Reviewer comments are italicized and our detailed explanation of changes are in regular text.

Associate Editor's comments:

Handling Editor Comments for Authors: Both reviewers felt this ms has potential to offer insights into whether functional traits of woody species can be used (in place of species identities) to predict the relationships between leaf budburst timing and abiotic conditions. However, the reviewers struggled to understand the framing, methods, results, and conclusions, in part due to a lack of clarity and sufficient detail, and in part because of contradictory information. The constructive comments from the reviewers should help this work reach its potential, but addressing these issues will likely mean the ms will change substantially.

Reviewer 1 – comments:

Major comment:

1. I find the approach they have chosen to analyse the data very exciting. However, I cannot follow all the conclusions (in particular from Figure 2). I wonder if the lines shown represent some kind of significance test indicating that there are significant linear correlations between x and y? You need to clarify this. Additionally, I wonder if the analyses from Figure 4 are relevant for your study?

Minor comments:

L9: You can shorten this sentence by only writing "spring growing".

We have made these changes (line 9).

L15: In my opinion, here and elsewhere, it is important to point out that you worked with trees. Later on, the introduction repeatedly refers to 'plants', suggesting that herbaceous species and trees were considered in this study.

Done (line 16).

L17: Please briefly describe what forcing and chilling mean.

L18: What do you mean with "earlier species"? Please be more precise here and for example write "earlier growing species".

We have changed this wording to improve clarity here and throughout the manuscript (e.g. line 18 and line 47)
L18-19: Is this actually the main findings of your study? Was the study not primarily about the different responses of species to environmental factors explained by traits?

L 25: "Budburst" is already used in the title. You should consider another keyword.

We have changed "budburst" to "leafout" to no longer replicate the terminology used in the title.
L40-42: Nice to know that egg laying is also shifting but is this relevant for your study? I Would rather name and describe more examples of plants and especially trees.

We have revised this sentence to be more focused on examples of phenological shifts in plants (line 41).

L45: As you show an example for animals a few s before writing "species" here is misleading. You should write "plants" or even more specific "trees".

Done (line 44).

L48: "later growing species"
Done (line 47).

L50: You can delete "For plants,".
Done (line 49).

L68-70: This statement is very strict. I would phrase it a bit more carefully (e.g. 'to the best of our knowledge'). For example, in a study cited by you (Sporbert et al. 2022), trait-phenology interactions in herbaceous plants from different sites were examined. Whether there are similar examples for trees, I cannot say.

We agree with the reviewer that this sentence was too strict. To address the reviewer comments we have revised this sentence to:

> To the best of our knowledge, studies have generally only examined the relationship between traits and phenology for a single or limited number of sites (as reviewed by Wolkovich & Cleland [46] and Wolkovich & Donahue [47], but see Sporbert *et al.* [37] and **?**) where the problem of proximate drivers causing phenological variation can be more easily ignored.

L75: What do you mean with "which do not generally vary strongly across space and time"? Is it important to mention this in the context of your study?

We see the reviewer's points and agree that this is not important to our study. We have removed this phrase from line 74.

L107: To my knowledge OSPREE focusses on temperate tree species. This information should be added.

Done (line 105)
L115: Articles can't manipulate something. Researchers manipulated the abiotic factors. Please reformulate.

We see the reviewer's points and appreciate the need to be accurate, we have revised this sentence as follows (line 113:

> We then scraped the data from all articles of woody species reporting experiments that manipulated both photoperiod and/or temperature cues of budburst, leafout, or flowering phenology, and for which there was a quantifiable response to each cue.

L121: Instead of "used" you can write "considered".

The appreciate the importance of word choice in methods and feel the word "considered" implies that we thought carefully about the subset of data where as "used" clearly implies what data was utilized in the analysis.

L130: Are the 96 species from TRY additional ones to the species extracted from BIEN meaning that you considered 190 species in total?

L132: How is this really specific threshold (1.38m) defined?

L136: What does this n refer to? Is it the duplicates that were removed? If so, is this information important?

L142: Please reformulate. For example "[...] and to adjust which traits should be included for further analyses."

L142: "the" instead of "our".

Done (line 144).

L149: Please delete "also".

Done (line 153).

L150-151: I appreciate that you have considered the different data available for the different species. However, I am not sure whether the approach chosen here is correct. What was the reason for you randomly selecting 3000 times? Could it be that, due to the 3000-fold random drawing, species with relatively few data points are always drawn with the same data points? This then leads to the fact that there are species with and without variability in the 3000 data points.

L153: Maybe better to write "traits" instead of "trait"?

Done (line 159).

L153: I think it would be very helpful for readers to visualise the structure of the data analysis in a graph. The data filtering process could also be included here.

L160: Please shorten: "[...] that determines the day of year of budburst".

L212: How do you conclude that SLA is not related to chilling or forcing, unlike to photoperiod? For example, the mean slope for chilling is 0.34 (see Table S4). Or have I not understood this table correctly? But also, if I look at Figure 2, I would not conclude that there are no effects for the two cues.

L222-231: These are definitely interesting analyses. But I don't understand how they are relevant to your study. If they are relevant, then these results would have to be taken up again in the discussion.

L244-245: The examples presented are studies with herbaceous species. I think it's important not to lump together the results of different growth forms. For example, Horbach et al. showed that phenology-trait relationships for herbaceous species and for trees can differ greatly (https://doi.org/10.1016/j.flora.2023.152218).

L261: Are these results not be visualised with even greater clarity in Figure 3?

L446: The authors are missing.

Figure 1: I really like this figure. You also analysed seed mass. Could you add this trait to the "early" and "late" boxes?

Figure 2: The y-axes names are really cryptic. It would be helpful if simpler formulations could be used here. In addition, the labels on both axes have to be larger.

Figure 3: How can the points reflect both the raw phenological data and individual trait measurement at the same time? This description of the figure is misleading. Finally, I would only write the name of the columns ("Forcing", "Chilling" and "Photoperiod") once at the top.

Figure 2 and 3: I am colour-blind, with a red-green deficiency. I can hardly distinguish the colours you have chosen. This should be adjusted.

Reviewer 2 – comments

This study investigates the role of functional traits for mediating phenological responses of bud burst timing to different environmental cues (forcing, chilling, photoperiod). The authors compiled data on phenological experiments on woody species from a database and combined them with trait data from other databases. While the manuscript is generally very interesting and well written I see major problems regarding the underlying assumptions and clarity of the manuscript.

The underlying assumptions and hypotheses presented in the introduction are not supported by suitable references and largely contradict the current literature (see comments below). Acquisitive plants relate to high SLA, nitrogen rich leaves and low resistance (e.g. against frost) but higher competitive ability. In contrast, conservative growth strategies relate to low SLA, nitrogen poor leaves with high resistance (e.g. against frost). With competition playing a minor role in spring but being more important later in the season as stated by the authors I would expect more conservative strategies to be present in spring to ensure frost resistance and expect more acquisitive strategies later in the season when the risk of frost events is low, but competition is high. The assumptions and hypotheses by the authors presented in Figure 1 and throughout the manuscript are in parts opposing to these assumptions. In addition, the authors define the terms "acquisitive" and "conservative" differently throughout the manuscript (e.g., Introduction: acquisitive = high SLA, Figure 1: acquisitive = low SLA), which leads to confusion and makes it difficult to understand what the authors mean.

The results should be presented more clearly, and the discussion should directly refer to the results instead of reiterating statements from the introduction or making claims that are not adequately supported by the presented results.

Summary

L10-12: Why do you expect that? I would expect exactly the opposite.

L16: Add information here whether the traits are coming from the same experiments or if they were derived by databases.

L18: This indicates a more conservative strategy of early species.

L19: Where did the high nitrogen leaves belong to if early and late species had low nitrogen?

L20: Not clear what the authors want to say here. Maybe be more precise about the patterns regarding late species.

Introduction

General comment: I find the wording for "proximate" and "ultimate" level a bit complicated even though it is well explained. Maybe it's a language related issue, but I would consider using more intuitive words to describe the different factors (abiotic cues vs. plant traits).

From the introduction I get the impression that the authors aim to understand not only interspecific but also intraspecific phenological variability through plant functional traits. However, as they used only database information for the traits they cannot account for intraspecific trait variability in relation to phenological variability, or do I get anything wrong?

L55-57: I find this statement quite counterintuitive. Also, in the two references (Reich et al., 1999; Gorne et al., 2020) I can't find any link to phenology or related statement that would support this sentence.

L68-70: Please reevaluate this statement. There are publications on phenology-trait relationships also across sites check e.g., (Rauschkolb et al., 2024; Sporbert et al., 2022).

L74-75: This sentence is not clear to me. Please improve clarity of the statement.

L76-80: This seems counterintuitive to me as early-flowering species are repeatedly shown to experience more pronounced phenological advancements in response to climate warming compared to species that flower later (Dunne et al., 2003; Fitter and Fitter, 2002; Lesica and Kittelson, 2010; Menzel et al., 2006; Pareja-Bonilla et al., 2023).
The Laube et al. (2014) reference does speak of early vs. late successional species while I understand this paragraph (and this manuscript overall) aiming at understanding differences between early and late bud bursting species. It needs to be defined more clearly what this the key message of this paragraph.

L84-85: Not clear what "strong gradients in potential selective environments" exactly means in this context.

L86-87: There are contradictions within this statement: Acquisitive leaves are linked to high SLA and high nitrogen content but are less resistant to frost damage. Thus, I would not expect acquisitive plants to show early budburst but rather conservative plants that have a higher resistance. For the relationship between SLA and frost resistance see: Bucher SF, Rosbakh S. Foliar summer frost resistance measured via electrolyte leakage approach as related to plant distribution, community composition and plant traits. Funct Ecol. 2021; 35: 590–600. https://doi.org/10.1111/1365-2435.13740

L87-88: But early species face the highest risk of frost damage so should rather invest in resistance (would be related to low SLA and a more conservative strategy). See also comment regarding L76-80.

L90-91: Wouldn't leaves that are suited for variable or less favorable light conditions be linked to higher SLA?

Methods

There are many different species numbers mentioned based on different databases. Please clarify for how many species information on all relevant variables was available and how many species were included in the analysis. Otherwise, it creates the impression that this study is based on 234 species while there is only trait information for about 95 species, in the end the authors speak of 37 species.

It would be also helpful to learn more about the experiments were the phenology data was derived from and to provide a map with the geographic locations of the included experiments. Only in the discussion the authors speak of a "global" study.

L153-162.: Why were not all traits included in one model to evaluate the influence of each of the traits in the phenological response? The model description is overall insufficient as it is not clear which methods or software was overall used.

L163-172: The model description is overall insufficient as it is not clear which methods or software was overall used.

I'm not very familiar with Bayesian statistics but the statistical analysis needs to be described in more detail to make them understandable for a broader readership. In the methods section it is not even mentioned that Bayesian statistics were applied. This is only stated within the introduction. Please provide all relevant details within the methods section.

Results

L204-205: Why were these species selected to represent typical acquisitive/ conservative species? Alnus incana ("acquisitive") and Quercus rubra ("conservative") are very closely together in terms of SLA that is a proxy for acquisitive/conservative growth strategies. Why not using two species that are more towards the ends of the acquisitive/conservative spectrum?

The result section is overall very short and written very complicated. Results part could be extended to describe the results in a little more detail.

Figure 1: The presented framework of early-budbursting species being related to 'acquisitive growth' and late-budbursting species being related to 'conservative growth' is not in line with literature and seems counterintuitive (see also comments regarding the introduction). Also, the presented trait expressions are not all correctly assigned to acquisitive and conservative (e.g. high SLA should be related to acquisitive growth and not conservative).

Figure 2: The y-axes should be unified to allow a better comparison between the figures and traits.

Figure 3: It is not clear what is exactly shown here. The legend does not help to understand what this figure shows.

According to the axes we again see the relationship between day of budburst and the three cues

as in Figure 2. Then it is not clear where the traits are incorporated in the figures. Also, it is not clear why only height and seed mass are presented.

"The effect of height on budburst timing was smaller in response to forcing cues, but larger in response to both b, chilling and c, photoperiod" I don't see any relationship for photoperiod. Please clarify.

Figure 4: What are the implications of the presented results? It should be discussed within the discussion section.

Discussion

The presented results are not sufficiently discussed. Rather it is my impression that new results are introduced that were not presented within the results section. In general, many arguments from the introduction are repeated without support by literature and also with no clear relationship to the presented results.

L235-236: From the presented results it is not clear to me how trait effects were associated with earlier or later phenology and how the results relate to the referenced growth strategy papers. Was there any test that included the phenological niche (timing of budburst or similar)?

L240-241: This is partly contradictory with results presented in the results section. Not all of the three traits showed large responses to all three cures.

L244: When speaking of 'more local studies' – what is the geographic range of the data that is the basis of this study? Maybe include a map with geographic locations of the included studies. Where do the authors present the relationship between plant height and bud burst? I only find the results of the traits with regard to the cue responses but no general relationships.

L251-253: Again, I can't find clear support for this statement in the presented results. Also, it is not clear why low LNC should be associated with an acquisitive strategy. This would rather imply a conservative strategy that could be expected for early species that face risk of frost.

L256-257: How does greater competitive ability and slower growth fit together? This doesn't make sense in my opinion.

L260: This is the first time that species successional position is mentioned but it seems to be one of the reasons for selecting the two exemplary species. I get the impression that the authors use early vs. late bud burst and early vs. late successional positions interchangeably throughout the manuscript, which are very different things.

L268-270: Contradicts statement from L240-241. Further the given explanation is not clear.

L282-285: Interesting but formulated too vague. Which traits and trait combinations and maybe more interesting which growth strategies would relate to stronger or less strong responses according to your results?