperiod not present for that method of chilling design (includes studies that did not clearly report chilling treatments), whereas a value of 'NA' indicates that we could not estimate a level of a particular cue because it was not clearly reported. Photoperiod treatments were generally only applied (or reported) during forcing conditions.

species' geographical range and projections considered. However, a simple analysis of two well-studied species, *Fagus sylvatica* and *B. pendula*, suggests experiments have generally bracketed the range of projected temperatures (Fig. 4), though there are a limited number of studies that directly manipulate chilling temperature (Fig. 3). Indeed, we found no studies with multiple chill temperatures for *F. sylvatica*, even though it is one of the most well-studied species (Fig. 4c).

Experimental treatments were generally larger than expected shifts due to climate change. This is not surprising given that certain forcing and photoperiod treatments were extremely common across studies (see earlier herein). It also makes sense from an experimental–statistical perspective: if the goal of an experiment is to identify whether a cue is present then larger treatment differences should yield larger effect sizes and higher statistical power. Such large shifts, however, may be risky to extrapolate from when predicting effects of warming.

Paths forward

We argue that controlled-environment experiments will be critical for accurate predictions of spring phenology given future warming. How accurate such predictions are will depend on the design of future experiments, breakthroughs in our physiological understanding of the major cues, and how well these two areas can be integrated with long-term data to improve models.

Improving controlled-environment studies

We expect the most useful future experiments for forecasting will be designed to identify threshold effects, optimal temperatures/photoperiods, and nonlinearities from interactive cues (Caffarra *et al.*, 2011). Identifying threshold effects and optimal temperatures or photoperiods generally translates into response function designs, which require many different levels of a single cue – making such experiments difficult to cross with other cues. Yet, where a threshold or optimum occurs is likely dependent on the level of other cues (Stearns & Olson, 1958; Flynn & Wolkovich, 2018), making studies of interactive cues critical for robust forecasting. Ideally, studies testing for nonlinearities due to interactive cues would be fully crossed (i.e. every combination of levels is present in treatments), resulting in a robust response-surface design. Such designs, however, can quickly require a large number of controlled environments. Researchers may be able to reduce such numbers by a modified response-surface design that aims to characterize the response function of one focal cue across several relevant levels of the other cue.

Controlled-environment studies may also be more readily applied to forecasting by exploring more realistic conditions. Though identifying thresholds, optima, and nonlinearities may involve considering informative extremes in levels of cues, most changes in cues due to climate change are and will be on a (relatively) smaller scale than the experiments currently study (Fig. 4). Whereas species distribution models generally assume

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.

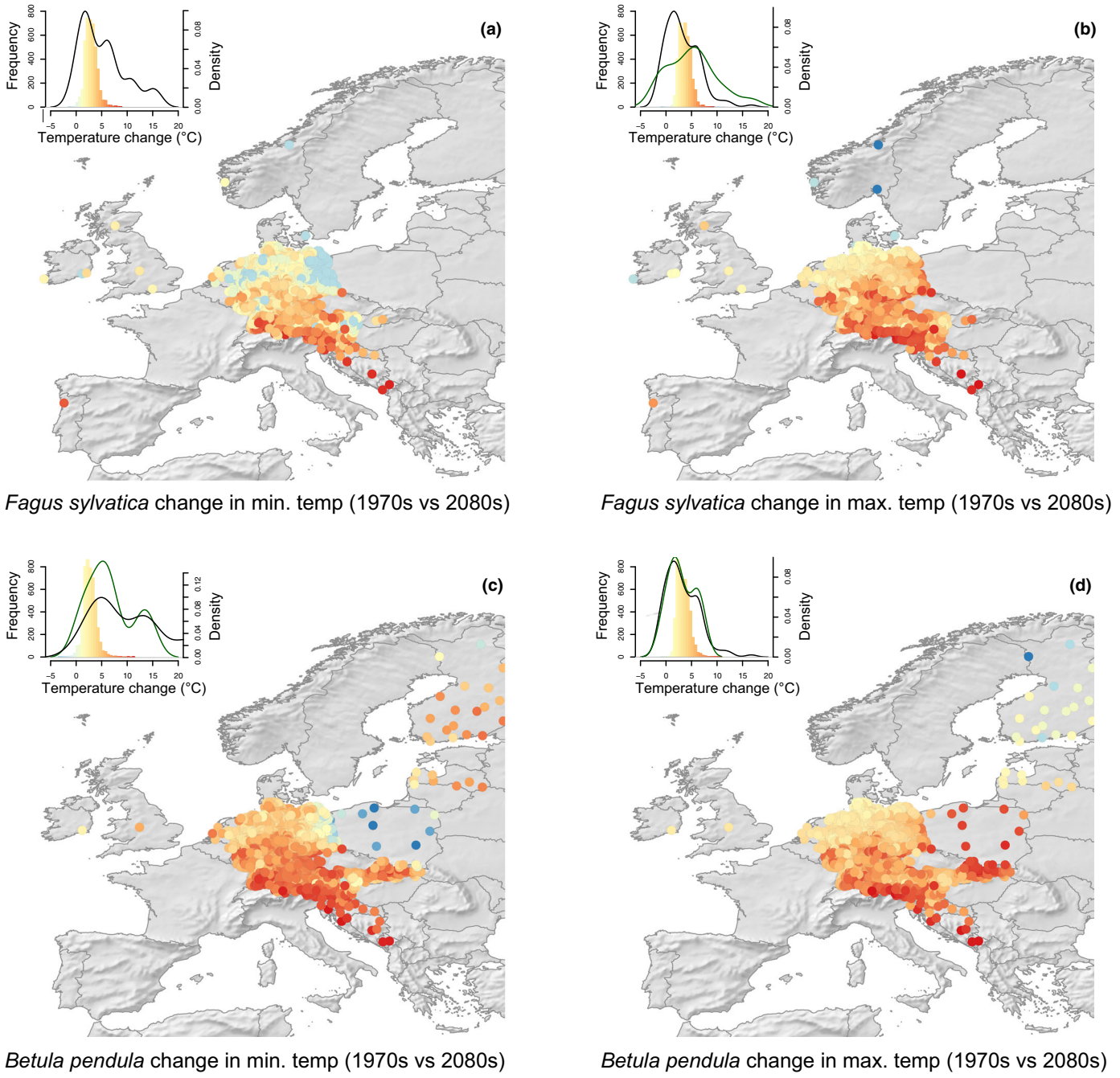

Fig. 4 Findings from controlled-environment experiments are relevant for forecasting applications based on comparison of two well-studied species: (a, b) *Fagus sylvatica* and (c, d) *Betula pendula*. Points represent a site with spring phenology data (from the PEP725 database; Templ *et al*., 2018) and show the predicted change (2080s average minus 1970s average) in (a, c) minimum and (b, d) maximum temperatures 60 d before leafout. We found that treatment differences (see lines in inset plots, where green lines show the treatments for that exact species and black lines show across all species in the OSPREE database) generally bracketed predicted changes (histograms in inset plots show the same predicted changes in temperature represented also by the mapped points), but were on average much larger in magnitude. We show (a, c) chilling and (b, d) forcing differences in treatments across studies (for *F. sylvatica* there are no chilling treatments of differing temperatures, so green line is absent in inset of (a)). For more details, see Supporting Information Methods S2.

species will remain in the same climatic conditions (Elith & Leathwick, 2009), most evidence suggests species will lag in their spatial responses, meaning shifts in cues in the current range may be important to the fate of trailing-edge populations (Bertrand *et al*., 2011; Lenoir & Svenning, 2015). Thus, researchers should design experiments that include values of cues within the current and projected future species' range, rather than focusing on extreme comparisons only (e.g. 12 vs 24 h photoperiods). In particular, our findings suggest forcing treatments (e.g. 20 and 25°C) are often much higher than the average temperatures that study species experience near budburst (Fu *et al*., 2015; Güsewell *et al*., 2017), and studies rarely consider multiple chilling temperatures.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.

Incorporating our understanding of physiology into forecasts of phenology

Even with all the improvements suggested herein, controlled-environment studies will still be fundamentally limited in their utility for prediction without a deeper understanding of how major phenological cues act physiologically (Bahuguna & Jagadish, 2015). This problem is most acute in our understanding of dormancy, which divides chilling and forcing (Chang *et al.*, 2021). Researchers may use the terms 'chilling' and 'forcing' for their treatments, but they rarely have physiological evidence that these are the actual conditions plants experience (Chuine *et al.*, 2016).

Chilling is generally defined as accumulated cool temperatures in the first phase of dormancy (endodormancy), after which plants enter ecodormancy, when accumulated forcing leads to budburst (Chuine *et al.*, 2016). Measuring endodormancy and its transition into ecodormancy is notoriously difficult (e.g. Junttila & Hänninen, 2012). In practice, most phenology studies use the terms 'chilling' and 'forcing' to mean 'cool temperatures' (either in the fall and winter or applied in experimental conditions) and 'warm temperatures' (either in the spring or applied after sufficient chilling) and generally hope they correspond to endo- and ecodormancy – without any evidence of this hoped-for correspondence (Ettinger *et al.*, 2020). Some studies use the sequential transfer of cuttings to warm conditions to estimate the transition from endo- to ecodormancy, with rapid and full budburst (e.g. > 90% of buds on a cutting) generally meaning a plant is ecodormant (e.g. Junttila & Hänninen, 2012), but, given that this is labor and space intensive, few studies include this method.

Physiologists have long recognized this issue, and recent breakthroughs provide new insights into what causes dormancy at the cellular level (van der Schoot *et al.*, 2014). Our understanding of the molecular pathways for dormancy and dormancy-break, however, is very poor (especially when compared with our understanding of flowering-time pathways; Azeez *et al.*, 2021), with most recent breakthroughs focused far downstream of the signal and receptor pathways. Thus far, research has been on a very limited number of species (Singh *et al.*, 2017; Rinne *et al.*, 2018), making extrapolation to other species difficult.

Our inchoate understanding of endo- and ecodormancy pathways is likely the greatest current impediment to improved models of spring phenology. Without markers or other methods to clearly identify the start or end of chilling – or whether there even is a clear start or end (as opposed to potential parallel models of chilling and forcing; Harrington & Gould, 2015) – experiments will continue to operate with a set of latent chilling and forcing treatments (i.e. the underlying values of chilling and forcing that the plants actually experience) some distance from the 'chilling' and 'forcing' treatments researchers discuss in their papers. In such a murky experimental world, we may be unlikely to see rapid advances from new experiments.

An improved understanding of endodormancy release, however, could revolutionize models of chilling and, in turn, alter estimates of forcing. Future estimates could more accurately define chilling and forcing using only temperatures during endodormancy and ecodormancy (respectively), assuming tractable tests of endo- and ecodormancy and the uptake of such tests in controlled-environment studies. With these experiments in hand, researchers could quickly build improved models of chilling and forcing and, for the first time, provide accurate estimates of how chilling has and will shift with climate change.

Improving integration of controlled environment and physiological studies with long-term data

Predicting phenological responses today is done at different scales by different groups of researchers to answer what appear to be vastly different questions. Yet the goal of forecasting phenology in the new environmental regimes of climate change have united these fields in many ways. Researchers at the macroscale, who traditionally use coarse metrics such as green-up and seasonal or annual climatic metrics, have turned increasingly to controlled-environment experiments and process-based models to understand whether multiple cues underlie trends over time (e.g. Fu *et al.*, 2019). At the same time, researchers building complex molecular pathways that use fine-scale data have started to scale up to make predictions of continent-scale common gardens (Wilczek *et al.*, 2009).

These examples provide hope to build from developmental models to predictions in natural conditions, but success will require far greater integration and more openness of the complexities (and failures) of different methods. Most studies of long-term observational phenology data still operate independently from advances in our physiological understanding and from controlled-environment studies. Similarly, controlled-environment studies generally do not use long-term data to help interpret results or define treatments. For major advances, more studies must attempt to use diverse methods and datasets together. Such attempts often highlight problems – large-scale trends in green-up that could be explained by multiple developmental pathways, or process-based models that work at local, controlled scales but fail spectacularly in larger or more natural settings – but without seeing the problems, it will be almost impossible to solve them.

Studies that have integrated results from both controlled-environment experiments and with long-term observations provide a path forward (Caffarra *et al.*, 2011; Satake *et al.*, 2013; Chuine & Regniere, 2017). Such research underlies progress towards model development that relies continuously on a back-and-forth process between developing models based on both long-term data and experiments, then testing predictions with new experiments and newly available observational data (Satake *et al.*, 2013). This research has been done only for a handful of crop and model organism species, however, and not extended to natural systems. Our synthesis suggests that *B. pendula* and *F. sylvatica* are good candidates for expanding to wild forest species across larger landscapes and represent potentially interesting extremes in leafout timing (early vs late) and photoperiod control of budbreak (weak vs strong).

Building population- and species-rich predictions

Given the efforts and data involved in models for a single population or ecotype, building up to multipopulation and

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.

multispecies predictions may appear daunting, but such models are crucial for accurate forecasts that can apply to diverse regions and large-scale vegetation models. Our current understanding of pathways for phenological events suggests strong population differences (Wilczek *et al.*, 2009; Tanino *et al.*, 2010): whereas population-level studies of spring budbreak suggest less local adaptation than budset (Aitken *et al.*, 2008), chilling is often strongly differentiated by population (Junttila & Hänninen, 2012). Thus, models that assume constant cues across a species range could make inaccurate predictions.

At the species level, molecular studies show temperature pathways diverge often across species (by contrast, photoperiodic control appears highly conserved; Satake *et al.*, 2022), suggesting a critical need for more molecular data across species. Long-term data are generally more species rich than controlled-environment studies (see Methods S1), suggesting more diverse controlled-environment studies may be the current major data limitation. Beyond data, however, new modeling approaches can help integrate current and future data more powerfully.

Bayesian hierarchical models are specifically designed to synthesize across multiple sources of data, attributing variation to the species studied, the levels of each cue, and remaining unmeasured variation in studies (i.e. unreported laboratory and chamber differences may be captured by including a parameter to estimate a 'study' effect). Such models are extremely powerful for building species-rich predictions, but (like all models) are more robust with more data. With greater efforts to publish data and improved treatment reporting (i.e. all cue conditions defined, even when not manipulated; Fig. 3) all studies – whether or not designed to improve models or forecasting – can be included in such models.

Right now: it is your tomorrow

Research on phenology had been conducted for centuries before anthropogenic climate change caused earlier budburst and leafout across much of the globe (Lamb, 1948; Sparks & Carey, 1995). Decades of controlled-environment studies contributed to our fundamental understanding of the drivers of spring plant phenology. Today, climate change requires leveraging these decades and centuries of research for more accurate predictions that can help humans adapt to warming.

We have outlined how researchers could better harness the power of controlled-environment experiments to transform our fundamental understanding of phenology and advance forecasting. Controlled-environment studies can critically rule out, or support, hypotheses to explain observed discrepancies in long-term data and open up new ways to use long-term data to understand current trends, helping the field move beyond trying to tease out cues using only long-term data where cues are inherently correlated. Though understanding, modeling, and predicting interactions among cues and their effects on phenology is challenging, it will yield more accurate predictions – with valuable implications to more realistically assess the effects of climate change on plant biodiversity.

Acknowledgements

Comments from A. Austin, R. Guy, L. Rieseberg, and three anonymous reviewers improved this paper. NSERC Discovery Award (RGPIN-05038 to EMW), Canada Research Chair in Temporal Ecology (EMW), and Canada Foundation for Innovation (38386 to EMW) provided funding. For IM-C, funding was additionally provided by the Spanish Ministry for Science and Innovation (grant no. PID2019-109711RJ-I00) and Comunidad de Madrid and University of Alcalá (funders of Iberian Future Wine through grant no. CM/2021-003).

Author contributions

EMW, CJC, DMB, AKE and IM-C prepared the database, conceived of the manuscript, edited it, and contributed to figures and analyses. EMW wrote the manuscript.

ORCID

D. M. Buonaiuto https://orcid.org/0000-0003-4022-2591
C. J. Chamberlain https://orcid.org/0000-0001-5495-3219
A. K. Ettinger https://orcid.org/0000-0002-6228-6732
I. Morales-Castilla https://orcid.org/0000-0002-8570-9312
E. M. Wolkovich https://orcid.org/0000-0001-7653-893X

Data availability

Data from a systematic literature review included in the paper are freely available at The Knowledge Network for Biocomplexity (doi: 10.5063/F1DF6PNQ at knb.ecoinformatics.org).

References

Aitken SN, Yeaman S, Holliday JA, Wang TL, Curtis-McLane S. 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* **1**: 95–111.

Azeez A, Zhao YC, Singh RK, Yordanov YS, Dash M, Miskolczi P, Stojkovic K, Strauss SH, Bhalerao RP, Busov VB. 2021. EARLY BUD-BREAK 1 and EARLY BUD-BREAK 3 control resumption of poplar growth after winter dormancy. *Nature Communications* **12**: 1123.

Bahuguna RN, Jagadish KS. 2015. Temperature regulation of plant phenological development. *Environmental and Experimental Botany* **111**: 83–90.

Basler D, Körner C. 2014. Photoperiod and temperature responses of bud swelling and bud burst in four temperate forest tree species. *Tree Physiology* **34**: 377–388.

Bertrand R, Lenoir J, Piedallu C, Riofrío-Dillon G, de Ruffray P, Vidal C, Pierrat JC, Gégout JC. 2011. Changes in plant community composition lag behind climate warming in lowland forests. *Nature* **479**: 517–520.

Caffarra A, Donnelly A, Chuine I. 2011. Modelling the timing of *Betula pubescens* budburst. II. Integrating complex effects of photoperiod into process-based models. *Climate Research* **46**: 159–170.

Chang CYY, Brautigam K, Huner NPA, Ensminger I. 2021. Champions of winter survival: cold acclimation and molecular regulation of cold hardiness in evergreen conifers. *New Phytologist* **229**: 675–691.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.

Chen L, Huang JG, Ma QQ, Hanninen H, Tremblay F, Bergeron Y. 2019. Long-term changes in the impacts of global warming on leaf phenology of four temperate tree species. *Global Change Biology* **25**: 997–1004.

Chew YH, Wilczek AM, Williams M, Welch SM, Schmitt J, Halliday KJ. 2012. An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytologist* **194**: 654–665.

Chuine I, Bonhomme M, Legave JM, García de Cortázar-Atauri I, Charrier G, Lacointe A, Améglio T. 2016. Can phenological models predict tree phenology accurately in the future? The unrevealed hurdle of endodormancy break. *Global Change Biology* **22**: 3444–3460.

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

Cleland EE, Chuine I, Menzel A, Mooney HA, Schwartz MD. 2007. Shifting plant phenology in response to global change. *Trends in Ecology & Evolution* **22**: 357–365.

Delpierre N, Dufrene E, Soudani K, Ulrich E, Cecchini S, Boe J, Francois C. 2009. Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France. *Agricultural and Forest Meteorology* **149**: 938–948.

Elith J, Leathwick JR. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* **40**: 677–697.

Ellwood ER, Temple SA, Primack RB, Bradley NL, Davis CC. 2013. Record-breaking early flowering in the Eastern United States. *PLoS ONE* **8**: e53788.

Elmendorf SC, Ettinger AK. 2020. Is photoperiod a dominant driver of secondary growth resumption? *Proceedings of the National Academy of Sciences, USA* **117**: 32861–32864.

Ettinger A, Chamberlain C, Morales-Castilla I, Buonaiuto D, Flynn D, Savas T, Samaha J, Wolkovich E. 2020. Winter temperatures predominate in spring phenological responses to warming. *Nature Climate Change* **10**: 1137–1142.

Ettinger AK, Buonaiuto DM, Chamberlain CJ, Morales-Castilla I, Wolkovich EM. 2021. Spatial and temporal shifts in photoperiod with climate change. *New Phytologist* **230**: 462–474.

Flynn DFB, Wolkovich EM. 2018. Temperature and photoperiod drive spring phenology across all species in a temperate forest community. *New Phytologist* **219**: 1353–1362.

Fu YH, Piao S, Zhou X, Geng X, Hao F, Vitasse Y, Janssens IA. 2019. Short photoperiod reduces the temperature sensitivity of leaf-out in saplings of *Fagus sylvatica* but not in horse chestnut. *Global Change Biology* **25**: 1696–1703.

Fu YSH, Zhao HF, Piao SL, Peaucelle M, Peng SS, Zhou GY, Ciais P, Huang MT, Menzel A, Uelas JP *et al.* 2015. Declining global warming effects on the phenology of spring leaf unfolding. *Nature* **526**: 104–107.

Gauzere J, Lucas C, Ronce O, Davi H, Chuine I. 2019. Sensitivity analysis of tree phenology models reveals increasing sensitivity of their predictions to winter chilling temperature and photoperiod with warming climate. *Ecological Modelling* **441**: 108805.

Gelman A, Hill J, Vehtari A. 2020. *Regression and other stories*. New York, NY, USA: Cambridge University Press.

Güsewell S, Furrer R, Gehrig R, Pietragalla B. 2017. Changes in temperature sensitivity of spring phenology with recent climate warming in Switzerland are related to shifts of the preseason. *Global Change Biology* **23**: 5189–5202.

Harrington CA, Gould PJ. 2015. Tradeoffs between chilling and forcing in satisfying dormancy requirements for Pacific Northwest tree species. *Frontiers in Plant Science* **6**: 120.

Heide O. 1993. Daylength and thermal time responses of budburst during dormancy release in some northern deciduous trees. *Physiologia Plantarum* **88**: 531–540.

Hereford J, Schmitt J, Ackerly DD. 2017. The seasonal climate niche predicts phenology and distribution of an ephemeral annual plant, *Mollugo verticillata*. *Journal of Ecology* **105**: 1323–1334.

Jochner S, Caffarra A, Menzel A. 2013. Can spatial data substitute temporal data in phenological modelling? A survey using birch flowering. *Tree Physiology* **33**: 1256–1268.

Junttila O, Hänninen H. 2012. The minimum temperature for budburst in Betula depends on the state of dormancy. *Tree Physiology* **32**: 337–345.

Lamb RC. 1948. Effect of temperatures above and below freezing on the breaking of rest in the Latham raspberry. *Proceedings of the American Society for Horticultural Science* **51**: 313–315.

Lenoir J, Svenning JC. 2015. Climate-related range shifts–a global multidimensional synthesis and new research directions. *Ecography* **38**: 15–28.

Luedeling E, Guo L, Dai J, Leslie C, Blanke MM. 2013. Differential responses of trees to temperature variation during the chilling and forcing phases. *Agricultural and Forest Meteorology* **181**: 33–42.

Maurya JP, Bhalerao RP. 2017. Photoperiod- and temperature-mediated control of growth cessation and dormancy in trees: a molecular perspective. *Annals of Botany* **120**: 351–360.

Menzel A, Sparks TH, Estrella N, Koch E, Aasa A, Ahas R, Alm-Kuebler K *et al.* 2006. European phenological response to climate change matches the warming pattern. *Global Change Biology* **12**: 1969–1976.

Partanen J, Koski V, Hänninen H. 1998. Effects of photoperiod and temperature on the timing of bud burst in Norway spruce (*Picea abies*). *Tree Physiology* **18**: 811–816.

Piao S, Liu Z, Wang T, Peng S, Ciais P, Huang M, Ahlstrom A, Burkhart JF, Chevallier F, Janssens IA *et al.* 2017. Weakening temperature control on the interannual variations of spring carbon uptake across northern lands. *Nature Climate Change* **7**: 359–363.

Polgar CA, Primack RB, Dukes JS, Schaaf C, Wang Z, Hoeppner SS. 2014. Tree leaf out response to temperature: comparing field observations, remote sensing, and a warming experiment. *International Journal of Biometeorology* **58**: 1251–1257.

Richardson AD, Hufkens K, Milliman T, Aubrecht DM, Furze ME, Seyednasrollah B, Krassovski MB, Latimer JM, Nettles WR, Heiderman RR *et al.* 2018. Ecosystem warming extends vegetation activity but heightens vulnerability to cold temperatures. *Nature* **560**: 368–371.

Richardson E. 1974. A model for estimating the completion of rest for 'Redhaven' and 'Elberta' peach trees. *HortScience* **9**: 331–332.

Rinne PLH, Paul LK, van der Schoot C. 2018. Decoupling photo- and thermoperiod by projected climate change perturbs bud development, dormancy establishment and vernalization in the model tree Populus. *BMC Plant Biology* **18**: 220.

Satake A, Kawagoe T, Saburi Y, Chiba Y, Sakurai G, Kudoh H. 2013. Forecasting flowering phenology under climate warming by modelling the regulatory dynamics of flowering-time genes. *Nature Communications* **4**: 2303.

Satake A, Nagahama A, Sasaki E. 2022. A cross-scale approach to unravel the molecular basis of plant phenology in temperate and tropical climates. *New Phytologist* **233**: 2340–2353.

Schnabel BJ, Wample RL. 1987. Dormancy and cold hardiness in *Vitis vinifera* L. cv. White Riesling as influenced by photoperiod and temperature. *American Journal of Enology and Viticulture* **38**: 265–272.

Schwartz MD. 1997. Spring index models: an approach to connecting satellite and surface phenology. In: Lieth H, Schwartz MD, eds. *Phenology in seasonal climates I*. Leiden, the Netherlands: Backhuys, 22–38.

Shen M, Piao S, Dorji T, Liu Q, Cong N, Chen X, An S, Wang S, Wang T, Zhang G. 2015. Plant phenological responses to climate change on the Tibetan Plateau: research status and challenges. *National Science Review* **2**: 454–467.

Singh RK, Svystun T, AlDahmash B, Jönsson AM, Bhalerao RP. 2017. Photoperiod-and temperature-mediated control of phenology in trees–a molecular perspective. *New Phytologist* **213**: 511–524.

Skuterud R, Dietrichson J. 1994. Budburst in detached birch shoots (*Betula pendula*) of different varieties winter-stored in darkness at three different temperatures. *Silva Fennica* **28**: 223–224.

Søgaard G, Johnsen Ø, Nilsen J, Junttila O. 2008. Climatic control of bud burst in young seedlings of nine provenances of Norway spruce. *Tree Physiology* **28**: 311–320.

Sønsteby A, Heide OM. 2014. Chilling requirements of contrasting black currant (*Ribes nigrum* L.) cultivars and the induction of secondary bud dormancy. *Scientia Horticulturae* **179**: 256–265.

Sparks TH, Carey PD. 1995. The responses of species to climate over two centuries – an analysis of the Marsham phenological record, 1736–1947. *Journal of Ecology* **83**: 321–329.

Stearns F, Olson J. 1958. Interactions of photoperiod and temperature affecting seed germination in *Tsuga canadensis*. *American Journal of Botany* **45**: 53–58.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.

Tanino KK, Kalcsits L, Silim S, Kendall E, Gray GR. 2010. Temperature-driven plasticity in growth cessation and dormancy development in deciduous woody plants: a working hypothesis suggesting how molecular and cellular function is affected by temperature during dormancy induction. *Plant Molecular Biology* **73**: 49–65.

Tansey CJ, Hadfield JD, Phillimore AB. 2017. Estimating the ability of plants to plastically track temperature-mediated shifts in the spring phenological optimum. *Global Change Biology* **23**: 3321–3334.

Templ B, Koch E, Bolmgren K, Ungersböck M, Paul A, Scheifinger H, Rutishauser T, Busto M, Chmielewski FM, Hájková L *et al.* 2018. Pan European phenological database (PEP725): a single point of access for European data. *International Journal of Biometeorology* **62**: 1109–1113.

Thielges B, Beck R. 1976. Control of bud break and its inheritance in *Populus deltoides. Tree Physiology and Yield Improvement* **14**: 253–259.

van der Schoot C, Paul LK, Rinne PLH. 2014. The embryonic shoot: a lifeline through winter. *Journal of Experimental Botany* **65**: 1699–1712.

Visser ME, Holleman LJ. 2001. Warmer springs disrupt the synchrony of oak and winter moth phenology. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**: 289–294.

Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin LJ, Muir CD, Sim S, Walker A, Anderson J *et al.* 2009. Effects of genetic perturbation on seasonal life history plasticity. *Science* **323**: 930–934.

Wolkovich EM, Cook BI, Allen JM, Crimmins TM, Betancourt JL, Travers SE, Pau S, Regetz J, Davies TJ, Kraft NJB *et al.* 2012. Warming experiments underpredict plant phenological responses to climate change. *Nature* **485**: 494–497.

Worrall J, Mergen F. 1967. Environmental and genetic control of dormancy in *Picea abies. Physiologia Plantarum* **20**: 733–745.

Yu HY, Luedeling E, Xu JC. 2010. Winter and spring warming result in delayed spring phenology on the Tibetan plateau. *Proceedings of the National Academy of Sciences, USA* **107**: 22151–22156.

Zohner CM, Benito BM, Svenning JC, Renner SS. 2016. Day length unlikely to constrain climate-driven shifts in leaf-out times of northern woody plants. *Nature Climate Change* **6**: 1120–1123.

Zohner CM, Renner SS. 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.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Map of locations of studies reviewed.

Fig. S2 Utah chill model and warming temperatures.

Fig. S3 Number of cues manipulated in studies over time.

Fig. S4 How experimental treatments correlate with latitude.

Methods S1 Overview of OSPREE database.

Methods S2 Comparing experimental treatments with forecasted trends.

Notes S1 The chilling enigma of overwinter temperatures.

Notes S2 Trends in experimental treatments over cue levels and space.

Table S1 Dataset names and references for papers in the OSPREE database.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

© 2022 The Authors
New Phytologist © 2022 New Phytologist Foundation.