Limiting cues: How spring warming, winter chilling and daylength shape climate change responses

THE LAB IN 2017[1–5]

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

[2] *Arnold Arboretum of Harvard University, 1300 Centre Street, Boston, Massachusetts, 02131, USA*

[3] *Organismic & Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts, 02138, USA*

[4] *The Nature Conservancy, 74 Wall Street, Seattle, Washington USA*

[5] *Global Change Ecology and Evolution Group, Department of Life Sciences, University of Alcalà, Alcalà de Henares 28805, Spain*

Corresponding author: Lizzie, see [3] above ; E-mail: e.wolkovich@ubc.ca

Abstract

Climate change has shifted plant phenology globally, with average shifts of 4-6 days per °C, and some species shifting much more. Such shifts have been some of the most reported and predictable biological impacts of rising temperatures. Yet as climate change has marched on, research has reported possible failures to predict observed changes; increasingly, phenological shifts appear more muted over recent decades, or in certain locations. The main hypotheses for these changing trends is that other cues of spring phenology—long-documented in lab 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 requires a greater integration of long-term data with results of lab experiments conducted in controlled environments (e.g., growth chambers). We highlight how seven decades of research in controlled environments can improve predictions for when, where and how the interactive effects of other cues will impact simple linear predictions. Finally, we discuss how a new generation of controlled environment experiments could rapidly improve our predictive capacity for woody plant phenology in coming decades.

Keywords: phenology, climate change, spring warming, chilling, forcing, daylength, photoperiod, non-linear responses, leafout, budburst

1 Main text

Shifts in spring plant phenology are one of the most reported and most predictable changes with climate change. Decades of research have documented advancing budburst, leafout and flowering across systems (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.*, 2003, 2006). Recently, however, these advances have appeared to slow (Fu *et al.*, 2015) or even reverse in some places (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 of spring phenology research on spring temperatures, increasing evidence suggests a more complicated underlying physiology for most plant species (e.g., Zohner *et al.*, 2016; Gauzere *et al.*, 2019; Ettinger *et al.*, 2020). For many species three major cues underlie spring phenology: forcing (warm temperatures, generally occurring in the late winter and early spring), chilling (cool temperatures, generally occurring in the fall and winter), and photoperiod (daylength).

Together, forcing, chilling and photoperiod may produce non-linear responses that many current methods do not predict, but observational studies appear to have seen recently (Fu *et al.*, 2015). Predicting these non-linearities is a common goal in plant phenology research today (Güsewell *et al.*, 2017; Martinez-Luscher *et al.*, 2017; Gauzere *et al.*, 2019; Chen *et al.*, 2019; Keenan *et al.*, 2020), but has been slowed by data gaps and the underlying complexity of spring phenology.

The first step towards improved phenological predictions is robust measurements of chilling, forcing and photoperiod cues. Recently, efforts have focused on estimating these cues from long-term observational data. Yet observational studies may struggle to robustly 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 chilling and photoperiod cues both yield similar predictions (i.e., both cues predict later spring phenology with warming). Approaches that attempt to break some of these correlations, such as leveraging elevational or latitudinal gradients, may run afoul of other correlations; for example, gradients in local adaptation can create shifts in the underlying cues along environmental gradients in complex ways (Tansey *et al.*, 2017). Second, most observational studies focus on linear models of each cue, often without interactions between cues.

Using simple linear models with observational data may make sense as phenological responses to most cues are expected to be linear except at the extremes. Natural conditions often see only a small slice of the range of values of each cue that are possible—and those values often appear to be in the middle—potentially linear—range (Gauzere *et al.*, 2017; Ettinger *et al.*, 2020). Further, given that interactions between cues are difficult to estimate—moreso when those cues are highly correlated in nature—focusing on the main effects of cues may provide more robust

estimates from observational data (main effects integrate over any interactions and require lower statistical power to robustly estimate compared to interactions, Gelman & Hill, 2006).

In contrast to the limitations of observational studies, one method is designed to measure the complexity of cues: controlled environment (e.g., growth chamber) studies (Nagano *et al.*, 2012; Satake *et al.*, 2013). Such studies have been conducted for over 70 years and are specifically designed to understand non-linearities both in individual cues and produced by interactions between cues. In contrast to observational studies, controlled environment studies can manipulate all three cues, extend to other cues that may be important in some species or biomes (e.g., humidity, drought conditions, light spectra), and can tease out interactions between cues by experimentally decoupling them.

Despite the prevalence of controlled environment studies on spring phenological cues, they have rarely been reviewed. Perhaps more surprisingly, they are often poorly integrated into the current phenological literature on climate change. This includes in debates where they are critical, such as about the importance of photoperiod (e.g., Fu *et al.*, 2015; Richardson *et al.*, 2018).

Here we review how the long-term literature on controlled environment studies can be integrated into current phenological research on climate change more fully. We begin by outlining how the three major phenological cues—forcing, chilling and photoperiod—can produce non-linearities and how they will shift in coming decades with anthropogenic climate change. We then review controlled environment studies to understand how much of the cue-space (possible natural range of each cue and interactions across cues) has been studied and how experimental treatments compare to shifts in cues caused by climate change. Based on this, we discuss how controlled environment studies can be best designed to build more robust predictions.

Given our aim to improve understanding of current trends and forecasts we focus on early vegetative phases (budburst and leafout) of mainly wild species, which are critical to plant growth and thus to models of carbon uptake and storage, and which have shifted most with climate change (Cleland *et al.*, 2007; IPCC, 2014). We touch on other areas of research, which have been important to our understanding of the cues underlying phenology. In particular, research has been especially strong in model systems (e.g., *Arabidopsis, Populus*) and crops (Cesaraccio *et al.*, 2004)—with the exact phenophase of interest varying (potentially by a species' life history: more focus on germination and flowering in *Arabidopsis*, and more on leafout and budset in *Populus*). Given our focus on budburst and leafout, our review concentrates on woody species phenology, where most research has been conducted. Most of our conclusions and suggested approaches, however, could be adapted to non-woody species and other phenophases with similar underlying interactive cues.

1.1   How do phenological cues produce non-linear responses?

Forcing, chilling and photoperiod cues together determine budburst each year for woody species in many temperate species (e.g., Chuine & Regniere, 2017; Ettinger *et al.*, 2020). Forcing and

chilling cues are generally understood to be accumulation processes, where plants must integrate chilling and forcing experienced over time to meet a threshold value at which they can budburst, leafout or flower (Chuine, 2000). In contrast, photoperiod is generally not considered an integral cue, but one evaluated daily (Singh *et al.*, 2017). In practice, these two types of processes are often abstracted into the effects of average experienced values (mean temperature or daylength), either over some window in long-term data (e.g., Wolkovich *et al.*, 2012a; Fu *et al.*, 2015), or through controlled environments that hold temperatures and light regimes more constant (e.g., Worrall & Mergen, 1967; Heide, 1993a,b; Skuterud & Dietrichson, 1994).

Controlled environment studies show two major ways these cues can produce non-linear responses—each cue alone, or through interactions between cues (Fig. reffig:intxncues). Each cue alone may produce a non-linear phenological response when examined across a sufficiently wide range of values (Fig. 1). Cues may be linear in the mid-range of values, while extremely high or low values of some cues may produce alternative response (Gauzere *et al.*, 2017). For example, at very low photoperiod (short days) plants often will budburst erratically (Heide, 1993a; Partanen *et al.*, 1998; Singh *et al.*, 2017; Rinne *et al.*, 2018), while at sufficiently long photoperiods maximum growth may occur, meaning photoperiods longer than some threshold will have no additional effect on budburst timing (e.g., Major, 1980). Similarly, extremely cool or very hot temperatures may limit forcing as plant developmental processes slow (Parent & Tardieu, 2012). Such extreme values of cues, however, are likely less common than interactions between cues that can produce non-linearities.

More commonly research has focused on how the interaction of cues may produce non-linearities. For example, multiple studies now show that the threshold of forcing needed for budburst depends on the sum of chilling over the fall and winter and by the photoperiod while forcing is accumulating in the spring (e.g., Zohner & Renner, 2014; Flynn & Wolkovich, 2018). Higher forcing is generally needed given lower chilling (Fig. 2) and short photoperiods (Basler & Körner, 2014; Fu *et al.*, 2019)—producing generally a subadditive effect of forcing x chilling and forcing x photoperiod (i.e., both cues together produce a more muted response than the addition of each cue's effect alone). This interaction of cues produces a non-linearity in environments where levels of cues are correlated over time or space—for example, if chilling declines are correlated with greater forcing (Fig. 1)—and may be critical to accurate forecasts with climate change.

1.2 How will chilling, forcing and photoperiod shift with climate change?

Translating when or if these non-linearities may be triggered by warming first requires understanding how climate change alters each of the three cues. All three cues are expected to shift with climate change, though the shifts will vary substantially across space and time. Most notably to date, warming increases the forcing plants experience each day, with more rapid shifts—and thus also greater shifts—at higher elevations and in the arctic (IPCC, 2014). Daily minima (generally night-time temperature) generally have and will warm more than maxima temperatures (though this effect varies spatially, Alexander *et al.*, 2006), making efforts to understand whether plants accumulate temperatures differently in the night or day more critical

(Prasad *et al.*, 2008; Shen *et al.*, 2018).

Warming across seasons is also variable (Alexander *et al.*, 2006), meaning warming's impact on forcing (generally accumulating in the late winter and spring) may not be equivalent to impacts on chilling (generally accumulating in the late fall through winter). Warming should translate into important shifts in chilling, which long-term observational studies have repeatedly suggested may already be occurring (Fu *et al.*, 2015; Piao *et al.*, 2017). Our poor understanding of chilling, however, makes current and predicted shifts in chilling complicated (Chuine *et al.*, 2016).

Research to date suggests chilling only accumulates in a certain range of temperatures with low (e.g., $<0°C$) temperatures generally not contributing to chilling accumulations and higher temperatures (e.g., $>12°C$) potentially decreasing previously accumulated chilling (see Fig. 3 and Richardson, 1974; Fishman *et al.*, 1987). Long-term studies generally focus on the warmer part of this chilling accumulation curve, suggesting that chilling should decrease with warming (Fu *et al.*, 2015; Piao *et al.*, 2017; Gauzere *et al.*, 2019). However, major shifts in accumulated chilling are also predicted where warming tips temperature regimes, which were were previously too low to accumulate chilling, to higher temperatures where chilling begins to accumulate (Guy, 2014). Areas with this shift would then expect much earlier budburst with warming, potentially far earlier than last frost dates.

Unfortunately, these predictions are based on models developed almost solely for agricultural crops (but see Harrington & Gould, 2015), especially stone fruits, and have rarely been robustly adapted to forest trees. While the development of classic models of chilling for peaches and related fruit trees benefited from data on these species being planted far outside their range into regions with extremely low or potentially no chilling, equivalent data on forest trees is almost never available (Dennis, 2003). Thus chilling models to date generally use limited observational and experimental data from forest trees to try to re-parametrize the basic stone fruit models (Chuine, 2000; Chuine *et al.*, 2016). This in turn makes any current observations of shifts in chilling—and all forecasts with warming—uncertain. Thus, we believe is important to consider both increases and decreases as potential outcomes of warming (Fig. 3).

Shifts in chilling and forcing with warming have been studied far more than shifts in photoperiod (but see Saikkonen *et al.*, 2012; Way & Montgomery, 2015). While an environment's photoperiod does not shift with climate change, the relevant photoperiod a plant experiences at critical physiological points may change dramatically with warming. In particular, increases in chilling and/or forcing, which could alone produce much earlier budburst, may be offset by short photoperiods that delay budburst (Gauzere *et al.*, 2019). Similarly, long photoperiods can lead to budburst sooner than predicted by solely low chilling or forcing conditions (Nienstaedt, 1966; Myking & Heide, 1995; Partanen *et al.*, 1998). Thus, changes in chilling and/or forcing correspond to changes in the relevant photoperiod with climate change.

These shifts—in forcing, chilling and photoperiod experienced near the time of an event—can produce non-linearities when they push a single cue across a critical threshold or inflection point

(Fig. 1). For example, if some species have a critical photoperiod for budburst and warming means forcing cues are met before the critical threshold, 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 other direction. For example, if pre-climate change conditions generally caused budburst to occur at the extreme levels of some cues and climate change has now pushed budburst into levels where responses are more linear. This is often the mechanism 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 for a minimal effect of chilling—making forcing the dominant cue—whereas warming has now reduced chilling such that more forcing is needed for budburst (producing an overall muted effect when estimated with current methods). 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 experimental studies.

We expect most non-linearities from climate change will come from the effects of interactive cues, as in the previous example where one cue pushed beyond an inflection point triggers shifts in other cues, and due to other covarying shifts caused by environmental change. While simple linear interactions between cues may not alone produce non-linearities (Fig. 1), they quickly become non-linear when changes occur together—for example if increased forcing also occurs in step with shorter photoperiods experienced at the time of an event (Fig. 1). Predicting these non-linearities, however, requires a refined understanding of the interaction between cues and whether there are critical inflection points that may be crossed with continued warming. These complexities highlight how difficult predictions may be without careful efforts to tease out how each cue works alone and interactively.

1.3   Forecasting non-linear responses

Controlled environment (generally growth chamber) studies can help predict non-linear 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 non-linearities occur, and are designed across a range of levels relevant to current versus future conditions. Indeed, one of the major advantages of experiments is that they allow treatments outside of the historical range—an option long-term observational data cannot provide.

We reviewed controlled environment studies over the last seven decades to understand the range of treatments already available, and how they compare to current and future conditions. We note that these studies were rarely conducted for climate change research, and most often done for fundamental science or other areas of applied science (e.g., horticulture or forestry). Yet they are some of the best available data for how plants respond to the environment and thus a critical resource for climate change research today.

How studies and their experimental treatments vary globally

Controlled environment studies have been conducted across 227 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. 4). Most studies manipulate one cue though studies of two or three cues have occurred in almost every decade (Fig. 5). Across study designs, chilling was the most commonly studied cue (69% manipulated chilling of 117 studies that manipulated forcing, chilling or photoperiod), followed by forcing and photoperiod (43% manipulated forcing; 40% manipulated photoperiod).

The levels of cues (e.g., 8 or 10 hours of photoperiod) varied across latitude with a general trend toward examining more extreme values at higher latitudes (see Supp). These shifts across space appear related to differences in extremes across latitudes (higher latitudes experience colder temperatures and longer photoperiods, and see similar shifts in their controlled environment study designs), but introduce a bias in results as any comparisons of studies from lower and higher latitudes are also comparing a different range of cues.

How studies manipulate cues

Studies can be broadly categorized as manipulating one, two or (rarely) three cues at once. Single cue studies were the most common (65 studies), with most manipulating chilling (41 studies), followed by 10 manipulating forcing, 14 manipulating photoperiod. Of the studies manipulating chilling 14 followed the design generally attributed to Weinberger *et al.* (1950) where tissue (e.g., cuttings of adult dormant trees) are taken progressively across the fall and/or winter seasons then exposed to controlled environment conditions. These studies often equate tissue removed later from the field as having received more chilling and thus often treat 'time of cutting' as interchangeable with 'chilling,' though forcing and photoperiod conditions also change. While valuable for defining potential non-linearities in one cue, single cue studies prevent understanding interactions among cues or comparisons of which cues dominate phenological responses—challenges that studies of multiple cues can overcome.

Of the studies manipulating at least one cue, roughly a third additionally manipulated another cue (37% or 43 studies). Study designs most often allowed examining whether cues were interactive (that is, whether the effect of one cue depends on the level of the other cue: 39 studies designed experiments to test this, while 4 studies did not), with the studies almost evenly split across all possible two-way interactions (14 studies tested for photoperiod x chilling, followed by 13 studies tested for forcing x chilling, followed by 12 studies tested for photoperiod x forcing).

Studies examining three cues directly were very rare: we identified only three studies examining all three cues at once (Worrall & Mergen, 1967; Skuterud & Dietrichson, 1994; Søgaard *et al.*, 2008), and two of which were on *Picea abies* (Worrall & Mergen, 1967; Søgaard *et al.*, 2008). A slightly larger set of studies (5 studies from 4 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,

7

1993a; Partanen *et al.*, 1998; Basler & Körner, 2014).

The paucity of studies examining multiple cues limits our fundamental understanding of each cue, as well as how—when combined—they 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. Knowing the level of each other cue is difficult both because they are often not reported (see 'NA' in Fig. 6), and also because they are somewhat impossible to know given our current understanding of endodormancy (when we understand chilling is accumulated) and ecodormancy (when we understand forcing is accumulated). Authors may use the terms 'chilling' and 'forcing' for their treatments, but they rarely have physiological evidence that these are the actual conditions plants experience. Studies using sequential removal from the field to estimate chilling are at perhaps the greatest disadvantage to estimate the cues applied: chilling must rely on field estimates from models that are currently only hypotheses of actual chilling (Dennis, 2003), and forcing and photoperiod treatments are most probably a mix conditions in the field and conditions applied in controlled environments. Though such studies also have the advantage of the most natural chilling conditions.

How relevant are treatments to current and future conditions?
The utility of controlled environment studies to forecasting also depends on how relevant treatments are to current and future conditions. Estimating such relevance is difficult as it depends on a species' range and projections considered. However, a, simple analysis of two well studied species, *Fagus sylvatica* and *Betula pendula*, suggests experiments have generally bracketed the range of projected temperatures (Fig. 7). Projected changes in maximum temperatures generally fit within the range of temperature differences conducted within forcing treatments in experiments (e.g., an experiment with both 16°C and 20°C forcing treatments would have a 4°C difference), and similarly matched differences in minimum temperatures in chilling treatments. As noted above, however, there is a limited number of chilling studies that directly manipulate chilling temperature—and thus allow a comparison of how differences in chilling temperatures impact phenology. Indeed, we found no studies with multiple chill temperatures tested for *Fagus sylvatica*, even though it is one of the most well-studied species.

Experimental shifts were generally larger than expected shifts due to climate change. This makes sense from an experimental-statistical perspective: if the goal of an experiment is to identify if a cue is present then larger treatment differences should yield larger effect sizes and thus higher statistical power. But such large shifts may be risky to extrapolate to smaller shifts due to warming. Further, experimental studies vary from natural settings in myriad ways. Different studies have ameliorated some of these differences. For example, most studies (34 of the 48 that manipulated forcing) have constant day/night temperatures, but many vary day and night temperatures (26 studies; 12 of those studies also include constant forcing) with nights generally being cooler, while some have introduced ramped temperature through the day and across an experiment's length (e.g., Basler & Körner, 2012; Laube *et al.*, 2014). Such ramped conditions are generally introduced across all treatments in experiments and thus provide little insight on how much such experimental artifacts matter (but see Erwin & Heins, 1995). This means

extrapolating from controlled environment studies should be done with care, and highlights a need for future experiments designed to improve forecasting effects of climate change.

1.4   Paths forward

We argue that controlled environment experiments will be critical for accurate predictions of 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 will be designed to improve phenological models. In particular, experiments designed to identify threshold effects and optimal temperatures/photoperiods, and non-linearities from interactive cues may be most useful. Identifying threshold effects and optimal temperatures or photoperiods generally requires many different levels of a single cue, which can make such experiments difficult to cross with other cues. Yet, understanding if findings are consistent across varying levels of other cues should be a follow-up step to confirm that findings can be applied across levels of other cues. Studies manipulating more than one cue also test for non-linearities due to interactive cues, as long as they are fully-crossed (i.e., every combination of levels is present in treatments). Such experiments can quickly require a large number of controlled environments, but provide critical information for models and to connect to long-term findings. As growing experimental results support that all cues are dependent on the level of other cues (Stearns & Olson, 1958; Flynn & Wolkovich, 2018) and long-term data hint at multiple cues (Fu *et al.*, 2015), we believe this should be a major research aim.

Controlled environment studies may also be more readily applied to forecasting by exploring more realistic conditions. While identifying thresholds, optima and non-linearities 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 (Fig. 7). Thus, when designing studies to contribute to improved forecasting of a particular species, experimentalists should examine cues within the current and projected future species' range. In most species distribution models, species are expected to remain in the same climatic conditions, suggesting there may be minimal changes—assuming such models of species distribution are accurate and that species track perfectly. Most evidence, however, 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; Savage & Vellend, 2015).

Beyond the absolute level of cues, controlled environment studies need more work on what attributes of the design are more or less critical for replicating responses from the field. For example, controlled environment studies have shown differing day/night temperatures are important for some species (Heuvelink, 1989; Abrol & Ingram, 1996; Thingnaes *et al.*, 2003; Pressman *et al.*, 2006), but comparison studies have not been conducted for most species. Equally, a few studies have attempted to replicate certain aspects of the environment, such as fluctuating tem-

9

peratures, ramped temperatures and the coincidence of temperature and sunrise (Erwin, 1998), but these are by no means widespread enough to understand how important these conditions are for extrapolation to models.

Understanding the physiology of phenology

Even with all the suggested above improvements, controlled environment studies will still be fundamentally limited in their utility for prediction without an improved understanding of how major phenological cues act physiologically. This problem is most apparent with chilling, but translates also to forcing—as the two are both defined by physiological phases of dormancy that are not easily measured (Singh *et al.*, 2019; Chang *et al.*, 2021).

Chilling is defined as what leads to break of endodormancy, after which plants enter ecodormancy, when accumulated forcing then leads to budburst (Chuine *et al.*, 2016). Measuring endodormancy and its transition into ecodormancy, however, is notoriously difficult. Thus, 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 eco-dormancy—without any evidence or tests of this hoped-for correspondence. Some studies use the sequential transfer of cuttings to warm conditions to estimate the transition from endo- to eco-dormancy, 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 of chilling include this.

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). Research suggests endodormancy may break when enzymes sufficiently remove the sugar (callose) that blocks plasmodesmata in bud cells (reviewed in Chang *et al.*, 2021), but work thus far has relied generally on cellular staining methods tested on a very limited subset of species (Rinne *et al.*, 2011; Singh *et al.*, 2019), making extrapolation to other species difficult. Such results, however, hold promise for a much improved physiological understanding of endodormancy release in the future.

An improved physiological understanding of endodormancy release could revolutionize models of chilling, and in turn, estimates of forcing. Forcing in controlled environment experiments is generally defined simply as warm temperatures (or warm temperatures after cool temperatures). Future estimates could be accurately defined simply as temperatures during ecodormancy, assuming tractable tests of endo- and eco-dormancy and the uptake of such tests in controlled environment studies. With these experiment studies in hand though, researchers could quickly build improved models of chilling, and forcing and—for the first time—provide accurate predictions of how chilling will shift with climate change.

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

A major need currently is improved integration between long-term observational, physiological and controlled environment studies. With important exceptions (e.g., Gauzere *et al.*, 2017),

10

studies of long-term observational phenology data have moved forward independently from advances in our physiological understanding and from controlled environment studies. Similarly, controlled environment studies, as we have reviewed here, generally do not use long-term data to help interpret results or define treatments.

While most studies of long-term data and controlled environments generally ignore one another, attempts to integrate the two provide a useful path forward that can happen now (Caffarra *et al.*, 2011; Nagano *et al.*, 2012; Satake *et al.*, 2013; Ford *et al.*, 2016; Chuine & Regniere, 2017), while we await physiological breakthroughs in defining endo- and eco-dormancy. Experiments that test for thresholds and the presence of important interactions have helped re-design models (Caffarra *et al.*, 2011; Chuine & Regniere, 2017), while other studies have used experiments to test extremes (e.g., extremely low chilling) combined with data from long-term provenance studies to understand how growth and phenology will combine to determine future ranges (Ford *et al.*, 2016). Further, some work has used controlled environments to test model predictions, especially in future climate scenarios where non-linearities are predicted (see Nagano *et al.*, 2012). Such work 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 as newer observational data are generated (i.e., more years and also data from new locations, Nagano *et al.*, 2012; Satake *et al.*, 2013). Such efforts of continual development take extensive data and thus have only been carried out for a very few species (e.g., *Arabidopsis thaliana*, *Oryza sativa* (rice), *Arabidopsis halleri*, Wilczek *et al.*, 2009; Nagano *et al.*, 2012; Satake *et al.*, 2013).

Building species-rich predictions

Given the efforts and data involved in models for a single species, building up to multi-species predictions may appear daunting, but multi-species models are crucial for accurate forecasts that can apply to diverse regions and large-scale vegetation models. Addressing this issue requires, of course, more data. Long-term data is generally more species-rich than controlled environment studies. For example long-term observational data in the PEP725 and NECTAR databases together have multi-site data on more than 2500 species (Wolkovich *et al.*, 2012b; Templ *et al.*, 2018), while our review of controlled environment studies found most studies focused on only one species with data on a total of 226 species. Thus, 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 for analysis of diverse datasets. With the right information and sufficient data, they can attribute variation across studies to the species studied, the cues (i.e., chilling, forcing and photoperiod levels in studies) and remaining unmeasured variation in studies (i.e., differences in chamber design may be captured by including a parameter to estimate a 'study' effect). Such models are extremely powerful for building species-rich predictions as they leverage data across all species into one model designed to capture both the cross-species and cross-study overall effects as well as species-level differences. Yet, like all models, they are more robust with more data. In particular, attributing variation due to study

11

versus species requires the same species to be studied across several studies, which is currently not the case for most species, according to our literature review (81% of the 226 species in our dataset appear in only one study; 32 species, or 14%, appear in two or three studies, and only 10 species, or 4%, appear in more than three studies). Thus, these models will be most useful given greater efforts to publish data. Given proper data reporting (i.e., all cue conditions defined, even when not manipulated, and controlled environment conditions should be fully described, including relative humidity and irradiance) all studies—whether designed to improve models or forecasting, or not—can be included in such models.

1.5 Right now: It's 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 pathways 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. While 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, including agricultural and forest species.

2 References

Abrol, Y.P. & Ingram, K.T. (1996) Effects of higher day and night temperatures on growth and yields of some crop plants. *Global climate change and agricultural production: Direct and indirect effects of changing hydrological, pedological and plant physiological processes* pp. 123–140.

Alexander, L.V., Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Tank, A., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Kumar, K.R., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D.B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M. & Vazquez-Aguirre, J.L. (2006) Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research-Atmospheres* **111**.

Basler, D. & Körner, C. (2012) Photoperiod sensitivity of bud burst in 14 temperate forest tree species. *Agricultural and Forest Meteorology* **165**, 73–81.

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., Riofr'ıo-Dillon, G., de Ruffray, P., Vidal, C., Pierrat, J.C. & G'egout, J.C. (2011) Changes in plant community composition lag behind climate warming in lowland forests. *Nature* **479**, 517.

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.

Cesaraccio, C., Spano, D., Snyder, R.L. & Duce, P. (2004) Chilling and forcing model to predict bud-burst of crop and forest species. *Agricultural and Forest Meteorology* **126**, 1–13.

Chang, C.Y.Y., Brautigam, K., Huner, N.P.A. & Ensminger, I. (2021) Champions of winter survival: cold acclimation and molecular regulation of cold hardiness in evergreen conifers. *New Phytologist* **229**, 675–691.

Chen, L., Huang, J.G., Ma, Q.Q., 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.

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

Chuine, I., Bonhomme, M., Legave, J.M., 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, E.E., Chuine, I., Menzel, A., Mooney, H.A. & Schwartz, M.D. (2007) Shifting plant phenology in response to global change. *Trends in Ecology & Evolution* **22**, 357–365.

Cronjé, P., Jacobs, G., Sadie, A. & Cook, N. (2003) Quantification of the dormancy progression in terminal apple buds. changes in growth rate and water status. *Advances in horticultural science* pp. 105–110.

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.

Dennis, F. (2003) Problems in standardizing methods for evaluating the chilling requirements for the breaking of dormancy in buds of woody plants. *HortScience* **38**, 347–350.

Ellwood, E.R., Temple, S.A., Primack, R.B., Bradley, N.L. & Davis, C.C. (2013) Record-breaking early flowering in the eastern united states. *Plos One* **8**, ellwood, Elizabeth R. Temple, Stanley A. Primack, Richard B. Bradley, Nina L. Davis, Charles C.

Elmendorf, S.C. & Ettinger, A.K. (2020) Is photoperiod a dominant driver of secondary growth resumption? *Proceedings of the National Academy of Sciences of the United States of America* **117**, 32861–32864.

Erwin, J.E. (1998) Temperature and light effects on stem elongation. *Journal of the Japanese Society for Horticultural Science* **67**, 1113–1120.

Erwin, J.E. & Heins, R.D. (1995) Thermomorphogenic responses in stem and leaf development. *Hortscience* **30**, 940–949.

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* pp. 1–6.

Fishman, S., Erez, A. & Couvillon, G. (1987) The temperature dependence of dormancy breaking in plants: mathematical analysis of a two-step model involving a cooperative transition. *Journal of Theoretical Biology* **124**, 473–483.

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

Ford, K.R., Harrington, C.A., Bansal, S., Gould, Peter, J. & St. Clair, J.B. (2016) Will changes in phenology track climate change? A study of growth initiation timing in coast Douglas–fir. *Global Change Biology* **22**, 3712–3723.

Fu, Y.H., Piao, S., Zhou, X., Geng, X., Hao, F., Vitasse, Y. & Janssens, I.A. (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, Y.S.H., Zhao, H.F., Piao, S.L., Peaucelle, M., Peng, S.S., Zhou, G.Y., Ciais, P., Huang, M.T., Menzel, A., Uelas, J.P., Song, Y., Vitasse, Y., Zeng, Z.Z. & Janssens, I.A. (2015) Declining global warming effects on the phenology of spring leaf unfolding. *Nature* **526**, 104–107.

Gauzere, J., Delzon, S., Davi, H., Bonhomme, M., de Cortazar-Atauri, I.G. & Chuine, I. (2017) Integrating interactive effects of chilling and photoperiod in phenological process-based models. A case study with two European tree species: *Fagus sylvatica* and *Quercus petraea*. *Agricultural and Forest Meteorology* **244**, 9–20.

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. (2006) *Data analysis using regression and multilevel/hierarchical models*. 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.

Guy, R.D. (2014) The early bud gets to warm. *New Phytologist* **202**, 7–9.

Harrington, C.A. & Gould, P.J. (2015) Tradeoffs between chilling and forcing in satisfying dormancy requirements for pacific northwest tree species. *Frontiers in Plant Science* **6**, 120.

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

Heide, O. (1993b) Dormancy release in beech buds (*Fagus sylvatica*) requires both chilling and long days. *Physiologia Plantarum* **89**, 187–191.

Hereford, J., Schmitt, J. & Ackerly, D.D. (2017) The seasonal climate niche predicts phenology and distribution of an ephemeral annual plant, mollugo verticillata. *Journal of Ecology* **105**, 1323–1334.

Heuvelink, E. (1989) Influence of day and night temperature on the growth of young tomato plants. *Scientia Horticulturae* **38**, 11–22.

IPCC (2014) *Climate Change 2014: Impacts, Adaptation, and Vulnerability*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

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.

Keenan, T.F., Richardson, A.D. & Hufkens, K. (2020) On quantifying the apparent temperature sensitivity of plant phenology. *New Phytologist* **225**, 1033–1040.

Lamb, R.C. (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*, vol. 51, pp. 313–315, AMER SOC HORTICULTURAL SCIENCE 701 NORTH SAINT ASAPH STREET, ALEXANDRIA, VA 22314-1998.

Laube, J., Sparks, T.H., Estrella, N., Höfler, J., Ankerst, D.P. & Menzel, A. (2014) Chilling outweighs photoperiod in preventing precocious spring development. *Global Change Biology* **20**, 170–182.

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

Major, D.J. (1980) Photoperiod response characteristics controlling flowering of 9 crop species. *Canadian Journal of Plant Science* **60**, 777–784.

Martinez-Luscher, J., Hadley, P., Ordidge, M., Xu, X.M. & Luedeling, E. (2017) Delayed chilling appears to counteract flowering advances of apricot in southern uk. *Agricultural and Forest Meteorology* **237**, 209–218.

Menzel, A., Jakobi, G., Ahas, R., Scheifinger, H. & Estrella, N. (2003) Variations of the climatological growing season (1951–2000) in Germany compared with other countries. *International Journal of Climatology* **23**, 793–812.

Menzel, A., Sparks, T.H., 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.

Myking, T. & Heide, O. (1995) Dormancy release and chilling requirement of buds of latitudinal ecotypes of *Betula pendula* and *B. pubescens*. *Tree Physiology* **15**, 697–704.

Nagano, A.J., Sato, Y., Mihara, M., Antonio, B.A., Motoyama, R., Itoh, H., Nagamura, Y. & Izawa, T. (2012) Deciphering and prediction of transcriptome dynamics under fluctuating field conditions. *Cell* **151**, 1358–1369.

Nienstaedt, H. (1966) Dormancy and dormancy release in white spruce. *Forest Science* **12**, 374–384.

Parent, B. & Tardieu, F. (2012) Temperature responses of developmental processes have not been affected by breeding in different ecological areas for 17 crop species. *New Phytologist* **194**, 760–774.

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, J.F., Chevallier, F., Janssens, I.A. *et al.* (2017) Weakening temperature control on the interannual variations of spring carbon uptake across northern lands. *Nature climate change* **7**, 359.

Prasad, P.V.V., Pisipati, S.R., Ristic, Z., Bukovnik, U. & Fritz, A.K. (2008) Impact of nighttime temperature on physiology and growth of spring wheat. *Crop Science* **48**, 2372–2380.

Pressman, E., Shaked, R. & Firon, N. (2006) Exposing pepper plants to high day temperatures prevents the adverse low night temperature symptoms. *Physiologia Plantarum* **126**, 618–626.

Richardson, A.D., Hufkens, K., Milliman, T., Aubrecht, D.M., Furze, M.E., Seyednasrollah, B., Krassovski, M.B., Latimer, J.M., Nettles, W.R., Heiderman, R.R., Warren, J.M. & Hanson, P.J. (2018) Ecosystem warming extends vegetation activity but heightens vulnerability to cold temperatures. *Nature* **560**, 368–+.

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

Rinne, P.L.H., Paul, L.K. & 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**.

Rinne, P.L.H., Welling, A., Vahala, J., Ripel, L., Ruonala, R., Kangasjarvi, J. & van der Schoot, C. (2011) Chilling of dormant buds hyperinduces FLOWERING LOCUS T and recruits GA-Inducible 1,3-beta-Glucanases to reopen signal conduits and release dormancy in *Populus*. *Plant Cell* **23**, 130–146.

Saikkonen, K., Taulavuori, K., Hyvönen, T., Gundel, P.E., Hamilton, C.E., Vänninen, I., Nissinen, A. & Helander, M. (2012) Climate change-driven species' range shifts filtered by photoperiodism. *Nature Climate Change* **2**, 239.

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**.

Savage, J. & Vellend, M. (2015) Elevational shifts, biotic homogenization and time lags in vegetation change during 40 years of climate warming. *Ecography* **38**, 546–555.

Schnabel, B.J. & Wample, R.L. (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, M.D. (1997) Spring index models: An approach to connecting satellite and surface phenology. *Phenology in Seasonal Climates I* (eds. H. Lieth & M.D. Schwartz), pp. 22–38, Backhuys Publisher, Leiden, The Netherlands.

Shen, X.J., Liu, B.H., Henderson, M., Wang, L., Wu, Z.F., Wu, H.T., Jiang, M. & Lu, X.G. (2018) Asymmetric effects of daytime and nighttime warming on spring phenology in the temperate grasslands of china. *Agricultural and Forest Meteorology* **259**, 240–249.

Singh, R.K., Miskolczi, P., Maurya, J.P. & Bhalerao, R.P. (2019) A tree ortholog of short vegetative phase floral repressor mediates photoperiodic control of bud dormancy. *Current Biology* **29**, 128–+.

Singh, R.K., Svystun, T., AlDahmash, B., Jönsson, A.M. & Bhalerao, R.P. (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, O.M. (2014) Chilling requirements of contrasting black currant (*Ribes nigrum* L.) cultivars and the induction of secondary bud dormancy. *Scientia Horticulturae* **179**, 256–265.

Sparks, T.H. & Carey, P.D. (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.

Tansey, C.J., Hadfield, J.D. & Phillimore, A.B. (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, F.M., Hájková, L., Hodzić, S., Kaspar, F., Pietragalla, B., Romero-Fresneda, R., Tolvanen, A., Vučetič, V., Zimmermann, K. & Zust, A. (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.

Thingnaes, E., Torre, S., Ernstsen, A. & Moe, R. (2003) Day and night temperature responses in arabidopsis: Effects on gibberellin and auxin content, cell size, morphology and flowering time. *Annals of Botany* **92**, 601–612.

van der Schoot, C., Paul, L.K. & Rinne, P.L.H. (2014) The embryonic shoot: a lifeline through winter. *Journal of Experimental Botany* **65**, 1699–1712.

Way, D.A. & Montgomery, R.A. (2015) Photoperiod constraints on tree phenology, performance and migration in a warming world. *Plant, Cell & Environment* **38**, 1725–1736.

Weinberger, J.H. *et al.* (1950) Chilling requirements of peach varieties. *Proceedings. American Society for Horticultural Science*, vol. 56, pp. 122–28.

Wilczek, A.M., Roe, J.L., Knapp, M.C., Cooper, M.D., Lopez-Gallego, C., Martin, L.J., Muir, C.D., Sim, S., Walker, A., Anderson, J., Egan, J.F., Moyers, B.T., Petipas, R., Giakountis, A., Charbit, E., Coupland, G., Welch, S.M. & Schmitt, J. (2009) Effects of genetic perturbation on seasonal life history plasticity. *Science* **323**, 930–934.

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

Wolkovich, E.M., Cook, B.I. & Regetz, J. (2012b) NECTAR: Network of Ecological and Climatological Timings Across Regions, http://knb.ecoinformatics.org/knb/metacat/nceas.988/knb.

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

Yu, H.Y., Luedeling, E. & Xu, J.C. (2010) Winter and spring warming result in delayed spring phenology on the Tibetan Plateau. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 22151–22156.

Zohner, C.M., Benito, B.M., Svenning, J.C. & Renner, S.S. (2016) Day length unlikely to constrain climate-driven shifts in leaf-out times of northern woody plants. *Nature Climate Change* **6**, 1120–1123.

Zohner, C.M. & Renner, S.S. (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.
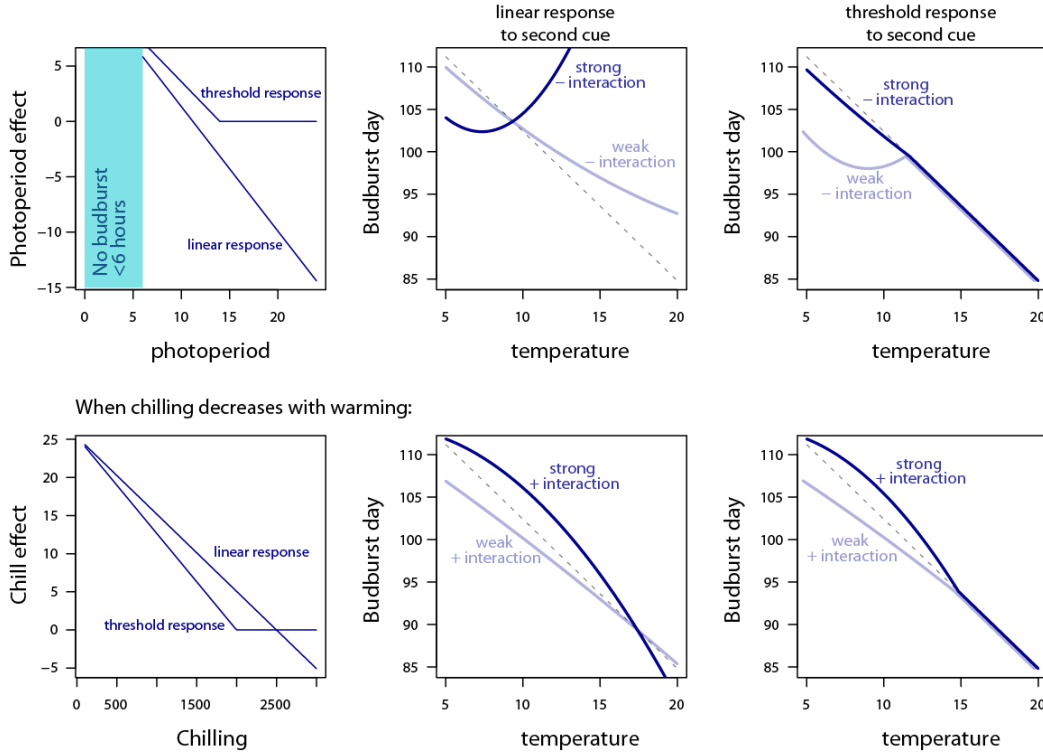
3   Figures

Figure 1: Interactions can produce nonlinearities, even in simple linear models if there are correlated shifts in cues. Here we show examples considering a forcing x photoperiod interaction (top) and forcing x chilling (bottom), with both a linear effect of photoperiod or chilling (middle panels) or a threshold effect (left panels) versus an effect of forcing only (dashed gray line). These non-linearities come from the levels of both cues shifting at once: for the photoperiod example (top), as temperatures rise forcing increases and the photoperiod experienced near the time of the event shortens (a similar response would be seen for lowering chilling), while for the chilling example (bottom), as temperatures rise forcing and chilling increases.
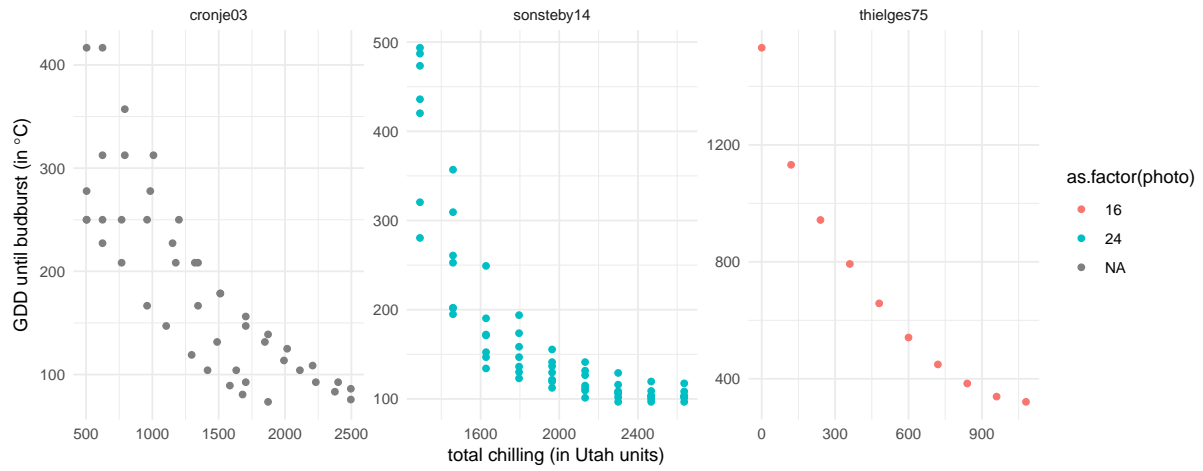
Figure 2: A common example of how the level of one cue can modify another is experiments finding that the amount of chilling affects of the amount of forcing needed for budburst. Here, we show this from Cronjé *et al.* (2003) which studied apple (*Malus sylvestris*), Sønsteby & Heide (2014) which studied black currant (*Ribes nigrum*) and Thielges & Beck (1976) which studied eastern cottonwood (*Populus deltoides*).
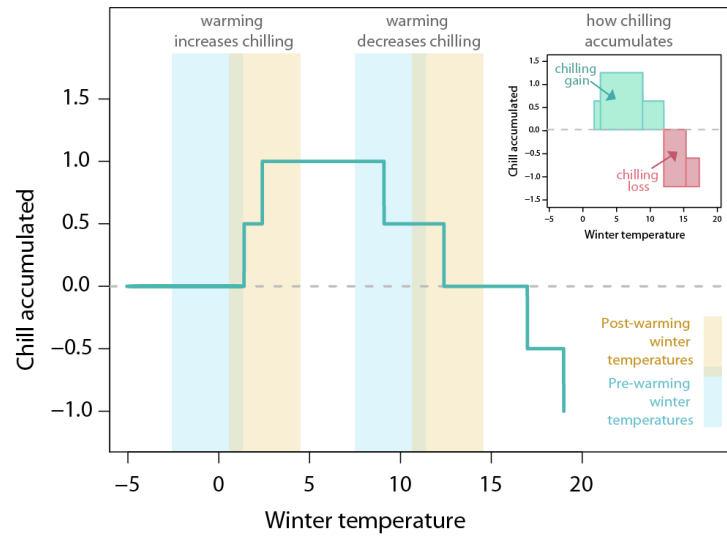
Figure 3: Current models of chilling suggest it may decrease or increase with winter warming. Here we show a common version of the Utah chilling model (top right inset and also turquoise line in main figure) with two conceptual scenarios of mean daily winter temperatures. When temperatures are generally below zero warming may increase accumulated chilling (left), while if pre-climate change temperatures are generally higher (near where chilling accumulates most per °C) then warming may decrease accumulated chilling (right).
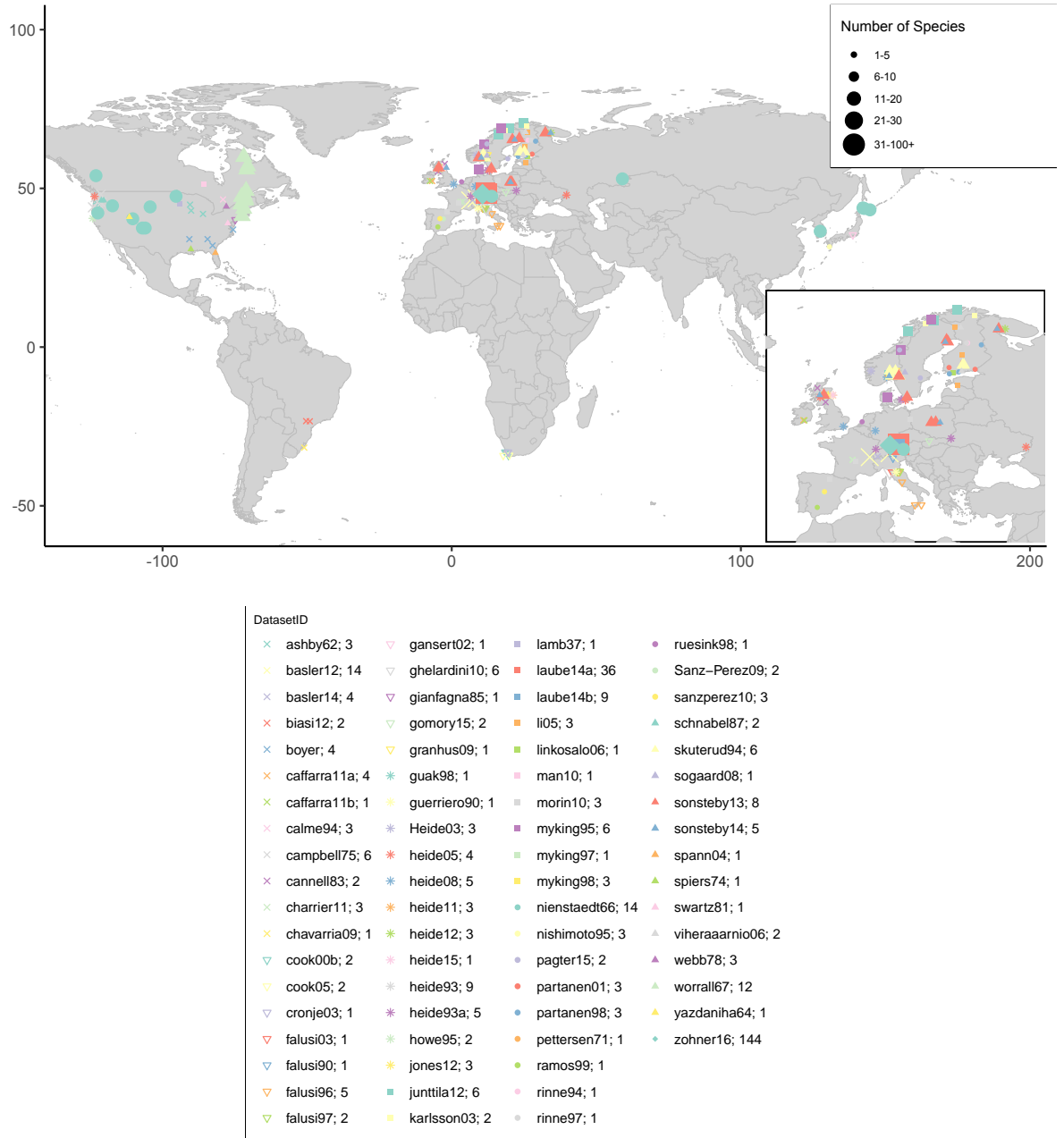
Figure 4: We review seven decades of controlled environment studies, from Lamb (1948) to Zohner *et al.* (2016), conducted across the globe generally on 1-3 species in each experiment (size of circles and exact number of species given after each each study).
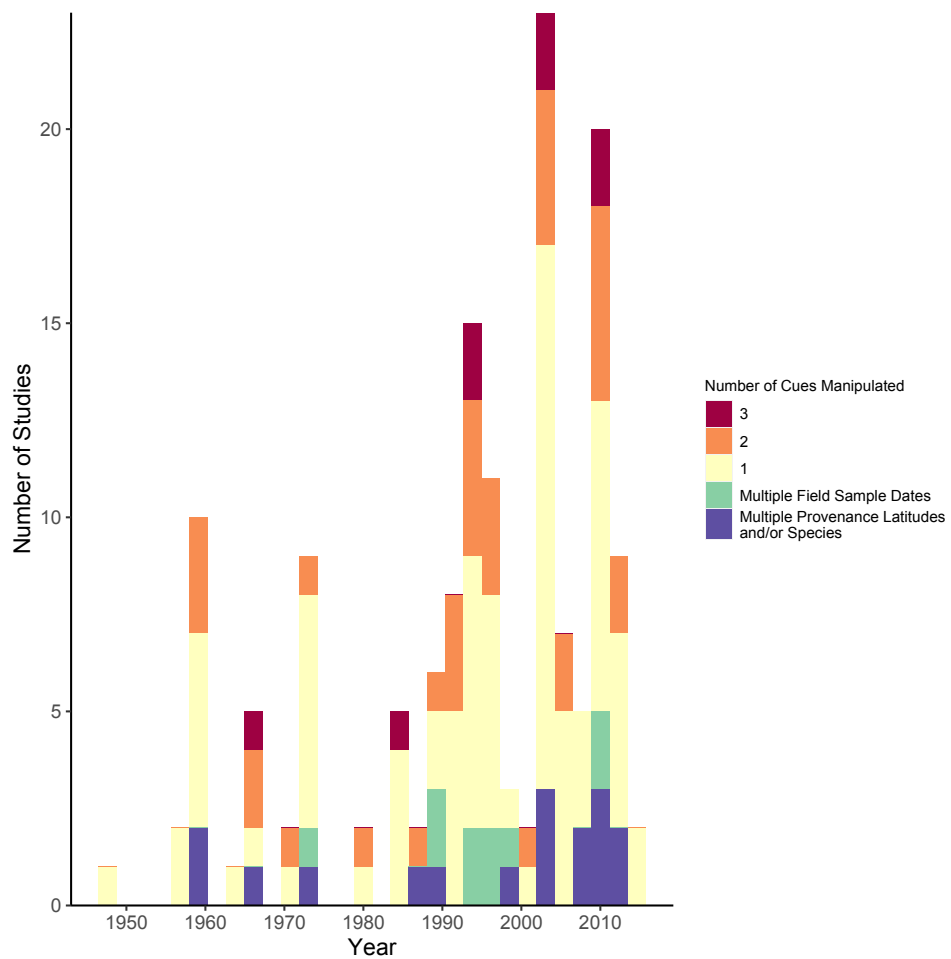
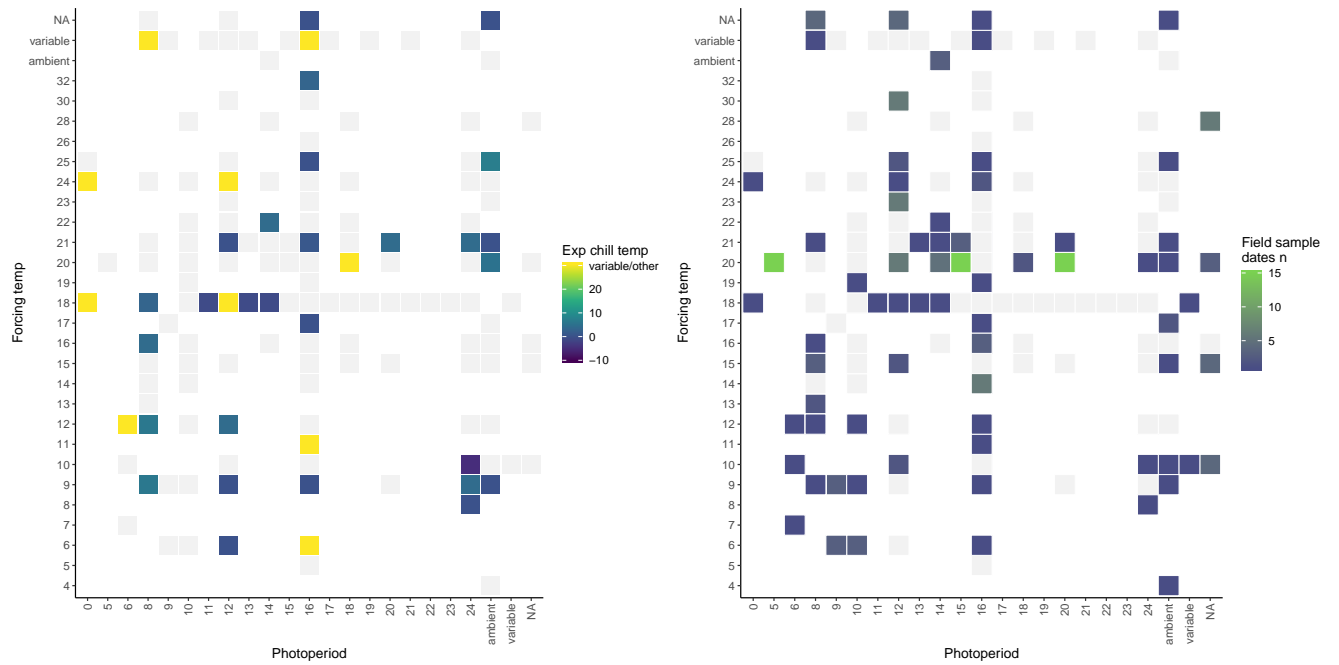Figure 5: Cues manipulated over time.

Figure 6: Studies have manipulated a range of forcing temperatures (vertical axis) and photoperiods (horizontal axis) across two methods to manipulate chilling: experimentally (left) or using multiple field sampling dates (right). A value of 'NA' means we could not estimate a level of a particular cue from the study because it was not clearly reported
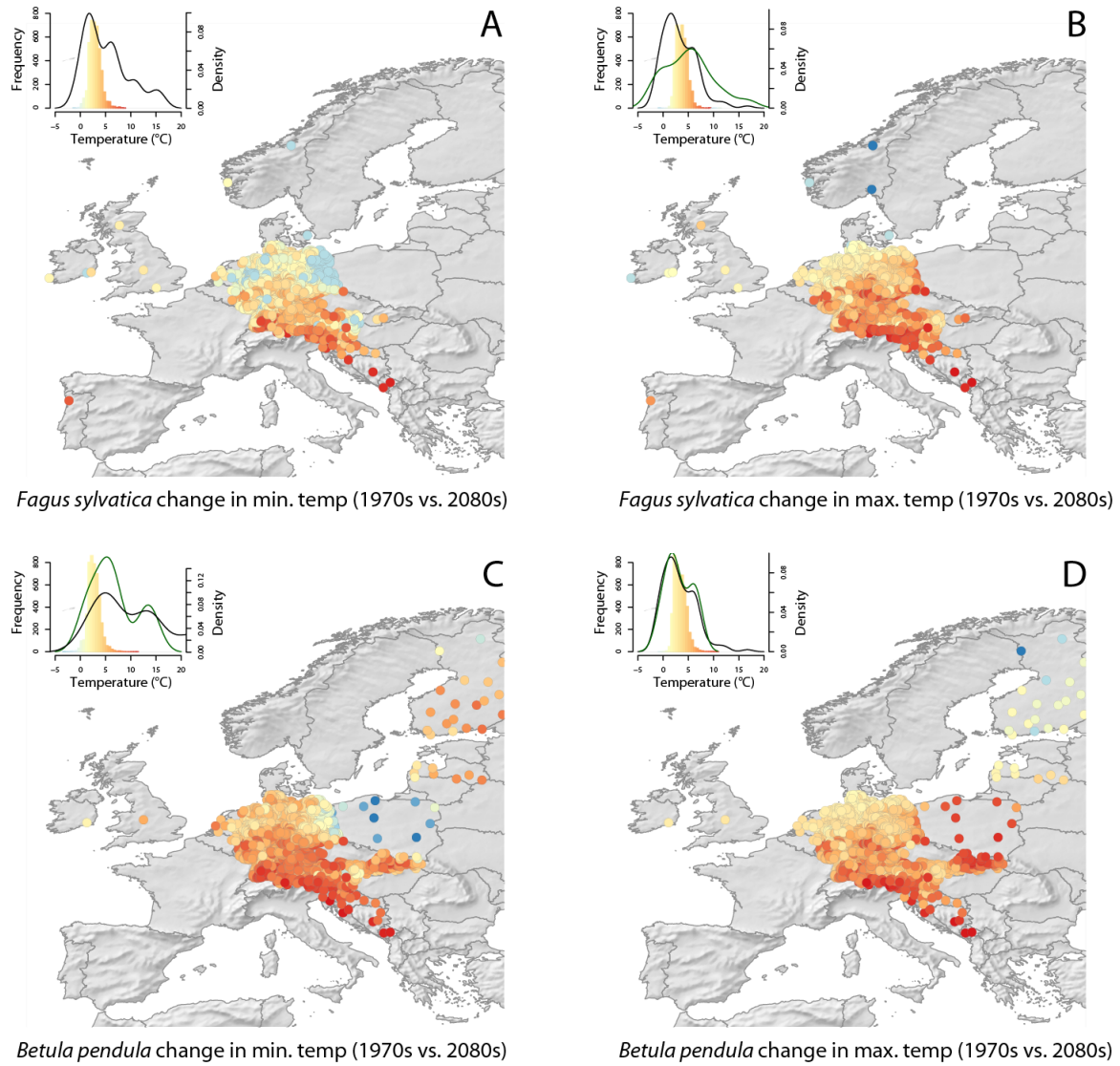
Figure 7: Comparing predicted changes in temperatures relevant to chilling (A, C) and forcing (B, D) with controlled environment studies for two species: *Fagus sylvatica* (A-B) and *Betula pendula*. Points represent a PEP725 site with spring phenology data for each species respectively and show the predicted change (2080s average - 1970s average) in minima winter temperatures (A,C) and maximum spring temperatures (B,D). Inlay plots in the upper-left corner of each plot show a histogram of these same predicted changes in temperature overlaid with densities of the chilling (A, C) and forcing (B, D) differences in treatments across 33 studies (green lines show the treatments for that exact species, while black lines show across all species; note that for *Fagus sylvatica* there are no chilling treatments of differing temperatures). Treatments differences were calculated as the differences within varying forcing and chilling treatments within a single study (e.g., a study with a 1 and 4°C chilling treatment would yield a value of 3°C).