UNIVERSITY OF CALIFORNIA

Santa Barbara

Patterns, processes, and emergent properties of flowering responses to climate across spatial and ecological scales

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Ecology, Evolution, and Marine Biology

by

Tadeo Hernan Ramirez Parada

Committee in charge:

Professor Susan J. Mazer, Chair

Professor Leander Anderegg

Professor Elsa E. Cleland

December 2024

The dissertation of Tadeo Hernan Ramirez Parada is approved.

_____

Leander D. L. Anderegg

_____

Elsa E. Cleland

_____

Susan J. Mazer, Committee Chair

December 2024

Patterns, processes, and emergent properties of flowering responses to climate across spatial and ecological scales

by

Tadeo Hernan Ramirez Parada

ACKNOWLEDGEMENTS

I am grateful to my advisor, Dr. Susan Mazer, for her insight and steadfast dedication to our work throughout these difficult years—no one will give feedback as reliably and with as much attention to detail as Susan. I also thank Dr. Isaac Park for his close collaboration, and for laying the foundation along with Susan for much of the research I conducted during my Ph.D. I am grateful for the support and good company of my lab mates present and past, Cameron Hannah-Bick, Devin Gamble, Helen Payne, Lisa Kim, Natalie Love, and Kristen Peach. I also thank my committee members, Dr. Elsa Cleland and Dr. Leander Anderegg, for their well-timed and encouraging feedback. I am grateful to An Bui, Michelle Lepori-Bui, Dr. Jasmine Childress, Dr. Halley Froehlich, Dr. Todd Oakley, Dr. Marisela Marquez, and Dr. Meredith Merchant. I cannot imagine what the challenge of serving in the D.E.I. committee through the depths of the pandemic would have been like without them. I would not be here without my parents Elia Parada and Hernan Ramirez, my siblings Maca, Luciana and Pascal, my brother-in-law Vila, my nephew Tomi, and my niece Isa. Their love and encouragement have rooted and steadied me throughout my life. I am incredibly lucky to have landed in the home of my amazing friends Paul Krsnik and Mercedes Murphy when I moved to Santa Barbara. They will forever be one of the best things the internet has done for me. Finally, and most importantly, I am blessed with the love and companionship of my remote co-worker and soon-to-be wife Cathy Balfe, whose sense of humor, kindness, intelligence, and optimism make every moment worthwhile. I will always think back to these pivotal years in our lives in images of our apartments, our plants, the doggies, and the many equal-parts-awful-and-amazing camping trips with the Balfes, that lovely and chaotic bunch that through you is now my family.

VITA OF TADEO HERNAN RAMIREZ PARADA

December 2024

EDUCATION

Bachelor of Science, Tulane University (*cum laude, with honors*)
Majors: Ecology and Evolutionary Biology, Mathematics

Doctor of Philosophy in Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, December 2024 (expected)

PUBLICATIONS

Ramirez-Parada, T. H., Park, I. W., Record, S., Davis, C. C., Ellison, A. M., & Mazer, S. J. (2024). Plasticity and not adaptation is the primary source of temperature-mediated variation in flowering phenology in North America. *Nature Ecology & Evolution*, *8*(3), 467–476. https://doi.org/10.1038/s41559-023-02304-5

Peng, S**., Ramirez-Parada, T. H**., Mazer, S. J., Record, S., Park, I., Ellison, A. M., & Davis, C. C. (2024). Incorporating plant phenological responses into species distribution models reduces estimates of future species loss and turnover. *New Phytologist*, *242*(5), 2338–2352. https://doi.org/10.1111/nph.19698

Park, I. W., **Ramirez-Parada, T**., Record, S., Davis, C., Ellison, A. M., & Mazer, S. J. (2024). Herbarium data accurately predict the timing and duration of population-level flowering displays. *Ecography*, e06961. https://doi.org/10.1111/ecog.06961

Larios, E., **Ramirez-Parada, T. H.**, & Mazer, S. J. (2023). Heritability and variance components of seed size in wild species: Influences of breeding design and the number of genotypes tested. *Heredity*, *130*(4), 251–258. https://doi.org/10.1038/s41437-023-00597-7

Ramirez-Parada, T. H., Park, I. W., & Mazer, S. J. (2022). Herbarium specimens provide reliable estimates of phenological responses to climate at unparalleled taxonomic and spatiotemporal scales. *Ecography*, *2022*(10), e06173. https://doi.org/10.1111/ecog.06173

Mazer, S. J., Love, N. L. R., Park, I. W., **Ramirez-Parada, T.**, & Matthews, E. R. (2021). Phenological sensitivities to climate are similar in two *Clarkia* congeners: Indirect evidence for facilitation, convergence, niche conservatism or genetic constraints. *Madroño*, *68*(4). https://doi.org/10.3120/0024-9637-68.4.388

Park, I. W., **Ramirez-Parada, T.**, & Mazer, S. J. (2021). Advancing frost dates have reduced frost risk among most North American angiosperms since 1980. *Global Change Biology*, *27*(1), 165–176. https://doi.org/10.1111/gcb.15380

Pearson, K. D., Love, N. L. R., **Ramirez-Parada, T**., Mazer, S. J., & Yost, J. M. (2021). Phenological trends in the California poppy (*Eschscholzia californica*): Digitized

herbarium specimens reveal intraspecific variation in the sensitivity of flowering date to climate change. *Madroño*, *68*(4). https://doi.org/10.3120/0024-9637-68.4.343

Cook*, R. N., **Ramirez-Parada*, T**., Browne, L., Ellis, M., & Karubian, J. (2020). Environmental correlates of richness, community composition, and functional traits of terrestrial birds and mammals in a fragmented tropical landscape. *Landscape Ecology*, *35*(12), 2825–2841. https://doi.org/10.1007/s10980-020-01123-4

* These authors contributed equally to the study.

Ramirez-Parada, T., Cabrera, D., Diaz-Martin, Z., Browne, L., & Karubian, J. (2020). Resource-related variables drive individual variation in flowering phenology and mediate population-level flowering responses to climate in an asynchronously reproducing palm. *Biotropica*, *52*(5), 845–856. https://doi.org/10.1111/btp.12792

Rivera-Hechem, M. I., Rodríguez-Sickert, C., Guzmán, R. A., **Ramírez-Parada, T**., Benavides, F., Landaeta-Torres, V., Aspé-Sánchez, M., & Repetto, G. M. (2020). No association between genetic variants in MAOA, OXTR, and AVPR1a and cooperative strategies. *PLOS ONE*, *15*(12), e0244189. https://doi.org/10.1371/journal.pone.0244189

AWARDS

2019 – 2024	Doctoral UC Chancellor Fellow (UCSB)
2013 – 2017	Dean's Honor Scholar (Tulane)
2017	Senior Scholar in Ecology and Evolution (Tulane)
2013 – 2017	Honors Program (Tulane)

ABSTRACT

Patterns, processes, and emergent properties of flowering responses to climate across spatial and ecological scales

by

Tadeo Hernan Ramirez Parada

Climate change can reshape the flowering season by shifting species distributions and phenology, leading to cascading ecological impacts through changes in temporal and spatial synchrony among species. However, data limitations and incomplete knowledge of species-level flowering responses to climate within and among communities have hindered our understanding of how environmental changes impact the structure of the flowering season across biomes. Here, I used millions of herbarium specimens and community-science records to assess: i) whether variation in flowering time reflects local adaptation to long-term climatic conditions or rapid plastic responses, and ii) how species distribution shifts and plastic phenological responses scale to alter the flowering season across North American biomes. I first validated estimates of phenology-climate relationships from herbarium specimens against field observations, confirming their reliability at vast taxonomic and spatiotemporal scales. Then, I measured flowering time sensitivity to temperature for 1,605 well-sampled species across the United States to infer how adaptation and plasticity historically influenced phenology along temperature gradients. Results suggest plasticity is the primary driver of temperature-mediated variation in flowering phenology, with adaptation playing a more context-dependent role. Simulations showed that interspecific variation in

plastic responses to temperature in flowering onset and duration could significantly redistribute floral resources and restructure patterns of flowering synchrony among co-occurring species. To evaluate this, I modeled the distribution and flowering phenology of 2,837 species across the U.S. under historical, current, and projected climate and land cover conditions, scaling duration, and termination responses from species to communities and from local to continental levels. Within species, flowering onset, duration, and termination responded differently to climate, with substantial variation in sensitivity among species both within and between communities. At the community level, climate change altered species composition and the timing and duration of flowering seasons, with ecoregion-specific shifts in co-flowering species diversity and flowering overlap networks that are projected to intensify with ongoing climate and land use changes. Together, these studies demonstrate that climate change is reassembling the flowering season across North America through plastic species-level responses, with sharp differences in the severity of changes across biomes. In doing so, these analyses highlight the emerging opportunities afforded by community-science records to study the impacts of climate change on plant phenology at unprecedented taxonomic, spatiotemporal, and ecological scales.

Table of Contents

I. Dissertation overview **1**

A. Motivation and knowledge gaps 1

B. Overview of chapters 4

C. Literature cited 8

II. Herbarium specimens provide reliable estimates of phenological responses to climate at unparalleled taxonomic and spatiotemporal scales **11**

A. Abstract 11

B. Introduction 12

C. Methods 16

D. Results 23

E. Discussion 27

F. Literature cited 31

III. Plasticity and not adaptation is the primary source of temperature-mediated variation in flowering phenology in North America **37**

A. Abstract 37

B. Introduction 38

C. Methods 43

D. Results 54

E. Discussion 59

F. Literature cited 66

IV. Interspecific variation in onset and duration sensitivity mediates flowering reassembly under warming **73**

A. Abstract 73

B. Introduction 74

C. Methods 87

D. Results 97

E. Discussion 105

F. Literature cited 113

V. Climate change restructures the flowering season across North America but effects vary among ecoregions **118**

A. Abstract 118

B. Introduction 119

C. Methods....122
D. Results....136
E. Discussion....150
F. Literature cited....156
VI. Appendix 1—Supplementary materials for **II**....**167**
VII. Appendix 2—Supplementary materials for **III**....**172**
VIII. Appendix 3—Supplementary materials for **IV**....**220**
IX. Appendix 4—Supplementary materials for **V**....**232**

I. Dissertation overview: motivation and objectives

A. Motivation and knowledge gaps

The timing of plant life-cycle events (i.e., phenology) is a key determinant of the structure of terrestrial ecosystems (Chuine 2010, Richardson et al. 2013). Phenology determines the abiotic conditions that individuals, or their seasonally produced organs, encounter throughout development and their temporal overlap with mutualists and antagonists, profoundly impacting their fitness (Elzinga et al. 2007). The phenology of co-occurring plant species also defines the temporal availability and abundance of leaves, flowers, and fruits within a community, with consequences for a diverse array of organisms that depend on plant resources for growth and reproduction (Iler et al. 2021). Consequently, understanding the patterns and drivers of spatiotemporal phenological variation across species, communities, and landscapes is essential for forecasting the potential demographic and ecological consequences of global change on terrestrial ecosystems (Wolkovich & Donahue 2021).

In plants, phenological variation over space and time is intimately linked to climate. Phenology is usually cued by seasonally and interannually variable climatic factors that enable individuals to adjust growth and reproduction to fluctuating environmental conditions (Forrest & Miller-Rushing 2010). In many species, phenology also varies intraspecifically due to evolutionary adaptation to local environments, which may select for different mean phenological timings among populations in space or within populations over time (Franks & Weis 2007, Geinapp et al. 2008, Hoffmann & Sgrò 2011, Wu & Colautti 2022). Accordingly, a trend towards earlier flowering and leaf out in response to warming in recent decades has demonstrated that phenology is highly sensitive to ongoing climatic trends, foreshadowing

alterations to species distributions (Chuine 2010), trophic interactions (Renner and Zohner 2018), species persistence (Cleland et al. 2012) and community structure (Miller-Rushing et al. 2008).

Although a vast body of research has firmly established the pervasiveness and variability of phenological shifts across species and biomes (Parmesan & Yohe 2003, Menzel et al. 2020), several knowledge gaps limit our understanding of historical and future phenological shifts due to climatic change. For example, most research on phenology-climate relationships has focused on relatively few species within temperate biomes in North America and Europe, limiting the taxonomic and biogeographic scope of documented patterns of phenological sensitivity and change (Wolkovich et al. 2014, Willis et al. 2017, Ettinger et al. 2021). Moreover, even across well-studied regions, phenological sensitivity to climate varies widely among plant assemblages and taxa (Cook et al. 2012, Park 2014, Menzel et al. 2020, limiting our ability to extrapolate documented patterns to unstudied systems. Therefore, predicting plant phenological responses to climate change and their impacts from species to communities, landscapes and biomes will require significant increases in the geographic and taxonomic coverage of phenoclimatic analysis.

The degree to which different biological processes drive intraspecific spatiotemporal variation in phenology is also largely unresolved. Specifically, both phenotypic plasticity and adaptation can drive phenological variation within a species over space and time (Anderson et al. 2012). However, their relative contributions rarely have been measured within the same system because doing so typically requires extensive experiments or genetic sampling in space and time (Chevin et al. 2010, Franks 2011, Anderson et al 2012, but see Wu & Colautti 2022). Accordingly, most studies have evaluated either plasticity or adaptation as drivers of

phenology and their relative importance across species and ecological contexts remains unresolved (Merilä & Hendry 2014). Elucidating the degree to which species have phenologically adjusted to historical climatic conditions through plasticity or adaptation could help to predict whether organismal plastic responses may be sufficient—or evolutionary change necessary—to maintain synchrony between plant development and climate in a warming world (Fox et al. 2019).

While variation among species in flowering onset, budburst, and leafout responses to climate has been extensively documented, how such patterns will scale up to alter the structure of the growing or flowering season within and among communities remains unresolved. The flowering season, for example, is the cumulative outcome—and an emergent property—of the population-level flowering patterns of its angiosperm species. Consequently, changes to a community's flowering season due to warming depend on how the flowering periods of co-occurring species—both their onset and duration— respond to temperature. The few studies that have assessed community-level changes to the flowering season have found temperature responses in season length differing in magnitude and direction (e.g., Diez et al. 2012, Caradonna et al. 2014, Jabis et al. 2020, Nam and Kim 2020, Chen et al. 2022, Zhou et al. 2022). However, temperature responses for more granular (and ecologically critical) characteristics of the season—such as the temporal distribution of floral diversity and abundance—remain understudied. Moreover, the temperature sensitivity of the termination (and therefore the duration) of flowering has been less intensively examined than that of flowering onset, limiting our ability to predict the temperature responses of entire flowering periods (Pearse et al. 2017). Unfortunately, addressing these gaps empirically is difficult because long-term phenological datasets that provide the flowering periods of

populations across enough species to assess community-level change over time are exceedingly rare (Willis et al. 2017, Ettinger et al. 2021).

B. Overview of chapters

This dissertation leveraged the vast taxonomic and spatiotemporal scope of herbarium collections—as well as computer simulations—to address these knowledge gaps in the context of flowering phenology. Herbarium specimens capture snapshots of the reproductive status of individuals in space and time, and with hundreds of millions of records worldwide, provide unique opportunities to study phenology-climate relationships at large spatiotemporal and taxonomic scales (Willis et al. 2017, Meineke et al. 2018). In recent years, researchers have increasingly leveraged specimens to study phenological responses to climate (Jones and Daehler 2018, Heberling et al. 2019), estimating responsiveness for thousands of species (Park and Mazer 2018) and generating results qualitatively consistent with those from field studies (Calinger et al. 2013).

However promising, the use of herbarium specimens in phenoclimatic research is relatively recent, and studies designed to validate herbarium-based estimates of phenology and its sensitivity to climate using field observations are few and limited in scope. Accordingly, Chapter II evaluated whether herbarium specimens can generate reliable estimates of phenological responsiveness by contrasting inferences derived from field observations from the U.S.A. National Phenology Network (USA-NPN) against those obtained from herbarium data. This chapter shows that estimates of flowering time variation due to climatic variation over space and time obtained from both data sources closely agree

in both magnitude and direction. In doing so, it demonstrates that herbarium-based sensitivity estimates are reliable among species spanning a wide diversity of life histories and biomes, highlighting their utility in a broad range of ecological contexts, and underscoring the potential of herbarium collections to enable phenoclimatic analysis at taxonomic and spatiotemporal scales not yet captured by observational data.

In Chapter III, I used over one million herbarium specimens from 1,605 species across the conterminous United States (CONUS) to examine how adaptation and plasticity influence phenology along temperature gradients and how their contributions vary among species with different phenological niches and native climates. Phillimore et al. (2010) proposed that the contributions of plasticity and local adaptation to spatial variation in phenology within a species can be estimated from the difference between the slopes of spatial and temporal phenology-climate relationships. Using this framework, I measured flowering time sensitivity to temperature variation over time and space to infer how adaptation and plasticity influence phenology along temperature gradients and how their contributions vary among species with different phenological niches and native climates. These analyses revealed that, across ecoregions, plasticity is the predominant driver of temperature variation over space and time, with adaptation playing a widespread but comparatively limited role. My findings suggest that plasticity historically has enabled flowering phenology to respond quickly to a wide range of temperature conditions among North American angiosperms, with adaptation playing a comparatively context-dependent role.

Chapter IV used computer simulations to provide a conceptual overview of how variation in the sensitivity of both flowering onset and duration among species can influence changes in a community's flowering season under warming. I focused on the effects of i)

differences in the mean sensitivity of flowering onset and duration among communities, and ii) sensitivity variation in species flowering sequentially through the season within a community. Simulations, based on a wide and empirically informed range of parameter values and combinations, examined how these forms of between-species sensitivity might independently and interactively affect the flowering season's structure. The analyses revealed that documented patterns of variation in onset and duration sensitivities can lead to diverse flowering season changes across biomes, emphasizing the need to account for both factors to fully understand community-level flowering dynamics in response to warming.

Finally, chapter V offers an empirical counterpart to the conceptual outline from chapter IV, and expands it outlook to assess how both changes in species distributions and flowering responses to climate change among species scale to restructure the flowering season across North America. Using millions of herbarium and occurrence records from community-science initiatives (e.g., iNaturalist), I modeled the distribution and flowering phenology of 2,837 species across the United States under historical, current, and projected climate and land cover conditions, scaling onset, duration, and termination responses from species to communities, and from local to continental levels. These analyses demonstrate that—within species—the onset, duration, and end of flowering respond differently to temperature and precipitation, with broad variation in mean sensitivity among and within communities. Climate change has resulted in widespread shifts in species composition and the timing and length of the flowering season across ecoregions, leading to region-specific changes in the seasonal distribution of flowering diversity and patterns of flowering overlap that are projected to intensify under ongoing climate trends. These findings outline broad

macroecological changes, revealing uneven impacts of climate change on the identity, diversity, and flowering patterns of co-occurring species across biomes.

Together, these projects demonstrate the enormous potential of herbarium collection and community-science record to expand the taxonomic and spatiotemporal scale of phenoclimatic analysis from species to communities and from local to continental scales.

C. Literature cited

Calinger, K. M., Queenborough, S., & Curtis, P. S. (2013). Herbarium specimens reveal the footprint of climate change on flowering trends across north-central North America. *Ecology Letters*, *16*(8), 1037–1044. https://doi.org/10.1111/ele.12135

Chen, Y., Collins, S. L., Zhao, Y., Zhang, T., Yang, X., An, H., Hu, G., Xin, C., Zhou, J., Sheng, X., He, M., Zhang, P., Guo, Z., Zhang, H., Li, L., & Ma, M. (2022). Warming reduced flowering synchrony and extended community flowering season in an alpine meadow on the Tibetan Plateau. *Ecology*, *n/a*(n/a), e3862. https://doi.org/10.1002/ecy.3862

Chuine, I. (2010). Why does phenology drive species distribution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1555), 3149–3160. https://doi.org/10.1098/rstb.2010.0142

Cleland, E. E., Allen, J. M., Crimmins, T. M., Dunne, J. A., Pau, S., Travers, S. E., Zavaleta, E. S., & Wolkovich, E. M. (2012). Phenological tracking enables positive species responses to climate change. *Ecology*, *93*(8), 1765–1771. https://doi.org/10.1890/11-1912.1

Cook, B. I., Wolkovich, E. M., Davies, T. J., Ault, T. R., Betancourt, J. L., Allen, J. M., Bolmgren, K., Cleland, E. E., Crimmins, T. M., Kraft, N. J. B., Lancaster, L. T., Mazer, S. J., McCabe, G. J., McGill, B. J., Parmesan, C., Pau, S., Regetz, J., Salamin, N., Schwartz, M. D., & Travers, S. E. (2012). Sensitivity of Spring Phenology to Warming Across Temporal and Spatial Climate Gradients in Two Independent Databases. *Ecosystems*, *15*(8), 1283–1294. https://doi.org/10.1007/s10021-012-9584-5

Diez, J. M., Ibáñez, I., Miller-Rushing, A. J., Mazer, S. J., Crimmins, T. M., Crimmins, M. A., Bertelsen, C. D., & Inouye, D. W. (2012). Forecasting phenology: From species variability to community patterns. *Ecology Letters*, *15*(6), 545–553. https://doi.org/10.1111/j.1461-0248.2012.01765.x

Elzinga, J. A., Atlan, A., Biere, A., Gigord, L., Weis, A. E., & Bernasconi, G. (2007). Time after time: Flowering phenology and biotic interactions. *Trends in Ecology & Evolution*, *22*(8), 432–439. https://doi.org/10.1016/j.tree.2007.05.006

Ettinger, A. K., Buonaiuto, D. M., Chamberlain, C. J., Morales-Castilla, I., & Wolkovich, E. M. (2021). Spatial and temporal shifts in photoperiod with climate change. *New Phytologist*, *230*(2), 462–474. https://doi.org/10.1111/nph.17172

Fox, R. J., Donelson, J. M., Schunter, C., Ravasi, T., & Gaitán-Espitia, J. D. (2019). Beyond buying time: The role of plasticity in phenotypic adaptation to rapid environmental change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1768), 20180174. https://doi.org/10.1098/rstb.2018.0174

Franks, S. J., Sim, S., & Weis, A. E. (2007). Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences*, *104*(4), 1278–1282. https://doi.org/10.1073/pnas.0608379104

Franks, S. J. (2011). Plasticity and evolution in drought avoidance and escape in the annual plant *Brassica rapa*. *New Phytologist*, *190*(1), 249–257. https://doi.org/10.1111/j.1469-8137.2010.03603.x

Heberling, J. M., Prather, L. A., & Tonsor, S. J. (2019). The Changing Uses of Herbarium Data in an Era of Global Change: An Overview Using Automated Content Analysis. *BioScience*, *69*(10), 812–822. https://doi.org/10.1093/biosci/biz094

Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*(7335), Article 7335. https://doi.org/10.1038/nature09670

Jabis, M. D., Winkler, D. E., & Kueppers, L. M. (2020). Warming acts through earlier snowmelt to advance but not extend alpine community flowering. *Ecology*, *101*(9), e03108. https://doi.org/10.1002/ecy.3108

Jones, C. A., & Daehler, C. C. (2018). Herbarium specimens can reveal impacts of climate change on plant phenology; a review of methods and applications. *PeerJ*, *6*, e4576. https://doi.org/10.7717/peerj.4576

Li, D., Barve, N., Brenskelle, L., Earl, K., Barve, V., Belitz, M. W., Doby, J., Hantak, M. M., Oswald, J. A., Stucky, B. J., Walters, M., & Guralnick, R. P. (2021). Climate, urbanization, and species traits interactively drive flowering duration. *Global Change Biology*, *27*(4), 892–903. https://doi.org/10.1111/gcb.15461

Meineke, E. K., Classen, A. T., Sanders, N. J., & Jonathan Davies, T. (2019). Herbarium specimens reveal increasing herbivory over the past century. *Journal of Ecology*, *107*(1), 105–117. https://doi.org/10.1111/1365-2745.13057

Menzel, A., Yuan, Y., Matiu, M., Sparks, T., Scheifinger, H., Gehrig, R., & Estrella, N. (2020). Climate change fingerprints in recent European plant phenology. *Global Change Biology*, *26*(4), 2599–2612. https://doi.org/10.1111/gcb.15000

Miller-Rushing, A. J., & Primack, R. B. (2008). Global warming and flowering times in Thoreau's Concord: a community perspective. *Ecology*, *89*(2), 332–341. https://doi.org/10.1890/07-0068.1

Nam, B. E., & Kim, J. G. (2020). Flowering season of vernal herbs is shortened at elevated temperatures with reduced precipitation in early spring. *Scientific Reports*, *10*(1), 17494. https://doi.org/10.1038/s41598-020-74566-z

Park, D. S., Breckheimer, I., Williams, A. C., Law, E., Ellison, A. M., & Davis, C. C. (2019). Herbarium specimens reveal substantial and unexpected variation in phenological sensitivity across the eastern United States. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1763), 20170394. https://doi.org/10.1098/rstb.2017.0394

Park, I. W. (2014). Impacts of differing community composition on flowering phenology throughout warm temperate, cool temperate and xeric environments. *Global Ecology and Biogeography*, *23*(7), 789–801. https://doi.org/10.1111/geb.12163

Park, I. W., & Mazer, S. J. (2018). Overlooked climate parameters best predict flowering onset: Assessing phenological models using the elastic net. *Global Change Biology*, *24*(12), 5972–5984. https://doi.org/10.1111/gcb.14447

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*(6918), Article 6918. https://doi.org/10.1038/nature01286

Pearse, W. D., Davis, C. C., Inouye, D. W., Primack, R. B., & Davies, T. J. (2017). A statistical estimator for determining the limits of contemporary and historic phenology. *Nature Ecology & Evolution*, *1*(12), 1876–1882. https://doi.org/10.1038/s41559-017-0350-0

Prevéy, J. S., Rixen, C., Rüger, N., Høye, T. T., Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., Ashton, I. W., Cannone, N., Chisholm, C. L., Clark, K., Cooper, E. J., Elberling, B., Fosaa, A. M., Henry, G. H. R., Hollister, R. D., Jónsdóttir, I. S., Klanderud, K., Kopp, C. W., … Wipf, S. (2019). Warming shortens flowering seasons of tundra plant communities. *Nature Ecology & Evolution*, *3*(1), 45–52. https://doi.org/10.1038/s41559-018-0745-6

Renner, S. S., & Zohner, C. M. (2018). Climate Change and Phenological Mismatch in Trophic Interactions Among Plants, Insects, and Vertebrates. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 165–182. https://doi.org/10.1146/annurev-ecolsys-110617-062535

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Willis, C. G., Ellwood, E. R., Primack, R. B., Davis, C. C., Pearson, K. D., Gallinat, A. S., Yost, J. M., Nelson, G., Mazer, S. J., Rossington, N. L., Sparks, T. H., & Soltis, P. S. (2017). Old Plants, New Tricks: Phenological Research Using Herbarium Specimens. *Trends in Ecology & Evolution*, *32*(7), 531–546. https://doi.org/10.1016/j.tree.2017.03.015

Wolkovich, E. M., Cook, B. I., & Davies, T. J. (2014). Progress towards an interdisciplinary science of plant phenology: Building predictions across space, time and species diversity. *New Phytologist*, *201*(4), 1156–1162. https://doi.org/10.1111/nph.12599

Wolkovich, E. M., & Donahue, M. J. (2021). How phenological tracking shapes species and communities in non-stationary environments. *Biological Reviews*, *96*(6), 2810–2827. https://doi.org/10.1111/brv.12781

Wu, Y., & Colautti, R. I. (2022). Evidence for continent-wide convergent evolution and stasis throughout 150 y of a biological invasion. *Proceedings of the National Academy of Sciences*, *119*(18), e2107584119. https://doi.org/10.1073/pnas.2107584119

Xie, Y., Thammavong, H. T., & Park, D. S. (2022). The ecological implications of intra- and inter-species variation in phenological sensitivity. *New Phytologist*, *236*(2), 760–773. https://doi.org/10.1111/nph.18361

II. Herbarium specimens provide reliable estimates of phenological responses to climate at unparalleled taxonomic and spatiotemporal scales[1]

A. Abstract

Understanding the effects of climate change on the phenological structure of plant communities will require measuring variation in sensitivity among thousands of co-occurring species across regions. Herbarium collections provide vast resources with which to do this, but may also exhibit biases as sources of phenological data. Despite general recognition of these caveats, validation of herbarium-based estimates of phenological sensitivity against estimates obtained using field observations remain rare and limited in scope. Here, I leveraged extensive datasets of herbarium specimens and of field observations from the USA National Phenology Network for 21 species in the United States and, for each species, compared herbarium- and field-based estimates of peak flowering dates expected under standardized temperature conditions, and of sensitivity of peak flowering time to geographic and interannual variation in mean spring minimum temperatures (TMIN). I found strong agreement between herbarium- and field-based estimates for standardized peak flowering time ($r=0.91$, $p<0.001$) and for the direction and magnitude of sensitivity to both geographic TMIN variation ($r=0.88$, $p<0.001$) and interannual TMIN variation ($r=0.82$, $p<0.001$). This agreement was robust to substantial differences between datasets in 1) the long-term TMIN conditions observed among collection and phenological monitoring sites and 2) the

[1] This chapter has been published in the following peer-reviewed article:
Ramirez-Parada, T. H., Park, I. W., & Mazer, S. J. (2022). Herbarium specimens provide reliable estimates of phenological responses to climate at unparalleled taxonomic and spatiotemporal scales. *Ecography*, *2022*(10), e06173. https://doi.org/10.1111/ecog.06173

interannual TMIN conditions observed in the time periods encompassed by both datasets for most species. These results show that herbarium-based sensitivity estimates are reliable among species spanning a wide diversity of life histories and biomes, demonstrating their utility in a broad range of ecological contexts, and underscoring the potential of herbarium collections to enable phenoclimatic analysis at taxonomic and spatiotemporal scales not yet captured by observational data.

B. Introduction

Widespread shifts in plant phenology (i.e., the timing of life cycle events) due to climate change have the potential to significantly alter species distributions (Chuine 2010), trophic interactions (Renner & Zhoner 2018), species persistence (Cleland et al. 2012), and community structure (Willis et al. 2008). A trend towards earlier flowering and leaf out in response to warming has demonstrated that phenology is highly sensitive to climate variation, but sensitivity varies widely among regions and taxa (Cook et al. 2012, Park 2014, Menzel et al. 2020), and even within species (Song et al. 2020, Love & Mazer 2021, Pearson et al. 2021), limiting our ability to extrapolate documented patterns to unstudied systems. Therefore, predicting plant phenological responses to climate change and their impact across communities, landscapes, and biomes will require significant increases to the geographic and taxonomic coverage of phenoclimatic analysis.

Regular field observations of individual plants allow precise records of the date of phenological events and are the gold-standard for the study of phenology-climate relationships. However, observational datasets spanning enough time to permit detection of

phenological shifts are scarce and predominantly consist of phenological records from North America and Western Europe (Cook et al. 2012, Templ et al. 2018), limiting their utility in assessing phenology-climate relationships across many unstudied taxa and biomes (Wolkovich et al. 2014, Tang et al. 2016). Moreover, field-based time series of phenology are usually available only at single sites for most species, constraining estimation of phenological responses to climate to small subsets of their ranges.

In contrast, herbarium specimens capture snapshots of the reproductive status of individual plants in space and time, and with hundreds of millions of records worldwide increasingly available digitally, provide unique opportunities to expand the taxonomic and spatiotemporal coverage of phenoclimatic studies (Willis et al. 2017, Meineke et al. 2018). In recent years, researchers have leveraged specimens to study phenology-climate relationships (Jones & Dahler 2018, Heberling et al. 2019), estimating phenological responsiveness for thousands of species (e.g., Park & Mazer 2018) and generating results qualitatively consistent with those from field studies (Calinger et al. 2013). However, potential biases in collection practices could yield inaccurate estimates of a species' phenology and its sensitivity to climate. For example, while field observations can pinpoint the timing of a phenological with known degrees of uncertainty, herbarium specimens may have been collected anytime between the onset and termination of a phenophase or botanists may preferentially collect individuals in specific phenophases (e.g., peak flowering, Panchen et al. 2019), potentially compromising collection dates as reliable proxies for the dates of phenological events, especially the onset and termination of a phenophase. Additionally, the opportunistic collection of specimens could result in sampling of early or late flowering individuals that may not accurately reflect the phenological behavior of their populations.

Despite these caveats, studies designed to validate herbarium-based estimates of phenology and its sensitivity to climate using field observations are few and limited in scope. Most validation studies have been restricted to areas with long records of field observations and specimen collections covering a small portion of species' ranges (Miller-Rushing et al. 2006, Robbirt et al. 2011, Davis et al. 2015). In turn, the only studies comparing herbarium- and field-based phenological records at large spatial scales have not aimed to validate phenological sensitivity estimates (Spellman and Mulder 2016, Park & Mazer 2018). Some studies have compared herbarium- vs. field-based estimates of sensitivity for a single species (Robbirt et al. 2011), or conducted pooled, multi-species analyses that do not enable validation of estimates for individual species (Miller-Rushing et al. 2006, Park 2012). As an exception, Davis et al. (2015) used herbarium and field data for 20 species collected in Middlesex County (Massachusetts, USA), finding overall agreement between data for the direction of phenological responses to spring temperature variation; however, sensitivity estimates derived from the two sources tended to differ in magnitude and were not positively correlated among species. Collectively, these studies have shown herbarium specimens are promising data sources for phenoclimatic analysis, but their limited scope and the mismatch in estimates between data types in Davis et al. (2015) make it difficult to establish to what extent specimens may represent generally valid resources for the study of phenology-climate relationships.

This study provides a multi-species comparison of herbarium- and field-based estimates of peak flowering sensitivity to spatiotemporal variation in mean minimum temperatures (TMIN). I used two geographically extensive datasets obtained from herbaria across the United States and from field observations aggregated by the USA National

Phenology Network (USA-NPN, hereafter, 'NPN'; Schwartz et al. 2012). These data included a total of 21 species spanning diverse life histories and biomes and included phenological observations across thousands of unique site-year combinations throughout the United States. These data substantially exceed the sample sizes of previous validation studies and enabled us to compare herbarium- and field-based estimates of sensitivity to climate variation over both space and time and across a broad range of ecological contexts.

I measured peak flowering time sensitivities to both geographic and interannual variation in TMIN because, among conspecifics distributed across large geographic scales, associations between phenology and climate might be driven both by phenotypic plasticity and by local adaptation to long-term climatic conditions among populations (Anderson et al. 2012). While associations between phenology and interannual climate variation are thought to predominantly reflect plastic responses, correlations between phenology and long-term, mean climatic conditions over space may be strongly influenced by local adaptation across populations (Delgado et al. 2020). To the extent that phenology-climate relationships over space and time have different drivers, they may also differ in magnitude or direction. Therefore, I leveraged the spatiotemporal scale of these datasets to partition observed variation in temperature across sites and years into interannual and geographic components, comparing herbarium- and field-based estimates of sensitivity to both sources of temperature variation in all examined species. In doing so, this study provides the first concurrent validation of herbarium-based estimates of phenological sensitivity to spatial and temporal variation in temperature.

C. Methods

Phenological data

Field observations consisted of all records of flowering onset and termination available in the USA National Phenology Network database, representing an initial 1,105,764 phenological observations. To ensure data quality, I retained only observations for which flowering onset and termination dates had an arbitrary maximum error of 14 days. Accordingly, I filtered the data to include only records for which the date of flowering onset was preceded by an observation of the same individual without flowers no more than 14 days prior, and for which the flowering termination date was followed by an observation of the same individual without flowers no more than 14 days later. The remaining field observations had an average maximum error of 6.4 days for flowering onset, and of 6.6 days for flowering termination.

Herbarium data consisted of an initial 894,392 digital specimen records archived by 72 herbaria across North America (see Appendix 1—Note S1 for a list). I removed all specimens not explicitly recorded as being in flower, or for which GPS coordinates or dates of collection were not available. I further filtered both datasets by only retaining species that were found in both datasets and that were represented by observations at an arbitrary minimum of 15 unique sites in both datasets. To better align the geographic range of each dataset for each species, I filtered herbarium observations to include only specimens within the range of latitudes and longitudes represented among field observations in the NPN data. Finally, I retained only species represented by 70 or more herbarium specimens to ensure sufficient sample sizes for phenoclimatic modelling (Park & Mazer 2018). This procedure identified a final set of 21 native species represented in 3,243 field observations across 1,406 unique site-year combinations, and a final sample of 5,405 herbarium specimens across 4,906

unique site-year combinations (Fig. 1). These species represented 15 families and 17 genera, spanning a diverse range of life history strategies and growth forms, including evergreen and deciduous shrubs and trees, as well as herbaceous perennials and annuals. The study's focal species covered a wide variety of biomes and regions including Western deserts, Mediterranean shrublands, oak woodlands, and Eastern deciduous forests (Appendix 1—Table S1).

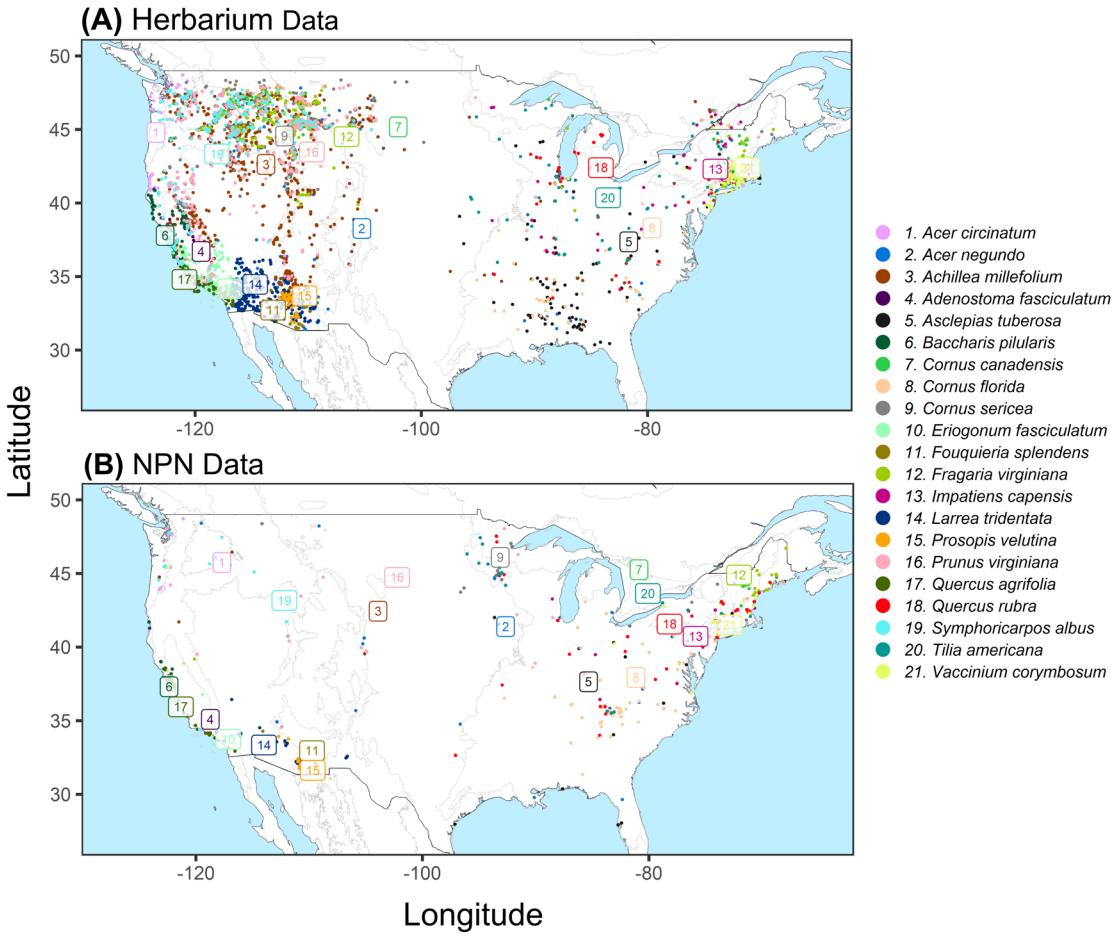

Figure 1 – Geographic distribution of herbarium specimens (A) and field observations from the USA-NPN (B) for 21 species in the continental United States. Numbered labels represent the centroid of the spatial distribution of observations for each species, obtained by calculating the average latitude and longitude among data points for each species.

I employed the day of year of collection (henceforth 'DOY') of each specimen collected while in flower as a proxy of flowering dates. Flowering specimens could have been collected at any point between onset and termination and botanists may preferentially collect individuals in their flowering peak for many species (Panchen et al. 2019). Therefore, specimen DOYs are more likely to reflect peak flowering dates than onset or termination dates (Primack et al. 2004). To increase the phenological equivalence of field and herbarium observations, I used the median date between flowering onset and termination for each observation in the NPN data as a proxy for peak flowering time. Median flowering dates also had a maximum error of 14 days, with an average maximum error among observations of 6.5 days. Because flowering spanned year ends for some species (e.g., *Quercus agrifolia*), I accounted for the artificial DOY discontinuity between December 31st (DOY = 365-366) to January 1st (DOY = 1) by converting DOY into a circular variable using an Azimuthal correction (e.g., Park & Mazer 2018).

Climate data

Daily minimum temperatures mediate key developmental processes including the break of dormancy, floral induction, and anthesis (Reeves & Coupland 2000). Therefore, I used minimum surface temperatures averaged over the three months leading up to (and including) the mean flowering month for each species (hereafter 'TMIN') as the climatic correlate of flowering time in this study; consequently, the specific months over which temperatures were averaged varied among species. Using TMIN calculated over different time periods instead (e.g., during spring for all species) did not qualitatively affect the results. Then, I partitioned variation among sites into spatial and temporal components, characterizing TMIN for each

observation by the long-term mean TMIN at its site of collection (henceforth 'TMIN normals'), and by the deviation between its TMIN in the year of collection (for the three-month window of interest) and its long-term mean TMIN (henceforth 'TMIN anomalies') (Appendix 1—Fig. S1; see Munson & Long 2017 for an example of this approach).

For each site, I obtained a monthly time series of TMIN from January, 1901, and December, 2016, using ClimateNA v6.30 (Wang et al. 2016), a software package that interpolates $4km^2$ resolution climate data from PRISM (PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu) to generate elevation-adjusted climate estimates. To calculate TMIN normals, I averaged observed TMIN for the three months leading up to the mean flowering date of each species across all years between 1901 and 2016 for each site. TMIN anomalies relative to long-term conditions were calculated by subtracting TMIN normals from observed TMIN conditions in the year of collection. Therefore, positive and negative values of the anomalies respectively reflect warmer-than-average and colder-than-average conditions in a given year (Appendix 1—Fig. S1).

Pooling across species, herbarium records showed slightly cooler TMIN normals than did NPN field observations, and spanned a wider envelope encompassing warmer and cooler long-term conditions in the months leading up to mean flowering dates (Fig. 2A). Specimen collection dates spanned a long period (1901-2016) largely preceding the onset of rapid warming trends while NPN observations were all conducted in recent years (2009-2020). Consequently, TMIN anomalies in the NPN dataset encompassed warmer conditions than those in the herbarium dataset both globally and for most species (Fig. 2A, C). Among species, differences between datasets in the width and median of TMIN normal and anomaly

envelopes varied substantially (Fig. 3B,C), but relative differences in TMIN envelopes among species were largely consistent in both datasets (Appendix 1—Fig. S2).

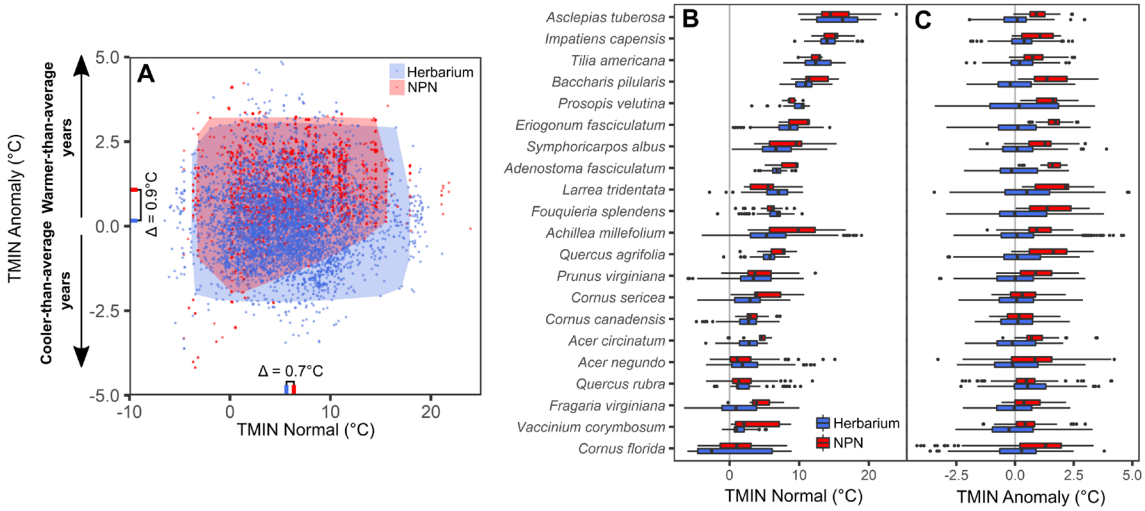

Figure 2 – 90% TMIN normal and TMIN anomaly envelopes for herbarium specimens vs. field observations for all species pooled (A), and by species (B, C). The small red and blue segments located on the x- and y-axes in (A) indicate the mean values of each climatic variable for herbarium (blue) vs. field (red) data.

Analyses

I compared estimates of sensitivity to spatiotemporal variation in TMIN derived from herbarium specimens and field observations, concurrently measuring the effects of TMIN normals and anomalies on peak flowering time for each species-by-dataset combination. I combined herbarium and field records in a single dataset, which I analyzed using a varying-intercepts, varying-slopes Bayesian mixed-effect model. The model fitted species-specific intercepts and slopes and treated them as random effects stemming from community-level distributions (defined by separate 'hyperparameters') for field and herbarium records. This

hierarchical structure improves estimation of parameters for species with low sample sizes by using community-level information and estimates from better-sampled species. In turn, the Bayesian inference framework enables direct measurement of uncertainty for all parameters.

I used peak flowering DOY for each observation *i* in the combined dataset as a response, which was assumed to be normally distributed, with mean μ_i and species-specific standard deviation σ_s (Equation 1).

Equation 1

$$DOY_i \sim N(\mu_i, \sigma_s)$$

I modeled μ_i as a linear function of TMIN normal (TMIN $Norm_i$), and TMIN anomaly (TMIN $Anom_i$) for each observation *i*. To obtain intercepts and slopes unique to each species-by-dataset combination, I used two dummy variables (with values of 0 or 1) respectively indicating whether each observation was obtained from field observations in the NPN (F_i) or from herbarium records (H_i). This resulted in the inclusion of only NPN or herbarium observations when a given parameter was estimated (i.e., model terms were turned "on and off"' depending on data type). For each data type, the model yielded species-specific intercepts representing standardized flowering dates expected under mean TMIN normal and mean TMIN anomaly conditions (herbarium: α_{1_s}; NPN: α_{2_s}), species-specific sensitivities (i.e., regression slopes) for TMIN normal (herbarium: β_{1_s}; NPN: β_{2_s}), and species-specific sensitivities for TMIN anomaly (herbarium: β_{3_s}; NPN: β_{4_s}) (Equation 2).

Equation 2

$$\mu_i = \alpha_{1_s} \times H_i + \alpha_{2_s} \times F_i + \beta_{1_s} \times TMIN\ Norm_i \times H_i + \beta_{2_s} \times TMIN\ Norm_i \times F_i +$$

$$\beta_{3_s} \times TMIN\ Anom_i \times H_i + \beta_{4_s} \times \text{TMIN}\ Anom_i \times F_i$$

Estimating peak flowering dates under standardized TMIN normal and TMIN anomaly enabled us to directly compare herbarium- and field-based predictions of peak flowering phenology. To account for co-variation among parameters (e.g., standardized flowering dates and temperature sensitivities, Mazer et al. 2013), I assumed that community-level distributions for intercepts and slopes were generated by a multivariate normal distribution with a vector of hyper-means μ and a variance-covariance matrix Σ.

Equation 3

$$(\alpha_{1_s}, \alpha_{2_s}, \beta_{1_s}, \beta_{2_s}, \beta_{3_s}, \beta_{4_s}) \sim N(\mu, \Sigma)$$

The diagonals in **Σ** correspond to community-level variances for each intercept and slope, whereas off-diagonal values correspond to the covariances between parameters among species.

Priors in the model were weakly informative, with wide, 0 centered normal distributions for intercepts, slopes, and rate parameters for exponential distributions (used to obtain species-specific variances). For the variance-covariance matrix **Σ,** I used a Lewandowski-Kurowicka-Joe (LKJ) Cholesky covariance prior, with ŋ = 1 to allow for high correlations among parameters. Posterior distributions were obtained using Hamiltonian Monte Carlo (HMC) in Stan (code provided in Appendix 1—Note S2), implemented in R Studio v1.4.1106 using the 'rstan' package v.2.21.2 (R Core Team 2020, Stan Development Team 2020). I implemented a non-centered parameterization to improve sampling of the parameter space. Sampling was done using two MCMC chains with training and sampling

periods of 1000 iterations each. All parameters had Gelman-Rubin statistics ('R-hat') values close to 1, and visual examination of trace plots confirmed convergence. Bulk and tail effective sample size were both high relative to the total number of samples.

To evaluate the correlation between herbarium- and field-derived estimates accounting for differences in the number of sampled sites between datasets, I calculated weighted Pearson correlation coefficients between Maximum *a posteriori* (MAP) estimates for field data and herbarium specimens generated by the model, using the minimum number of unique sites in the NPN or the herbarium dataset for each species as weights. Alternative weighing schemes (e.g., using total sample sizes instead) yielded nearly identical results.

I assessed whether, among species, mismatches between herbarium- vs. field-based estimates could be explained by differing climate conditions captured by each dataset. For each species, I calculated the absolute difference between herbarium- and field-based estimates of standardized flowering time and TMIN sensitivities, and the absolute differences in mean TMIN normal and mean TMIN anomaly for both datasets. Finally, I calculated weighted Pearson correlations between absolute differences in parameter estimates and absolute differences in TMIN normal and TMIN anomaly between datasets, using the number of unique sites in the NPN or the herbarium dataset for each species as weights. All p-values were adjusted for multiple hypothesis testing using a Holm-Bonferroni correction.

D. Results

I found strong correlations between herbarium- and field-derived estimates for all phenological parameters. Standardized flowering times ranged from mid Spring (early April,

Cornus florida) to late Summer (mid-September, *Baccharis pilularis*), with very high correlation between herbarium- and field-derived estimates ($r = 0.91$, $p < 0.001$; Fig. 3A). Absolute differences between herbarium- and field-based flowering dates ranged from 0 days for *Fouquieria splendens* to 64 days for *Quercus rubra* (Appendix—Table S1). Overall, estimates from both datasets differed by a mean of 14 days among species, with herbarium specimens generating estimates that were, on average, 11 days later than NPN-derived estimates across species.

Estimates of sensitivity to TMIN normals and to TMIN anomalies were consistent between data types. Field- and herbarium-based estimates of sensitivity to TMIN normal were highly correlated ($r = 0.88$, $p < 0.001$) and largely co-varied along a one-to-one line, indicating agreement in the magnitude of species-specific sensitivities (Fig. 3B). TMIN normal sensitivities agreed in direction (i.e., the sign of the slope coefficient; Equation 2) for 20 out of 21 species (95%), and the only species showing discrepancies between data types *(Asclepias tuberosa)* showed a non-significant estimate of TMIN normal sensitivity for field observations (Appendix—Table S2). On average, estimates of TMIN normal sensitivity differed by 1.5 d/°C among species between data types, with absolute differences ranging from 0.1 d/°C for *Cornus florida* and *Tilia americana* to 4.5 d/°C for *Asclepias tuberosa* (Appendix—Table S1). Collectively, herbarium-based estimates were an average of 0.1 d/°C more negative than field-based estimates.

Similarly, sensitivities to TMIN anomalies were significantly correlated between data types ($r = 0.82$, $p < 0.001$) and tended to agree in both direction and magnitude (Fig. 3C). Sensitivities to TMIN anomalies agreed in direction for 19 out of 21 species (90%), and the two species with mismatches in direction between data types *(Eriogonum fasciculatum* and

Tilia americana) had non-significant estimates that were very close to 0 for both data types (Appendix 1—Table S2, Fig. 3C). Herbarium- and field-based estimates of sensitivity to TMIN anomaly differed by an average of 1.3 d/°C among species, with absolute differences ranging from 0.0 d/°C for *Acer negundo* to 3.9 d/°C for *Fouquieria splendens* (Appendix—Table S1). Herbarium-based estimates of sensitivity to TMIN anomaly were, on average, 0.5 d/°C more positive than field-based estimates.

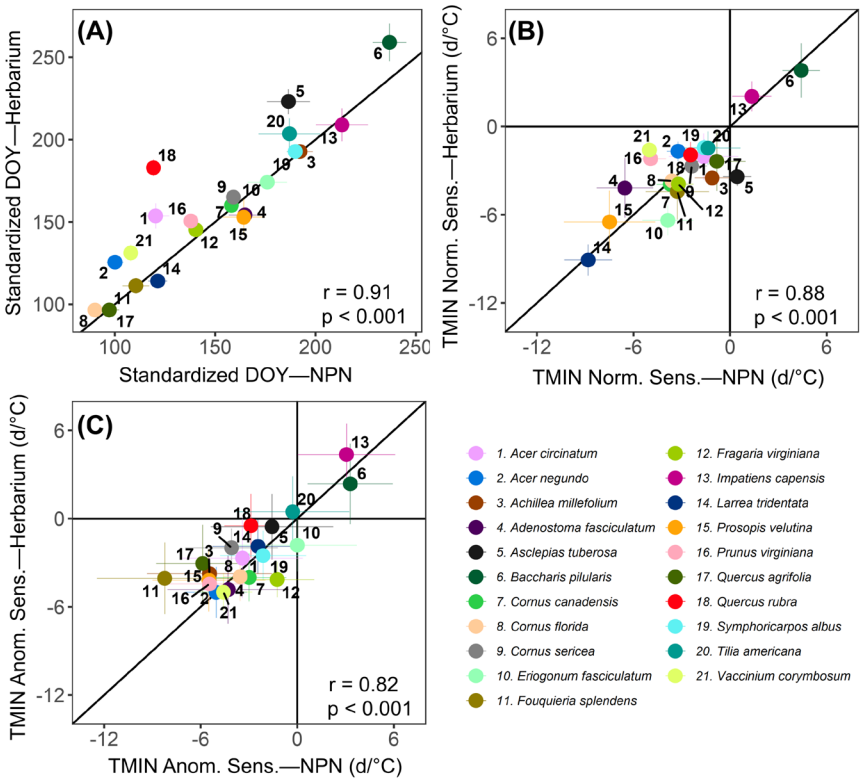

Figure 3 – Comparison of species-specific estimates derived from herbarium vs. USA-NPN field observations for standardized flowering dates (A), flowering sensitivity to TMIN normal (B), and flowering sensitivity to TMIN anomaly (C). Vertical and horizontal lines around each point correspond to the standard deviation of each species-specific parameter and each dataset, and were obtained from the posterior distribution of each coefficient. Reported correlation coefficients and p-

values were obtained using the minimum between the number of unique locations for each species in either the herbarium or the NPN dataset as weights.

Among species, absolute differences in mean TMIN normal and in TMIN anomaly between datasets were not significantly correlated to mismatches between herbarium- vs. field-derived phenological estimates (Fig. 4). While I detected a marginally significant negative relationship between mismatches in standardized flowering dates and differences in TMIN normal between datasets, which would nonsensically indicate higher agreement between herbarium-vs field-based estimates for species showing greater differences in TMIN normal (Fig. 4A, C), such relationship was driven by a single outlier, *Quercus rubra*, exhibiting the greatest mismatch in estimated flowering time between datasets (64 days), and one of the lowest absolute differences in TMIN normal among species (0.1 °C). Excluding *Q. rubra* yielded a non-significant relationship instead (p = 0.50).

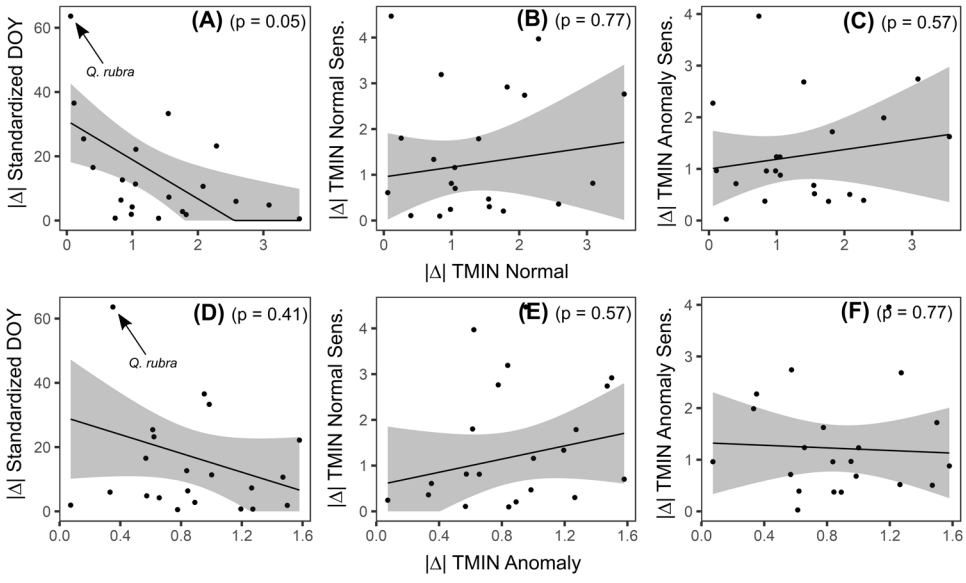

Figure 4 – Correlation among 21 species in the continental United States between absolute differences (|Δ|) between herbarium- vs. field-derived estimates of standardized DOY, of sensitivity to TMIN normals, and of sensitivity to TMIN anomalies and absolute differences in mean TMIN normal (A, B, C) and TMIN anomaly (D, E, F). *Quercus rubra* is labeled in panels showing absolute differences in standardized DOY, for which it was a clear outlier. All P values displayed were adjusted for multiple hypothesis testing using a Holm-Bonferroni correction.

E. Discussion

This study extends these results by demonstrating strong quantitative agreement between field- and herbarium-based estimates of sensitivity to climate over both space and time and across multiple species. Moreover, while Robbirt et al. (2011) focused on a single species and Davis et al. (2015) analyzed only herbaceous species with ephemeral spring and summer flowering in New England, the 21 focal species spanned a wide diversity of growth forms, life histories, and native biomes, suggesting that herbarium-based estimates of phenology-climate relationships may be reliable across a wide spectrum of ecological contexts.

Despite a strong correlation, herbarium specimens produced later standardized estimates of flowering dates than did NPN observations. For most species, herbarium specimens encompassed colder TMIN conditions than NPN observations (Fig. 2). However, while this difference would predict later flowering dates in the cooler herbarium dataset (Fig. 2), differences in TMIN normal and anomaly did not explain mismatches between datasets (Fig. 4). Herbarium specimens are predominantly collected opportunistically or during sporadic botanical expeditions, which might make them more likely to represent median rather than early or late flowering individuals within a population. In turn, NPN records are

assembled from regular visits to sites or individuals, which may result in capture of early flowering plants for herbaceous species for which monitoring the same individuals across years may not be possible, and to the extent that observers might choose to monitor large and healthy trees or shrubs (which may flower early), this could be the case for woody taxa as well. Nevertheless, estimates from both datasets showed modest differences and high correlation despite marked differences in collection periods and climatic conditions.

Similarly, herbarium- and field-based estimates of sensitivities to spatiotemporal TMIN variation overwhelmingly agreed in direction and magnitude despite differences in TMIN conditions between datasets (Fig. 2B-C; Fig. 3B-C). While recent studies shows that species can exhibit variation in phenological sensitivity among areas characterized by different long-term climatic conditions (e.g., Song et al. 2020; Love and Mazer, 2021, Pearson et al. 2021), the results suggest that such intraspecific differences might not be substantial enough to mask patterns of among-species variation in sensitivity to TMIN in this case. Similarly, plastic phenological responses to interannual climate variation can vary intraspecifically between cool and warm periods due to non-linearities in the underlying phenology-temperature relationship (Fu et al. 2015, Güsewell et al. 2017). However, the lack of associations between mismatches in TMIN conditions and in TMIN sensitivity suggests that phenology-temperature relationships among the focal species might be stable within the range of interannual variation encompassed here.

While I lacked enough taxa to test this statistically, no apparent relationships between species-level characteristics and the degree of mismatch between herbarium vs. field estimates was detected. For example, while the species that showed the greatest mismatches for different phenological parameters consisted of a mix of evergreen and deciduous woody

species (and a few herbs) from various Western and Eastern ecoregions, so did groups of species showing the smallest mismatches (Appendix 1—Table S1). Likewise, no clear taxonomic patterns in mismatches between datasets were apparent. For example, while species in the genera *Cornus* and *Quercus*, respectively, showed some of the smallest mismatches in TMIN normal sensitivity and the greatest mismatches in TMIN anomaly sensitivity, congeners *Cornus florida* and *Cornus sericea* were respectively among the species showing the smallest and greatest mismatches in TMIN anomaly, obfuscating whether the reliability of herbarium-derived estimates may vary taxonomically.

NPN observations and herbarium collections might exhibit similar biases not examined in this study. For example, specimens might be collected and NPN observations conducted at easily accessible sites near roads or at low elevations that may inaccurately represent the overall environmental conditions and phenology observed throughout a species' range (Daru et al. 2018, Meineke et al. 2021). Additionally, I detected large differences in sensitivity estimates for some species and substantial uncertainty in parameter estimation (especially for sensitivity to TMIN anomaly; Fig 3C, Appendix 1—Table S2), suggesting that herbarium-derived sensitivities for some species may lead to different conclusions from field observations or require much greater sample sizes than employed here for accurate estimation. Nevertheless, within the geographic and climatic space and the ecological diversity sampled in this study, this study demonstrates that herbarium specimens can uncover patterns of variation in phenology-climate relationships largely equivalent to those generated using field observations, suggesting that herbarium-based estimates may be generally robust to potential error or bias in specimen collection dates as proxies of peak flowering time.

Future directions

This study provides strong evidence of the reliability of herbarium specimens as resources with which to study phenological responses to spatiotemporal climate variation among species. However, the study was constrained by the availability of well-represented species in the NPN and herbarium datasets, preventing statistical comparison of the reliability of herbarium-based estimates among, for example, species with different life history traits. Future studies could leverage the growing number of digitized collections across the United States to identify additional species that are well represented in the NPN or other observational datasets and that might facilitate such analyses. Additionally, the study focused on a single component of the flowering phenology of a species (peak flowering); further research could determine whether specimens can generate reliable estimates of sensitivity for flowering onset or termination (which can show differing responses to climate; Caradonna et al. 2014), or for different life-cycle stages altogether. Phenological data from herbarium specimens is usually limited to presence-absence of reproductive structures, providing coarse information on the reproductive stage of specimens. Ongoing efforts to automate scoring of reproductive structures in herbarium sheets (Pearson et al. 2020) combined with new metrics that provide fine grained information of the reproductive status of herbarium specimens (Love et al. 2019, Goëau et al. 2020) might eventually enable sensitivity analyses for a wide range of phenological events and stages.

F. Literature cited

Anderson, J. T., Inouye, D. W., McKinney, A. M., Colautti, R. I., & Mitchell-Olds, T. (2012). Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. Proceedings of the Royal Society B: Biological Sciences, 279(1743), 3843–3852. https://doi.org/10.1098/rspb.2012.1051

Calinger, K. M., Queenborough, S., & Curtis, P. S. (2013). Herbarium specimens reveal the footprint of climate change on flowering trends across north-central North America. *Ecology Letters*, *16*(8), 1037–1044. https://doi.org/10.1111/ele.12135

Chuine, I. (2010). Why does phenology drive species distribution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1555), 3149–3160. https://doi.org/10.1098/rstb.2010.0142

Cleland, E. E., Allen, J. M., Crimmins, T. M., Dunne, J. A., Pau, S., Travers, S. E., Zavaleta, E. S., & Wolkovich, E. M. (2012). Phenological tracking enables positive species responses to climate change. Ecology, 93(8), 1765–1771. https://doi.org/10.1890/11-1912.1

Cook, B. I., Wolkovich, E. M., Davies, T. J., Ault, T. R., Betancourt, J. L., Allen, J. M., Bolmgren, K., Cleland, E. E., Crimmins, T. M., Kraft, N. J. B., Lancaster, L. T., Mazer, S. J., McCabe, G. J., McGill, B. J., Parmesan, C., Pau, S., Regetz, J., Salamin, N., Schwartz, M. D., & Travers, S. E. (2012). Sensitivity of Spring Phenology to Warming Across Temporal and Spatial Climate Gradients in Two Independent Databases. *Ecosystems*, *15*(8), 1283–1294. https://doi.org/10.1007/s10021-012-9584-5

Daru, B. H., Park, D. S., Primack, R. B., Willis, C. G., Barrington, D. S., Whitfeld, T. J. S., Seidler, T. G., Sweeney, P. W., Foster, D. R., Ellison, A. M., & Davis, C. C. (2018). Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytologist*, *217*(2), 939–955. https://doi.org/10.1111/nph.14855

Davis, C. C., Willis, C. G., Connolly, B., Kelly, C., & Ellison, A. M. (2015). Herbarium records are reliable sources of phenological change driven by climate and provide novel insights into species' phenological cueing mechanisms. *American Journal of Botany*, *102*(10), 1599–1609. https://doi.org/10.3732/ajb.1500237

Delgado, M. del M., Roslin, T., Tikhonov, G., Meyke, E., Lo, C., Gurarie, E., Abadonova, M., Abduraimov, O., Adrianova, O., Akimova, T., Akkiev, M., Ananin, A., Andreeva, E., Andriychuk, N., Antipin, M., Arzamascev, K., Babina, S., Babushkin, M., Bakin, O., … Ovaskainen, O. (2020). Differences in spatial versus temporal reaction norms for spring and autumn phenological events. *Proceedings of the National Academy of Sciences*, *117*(49), 31249–31258. https://doi.org/10.1073/pnas.2002713117

Diez, J. M., Ibáñez, I., Miller-Rushing, A. J., Mazer, S. J., Crimmins, T. M., Crimmins, M. A., Bertelsen, C. D., & Inouye, D. W. (2012). Forecasting phenology: From species variability to community patterns. *Ecology Letters*, *15*(6), 545–553. https://doi.org/10.1111/j.1461-0248.2012.01765.x

Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., Huang, M., Menzel, A., Peñuelas, J., Song, Y., Vitasse, Y., Zeng, Z., & Janssens, I. A. (2015). Declining global warming effects on the phenology of spring leaf unfolding. *Nature*, *526*(7571), 104–107. https://doi.org/10.1038/nature15402

Goëau, H., Mora-Fallas, A., Champ, J., Love, N. L. R., Mazer, S. J., Mata-Montero, E., Joly, A., & Bonnet, P. (2020). A new fine-grained method for automated visual analysis of herbarium specimens: A case study for phenological data extraction. *Applications in Plant Sciences*, *8*(6), e11368. https://doi.org/10.1002/aps3.11368

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*(12), 5189–5202. https://doi.org/10.1111/gcb.13781

Heberling, J. M., Prather, L. A., & Tonsor, S. J. (2019). The Changing Uses of Herbarium Data in an Era of Global Change: An Overview Using Automated Content Analysis. *BioScience*, *69*(10), 812–822. https://doi.org/10.1093/biosci/biz094

Jones, C. A., & Daehler, C. C. (2018). Herbarium specimens can reveal impacts of climate change on plant phenology; a review of methods and applications. *PeerJ*, *6*, e4576. https://doi.org/10.7717/peerj.4576

Love, N. L. R., & Mazer, S. J. (2021). Region-specific phenological sensitivities and rates of climate warming generate divergent temporal shifts in flowering date across a species'

range. *American Journal of Botany*, *108*(10), 1873–1888. https://doi.org/10.1002/ajb2.1748

Love, N. L. R., Park, I. W., & Mazer, S. J. (2019). A new phenological metric for use in pheno-climatic models: A case study using herbarium specimens of Streptanthus tortuosus. *Applications in Plant Sciences*, *7*(7), e11276. https://doi.org/10.1002/aps3.11276

Meineke, E. K., & Daru, B. H. (2021). Bias assessments to expand research harnessing biological collections. *Trends in Ecology & Evolution*, S0169534721002214. https://doi.org/10.1016/j.tree.2021.08.003

Meineke, E. K., Davis, C. C., & Davies, T. J. (2018). The unrealized potential of herbaria for global change biology. *Ecological Monographs*, *88*(4), 505–525. https://doi.org/10.1002/ecm.1307

Menzel, A., Yuan, Y., Matiu, M., Sparks, T., Scheifinger, H., Gehrig, R., & Estrella, N. (2020). Climate change fingerprints in recent European plant phenology. *Global Change Biology*, *26*(4), 2599–2612. https://doi.org/10.1111/gcb.15000

Miller-Rushing, A. J., & Primack, R. B. (2008). GLOBAL WARMING AND FLOWERING TIMES IN THOREAU'S CONCORD: A COMMUNITY PERSPECTIVE. *Ecology*, *89*(2), 332–341. https://doi.org/10.1890/07-0068.1

Miller-Rushing, A. J., Primack, R. B., Primack, D., & Mukunda, S. (2006). Photographs and herbarium specimens as tools to document phenological changes in response to global warming. *American Journal of Botany*, *93*(11), 1667–1674. https://doi.org/10.3732/ajb.93.11.1667

Munson, S. M., & Long, A. L. (2017). Climate drives shifts in grass reproductive phenology across the western USA. *New Phytologist*, *213*(4), 1945–1955. https://doi.org/10.1111/nph.14327

Omernik, J. M., & Griffith, G. E. (2014). Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework. Environmental Management, 54(6), 1249–1266. https://doi.org/10.1007/s00267-014-0364-1

Panchen, Z. A., Doubt, J., Kharouba, H. M., & Johnston, M. O. (2019). Patterns and biases in an Arctic herbarium specimen collection: Implications for phenological research. *Applications in Plant Sciences*, *7*(3), e01229. https://doi.org/10.1002/aps3.1229

Park, I. W. (2012). Digital herbarium archives as a spatially extensive, taxonomically discriminate phenological record; a comparison to MODIS satellite imagery. *International Journal of Biometeorology*, *56*(6), 1179–1182. https://doi.org/10.1007/s00484-012-0521-2

Park, I. W. (2014). Impacts of differing community composition on flowering phenology throughout warm temperate, cool temperate and xeric environments. *Global Ecology and Biogeography*, *23*(7), 789–801. https://doi.org/10.1111/geb.12163

Park, I. W., & Mazer, S. J. (2018). Overlooked climate parameters best predict flowering onset: Assessing phenological models using the elastic net. *Global Change Biology*, *24*(12), 5972–5984. https://doi.org/10.1111/gcb.14447

Pearson, K. D., Love, N. L. R., Ramirez-Parada, T., Mazer, S. J., & Yost, J. M. (2021). Phenological trends in the California poppy (*Eschscholzia californica*): digitized herbarium specimens reveal intraspecific variation in the sensitivity of flowering date to climate change. Madroño, 68(4), 343–359. https://doi.org/10.3120/0024-9637-68.4.343

Pearson, K. D., Nelson, G., Aronson, M. F. J., Bonnet, P., Brenskelle, L., Davis, C. C., Denny, E. G., Ellwood, E. R., Goëau, H., Heberling, J. M., Joly, A., Lorieul, T., Mazer, S. J., Meineke, E. K., Stucky, B. J., Sweeney, P., White, A. E., & Soltis, P. S. (2020). Machine Learning Using Digitized Herbarium Specimens to Advance Phenological Research. *BioScience*, *70*(7), 610–620. https://doi.org/10.1093/biosci/biaa044

Primack, D., Imbres, C., Primack, R. B., Miller-Rushing, A. J., & Tredici, P. D. (2004). Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. *American Journal of Botany*, *91*(8), 1260–1264. https://doi.org/10.3732/ajb.91.8.1260

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reeves, P. H., & Coupland, G. (2000). Response of plant development to environment: Control of flowering by daylength and temperature. *Current Opinion in Plant Biology*, *3*(1), 37–42. https://doi.org/10.1016/S1369-5266(99)00041-2

Renner, S. S., & Zohner, C. M. (2018). Climate Change and Phenological Mismatch in Trophic Interactions Among Plants, Insects, and Vertebrates. *Annual Review of Ecology,*

Evolution, and Systematics, *49*(1), 165–182. https://doi.org/10.1146/annurev-ecolsys-110617-062535

Robbirt, K. M., Davy, A. J., Hutchings, M. J., & Roberts, D. L. (2011). Validation of biological collections as a source of phenological data for use in climate change studies: A case study with the orchid Ophrys sphegodes. *Journal of Ecology*, 235–241. https://doi.org/10.1111/j.1365-2745.2010.01727.x@10.1111/(ISSN)1365-2745.orchidVI

Schwartz, M. D., Betancourt, J. L., & Weltzin, J. F. (2012). From Caprio's lilacs to the USA National Phenology Network. *Frontiers in Ecology and the Environment*, *10*(6), 324–327. https://doi.org/10.1890/110281

Song, Z., Fu, Y. H., Du, Y., Li, L., Ouyang, X., Ye, W., & Huang, Z. (2020). Flowering phenology of a widespread perennial herb shows contrasting responses to global warming between humid and non-humid regions. *Functional Ecology*, *n/a*(n/a). https://doi.org/10.1111/1365-2435.13634

Spellman, K. V., & Mulder, C. P. H. (2016). Validating Herbarium-Based Phenology Models Using Citizen-Science Data. *BioScience*, *66*(10), 897–906. https://doi.org/10.1093/biosci/biw116

Stan Development Team (2020). "RStan: the R interface to Stan." R package version 2.21.2, http://mc-stan.org/

Tang, J., Körner, C., Muraoka, H., Piao, S., Shen, M., Thackeray, S. J., & Yang, X. (2017). Emerging opportunities and challenges in phenology: A review. *Ecosphere*, e01436. https://doi.org/10.1002/ecs2.1436@10.1002/(ISSN)2150-8925.I_LTER

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(6), 1109–1113. https://doi.org/10.1007/s00484-018-1512-8

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally Downscaled and Spatially Customizable Climate Data for Historical and Future Periods for North America. *PLOS ONE*, *11*(6), e0156720. https://doi.org/10.1371/journal.pone.0156720

Wolkovich, E. M., Cook, B. I., & Davies, T. J. (2014). Progress towards an interdisciplinary science of plant phenology: Building predictions across space, time and species diversity. *New Phytologist*, *201*(4), 1156–1162. https://doi.org/10.1111/nph.12599

III. Plasticity and not adaptation is the primary source of temperature-mediated variation in flowering phenology in North America[2]

A. Abstract

Phenology varies widely over space and time because of its sensitivity to climate. However, whether phenological variation is primarily generated by rapid organismal responses (i.e., plasticity) or local adaptation remains unresolved. Here, I used 1,038,027 herbarium specimens representing 1,605 species from the continental United States to measure flowering time sensitivity to temperature over time ('S_{time}') and space ('S_{space}'). By comparing these estimates, I inferred how adaptation and plasticity historically influenced phenology along temperature gradients and how their contributions vary among species with different phenology and native climates, and among ecoregions differing in species composition. S_{space} and S_{time} were positively correlated ($r = 0.87$), of similar magnitude, and more frequently consistent with plasticity than adaptation. Apparent plasticity and adaptation generated earlier flowering in spring, limited responsiveness in late summer, and delayed flowering in fall in response to temperature increases. Nonetheless, ecoregions differed in the relative contributions of adaptation and plasticity, from consistently greater importance of plasticity (e.g., Southeastern USA Plains) to their nearly equal importance throughout the season (e.g., Western Sierra Madre Piedmont). These results support the hypothesis that plasticity is the primary driver of flowering time variation along temperature gradients, with local adaptation having a widespread but comparatively limited role.

---

[2]This chapter has been published in the following peer-reviewed article:
Ramirez-Parada, T. H., Park, I. W., Record, S., Davis, C. C., Ellison, A. M., & Mazer, S. J. (2024). Plasticity and not adaptation is the primary source of temperature-mediated variation in flowering phenology in North America. *Nature Ecology & Evolution*, *8*(3), 467–476. https://doi.org/10.1038/s41559-023-02304-5

B. Introduction

The timing of life-cycle events ('phenology') determines the environmental conditions that organisms encounter throughout development and often mediates their fitness (Elzinga et al. 2007). Phenology usually is cued by seasonally and interannually variable climatic factors—such as temperature—that enable individuals to adjust growth and reproduction plastically in response to fluctuating environmental conditions (Bradshaw et al. 1965, Elzinga et al. 2007). Phenology also varies within species as a result of evolutionary adaptation to local environments, which may select for different mean phenological timings among or within populations in space and time (Franks et al. 2007, Gienapp et al. 2008, Hoffman and Sgrò 2011, Wu and Colautti 2022). Although both plasticity and adaptation alter phenology, their relative contributions rarely have been measured within the same system largely because doing so requires experiments or spatiotemporally extensive genetic sampling (Phillimore et al. 2010, Merilä and Hendry 2011, Fox et al. 2019) (but see Wu and Colautti 2022). Accordingly, most studies have highlighted either plasticity or adaptation as mechanisms of phenological variation attributable to environmental change[7], but their relative importance across species and ecological contexts remains unresolved. Elucidating the degree to which species have phenologically responded to historical climatic variation through plasticity or adaptation could provide important context for predicting whether organismal responses may be sufficient—or evolutionary change necessary—to maintain development synchronized with suitable climatic conditions in a warming world (Fox et al. 2019).

Phillimore et al. (2010) proposed that the relative and joint contributions of plasticity and local adaptation to spatial variation in phenology within a species can be estimated from the difference between the slopes of spatial and temporal phenology-climate relationships.

This proposition rests on several observations. The effects of interannual climatic variation on phenology generally reflect plastic responses, especially among long-lived species less liable to experience microevolutionary changes from year to year (Bonamour et al. 2019). Phenological variation over space also can be caused by phenotypic plasticity where, for example, growing-degree day (GDD) thresholds that trigger life-cycle events occur on different dates across sites (Ensing and Eckert 2019). However, among populations, local adaptation also can generate phenological variation along climatic gradients (Stinchcombe et al. 2004, Montague et al. 2008). Therefore, assuming no confounding factors, and absent significant variation in phenological plasticity within and among populations, phenological variation along spatial climate gradients should reflect the joint effects of plasticity and adaptation (Anderson et al. 2012).

Given these observations and assumptions, plasticity and adaptation can generate five empirical patterns of sensitivity to temporal climatic variation (hereafter 'S_{time}') and to spatial climatic variation (hereafter 'S_{space}') (Fig. 1). First, if a species does not show phenological plasticity but population-level phenological means are locally adapted across a climatic gradient, we should observe negligible sensitivity to temporal climatic variation (i.e., no plasticity; $S_{time} = 0$) and a biologically significant difference between the slopes of the temporal and spatial relationships ($S_{space} - S_{time} \neq 0$ attributable to adaptation along the gradient; Figs. 1a,b). Alternatively, a phenologically plastic species whose populations are not locally adapted along the gradient should show biologically significant sensitivity to interannual climatic variation (i.e., $S_{time} \neq 0$) and no differences between temporal and spatial slopes ($S_{space} - S_{time} = 0$; Figs. 1c,d), implying that variation along the gradient can be attributed to plastic responses (i.e., $S_{space} = S_{time}$). When both adaptation and plasticity drive

phenological variation along the climate gradient (i.e., $S_{time} \neq 0$ and $S_{space} - S_{time} \neq 0$), the resulting empirical pattern should depend on the relative direction of plastic and adaptive responses. Specifically, when adaptation operates in the same direction as plasticity (i.e., "co-gradient adaptation"; Conover and Schultz 1995, Nylin and Gotthard 1998), we should observe a greater spatial than temporal sensitivity (e.g., $S_{time} < 0$ and $S_{space} - S_{time} < 0$ implies that $S_{space} < S_{time}$, so S_{space} is more negative; Figs. 1e, f). In turn, when adaptation operates in the opposite direction as plasticity (i.e., "counter-gradient adaptation"[15,16]), we should observe a lesser spatial sensitivity or one of opposite direction to the temporal relationship (e.g., $S_{time} < 0$ and $S_{space} - S_{time} > 0$ implies that $S_{space} > S_{time}$, so S_{space} is either less steep, or positive; Figs. 1g, h). Finally, if a species shows no plasticity or local adaptation along a climate gradient, we would expect biologically non-significant temporal and spatial sensitivities (Figs. 1i, j).

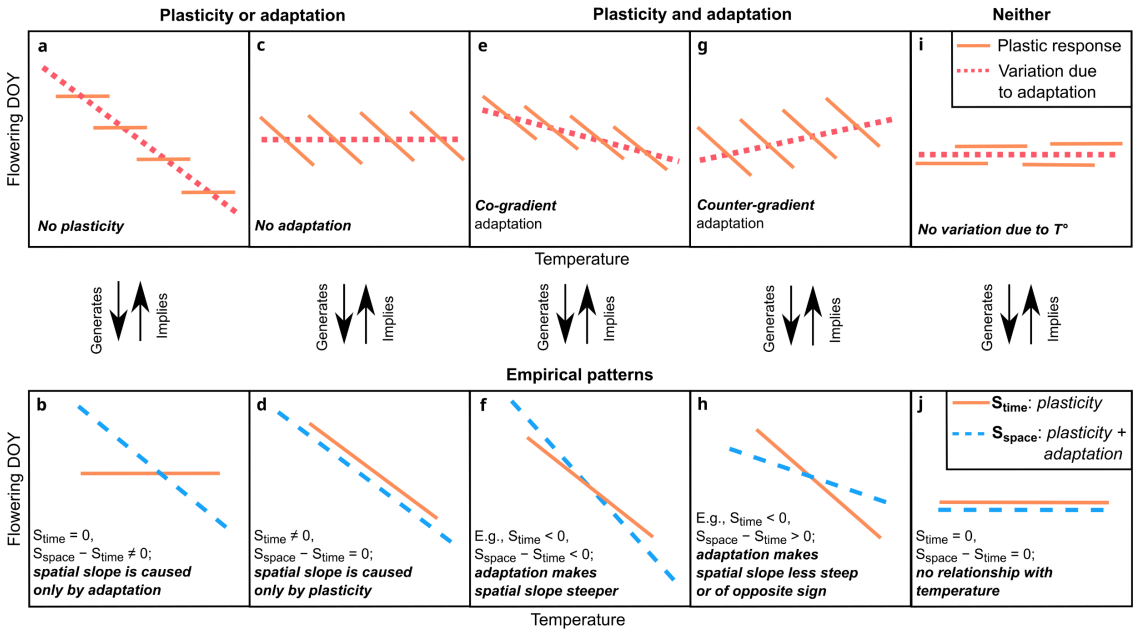

Figure 1—Spatial and temporal relationships between flowering time and temperature resulting from plasticity and adaptation. (a) Local adaptation acting as the sole driver of flowering time along the

gradient (i.e., no phenological plasticity) should result in (b) a negligible temporal relationship and a biologically significant difference between temporal and spatial slopes. In contrast, (c) plasticity acting as the sole driver of flowering time variation along the gradient (i.e., no adaptation) should result in (d) a biologically significant temporal relationship and negligible differences between spatial and temporal slopes. Local adaptation and plasticity jointly influencing flowering time should result in different empirical patterns depending on the direction of their effects. (e) Plasticity and adaptation operating in the same direction (e.g., both negative) should result in (f) a biologically significant temporal relationship and a spatial relationship of significantly greater magnitude. In contrast, (g) plasticity and adaptation operating in opposite directions (e.g., plasticity negative, adaptation positive) should result in (h) a biologically significant temporal relationship and a spatial relationship of significantly lesser magnitude (or having a different sign altogether). (i) Species exhibiting no plasticity or adaptation along the gradient would generate (j) biologically non-significant temporal and spatial slopes. Orange lines in a, c, e, and g illustrate phenological responses of spatially separated populations to temporal temperature variation, which spans a narrower temperature range than spatial temperature variation across the entire species range (segmented red lines). The biological processes in a, c, e, and g generate the empirical patterns in b, d, f, and h. In turn, the empirical patterns imply the processes that generated them. See "Methods – *Exploring Assumptions*" for an overview of the assumptions of this approach and the degree to which they were met by the data. For examples of species exhibiting each of these patterns, see Appendix 2—Fig. S1.

Phenological sensitivity to temperature often varies among species occurring in different regions or that initiate phenological events at different times throughout the growing season (Fitter and Fitter 2002, Cook et al. 2012, Lapenis et al. 2014, Park et al. Zhang et al. 2015, Park et al. 2019, Prevéy et al. 2019, Delgado et al. 2020, Li et al. 2021). However, comparisons of phenological sensitivity to climate over space and time—which are necessary

to evaluate the apparent contributions of plasticity and adaptation (Fig. 1)—across species differing in phenology and occupying different climates require spatiotemporally extensive datasets and therefore remain rare. Herbaria provide abundant and increasingly available data to conduct these analyses at unprecedented taxonomic, temporal, and spatial scales (Davis et al. 2015, Willis et al. 2017, Park and Mazer 2018, Park I. et al. 2021 Park D. et al. 2021, Park et al. 2022, Ramirez-Parada et al. 2022). However, few studies have separately estimated sensitivity to spatial versus temporal climate variation using specimens (but see Kharouba and Vellend 2015, Munson and Long 2017, Kopp et al. 2020, Pearson et al. 2021, Mazer et al. 2021, Ramirez-Parada et al. 2022, Park et al. 2023), and none have leveraged their unique scope to determine the ecological contexts in which plasticity or adaptation might contribute more strongly to spatial variation in phenology.

Here, I analyzed a dataset of over a million flowering specimens from 1,605 species across the continental United States to compare phenological sensitivities to spatial and temporal variation in temperature ('S_{space}' and 'S_{time}', respectively). For each species, I assessed whether its empirical sensitivity patterns were consistent with the effects of plasticity, adaptation, or both along temperature gradients (Fig. 1). Additionally, I evaluated how apparent temperature-related plasticity and adaptation of flowering time varied among species with different native climates, phenological niches, and occurring within different regional floras. Together, these analyses identified ecological contexts in which plasticity or adaptation appear to have most strongly influenced spatial phenological variation, providing the most taxonomically and geographically extensive assessment of temperature-mediated variation in flowering time among North American angiosperms conducted to date.

C. Methods

Specimen data

I assembled specimen records from 220 herbaria made available digitally through 16 consortia from Mexico, the United States, and Canada (accessed during July and August of 2022; Appendix 2—Note S1). Only specimens explicitly recorded as bearing flowers were retained, which were identified by summarizing all unique entries in the DarwinCore 'reproductiveCondition' column and identifying those that unambiguously indicated presence of flowers. After harmonizing species names using the Taxonomic Name Resolution Service (Boyle et al. 2013), I removed specimens lacking species-level identification, GPS coordinates, or dates of collection. To match the spatial and temporal coverage of the climate data (see '*Climate data*' section below), I retained only specimens collected from 1896 to 2020 within the United States. I considered as duplicates any conspecific specimens collected within 111m (i.e., 0.001 of a decimal degree) of one another on the same date. For subsequent analysis, species represented by at least 300 specimens were selected to ensure that the model was computationally tractable and that sufficient sample were available for estimating temperature responses in space and time. This filtering yielded a sample of 1,038,047 specimens from 1,605 species (Appendix 2—Fig. S2) (see Ramirez-Parada et al. 2023 for additional methodological detail).

I used day of year ('DOY') of collection of each specimen as a proxy for flowering date. Because flowering spanned year-ends for many species, I accounted for the DOY discontinuity between December 31st and January 1st using an azimuthal correction, whereby DOYs from the year prior become negative values (Park and Mazer 2018).

Climatic data

Temperature conditions preceding and leading up to anthesis can mediate flowering time through their effects on developmental rates of preceding phenophases or by cueing floral development and anthesis. Accordingly, I used mean surface temperatures averaged over a standard period of three months (Cook et al. 2012, Calinger et al. 2013, Mazer et al. 2013, Park et al. 2019) leading up to (and including) the mean flowering month for each species (hereafter 'TMEAN') as a predictor. For each collection site, I obtained monthly TMEAN time series (January 1896 – December 2020) at a 16-km^2 spatial resolution from the Parameter-elevation Regressions on Independent Slopes Model (PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu). Each collection site was characterized by its long-term mean temperature (hereafter '$TMEAN_{Normal}$'), obtained by averaging observed TMEAN across all years between 1896 and 2020. Annual deviations from long-term TMEAN conditions (hereafter '$TMEAN_{Anomaly}$') at each site and in each year were calculated by subtracting the $TMEAN_{Normal}$ from the observed TMEAN conditions in the year of collection. Positive and negative $TMEAN_{Anomaly}$ values respectively reflect warmer-than-average and colder-than-average years. $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ were uncorrelated irrespective of the latitudinal and elevational range spanned by a species (median $r = -0.04$), thus representing independent axes of climatic variation (Appendix 2—Fig. S3). $TMEAN_{Normal}$ spanned a wider temperature range than $TMEAN_{Anomaly}$ for most species, with respective median ranges of 13.7 °C and 5.4 °C (Appendix 2—Fig. S4). Species occurring in cold climates tended to show later mean flowering dates than species occupying warmer regions (Appendix 2—Fig. S5a); consequently, average $TMEAN_{Normal}$ values were well

above 0°C leading up to the mean flowering dates of all species in the data (Appendix 2—Fig. S5b).

To assess how sensitivities varied across climatic gradients (see *Analyses*, below), I first characterized long-term precipitation and temperature at each site of collection using a Principal Component Analysis (PCA), with mean annual temperature normal (MAT_{Normal}), mean annual precipitation normal (PPT_{Normal}), temperature seasonality, and precipitation seasonality as input features. I obtained precipitation (hereafter 'PPT') data from PRISM and calculated PPT and temperature seasonality for each collection site as the difference between the months with the highest and lowest PPT and mean temperature normal, respectively. PPT seasonality was calculated proportional to local levels of precipitation by dividing differences in maximum versus minimum monthly precipitation normal by PPT_{Normal} at each site. The PCA identified 2 principal components accounting for more variance than its input features, jointly explaining 78% of observed variation. PC1 was associated with increasing PPT seasonality (36%), decreasing temperature seasonality (31%) and increasing MAT_{Normal} (28%) (Appendix 2—Fig. S2). PC2 represented a gradient of decreasing PPT_{Normal} (74%) and increasing temperature seasonality (22%).

Analyses

Estimating apparent plasticity and adaptation

Estimates of flowering time sensitivity to $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ were generated using a Bayesian mixed-effects model. The model fitted species-specific intercepts and slopes and treated them as random effects sampled from 'community-level' distributions

(defined by among-species mean and standard deviation of intercepts and slopes). This hierarchical structure improved estimation of parameters by using information and estimates from all species in the data. In turn, the Bayesian inference framework allowed for estimation of the correlations between TMEAN sensitivities over space and time and their differences for each species while propagating parameter uncertainty.

The model used DOY for each observation *i* as a response, assuming a normal distribution with mean μ_i and species-specific standard deviation σ_{sp}:

$$DOY_i \sim N(\mu_i, \sigma_{sp}) \quad (1)$$

μ_i was modeled as a linear function of $TMEAN_{Normal}$ (TMEAN $Norm_i$), and $TMEAN_{Anomaly}$ (TMEAN $Anom_i$) for each observation *i*.

$$\mu_i = \alpha_{sp} + S_{space_{sp}} \times TMEAN\ Norm_i + S_{time_{sp}} \times TMEAN\ Anom_i \quad (2)$$

For each species *sp*, the model yielded intercepts representing mean flowering dates (α_{sp}), sensitivities (i.e., regression slopes) for TMEAN normal ($S_{space_{sp}}$), and sensitivities for TMEAN anomaly ($S_{time_{sp}}$).

To assess the correlation between S_{space} and S_{time}, I modeled community-level distributions for intercepts and slopes as generated by a multivariate normal distribution with a vector of hypermeans μ and a variance-covariance matrix Σ:

$$\left(\alpha_{sp}, S_{N_{sp}}, S_{A_{sp}}\right) \sim N(\mu, \Sigma) \quad (3)$$

I also calculated the difference between sensitivity types ($S_{space_{sp}} - S_{time_{sp}}$) as a derived quantity within the model, interpreted as the degree of apparent local adaptation in

DOY observed across the TMEAN normal gradient (Fig. 1), with negative and positive values respectively indicating advances and delays in flowering DOY across warmer locations.

I used weakly informative priors, with wide, 0-centered normal distributions for intercepts, slopes, and rate parameters for exponential distributions (used to obtain species-specific variances). For the variance-covariance matrix **Σ,** I used a Lewandowski-Kurowicka-Joe (LKJ) Cholesky covariance prior, with $\eta = 1$ to allow for high correlations among parameters. Posterior distributions were obtained using Hamiltonian Monte Carlo (HMC) in Stan as implemented in R v.4.2.1 using the 'rstan' package v.2.21.2 (Carpenter et al. 2017). A non-centered parameterization was implemented to improve sampling of the parameter space. Sampling was done using three MCMC chains with a training period of 1000 iterations and sampling of 4000 iterations. All S_{space}, S_{time}, and $S_{space} - S_{time}$ estimates had Gelman-Rubin statistics ('R-hat') of less than 1.002, and visual examination of trace plots confirmed convergence.

Fitting the model on simulated data (Appendix 2—Note S2), which emulated the average range of TMEAN conditions and the signal-to-noise ratio of DOY vs. TMEAN observed within species in the data, confirmed that the model could accurately recover the parameters of interest (S_{time}, S_{space}, and $S_{space} - S_{time}$) for a range of sample and effect sizes (Appendix 2—Note S2; Appendix 2—Figs. S6–8). Moreover, apparent plasticity (S_{time}) and apparent adaptation ($S_{space} - S_{time}$) could be estimated with similar degrees of precision (Appendix 2—Fig. S10).

Because the model did not include an explicit temporal predictor, it may appear to ignore widespread trends in phenology and temperature reported in recent decades. However,

additional simulation analyses (Appendix 2—Note S3) showed that the model does account for temporal trends in phenology among species that experience trends in $TMEAN_{Anomaly}$ over time and that are responsive to $TMEAN_{Anomaly}$ (i.e., non-zero S_{time}) (Appendix 2—Fig. S10a). To evaluate the model's implicit assumption that trends in $TMEAN_{Anomaly}$ cause observed trends in phenology, I used the herbarium dataset to determine empirically whether observed temporal trends in $TMEAN_{Anomaly}$ and a species' S_{time} indeed explain observed trends in DOY. The same patterns observed in the simulation were recovered (Appendix 2—Fig. S10b), suggesting that phenology and $TMEAN_{Anomaly}$ trends are causally related. Moreover, detrending DOY and $TMEAN_{Anomaly}$ prior to fitting the model did not affect the results, suggesting that omitting time as a covariate was unlikely to generate bias (Appendix 2—Fig. S11).

Finally, I evaluated the impact on model estimates of choosing alternative reference periods to calculate $TMEAN_{Normal}$ (i.e., 1901–2020 vs. 1901–1930, 1931–1960, 1961–1990, 1991–2020) (Appendix 2—Note S4, Appendix 2—Figs. S12–14). These analyses confirmed that period selection was unlikely to have affected the results.

Exploring assumptions

Herbarium specimens rarely are collected repeatedly at the same location across years. Accordingly, few collections over time were obtained in close enough proximity to represent single populations. Because of this, S_{space} and S_{time} were estimated using statistical methods different from Phillimore et al. (2010) and Delgado et al. (2020) (Appendix 2—Note S5). Nevertheless, the interpretation of the model relied on the same simplifying assumptions:

spatial slopes reflect variation in DOY among populations along a temperature gradient, temporal slopes reflect plasticity, plasticity does not vary within and among populations, and the temporal and spatial relationships between phenology and climate are not biased by confounding factors.

I evaluated the plausibility of many of these assumptions. S_{space} likely represented phenological variation among populations because conspecific specimens were collected over vast regions spanning median latitudinal and longitudinal ranges of 1,356 km and 1,819 km (removing outliers), respectively. In turn, S_{time} likely reflected the effects of plasticity and not adaptation: analyses including only long-lived perennials (unlikely to show microevolutionary changes over short periods) yielded very similar results to those presented below (Appendix 2—Fig. S15); moreover, detrending DOY and $TMEAN_{Anomaly}$ prior to fitting the model—which may account for temporal confounds or microevolution (Iler et al. 2017)—yielded nearly identical estimates (Appendix 2—Fig. S11). Furthermore, I generated a single estimate of S_{time} per species, thus assuming uniform plastic responses within and among populations. This assumption was supported by the observation that, for a large majority of species, S_{time} did not vary along geographic gradients of long-term TMEAN, long-term PPT, TMEAN seasonality, PPT seasonality, or the joint gradients described by PC1 and PC2 (Appendix 2—Fig. S16). Cumulative precipitation and photoperiod are unlikely to confound S_{space} and S_{time}: accounting for cumulative PPT yielded nearly identical estimates in single-species models (Appendix 2—Fig. S17), and an analysis of 120 species collected withing geographic ranges restricted to narrower latitudinal bands (≤1°)—and therefore to limited geographically-driven variation in photoperiod—yielded results very similar to those based on the entire dataset (Appendix 2—Fig. S18). Finally, I detected no biases in S_{space} or

S_{time} due to differences in sample size among species (Appendix 2—Fig. S19a, b), phylogeny (Appendix 2—Fig. S19c, d), spatial autocorrelation (Appendix 2—Fig. S19e, f), non-linear phenology-temperature relationships (Appendix 2—Fig. S20), or difference in range size among species (Appendix 2—Fig. S21).

Although herbarium data has many spatial and temporal collection biases and limitations—including preferential collection near roads and urban areas, and sharp decreases in collection intensity in recent decades (Daru et al. 2018)—such biases are likely not severe in the data (Appendix 2—Note S6, 7, Appendix 2—Figs. 22–29). Estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$ were robust to inclusion in the models of factors such as urbanization (Appendix 2—Fig. S23) and proximity to major roads (Appendix 2—Fig. S26, 27), and showed no evidence of various forms of temporal non-independence (Appendix 2—Fig. S29). Collector preferences can result in overrepresentation of certain taxa or traits among specimens[65]. While these biases cannot be ruled out, this study encompassed species from 106 families and 740 genera, capturing vast functional, evolutionary, and life history diversity. Therefore, I consider it unlikely that these results were driven by overrepresentation of taxa or traits. Finally, some herbaria obscure location data for endangered or heavily poached species. However, since I only included georeferenced specimens from well-represented species—of which only 12 (0.7% of the total) are listed as endangered by the United States Department of Agriculture (USDA, 2023)—it is unlikely that this species list includes many such taxa.

Categorizing sensitivity patterns

To assess the prevalence of apparent plasticity and adaptation among species, I categorized each species' S_{space} versus S_{time} patterns as consistent with the effects of plasticity alone (Figs. 1a,b), adaptation alone (Figs. 1c,d), the joint effects of plasticity and adaptation (co- or counter-gradient adaptation; Figs. 1e–h), or neither. Classifications were based on the proportion of the posterior probability distribution of S_{time} and $S_{space} - S_{time}$ lying in the direction of their *maximum a posteriori* (MAP) estimate (i.e., their "probability of direction", henceforth 'PD'). PD is bound by 0.5 (maximum uncertainty about the effect of the predictor) and 1 (certainty of an effect in the direction of the MAP estimate). I subjectively considered apparent plasticity (S_{time}) and adaptation ($S_{space} - S_{time}$) as significant when their PD was ≥ 0.95 (Table 1). Apparent plasticity and adaptation showed similar levels of estimation uncertainty both empirically (SD = 0.87 ± 0.34 d/°C for S_{time}; SD = 0.93 ± 0.32 d/°C for $S_{space} - S_{time}$) and in simulation analyses (Appendix 2—Note S2), suggesting sensitivity patterns were not substantially more likely to be classified as consistent with plasticity than with adaptation (and vice versa) due to estimation uncertainty.

Biological Process		Empirical Sensitivity Pattern
Plasticity only		1. Probability of direction for $S_{time} \geq 0.95$ 2. Probability of direction for $S_{space} - S_{time} < 0.95$
Adaptation only		1. Probability of direction for $S_{space} - S_{time} \geq 0.95$ 2. Probability of direction for $S_{time} < 0.95$
Plasticity and Adaptation	*Co-gradient*	1. Probability of direction for $S_{time} \geq 0.95$ 2. Probability of direction for $S_{space} - S_{time} \geq 0.95$ 3. S_{space} and S_{time} have the same direction 4. $\lvert S_{space} \rvert > \lvert S_{time} \rvert$
	Counter-gradient	1. Probability of direction for $S_{time} \geq 0.95$ 2. Probability of direction for $S_{space} - S_{time} \geq 0.95$ **Case 1:** 3. S_{space} and S_{time} have opposite direction **Case 2:** 4. S_{space} and S_{time} have the same direction 5. $\lvert S_{space} \rvert < \lvert S_{time} \rvert$
Neither		1. Probability of direction for $S_{time} < 0.95$ 2. Probability of direction for $S_{space} - S_{time} < 0.95$

Table 1—Criteria for classifying the sensitivity pattern of each species. Patterns were classified as consistent with the role of plasticity only, adaptation only, the joint effects of plasticity and adaptation in a co- or counter-gradient adaptation pattern, or neither adaptation nor plasticity. The probability that S_{time} or $S_{space} - S_{time}$ differed from 0 in the direction of its maximum a posteriori (MAP) estimate (i.e., their probability of direction) was obtained from the posterior distribution of these parameters for each species.

Phenological niches, local climates, and ecoregions

To assess how apparent plasticity and adaptation varied with native climate and phenological niche among species, I first calculated the mean flowering DOY and the mean coordinates along the climate gradients described by PC1 and PC2 among specimens of each species. Two generalized additive models (GAMs) were then using S_{time} or $S_{space} - S_{time}$ as responses—assumed to be normally distributed—and a three-variable tensor-product smooth of mean flowering DOY, mean PC1, and mean PC2 as a predictor. This design allowed us to

assess how native climate and phenological niche jointly determined the apparent roles of plasticity and adaptation while accounting for possible interactions and non-linearities. Because S_{time} and $S_{space} - S_{time}$ are estimates, I accounted for parameter uncertainty by weighting each observation by the inverse of its posterior variance (i.e., its precision).

Additionally, I assessed the relative contributions of apparent plasticity and adaptation throughout the season within ecoregions of the contiguous United States. To do so, I identified the Level II Ecoregion—as classified by the USA Environmental Protection Agency (EPA) (Omernik 1987, Omernik and Griffith 2014)—within which each specimen was collected. I used Level II Ecoregions because they provide sufficient ecological detail to distinguish regional floras while encompassing areas broad enough for each to capture multiple species in the data. To avoid inflating species overlap among regions or the influence of species that were rarely sampled within an ecoregion, I arbitrarily considered a species as present within an ecoregion if at least 10% of its collections occurred within it. Only ecoregions represented by a minimum of 8 species were retained. Under this scheme, the median species was classified as occurring within 2 ecoregions (range = 1–7), the median ecoregion was represented by 156 species (range = 17–956 for Atlantic Highlands and Western Cordilleras, respectively), and pairs of ecoregions shared, on average, 4% of their species (range = 0–39%; Appendix 2—Fig. S30). Of the 120 ecoregion pairs examined, 57 shared less than 1% of species, 100 shared less than 10% of species, and 114 shared less than 20% of species.

Once species × ecoregion combinations were identified (n = 3,570), two GAMs were fitted including apparent plasticity (S_{time}) or apparent adaptation ($S_{space} - S_{time}$) as a response, ecoregion as a categorical predictor, mean flowering DOY as continuous predictor, and a

mean flowering DOY × ecoregion spline assessing the ecoregion-specific effects of mean DOY on apparent plasticity or adaptation. Again, I accounted for parameter uncertainty by weighing each observation by the precision of its corresponding apparent plasticity or adaptation estimate. Collection locations in different ecoregions differed substantially in their long-term climatic conditions (Appendix 2—Fig. S31). However, I assumed no intraspecific variation in S_{time} across ecoregions, an assumption partially supported by the observation that S_{time} did not tend to vary along climatic gradients within species (Appendix 2—Fig. S16). All GAMs were implemented using the 'mgcv' package v.1.8-40 in R (R Core Team 2013, Wood 2017).

D. Results

Plasticity vs. adaptation as determinants of phenology

S_{space} and S_{time} of 93% and 79% of species, respectively, differed from 0 with at least 95% probability. S_{space} and S_{time} agreed in direction for 94% of species and estimates of both S_{time} and S_{space} were negative for 89% and 91% of species, indicating earlier flowering across increasingly warmer locations and in warmer-than-average years (Fig. 2a).

Both apparent plasticity and adaptation were associated with clinal variation in flowering time along temperature gradients, with plasticity playing a predominant role among species. S_{space} and S_{time} were highly positively correlated, and their magnitude tended to correspond 1-to-1 for many species (Fig. 2b). Therefore, flowering shifts in warmer-than-average years typically had similar direction and magnitude (in days/°C) as those observed

across increasingly warmer locations, consistent with a scenario of plasticity as the cause of phenological variation along spatial temperature gradients (Figs. 1c,d; Table 1).

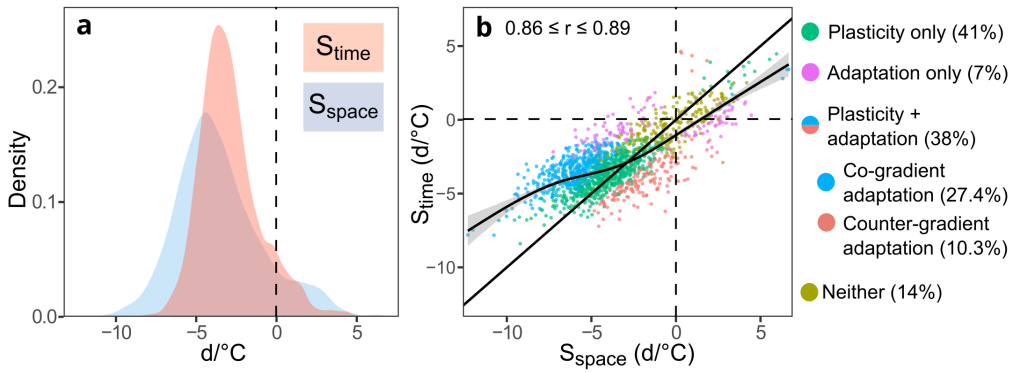

Figure 2—Distributions of, and relationship between S_{space} and S_{time} among 1,605 North American angiosperms. Shaded regions in (a) correspond to the kernel density distributions of S_{time} (red) and S_{space} (blue) among species. Each point in (b) represents a species whose *x, y* coordinates are given by the maximum a posteriori (MAP) estimates for S_{space} and S_{time}, respectively. Colors in (b) indicate whether sensitivity patterns were consistent with plasticity (green) or adaptation (magenta) as the sole drivers of flowering time variation along the temperature gradient, with both plasticity and adaptation in a co- or counter-gradient adaptation pattern (blue, orange), or neither (dark yellow). The straight, solid black line in (b) indicates a 1:1 relationship (i.e., $S_{space} = S_{time}$), whereas the curved solid line shows the observed relationship estimated from a generalized additive model (GAM). The shaded region along the curved solid line in (b) corresponds to the standard error of the predicted value of S_{time}. The percentage of species showing each pattern is shown in parentheses in the legend. The 95% credible interval for the correlation between S_{space} and S_{time} is provided as a text inset in (b).

More species showed sensitivity patterns consistent with plasticity (79%) than with adaptation (45%) (see Fig. 1, and a detailed classification scheme in Table 1). Apparent

plasticity explained approximately 52% of the variance in flowering-time clines along temperature gradients among species (Fig. 2b). Fourty-one percent of species showed sensitivity patterns consistent with plasticity as the sole driver of phenological variation across gradients. In contrast, only 7% of species showed sensitivity patterns consistent solely with adaptation (see Figs. 1a,b). Thirty-eight percent of the species showed both apparent local adaptation and evidence of plasticity. Among these, a greater proportion showed flowering advances (and co-gradient patterns; 27%) than flowering delays (and counter-gradient patterns; 10%) resulting from apparent adaptation along temperature gradients (Fig. 2b). Fourteen percent of species showed patterns that were consistent neither with temperature-related plasticity nor with adaptation. These patterns remained consistent when analyzing only long-lived species (whose responses to yearly temperature anomalies are certain to be plastic) (Appendix 2—Fig. S32).

Plasticity and adaptation across ecological contexts

Apparent plasticity (S_{time}) varied substantially among species with different phenological niches and across local climates ($R^2 = 0.55$; Fig. 3a,c). Species flowering during late winter and spring tended to show flowering advances in warmer-than-average years. Such advances decreased in magnitude throughout the season, typically reversing to flowering delays during late summer and fall (Fig. 3a,c). The timing of the transition from positive values was consistent throughout PC1 (Fig. 3a), but occurred much earlier in arid regions with high temperature seasonality along PC2 (Fig. 3c). Apparent adaptation ($S_{space} - S_{time}$) also varied with phenological niche and native climate ($R^2 = 0.47$, Figs. 3b,d). Apparent adaptation varied from negative to positive values throughout the growing season, indicating a transition

from flowering advances to delays attributable to local adaptation. Such transitions occurred much earlier in cool, thermally seasonal regions (i.e., the low range of PC1) (Fig. 3b). Apparent adaptation also varied throughout the growing season along PC2, with transition from advances to delays under warmer conditions occurring earlier in regions with high precipitation (Fig. 3d).

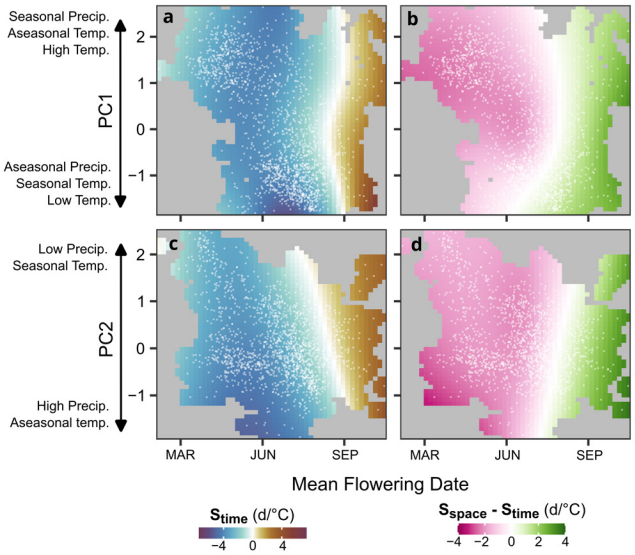

Figure 3—Variation in apparent plasticity (S_{time}) and apparent adaptation ($S_{space} - S_{time}$) attributable to differences in phenological niche and native climate among species. PC1 (a, b) represents a climate gradient of increasing precipitation seasonality, decreasing temperature seasonality, and increasing mean annual temperature, whereas PC2 (c, d) corresponds to a gradient of decreasing mean annual precipitation and increasing temperature seasonality. The color gradients in each panel represents the predicted magnitude of S_{time} or $S_{space} - S_{time}$ (in days/°C) for a combination of mean flowering DOY and PC1 or PC2 values. The predicted surfaces represented by the color gradients were obtained using three-variable tensor smooths in a generalized additive modelling (GAM) framework. In each panel, the value of the third variable (the one not plotted) was fixed at its mean.

These patterns were mirrored at the regional level: throughout the season, average apparent plasticity and adaptation among species transitioned from generating flowering advances to generating delays in response to higher temperatures in all sampled ecoregions (R^2 for S_{time} = 0.44; R^2 for $S_{space} - S_{time}$ = 0.35; Fig. 4). This transition invariably occurred during the summer months. The magnitude of apparent adaptation tended to be lower than that of apparent plasticity during most of spring and early summer for all ecoregions, but their difference tended to be less among species flowering during early spring and the magnitude of adaptation was often greater among species flowering during late summer and early fall (Figs. 4a–n). Nonetheless, ecoregions differed in the relative contributions of apparent adaptation and plasticity among species throughout the season. For example, apparent adaptation and plasticity had similar magnitudes within the Western Sierra Madre Piedmont (Fig. 4g). In contrast, mean apparent plasticity was consistently greater than adaptation among species in the Southeastern USA Plains (Fig. 4j). The difference in magnitude between apparent plasticity and adaptation was greatest among early- to mid-summer flowering species in the Western Cordilleras and Cold Deserts (Figs. 4b, c).

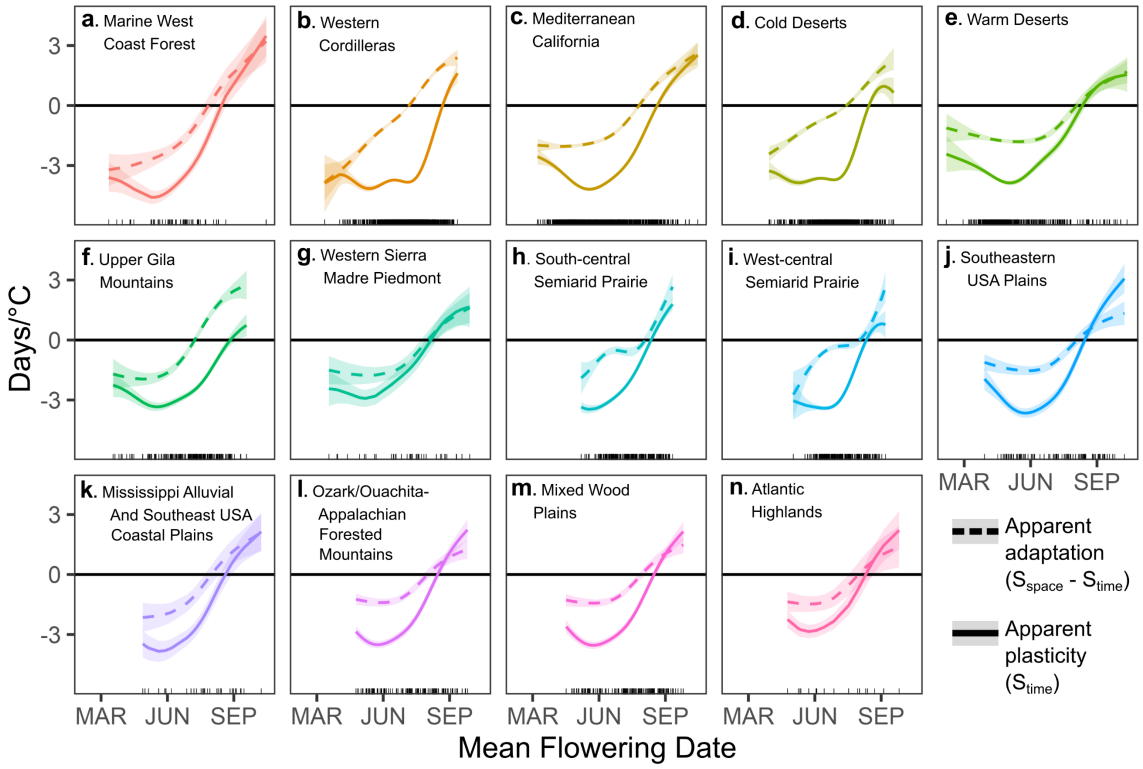

Figure 4—Variation in apparent plasticity and apparent adaptation among species with varying phenological niches across ecoregions of the United States. Shaded regions in each panel represent the 95% confidence interval for the mean apparent plasticity or apparent adaptation among species predicted for a given mean flowering date. The predicted mean values for apparent plasticity and adaptation were obtained using generalized additive models (GAMs).

E. Discussion

This study provides evidence that, for 1,605 North American plant species, phenotypic plasticity historically has been the primary mechanism generating flowering-time variation along temperature gradients. Nonetheless, apparent adaptation and plasticity jointly generated phenological variation in a large proportion of species. Both apparent plasticity and adaptation consistently generated flowering advances in spring, lesser advances during

summer, and flowering delays during early fall, and this pattern was consistent across climates and ecoregions. Whether phenological reaction norms to historical climatic conditions will remain adaptive under future climatic regimes is unclear (Bonamour et al. 2019). Nonetheless, these results suggest that plasticity historically has enabled flowering phenology to respond quickly to a wide range of temperature conditions among North American angiosperms, with adaptation frequently playing an important but context-dependent role.

Plasticity causes clinal variation in flowering time

Extensive research has documented phenological plasticity to spatial climatic variation in plants (Kramer 1995, Levin 2009, De Frenne et al. 2011, Franks et al. 2014) that can result in clinal phenological variation even among short-lived taxa (Vitasse et al. 2013, Ensing and Eckert 2019). This study extends these results by showing that the predominance of plasticity over adaptation associated with temperature-related variation in phenology over space might be the norm among North American species.

The greater importance of plasticity found in this study does not contradict the well-established role of phenological adaptation in space and time (Franks et al. 2014), which can mediate rapid temporal shifts in phenology (Franks et al. 2007) or facilitate ecological invasions (Colautti and Barret 2013, Wu and Colautti 2022). Indeed, 45% of species in the data showed evidence of adaptation-driven phenological variation along temperature gradients (Fig. 2b). It is also possible non-linear or patchy adaptation patterns were not detected, or that the contributions of apparent adaptation and plasticity may be different in

regions underrepresented in the data (e.g., the Great Plains and prairies; Appendix 2—Fig. S2). Crucially, I only assessed the apparent contributions of plasticity and adaptation to observed variation in flowering time over temperature gradients, so these results do not rule out the possibility that adaptation is the primary driver of phenological variation along gradients of different climatic variables. Finally, determining the exact environmental conditions within microsites where herbarium specimens were collected is impossible because continental-scale climate products have relatively coarse spatial resolution and because specimen coordinates typically are inexact. Climatic variation at the microsite level could confound estimates of S_{space} and the assessment of the prevalence of local adaptation if, for example, different populations along the gradient occupied distinct microsites that maintained temperatures more constant than apparent when looking at coarser pixel-level averages. However, to my knowledge, such microsite sorting across species ranges has only been reported at their trailing edges where climate is most limiting (Ackerly et al. 2020). Nonetheless, these potential complexities underscore the ultimate need for molecular or quantitative genetic studies to corroborate the broad correlational patterns outlined in this study.

Still, the strong correlation between S_{space} and S_{time} has important implications for phenoclimatic research. For example, it suggests that temperature-related variation in flowering time among conspecific populations is a good proxy of responsiveness to interannual temperature variation. Therefore, space-for-time substitutions might be viable approaches for quantifying plastic flowering responsiveness to temperature in North American angiosperms, for most of which we lack long-term phenological records (Wolkovich et al. 2014, Willis et al. 2017). Specifically, the match between S_{space} and S_{time}

shows that substituting space for time might reveal the direction and approximate magnitude on flowering sensitivity to temperature over time within species, or relative differences in sensitivity among species. However, co-gradient adaptation frequently generated spatial sensitivities of greater magnitudes than those over time, demonstrating that S_{space} might overestimate S_{time} in many species.

This study's results also indicate that plasticity may have generated phenological variation across a temperature range (i.e., a median range of 13.7 °C) exceeding the degree of warming forecasted for most regions in coming decades. However, such historical plastic flowering shifts over space will not necessarily be mirrored by temporal shifts within populations as warming trends continue. For example, historical temperature cues may become uncorrelated from the factors mediating the fitness consequences of phenology, rendering plastic reaction norms maladaptive (Gienapp et al. 2008, Bonamour et al. 2019). Plastic phenological shifts associated with warming also may be constrained by physiology (Chown et al. 2010) or by other competing cueing mechanisms such as photoperiod or winter chilling that may be disrupted by phenological shifts associated with higher temperatures (Fu et al. 2015, Güsewell et al. 2017, Wolkovich et al. 2022). These complexities highlight the need for research on the fitness consequences of recent and ongoing phenological shifts (Iler et al. 2021, De Lisle et al. 2022), and on the interrelated mechanisms underpinning associations between multiple abiotic cues (e.g., chilling, forcing, photoperiod, resources) and seasonal development beyond model systems (Amasino 2010, Wolkovich et al. 2022).

Plasticity and adaptation vary across ecological contexts

Sensitivities transitioned from flowering advances under warming in spring to reduced or no responsiveness during summer and even flowering delays in early fall (Figs. 3, 4). This pattern implies that temperature trends will likely drive changes to the structure of the flowering season during spring and fall under global change, but that other environmental factors might play predominant roles during summer.

These results support studies showing decreases in phenological sensitivity to temperature among species throughout the season in temperate biomes (Cook et al. 2012, Wolkovich et al. 2012, Mazer et al. 2013, Park et al. 2019), and others showing flowering delays among autumn-flowering species or lengthening of the growing and flowering seasons under warming (Delgado et al. 2020, Beil et al. 2021, Roslin et al. 2021, Zhou et al. 2022). While the causes of this pattern cannot be unambiguously identified, studies have shown that warming typically advances phenology during spring due to accelerated developmental rates, while phenophases occurring during fall are cued directly by seasonal cooling (Sherry et al. 2007, Chen et al. 2020, Zohner et al. 2023). This difference would explain why fall-flowering species showed phenological delays under warming (i.e., fall cooling occurs later in warmer-than-average years), or why the transition from advances to delays was more pronounced within cool regions with high temperature seasonality (i.e., those showing more pronounced cooling during fall; Fig. 3). Regardless of its causes, this study corroborates that transitions from spring flowering advances to fall delays because of climatic warming are consistent across thousands of species and diverse climate zones and biomes in the continental United States.

Likewise, apparent adaptation throughout the season typically transitioned from generating mean flowering advances to generating delays along temperature gradients. The

results are consistent with those reported by Delgado et al. (2020), who found changes in the direction of apparent plasticity and adaptation throughout the growing season for multiple trophic levels (i.e., saprotrophs, primary producers, and primary and secondary consumers) in Eastern Europe. That changes in apparent plastic and adaptive responses to warming throughout the year might be robust across different phenophases, taxa, trophic levels, or climatic regimes across the temperate zone may reflect shared cueing mechanisms or selective pressures for different phenological events occurring during the same seasons (Roslin et al. 2021), with factors other than temperature (e.g., resources or photoperiod) likely driving phenological variation for developmental events in summer. Additionally, the greater prevalence of co-gradient adaptation as opposed to counter-gradient adaptation suggests that adaptation typically operates to generate greater variation in phenology along temperature gradients than generated by plasticity alone.

Conclusions

This study indicates that phenotypic plasticity is the predominant historical mechanism of spatial phenological variation across a wide range of temperature conditions in the continental United States; adaptation plays more context-specific roles. Whether and how species-level attributes such as functional traits and life history may mediate these relative contributions or whether historical responses will tend to be adaptive under non-analog climatic conditions remain open questions and important directions for future research. These results outline broad correlational patterns whose verification will require direct measurements of plasticity and adaptation across species and climate regions. Nonetheless, the data—across many biomes and thousands of species—confirmed patterns of plastic and

adaptive phenological advances in spring and delays in fall in response to warming observed in detailed empirical studies, highlighting the increasing utility of biological collections for studying plant responses to global change at vast taxonomic and spatiotemporal scales.

F. Literature cited

Ackerly, D. D., Kling, M. M., Clark, M. L., Papper, P., Oldfather, M. F., Flint, A. L., & Flint, L. E. (2020). Topoclimates, refugia, and biotic responses to climate change. *Frontiers in Ecology and the Environment*, *18*(5), 288–297. https://doi.org/10.1002/fee.2204

Amasino, R. (2010). Seasonal and developmental timing of flowering. *The Plant Journal*, *61*(6), 1001–1013. https://doi.org/10.1111/j.1365-313X.2010.04148.x

Anderson, J. T., Inouye, D. W., McKinney, A. M., Colautti, R. I., & Mitchell-Olds, T. (2012). Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1743), 3843–3852. https://doi.org/10.1098/rspb.2012.1051

Beil, I., Kreyling, J., Meyer, C., Lemcke, N., & Malyshev, A. V. (2021). Late to bed, late to rise—Warmer autumn temperatures delay spring phenology by delaying dormancy. *Global Change Biology*, *27*(22), 5806–5817. https://doi.org/10.1111/gcb.15858

Bonamour, S., Chevin, L.-M., Charmantier, A., & Teplitsky, C. (2019). Phenotypic plasticity in response to climate change: The importance of cue variation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1768), 20180178. https://doi.org/10.1098/rstb.2018.0178

Boyle, B., Hopkins, N., Lu, Z., Raygoza Garay, J. A., Mozzherin, D., Rees, T., Matasci, N., Narro, M. L., Piel, W. H., Mckay, S. J., Lowry, S., Freeland, C., Peet, R. K., & Enquist, B. J. (2013). The taxonomic name resolution service: An online tool for automated standardization of plant names. *BMC Bioinformatics*, *14*(1), 16. https://doi.org/10.1186/1471-2105-14-16

Bradshaw, A. D. (1965). Evolutionary Significance of Phenotypic Plasticity in Plants. In E. W. Caspari & J. M. Thoday (Eds.), *Advances in Genetics* (Vol. 13, pp. 115–155). Academic Press. https://doi.org/10.1016/S0065-2660(08)60048-6

Calinger, K. M., Queenborough, S., & Curtis, P. S. (2013). Herbarium specimens reveal the footprint of climate change on flowering trends across north-central North America. *Ecology Letters*, *16*(8), 1037–1044. https://doi.org/10.1111/ele.12135

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, *76*, 1. https://doi.org/10.18637/jss.v076.i01

Chen, L., Hänninen, H., Rossi, S., Smith, N. G., Pau, S., Liu, Z., Feng, G., Gao, J., & Liu, J. (2020). Leaf senescence exhibits stronger climatic responses during warm than during cold autumns. *Nature Climate Change*, *10*(8), 777–780. https://doi.org/10.1038/s41558-020-0820-2

Chown, S. L., Hoffmann, A. A., Kristensen, T. N., Jr, M. J. A., Stenseth, N. C., & Pertoldi, C. (2010). Adapting to climate change: A perspective from evolutionary physiology. *Climate Research*, *43*(1–2), 3–15. https://doi.org/10.3354/cr00879

Colautti, R. I., & Barrett, S. C. H. (2013). Rapid Adaptation to Climate Facilitates Range Expansion of an Invasive Plant. *Science*, *342*(6156), 364–366. https://doi.org/10.1126/science.1242121

Conover, D. O., & Schultz, E. T. (1995). Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution*, *10*(6), 248–252. https://doi.org/10.1016/S0169-5347(00)89081-3

Cook, B. I., Wolkovich, E. M., Davies, T. J., Ault, T. R., Betancourt, J. L., Allen, J. M., Bolmgren, K., Cleland, E. E., Crimmins, T. M., Kraft, N. J. B., Lancaster, L. T., Mazer, S. J., McCabe, G. J., McGill, B. J., Parmesan, C., Pau, S., Regetz, J., Salamin, N., Schwartz, M. D., & Travers, S. E. (2012). Sensitivity of Spring Phenology to Warming Across Temporal and Spatial Climate Gradients in Two Independent Databases. *Ecosystems*, *15*(8), 1283–1294. https://doi.org/10.1007/s10021-012-9584-5

Daru, B. H., Park, D. S., Primack, R. B., Willis, C. G., Barrington, D. S., Whitfeld, T. J. S., Seidler, T. G., Sweeney, P. W., Foster, D. R., Ellison, A. M., & Davis, C. C. (2018). Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytologist*, *217*(2), 939–955. https://doi.org/10.1111/nph.14855

Davis, C. C., Willis, C. G., Connolly, B., Kelly, C., & Ellison, A. M. (2015). Herbarium records are reliable sources of phenological change driven by climate and provide novel insights into species' phenological cueing mechanisms. *American Journal of Botany*, *102*(10), 1599–1609. https://doi.org/10.3732/ajb.1500237

De Frenne, P., Brunet, J., Shevtsova, A., Kolb, A., Graae, B. J., Chabrerie, O., Cousins, S. A., Decocq, G., De Schrijver, A., Diekmann, M., Gruwez, R., Heinken, T., Hermy, M., Nilsson, C., Stanton, S., Tack, W., Willaert, J., & Verheyen, K. (2011). Temperature effects on forest herbs assessed by warming and transplant experiments along a latitudinal gradient. *Global Change Biology*, *17*(10), 3240–3253. https://doi.org/10.1111/j.1365-2486.2011.02449.x

De Lisle, S. P., Mäenpää, M. I., & Svensson, E. I. (2022). Phenotypic plasticity is aligned with phenological adaptation on both micro- and macroevolutionary timescales. *Ecology Letters*, *25*(4), 790–801. https://doi.org/10.1111/ele.13953

Delgado, M. del M., Roslin, T., Tikhonov, G., Meyke, E., Lo, C., Gurarie, E., Abadonova, M., Abduraimov, O., Adrianova, O., Akimova, T., Akkiev, M., Ananin, A., Andreeva, E., Andriychuk, N., Antipin, M., Arzamascev, K., Babina, S., Babushkin, M., Bakin, O., … Ovaskainen, O. (2020). Differences in spatial versus temporal reaction norms for spring and autumn phenological events. *Proceedings of the National Academy of Sciences*, *117*(49), 31249–31258. https://doi.org/10.1073/pnas.2002713117

Elzinga, J. A., Atlan, A., Biere, A., Gigord, L., Weis, A. E., & Bernasconi, G. (2007). Time after time: Flowering phenology and biotic interactions. *Trends in Ecology & Evolution*, *22*(8), 432–439. https://doi.org/10.1016/j.tree.2007.05.006

Ensing, D. J., & Eckert, C. G. (2019). Interannual variation in season length is linked to strong co-gradient plasticity of phenology in a montane annual plant. *New Phytologist*, *224*(3), 1184–1200. https://doi.org/10.1111/nph.16009

Fitter, A. H., & Fitter, R. S. R. (2002). Rapid Changes in Flowering Time in British Plants. *Science*, *296*(5573), 1689–1691. https://doi.org/10.1126/science.1071617

Fox, R. J., Donelson, J. M., Schunter, C., Ravasi, T., & Gaitán-Espitia, J. D. (2019). Beyond buying time: The role of plasticity in phenotypic adaptation to rapid environmental change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1768), 20180174. https://doi.org/10.1098/rstb.2018.0174

Franks, S. J., Sim, S., & Weis, A. E. (2007). Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences*, *104*(4), 1278–1282. https://doi.org/10.1073/pnas.0608379104

Franks, S. J., Weber, J. J., & Aitken, S. N. (2014). Evolutionary and plastic responses to climate change in terrestrial plant populations. *Evolutionary Applications*, *7*(1), 123–139. https://doi.org/10.1111/eva.12112

Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., Huang, M., Menzel, A., Peñuelas, J., Song, Y., Vitasse, Y., Zeng, Z., & Janssens, I. A. (2015). Declining global warming effects on the phenology of spring leaf unfolding. *Nature*, *526*(7571), 104–107. https://doi.org/10.1038/nature15402

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A., & Merilä, J. (2008). Climate change and evolution: Disentangling environmental and genetic responses. *Molecular Ecology*, *17*(1), 167–178. https://doi.org/10.1111/j.1365-294X.2007.03413.x

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*(12), 5189–5202. https://doi.org/10.1111/gcb.13781

Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*(7335), 479–485. https://doi.org/10.1038/nature09670

Iler, A. M., CaraDonna, P. J., Forrest, J. R. K., & Post, E. (2021). Demographic Consequences of Phenological Shifts in Response to Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, *52*(Volume 52, 2021), 221–245. https://doi.org/10.1146/annurev-ecolsys-011921-032939

Iler, A. M., Inouye, D. W., Schmidt, N. M., & Høye, T. T. (2017). Detrending phenological time series improves climate–phenology analyses and reveals evidence of plasticity. *Ecology*, *98*(3), 647–655. https://doi.org/10.1002/ecy.1690

Kharouba, H. M., & Vellend, M. (2015). Flowering time of butterfly nectar food plants is more sensitive to temperature than the timing of butterfly adult flight. *Journal of Animal Ecology*, *84*(5), 1311–1321. https://doi.org/10.1111/1365-2656.12373

Kopp, C. W., Neto-Bradley, B. M., Lipsen, L. P. J., Sandhar, J., & Smith, S. (2020). Herbarium records indicate variation in bloom-time sensitivity to temperature across

a geographically diverse region. *International Journal of Biometeorology*, *64*(5), 873–880. https://doi.org/10.1007/s00484-020-01877-1

Kramer, K. (1995). Phenotypic plasticity of the phenology of seven European tree species in relation to climatic warming. *Plant, Cell & Environment*, *18*(2), 93–104. https://doi.org/10.1111/j.1365-3040.1995.tb00356.x

Lapenis, A., Henry, H., Vuille, M., & Mower, J. (2014). Climatic factors controlling plant sensitivity to warming. *Climatic Change*, *122*(4), 723–734. https://doi.org/10.1007/s10584-013-1010-2

Levin, D. A. (2009). Flowering-time plasticity facilitates niche shifts in adjacent populations. *New Phytologist*, *183*(3), 661–666. https://doi.org/10.1111/j.1469-8137.2009.02889.x

Li, D., Barve, N., Brenskelle, L., Earl, K., Barve, V., Belitz, M. W., Doby, J., Hantak, M. M., Oswald, J. A., Stucky, B. J., Walters, M., & Guralnick, R. P. (2021). Climate, urbanization, and species traits interactively drive flowering duration. *Global Change Biology*, *27*(4), 892–903. https://doi.org/10.1111/gcb.15461

Mazer, S. J., Love, N. L. R., Park, I. W., Ramirez-Parada, T., & Matthews, E. R. (2021). Phenological sensitivities to climate are similar in two Clarkia congeners: Indirect evidence for facilitation, convergence, niche conservatism or genetic constraints. *Madroño*, *68*(4). https://doi.org/10.3120/0024-9637-68.4.388

Mazer, S. J., Travers, S. E., Cook, B. I., Davies, T. J., Bolmgren, K., Kraft, N. J. B., Salamin, N., & Inouye, D. W. (2013). Flowering date of taxonomic families predicts phenological sensitivity to temperature: Implications for forecasting the effects of climate change on unstudied taxa. *American Journal of Botany*, *100*(7), 1381–1397. https://doi.org/10.3732/ajb.1200455

Merilä, J., & Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: The problem and the evidence. *Evolutionary Applications*, *7*(1), 1–14. https://doi.org/10.1111/eva.12137

Montague, J. L., Barrett, S. C. H., & Eckert, C. G. (2008). Re-establishment of clinal variation in flowering time among introduced populations of purple loosestrife (*Lythrum salicaria* , Lythraceae). *Journal of Evolutionary Biology*, *21*(1), 234–245. https://doi.org/10.1111/j.1420-9101.2007.01456.x

Munson, S. M., & Long, A. L. (2017). Climate drives shifts in grass reproductive phenology across the western USA. *New Phytologist*, *213*(4), 1945–1955. https://doi.org/10.1111/nph.14327

Nylin, S., & Gotthard, K. (1998). Plasticity in Life-History Traits. *Annual Review of Entomology*, *43*(Volume 43, 1998), 63–83. https://doi.org/10.1146/annurev.ento.43.1.63

Omernik, J. M. (1987). Ecoregions of the Conterminous United States. *Annals of the Association of American Geographers*, *77*(1), 118–125. https://doi.org/10.1111/j.1467-8306.1987.tb00149.x

Omernik, J. M., & Griffith, G. E. (2014). Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework. *Environmental Management*, *54*(6), 1249–1266. https://doi.org/10.1007/s00267-014-0364-1

Park, D. S., Breckheimer, I. K., Ellison, A. M., Lyra, G. M., & Davis, C. C. (2022). Phenological displacement is uncommon among sympatric angiosperms. *New Phytologist*, *233*(3), 1466–1478. https://doi.org/10.1111/nph.17784

Park, D. S., Breckheimer, I., Williams, A. C., Law, E., Ellison, A. M., & Davis, C. C. (2019). Herbarium specimens reveal substantial and unexpected variation in phenological sensitivity across the eastern United States. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1763), 20170394. https://doi.org/10.1098/rstb.2017.0394

Park, D. S., Xie, Y., Ellison, A. M., Lyra, G. M., & Davis, C. C. (2023). Complex climate-mediated effects of urbanization on plant reproductive phenology and frost risk. *New Phytologist*, *239*(6), 2153–2165. https://doi.org/10.1111/nph.18893

Park, I. W., & Mazer, S. J. (2018). Overlooked climate parameters best predict flowering onset: Assessing phenological models using the elastic net. *Global Change Biology*, *24*(12), 5972–5984. https://doi.org/10.1111/gcb.14447

Park, I. W., Ramirez-Parada, T., & Mazer, S. J. (2021). Advancing frost dates have reduced frost risk among most North American angiosperms since 1980. *Global Change Biology*, *27*(1), 165–176. https://doi.org/10.1111/gcb.15380

Pearson, K. D., Love, N. L. R., Ramirez-Parada, T., Mazer, S. J., & Yost, J. M. (2021). Phenological trends in the California poppy (Eschscholzia californica): Digitized herbarium specimens reveal intraspecific variation in the sensitivity of flowering date to climate change. *Madroño*, *68*(4). https://doi.org/10.3120/0024-9637-68.4.343

Phillimore, A. B., Hadfield, J. D., Jones, O. R., & Smithers, R. J. (2010). Differences in spawning date between populations of common frog reveal local adaptation. *Proceedings of the National Academy of Sciences*, *107*(18), 8292–8297. https://doi.org/10.1073/pnas.0913792107

Prevéy, J. S., Rixen, C., Rüger, N., Høye, T. T., Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., Ashton, I. W., Cannone, N., Chisholm, C. L., Clark, K., Cooper, E. J., Elberling, B., Fosaa, A. M., Henry, G. H. R., Hollister, R. D., Jónsdóttir, I. S., Klanderud, K., Kopp, C. W., … Wipf, S. (2019). Warming shortens flowering seasons of tundra plant communities. *Nature Ecology & Evolution*, *3*(1), 45–52. https://doi.org/10.1038/s41559-018-0745-6

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Ramirez-Parada, T. H., Park, I. W., & Mazer, S. J. (2022). Herbarium specimens provide reliable estimates of phenological responses to climate at unparalleled taxonomic and spatiotemporal scales. *Ecography*, *2022*(10), e06173. https://doi.org/10.1111/ecog.06173

Ramirez-Parada, T. H. et al. (2023). Data and Code for: Plasticity and not adaptation is the primary source of temperature-mediated variation in flowering phenology in North America (Version 1) [Data set]. *Zenodo*. https://doi.org/10.5281/zenodo.8310387

Roslin, T., Antão, L., Hällfors, M., Meyke, E., Lo, C., Tikhonov, G., Delgado, M. del M., Gurarie, E., Abadonova, M., Abduraimov, O., Adrianova, O., Akimova, T., Akkiev, M., Ananin, A., Andreeva, E., Andriychuk, N., Antipin, M., Arzamascev, K., Babina, S., … Ovaskainen, O. (2021). Phenological shifts of abiotic events, producers and consumers across a continent. *Nature Climate Change*, *11*(3), 241–248. https://doi.org/10.1038/s41558-020-00967-7

Sherry, R. A., Zhou, X., Gu, S., Arnone, J. A., Schimel, D. S., Verburg, P. S., Wallace, L. L., & Luo, Y. (2007). Divergence of reproductive phenology under climate warming. *Proceedings of the National Academy of Sciences*, *104*(1), 198–202. https://doi.org/10.1073/pnas.0605642104

Stinchcombe, J. R., Weinig, C., Ungerer, M., Olsen, K. M., Mays, C., Halldorsdottir, S. S., Purugganan, M. D., & Schmitt, J. (2004). A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proceedings of the National Academy of Sciences*, *101*(13), 4712–4717. https://doi.org/10.1073/pnas.0306401101

USDA, NRCS. (2023). The PLANTS Database. http://plants.usda.gov

Vitasse, Y., Hoch, G., Randin, C. F., Lenz, A., Kollas, C., Scheepens, J. F., & Körner, C. (2013). Elevational adaptation and plasticity in seedling phenology of temperate deciduous tree species. *Oecologia*, *171*(3), 663–678. https://doi.org/10.1007/s00442-012-2580-9

Willis, C. G., Ellwood, E. R., Primack, R. B., Davis, C. C., Pearson, K. D., Gallinat, A. S., Yost, J. M., Nelson, G., Mazer, S. J., Rossington, N. L., Sparks, T. H., & Soltis, P. S. (2017). Old Plants, New Tricks: Phenological Research Using Herbarium Specimens. *Trends in Ecology & Evolution*, *32*(7), 531–546. https://doi.org/10.1016/j.tree.2017.03.015

Wolkovich, E. M., Chamberlain, C. J., Buonaiuto, D. M., Ettinger, A. K., & Morales-Castilla, I. (2022). Integrating experiments to predict interactive cue effects on spring phenology with warming. *New Phytologist*, *235*(5), 1719–1728. https://doi.org/10.1111/nph.18269

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. (2012). Warming experiments underpredict plant phenological responses to climate change. *Nature*, *485*(7399), 494–497. https://doi.org/10.1038/nature11014

Wolkovich, E. M., Cook, B. I., & Davies, T. J. (2014). Progress towards an interdisciplinary science of plant phenology: Building predictions across space, time and species diversity. *New Phytologist*, *201*(4), 1156–1162. https://doi.org/10.1111/nph.12599

Wood, S. N. (2017). Generalized additive models: An introduction with R. CRC Press/Taylor & Francis Group.

Wu, Y., & Colautti, R. I. (2022). Evidence for continent-wide convergent evolution and stasis throughout 150 y of a biological invasion. *Proceedings of the National Academy of Sciences*, *119*(18), e2107584119. https://doi.org/10.1073/pnas.2107584119

Zhang, H., Yuan, W., Liu, S., Dong, W., & Fu, Y. (2015). Sensitivity of flowering phenology to changing temperature in China. *Journal of Geophysical Research: Biogeosciences*, *120*(8), 1658–1665. https://doi.org/10.1002/2015JG003112

Zhou, Z., Zhang, K., Sun, Z., Liu, Y., Zhang, Y., Lei, L., Li, Y., Wang, D., Hu, M., Wang, S., Lu, Q., Cui, Y., Zhong, M., Han, S., & Miao, Y. (2022). Lengthened flowering season under climate warming: Evidence from manipulative experiments. *Agricultural and Forest Meteorology*, *312*, 108713. https://doi.org/10.1016/j.agrformet.2021.108713

Zohner, C. M., Mirzagholi, L., Renner, S. S., Mo, L., Rebindaine, D., Bucher, R., Palouš, D., Vitasse, Y., Fu, Y. H., Stocker, B. D., & Crowther, T. W. (2023). Effect of climate warming on the timing of autumn leaf senescence reverses after the summer solstice. *Science*, *381*(6653), eadf5098. https://doi.org/10.1126/science.adf5098

IV. Interspecific variation in onset and duration sensitivity mediates flowering reassembly under warming[3]

A. Abstract

Global warming has caused widespread shifts in plant phenology among species in the temperate zone, but it is unclear how population-level responses will scale to alter the structure of the flowering season at the community level. This knowledge gap exists largely because—while the climatic sensitivity of first flowering within populations has been studied extensively—little is known about the responsiveness of the duration of a population's flowering period. This limits our ability to anticipate how the entire flowering periods of co-occurring species may continue to change under warming. Nonetheless, flowering sensitivity to temperature often varies predictably among species between and within communities, which may help forecast temperature-related changes to a community's flowering season. However, no studies—empirical or theoretical—have assessed how patterns of variation in flowering sensitivity among species could scale to alter community-level flowering changes under warming. Here, I provide a conceptual overview of how variation in the sensitivity of flowering onset and duration among species can mediate changes to a community's flowering season due to warming trends. Specifically, I focused on the effects of differences in i) the mean sensitivity of flowering onset and duration among communities, and ii) the sensitivity of flowering onsets and durations among species flowering sequentially through the season

[3]This chapter has been accepted for publication:
Ramirez-Parada, T. H., Park, I. W., Record, S., Davis, C. C., & Mazer, S. J. (in press). Scaling flowering onset and duration responses among species predicts phenological community reassembly under warming. *Ecosphere*.

within a community. I evaluated the manner and degree in which these forms of between-species variation in sensitivity might affect the structure of the flowering season—both independently and interactively—using simulations, which covered a wide but empirically-informed range of parameter values and combinations representing distinct community-level patterns. These findings predict that communities across the temperate zone will exhibit varied and often contrasting flowering responses to warming across biomes, underscoring that accounting for the temperature sensitivity of both phenological onset and duration among species is essential for understanding community-level flowering dynamics in a warming world.

B. Introduction

Community-level flowering responses to warming are ecologically critical but poorly understood

The flowering season (the portion of the year during which the flowering of most co-occurring species within communities in the temperate zone occurs) is an ecologically critical period that mediates the fitness of plants and of the diverse organisms that depend on floral resources for survival and reproduction. The flowering season is the cumulative product (and an emergent property) of the blooming period of co-occurring plant species, and its structure determines the temporal distribution of floral diversity and abundance within a community. This temporal flowering structure mediates several ecological processes. For example, the local abundance of floral resources determines pollinator flight distances and the amount of pollen or nectar they collect (Zurbuchen et al. 2010, Pope and Jha 2018). Accordingly, floral

resource availability can mediate population growth rates among pollinator species (Roulston and Goodell 2011). For plants, the density of co-flowering species can mediate various ecological processes, including attraction and competition for pollen and seed dispersers, herbivory, flower and fruit predation, or gene flow among and within populations (Aide 1988, Devaux and Lande 2009, Jones and Comita 2010, Gavini et al. 2021). Consequently, the seasonal diversity, distribution, and abundance of floral resources influences a suite of ecological outcomes that can shape the evolution of life-history strategies (Elzinga et al. 2007). Importantly, many wild plants also support important crop pollinators (Morandin and Kremen 2013, Reilly et al. 2020), or produce pollen that is allergenic to humans (Stinson et al. 2018, Oh 2022). Accordingly, changes to the structure of the flowering season could have important implications for agriculture and human health.

In recent decades, warming trends have resulted in widespread shifts in flowering time across the temperate zone that could profoundly alter the network of ecological interactions within communities (Renner and Zohner 2018) and the overall structure of their flowering seasons (Diez et al. 2012, Caradonna et al. 2014, Theobald et al. 2017). Although the potential effects of warming on species interactions are well appreciated, it is unclear how the flowering responses of populations from different species will scale to jointly alter the flowering season at the community level. Few studies have assessed community-level changes to the flowering season (e.g., Jabis et al. 2020, Zhou et al. 2022), and they have largely focused on simple directional changes to season length. Consequently, existing studies encompass few regional floras, and the responses to temperature of more granular (and ecologically critical) characteristics of the season—such as the temporal distribution of the diversity and abundance of species in flower throughout the season—remain unexamined

aside from select study systems (e.g., Caradonna et al. 2014). Addressing these knowledge gaps is difficult because long-term phenological datasets representing enough species to characterize the flowering season of a community, or to measure floral resources throughout the season, are exceedingly rare (Caradonna et al. 2014, Willis et al. 2017). In turn, although remotely sensed spectral data can capture the onset and duration of the growing season across much of the temperate zone, a community's flowering season generates a much weaker spectral signal that is typically undetectable through satellite imagery despite successes in a few systems (e.g., Chen et al. 2019, Dixon et al. 2021).

In principle, changes to community-level flowering patterns caused by warming could be inferred by aggregating the flowering responses of co-occurring species. However, our ability to do so is severely limited by an incomplete understanding of how warming affects the blooming periods of populations within species. To date, most phenological research has focused on the sensitivity to temperature of the onset of the flowering period (i.e., first flowering date, or 'FFD'), with most studies measuring the effects of temperature on the emergence of the first flowering individuals within a population or site (hereafter 'S_{FFD}'; Fig. 1) (Fitter and Fitter 2002, Miller-Rushing and Primack 2008, Wolkovich et al. 2012, Prevéy et al. 2019). In contrast, comparatively little is known about the effects of temperature on the duration of the period over which individuals are observed in flower (i.e., the temperature sensitivity of population-level flowering duration, hereafter 'S_D') (Fig. 1b). At the individual level, the termination of flowering—whose timing relative to flowering onset determines flowering duration—can be induced by different endogenous and environmental factors than those triggering its onset (Nagahama et al. 2018, Gonzalez-Suarez et al. 2020). Consequently, S_{FFD} might not predict the temperature sensitivity of flowering

termination (and by extension, S_D), which has been measured directly in a relatively narrow range of species and biomes (Li et al. 2021, Zhou et al. 2022). Moreover, population-level flowering duration is also driven by variation in flowering onset and duration among individuals, and the drivers of such variation (e.g., microenvironmental differences, genetic variation in flowering plasticity, resources status) are largely unresolved across regions and taxa. These limitations—compounded by a scarcity of long-term observational records of flowering duration of individual populations—preclude inferences of the way in which blooming period responses to temperature among co-occurring species will scale to alter a community's flowering season (Park et al. 2024).

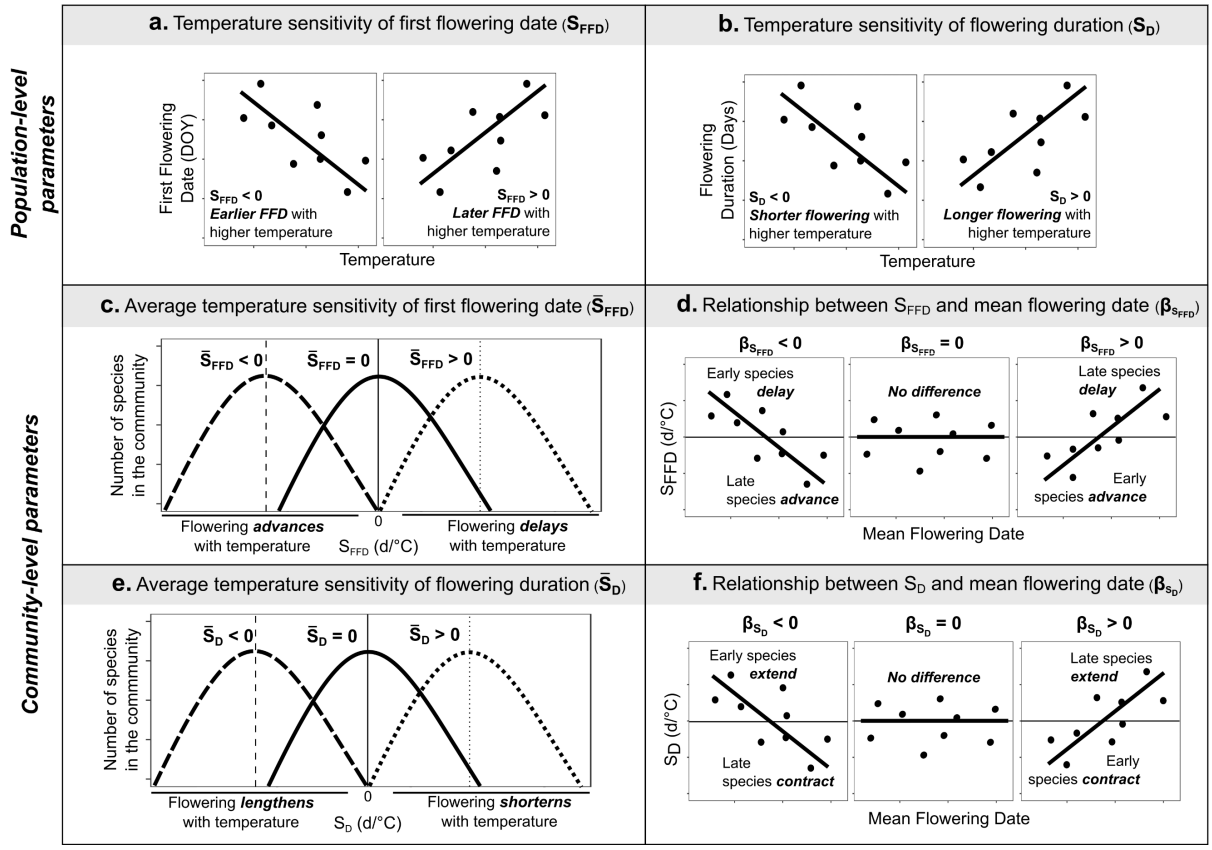

Figure 1—Population-level sensitivities (**a**, **b**) and patterns of variation in sensitivity among co-occurring species (**c**-**f**) that determine the responses of a community's flowering season to

temperature. (**a**) and (**b**) respectively show how a population's first flowering date ('FFD') and the duration of its flowering period change with temperature, with points respectively representing FFDs and durations of a single population across years with different temperatures in the season leading up to flowering. The temperature sensitivity of first flowering dates ('S_{FFD}') is defined as the slope of the relationship between FFD and temperature (**a**), with negative values indicating earlier FFD under higher temperatures, and positive values indicating later FFD under higher temperatures. Similarly, the temperature sensitivity of flowering duration ('S_D';) is measured as the slope of the relationship between duration and temperature (**b**), with negative values indicating shorter flowering periods under higher temperature, and positive values indicating longer flowering under higher temperatures. As population-level sensitivity metrics, S_{FFD} and S_D are mediated by factors affecting the mean and variability of phenological plasticity among individuals, such as genetic or microsite differences. In turn, **c-f** show patterns of among-species variation in S_{FFD} and S_D. (**c**) and (**e**) respectively depict variation among communities in average S_{FFD} and S_D among species ('$\bar{S}_{FFD}$' and '$\bar{S}_D$'), with each curve representing the among-species distribution of sensitivities within communities with different means. Negative values of $\bar{S}_{FFD}$ or $\bar{S}_{FFD}$ respectively indicate that species tend to advance or contract their flowering under warming, whereas positive values respectively indicate that species tend to delay or extend their flowering. Values of 0 for these parameters indicate that, on average, first flowering dates and flowering durations do not tend to change under warming. (**d**) and (**f**) respectively show variation among communities in the relationship between the mean flowering date and both S_{FFD} ($\beta_{S_{FFD}}$) and S_D (β_{S_D}) for species flowering successively throughout the season. Accordingly, points in (**d**) and (**f**) represent species within a community. Negative values of $\beta_{S_{FFD}}$ indicate that S_{FFD} decreases among species as the season progresses. Depending on the community's mean S_{FFD}, these patterns may respectively correspond to a transition from delays to advances between early and late flowering species (i.e., S_{FFD} switches sign), a decrease in the degree of delay under higher temperatures (i.e., S_{FFD} closer to 0), or increases in the degree of advancement under

higher temperatures (i.e., S_{FFD} further from 0 in the negative direction). In turn, positive $\beta_{S_{FFD}}$ indicate that S_{FFD} increases among species as the season progresses (with an interpretation opposite to that of $\beta_{S_{FFD}}$<0). Similarly, negative values of β_{S_D} indicate that S_D decreases among species as the season progresses, which can correspond to a transition from flowering extensions to contractions among species throughout the season (i.e., S_D switches sign), decreases in the degree of flowering extension (i.e., S_D becomes less positive), or increases in the degree of flowering contraction (i.e., S_D becomes more negative). Positive $\beta_{S_{FFD}}$ then indicate increases in S_D as the season progresses. $\beta_{S_{FFD}}$ or β_{S_D} equal to 0 respectively indicate no difference in average S_{FFD} and S_D among early- and late-flowering species. For simplicity, (**d**) and (**f**) respectively depict scenarios in which $\bar{S}_{FFD}$ and $\bar{S}_D$ equal 0.

Community-level flowering responses to warming can be predicted from non-random variation in sensitivity among co-occurring species

A recent surge of phenological research has established that flowering sensitivity typically varies non-randomly among species within and among communities, which can be leveraged to predict how flowering at the community level will change under ongoing warming trends (Figs. 1c-f). For example, due to species turnover across regions or variation in sensitivity within species ranges, the mean S_{FFD} observed among co-occurring species (hereafter '$\bar{S}_{FFD}$') often differs between communities (Zhang et al. 2015, Prevéy et al. 2017, Park et al. 2019) (Fig. 1c). Moreover, many studies have established that co-occurring species that flower at different times within a season tend to show marked differences in the climate sensitivity of their first (and peak) flowering dates (i.e., in S_{FFD}) (e.g., Wolkovich et al. 2012, Mazer et al. 2013, Ramirez-Parada et al. 2023). In relatively mesic communities, such directional

variation in S_{FFD} among species throughout the season (hereafter '$\beta_{S_{FFD}}$'; Fig. 1d) typically consists of greater advances under warming among early-flowering species than among late-flowering species (Cook et al. 2012). In turn, high-latitude and high-elevation communities show more variable patterns that include greater advances among late-flowering species (Prevéy et al. 2019). Meanwhile—and although still rare compared to research on flowering onset—studies of population-level flowering duration have reported substantial variation in mean S_D among species in different communities (hereafter '$\bar{S}_D$'; Fig. 1e), with some communities showing average increases and others average decreases in flowering duration among species in response to warming (Nagahama et al. 2018, Chen et al. 2020, Hu et al. 2020, Huang et al. 2020, Jabis et al. 2020, Nam and Kim 2020). Some communities also exhibit variation in S_D among species that flower at different times throughout the season (hereafter 'β_{S_D}'; Fig. 1f), with some exhibiting greater sensitivity among early-flowering and others among late-flowering species (Li et al. 2020, Chen et al. 2022).

Such forms of non-random variation in S_{FFD} and S_D among species should have distinct impacts on the structure of a community's flowering season under warming (depicted in simplified hypothetical scenarios in Fig. 2). For example, for communities with an identical flowering structure before warming (Fig. 2a)—and assuming no other forms of variation in S_{FFD} and S_D among species (i.e., $\bar{S}_D$, $\beta_{S_{FFD}}$, and β_{S_D} all equal 0)—$\bar{S}_{FFD}$ should determine whether the FFDs of co-occurring species tend to advance or delay under higher temperatures. In such cases, warming would significantly alter the start and end of the season but not the relative distribution of the flowering periods of different species (Fig. 2b). In contrast, directional variation in S_{FFD} as the season progresses ($\beta_{S_{FFD}}$) should mediate the degree to which the blooming periods of species tend to converge or diverge within the

season (Fig. 2c). For example, $\beta_{S_{FFD}} > 0$ would tend to spread the flowering periods of co-occurring species, advancing the start and delaying the end of the season (thus lengthening it), decreasing flowering overlap among species, and therefore the mean richness of flowering species throughout the season. In turn, greater advances among late-flowering species would concentrate the blooming periods of different species, shortening the season and increasing both the degree of flowering overlap and the average richness of co-flowering species throughout the season. Among-species variation in S_D should have distinct effects from those generated by among-species variation in S_{FFD}. A community's $\bar{S}_D$ would affect the timing of flowering termination among species, predominantly altering their degree of flowering overlap, the richness of co-flowering species (or the cumulative flowering intensity of the community), and the duration of the season (Fig. 2d). In turn, directional variation in S_D throughout the season (β_{S_D}) would most strongly impact the degree of flowering overlap among species and the richness of co-flowering species within it (Fig. 2e).

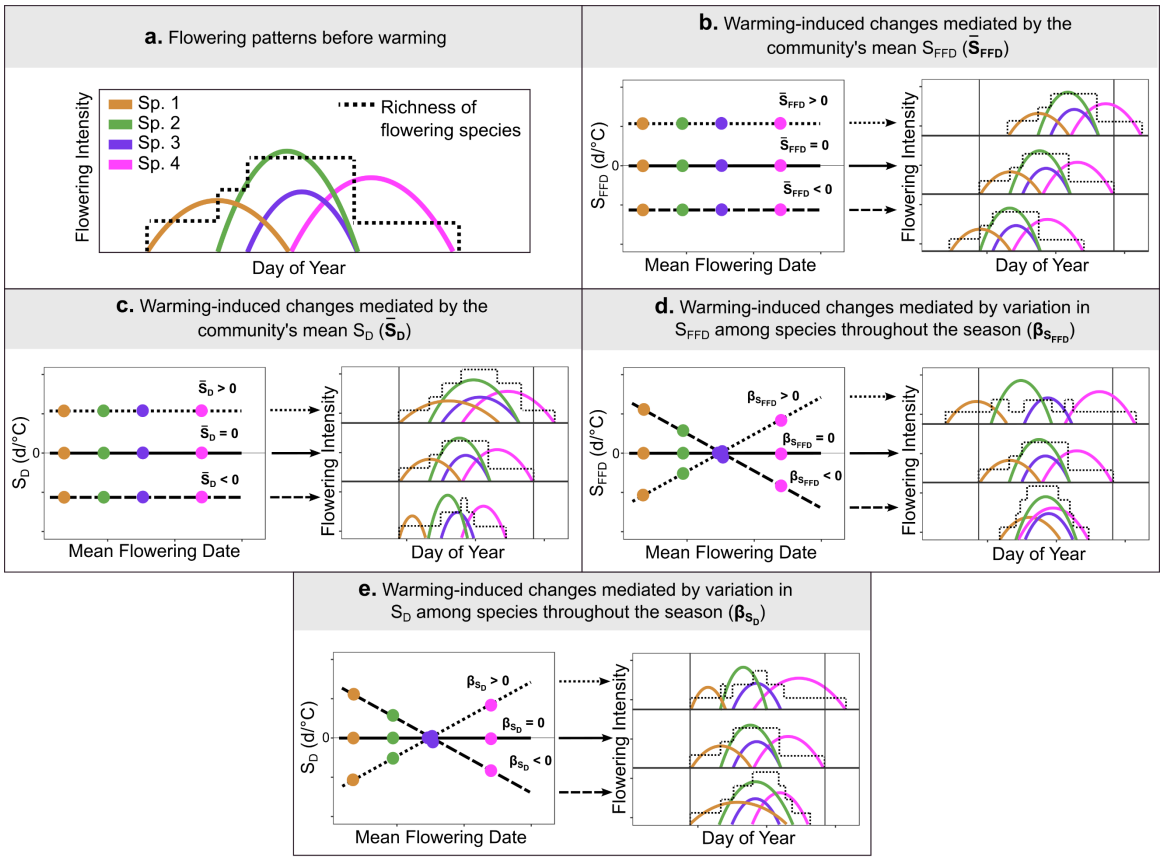

Figure 2—Community-level flowering responses to warming mediated by different forms of non-random variation in S_{FFD} and S_D among co-occurring species. (**a**) depicts a hypothetical community of 4 species, with each colored line representing the flowering intensity (interpretable as the number of individuals flowering or the cumulative flowering output of the population) each day of the season. In turn, **b-e** show various forms of non-random variation in S_{FFD} or S_D among species, and the responses to warming that they would generate for the community depicted in (**a**). Arrows between panels in **b-e** connect each form of among-species variation in S_{FFD} and S_D to the community-level flowering pattern that would emerge under warming. In each scenario, the community is assumed to exhibit only the focal form of among=species variation in S_{FFD} and S_D (i.e., other parameters describing non-random variation among species are assumed to equal 0). Vertical, solid lines in the right-hand panels of **b-e** indicate the start and end of the flowering season for the community before warming (for reference).

These forms of interspecific variation in sensitivity can generate non-analog flowering events in response to warming, with previously asynchronous species now overlapping (e.g., spp. 1 and 4, Figs. 2c, d, e), previously synchronous species no longer overlapping (e.g., spp. 1 and 3; Figs. 2c, d, e), and species flowering partially outside of (or no longer in) the community's historical flowering season (Figs. 2b-e). Such non-analogous synchrony patterns may therefore reassemble the network of interactions mediated by flowering (e.g., pollinator attraction or competition) among species within a community (Theobald et al. 2017).

While predicting community-level responses is simple when considering these patterns of variation in isolation, realized changes to the structure of the flowering season will depend on how these community-level attributes jointly mediate variation in S_D and S_{FFD} within communities. Moreover, the direction and magnitude of such changes are not intuitively obvious. For example, if late-flowering species advance their flowering but delay flowering termination, it is unclear whether the end of the season would delay or advance. Moreover, directional changes in season length or flowering peaks do not necessarily indicate whether floral diversity and abundance—or the network of flowering overlap among species—might increase or decrease during different parts of the season. It is also difficult to anticipate how these types of non-random variation in sensitivity among species might operate in more complex communities (e.g., temperate communities can harbor hundreds of species) or what their impact would be for more granular attributes of the flowering season (such as the community's cumulative flowering output each day of the season). To my knowledge, no studies—empirical or theoretical—have assessed whether and how non-

random variation in S_{FFD} and S_D determines temperature-related changes to the structure of the flowering season at the community level. This is largely because few datasets document the duration of the flowering period of plant populations across multiple years—which is needed for measuring the sensitivity of flowering duration to interannual temperature variation—and to my knowledge, none have done so comprehensively among co-occurring species within a community.

Simulations enable scaling of flowering responses from species to communities

Studies examining community-level flowering responses to climate are rare and have predominantly measured simple attributes of the season (e.g., season length), limiting our understanding of how the seasonal distribution of floral diversity and abundance will change in response to ongoing warming across regions. Assessing this knowledge gap empirically is difficult due to a scarcity of long-term datasets on the phenological sensitivity of both flowering onset and termination that also include enough species to reconstruct community-level patterns.

To circumvent these limitations, I used computer simulations to evaluate how patterns of non-random variation in S_{FFD} and S_D among species—some of which have been extensively reported in the literature—independently and jointly mediate the seasonal distribution of flowering species and the structure of flowering overlaps among species within a community. I modeled three forms of among-species variation in sensitivity that have been documented in the literature: (1) the degree and direction in which S_{FFD} varies among species flowering successively throughout the season (i.e., $\beta_{S_{FFD}}$), (2) the degree to

which co-occurring species tend to shorten or extend their flowering periods under warming (i.e., mean S_D among species, or $\bar{S}_D$), and (3) the degree and direction of variation in S_D among species flowering successively throughout the season (i.e., β_{S_D}) (Fig. 1). I excluded $\bar{S}_{FFD}$ from the simulations because, although it affects the overall timing of the flowering season within the year, it does not alter the timing of each species' flowering period relative to the rest of the community, nor the emergent community structure (Fig. 2b, Appendix 3—Fig. S1). By simulating communities under a range of values for $\beta_{S_{FFD}}$, $\bar{S}_D$, and β_{S_D}, I explored three general scenarios: one in which S_{FFD} and S_D vary independently among species (Scenario 1; Fig. 3a), and two others in which S_D and S_{FFD} are correlated either positively (Scenario 2; Fig. 3b) or negatively (Scenario 3) due to their shared covariation with the timing of flowering of a species within the season (Fig. 3; Fig. 3c).

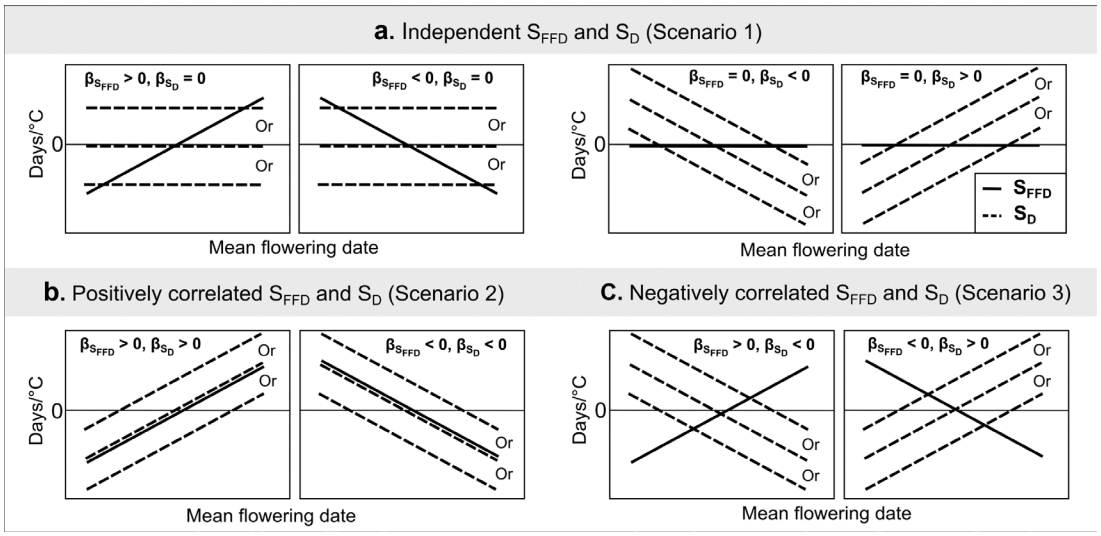

Figure 3—Combinations of $\beta_{S_{FFD}}$ and β_{S_D} values generating scenarios in which S_{FFD} and S_D are (**a**) independent (scenario 1) (**b**) positively correlated (scenario 2) or (**c**) negatively correlated (scenario 3) across species. Across panels, multiple dashed lines depicting the relationship between S_D and mean flowering date are shown because communities may differ in the mean sensitivity of S_D observed among species (i.e., $\bar{S}_D$), which changes the y-intercept of the relationship without altering its slope.

The goals of using simulations were multifold. First, I evaluated how species-level responses scale to alter community-level attributes beyond simple directional responses to season length. Such attributes included changes to the diversity and abundance of flowering each day of the flowering season, and to the network of flowering overlaps within the community (which mediates the potential for flowering-mediated species interactions). Second, I was able to concurrently assess the effects of flowering onset and duration sensitivities, identifying their distinct impacts during different portions of the season and on specific types of community-level flowering responses. Finally, simulating empirically documented forms of among-species variation in sensitivity for a wide range of parameter combinations enabled us to explore a much wider range of scenarios—or types of communities—than that captured by the few empirical studies conducted to date.

Overall, the simulations demonstrate that non-random variation in S_{FFD} and in S_D among species (modeled by $\beta_{S_{FFD}}$, $\bar{S}_D$, and β_{S_D}) is likely to have profound and predictable impacts on the structure of the flowering season, underscoring i) the importance of measuring the climatic responsiveness of phenological onsets and durations and ii) that characterizing patterns of among-species variation will help predict changes in community-level flowering patterns and their potential ecological consequences under a changing climate. Importantly, recent research suggests that directional variation in temperature sensitivity among species active at different times throughout the season—such as that described in this paper—might be common among producers, primary and secondary consumers, and saprotrophs (Roslin et al. 2021). Therefore, the modes of change described here might have implications for

understanding community-level phenological changes beyond plants and flowering phenology.

C. Methods

Simulation Design

Assigning community-level parameters

The simulations assessed changes to the flowering season under warming across communities differing in the structure of among-species variation in S_{FFD} and S_D, which I modeled using three parameters ($\beta_{S_{FFD}}$, $\bar{S}_D$, and β_{S_D}; Fig. 1). $\beta_{S_{FFD}}$ modeled the degree and direction in which S_{FFD} (measured in days/°C) varied among species flowering successively throughout the season as a linear relationship between each species-specific S_{FFD} and mean flowering date (Fig. 4a). Therefore, a negative $\beta_{S_{FFD}}$ indicates decreasing S_{FFD} as the season progresses, whereas a positive $\beta_{S_{FFD}}$ indicates increasing S_{FFD}. Because mean S_{FFD} was set to 0 for all communities (as it did not affect the structure of the flowering season; Fig 2b, Appendix 3—Fig. S1), negative $\beta_{S_{FFD}}$ implies a transition from positive S_{FFD} (i.e., flowering delays under warming) to negative S_{FFD} (i.e., flowering advances under warming) between early- and late-flowering species in a community. Each simulated community (described below) was assigned one of five $\beta_{S_{FFD}}$ values ranging from -0.1 (i.e., S_{FFD} decreases by 1 day/°C for every 10-day increase in mean flowering date among species) to 0.1 (S_{FFD} increases by 1 day/°C for every 10-day increase in mean flowering date) in 0.05 increments. I approximately set the mid-range of values for $\beta_{S_{FFD}}$ (-0.05 to 0.05) based on studies that have evaluated the relationship between mean flowering dates and S_{FFD} empirically (e.g., Fitter

and Fitter 2002, Cook et al. 2012, Mazer et al. 2013, Wolkovich et al. 2012, Prevéy et al. 2019, Park et al. 2019). As most studies comparing phenological sensitivities have been conducted within mesic regions of North America and Europe, I included more extreme values of $\beta_{S_{FFD}}$ (-0.1 and 0.1) to account for the possibility of more extreme relationships between S_{FFD} and mean flowering date in unstudied systems.

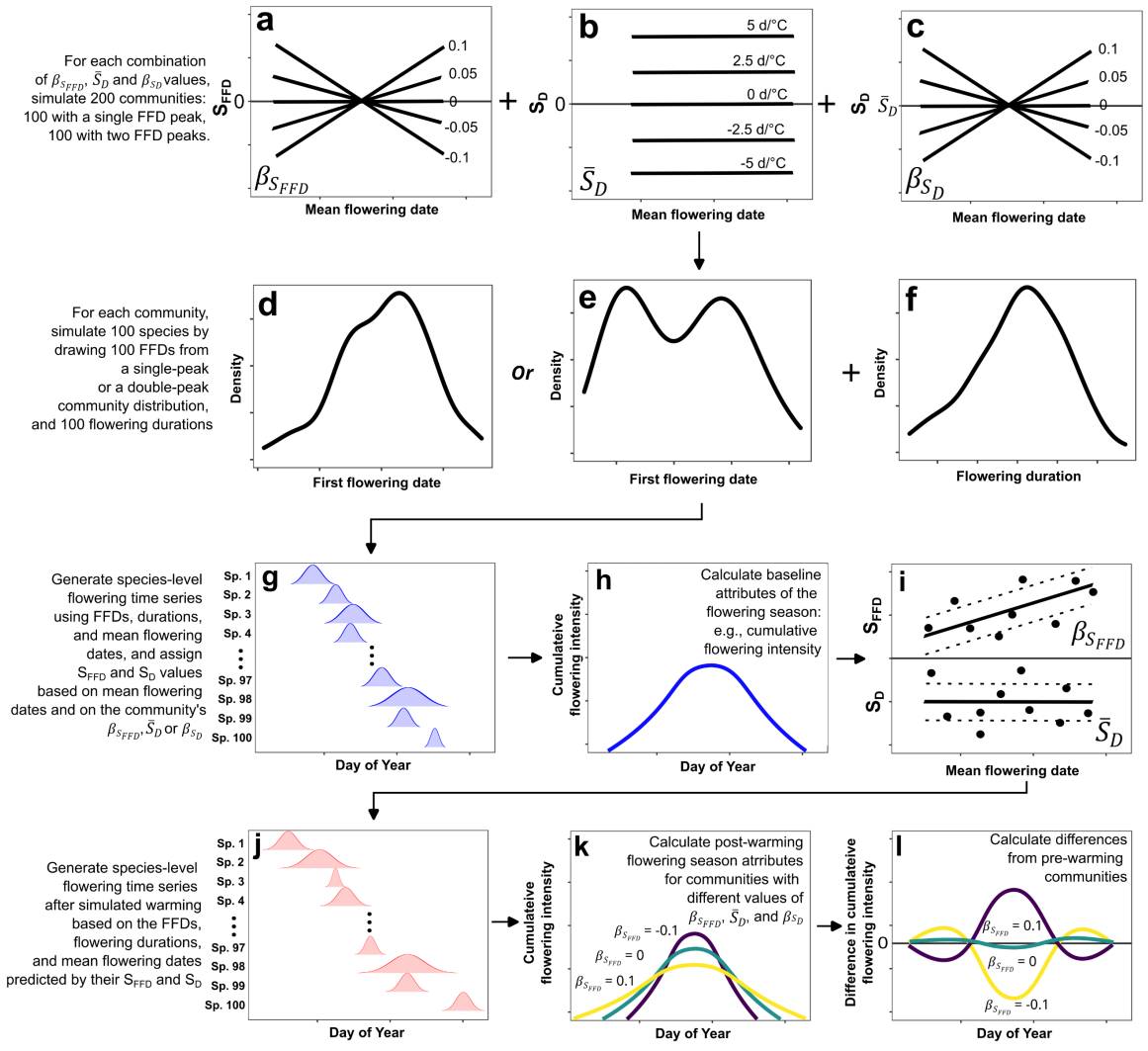

Figure 4—Diagram of the simulation design. For each combination of community-level parameter values describing how the temperature sensitivity of first flowering date (S_{FFD}) and of flowering duration (S_D) varies among species in a community (**a**, **b**, **c**), 200 communities were simulated: half

showing a single seasonal peak of first flowering dates (FFDs) among species, and half showing two peaks. For each community, 100 species were simulated by combining random first flowering dates (drawn from either unimodal or bimodal distributions) and flowering durations (**d**, **e**, **f**). For each species, population-level flowering time series were generated from their FFDs and durations assuming a bell-shaped flowering curve (**g**), and the flowering season attributes were calculated for each community (**h**). Then, each species was assigned an S_{FFD} and an S_D based on the relationship between these sensitivities and mean flowering date determined by its community's $\beta_{S_{FFD}}$, $\bar{S}_D$, and β_{S_D} values (**i**) prior to generating new flowering periods predicted by each species' S_{FFD} and S_D under 2 °C warming (**j**). Flowering season attributes were recalculated for each post-warming community (**k**) and compared to the preseason baseline (**l**).

I modeled variation among communities in both the mean S_D observed among species (i.e., $\bar{S}_D$) and in the degree to which S_D varied among species flowering successively throughout the season. $\bar{S}_D$ indicates whether (and to what degree) a community shows average decreases or increases in flowering duration under warming (Fig. 4b). Each community was assigned one of five $\bar{S}_D$ values ranging from -5 d/°C (i.e., mean decreases in the duration of the flowering period of 5 days per °C among species) to 5 d/°C (i.e., mean duration increases of 5 days per °C) in 2.5 d/°C increments. In this scenario, S_{FFD} and S_D are assumed to be uncorrelated among species. In turn, I modeled $\beta_{S_{FFD}}$ as a linear function of mean flowering date throughout the season, with values ranging from -0.1 (i.e., S_D decreases by 1 day/°C for every 10-day increase in the mean flowering date of a species) to 0.1 (S_D increases by 1 day/°C for every 10-day increase in mean flowering date) in 0.05 increments (Fig. 4c). Again, the range of possible values for β_{S_D} was roughly approximated from the few

studies that have quantified among-species variation in S_D throughout the season (e.g., Chen et al. 2020, Li et al. 2020), including more extreme values to acknowledge the possibility of stronger relationships (values of -0.1 and 0.1) than those documented to date. As communities differed both in β_{S_D} and $\bar{S}_D$, negative β_{S_D} can indicate a transition between early- and late-flowering species from positive to negative S_D (i.e., flowering lengthening to contraction), decreases in the degree of flowering lengthening (i.e., lower but still positive S_D among late-flowering species), or increases in the degree of contraction among late-flowering species (i.e., more negative S_D). In turn, positive β_{S_D} can indicate the opposite of each of these patterns (i.e., a switch from contractions to lengthening, increases in the degree of lengthening, or decreases in the degree of contraction between early- and late-flowering species).

Different combinations of $\beta_{S_{FFD}}$ and β_{S_D} generated distinct relationships between S_{FFD} and S_D, with a value of 0 for $\beta_{S_{FFD}}$ or β_{S_D} generating independent S_{FFD} and S_D (Scenario 1; Fig. 3a), non-zero $\beta_{S_{FFD}}$ and β_{S_D} of the same sign generating positively correlated S_{FFD} and S_D through their congruent relationship with mean flowering date (Scenario 2; Fig. 3b), and opposite signs generating negatively correlated S_{FFD} and S_D through their discordant variation throughout the season (Scenario 3; Fig. 3c).

Simulating pre-warming communities—For each of 125 combinations of parameter values (5 × 5 × 5 combinations of $\beta_{S_{FFD}}$, $\bar{S}_D$, and β_{S_D} values), I simulated 200 initial pre-warming communities for a total of 25,000 simulated communities. In each community, 100 species were simulated by first generating a random sample of first flowering dates (hereafter,

"FFD") drawn from one of two alternative types of community: those with a single flowering peak (i.e., unimodal distribution; 100 communities per parameter value combination; Fig. 4d) or those with two weaker flowering peaks (i.e., bimodal distribution; 100 communities per combination; Fig. 4e). These community types emulate flowering patterns known to occur in North American plant assemblages; unimodal patterns have been reported in mesic communities, while bi-modal patterns have been described in semi-arid assemblages experiencing summer monsoons and in subalpine communities (e.g., Diez et al. 2012, Caradonna et al. 2014). FFDs for the single-peak communities were obtained from a normal distribution centered on May 31st, with an SD of 30 days. In turn, FFDs for the bimodal communities were obtained from a joint distribution combining two truncated normal distributions: one bound between Jan 1st—May 21st, with a mean (or peak) on April 30th and an SD of 30 days, and another bound between May 22nd—Dec 31st with mean DOY on Jun 20th, and an SD of 30 days. Using normal distributions assumes that species within each community do not exhibit significant skew in FFD. Although many communities across the temperate zone likely do not conform to this assumption, the degree of skew in the distribution of FFDs among co-occurring species has not been widely characterized across biomes. The choice of normal distributions is agnostic about the prevalence of right- vs. left-skewed species-level FFD distributions across communities (or of greater first vs. second flowering peaks in the case of bimodal communities), and I thus consider it an appropriately conservative assumption.

Once FFDs were generated, I randomly assigned a flowering duration (in days) to each species by drawing values from a truncated normal distribution (lower bound = 14 days, mean = 60 days, SD = 10 days; Fig. 4f). Parameters for this truncated normal distribution

were chosen to generate a plausible range of variation in population-level flowering durations among species, which typically ranged from 14 to 90 days among species within a community. I then calculated flowering termination dates by adding the flowering duration of each species to its FFD (subtracting one). This approach assumes that the duration of the flowering period of a population is independent of its date of onset.

Next, the intensity of flowering of each species' population throughout its flowering period was modeled under the simplifying assumption that the time series of flowering intensity—interpretable either as the proportion of individuals in flower, the mean number of flowers produced per individual, or the floral output of the population each day relative to its peak—was bell-shaped with a peak at the median flowering date of each species (Fig. 4g). I implemented this by modeling the amplitude of each species' flowering period using truncated normal distributions bound by the flowering onset and termination dates for each species, with the median date between onset and termination as the mean, and SD equal to a third of the duration of the flowering period. To make the amplitude of the time series interpretable—and to make the range of variation in the amplitude of the flowering period among species comparable—I scaled each time series to a maximum amplitude of 1 at its peak. Consequently, for a given species, each point in the time series indicates the species' proportional flowering intensity relative to its flowering peak. As with the choice of distribution to model among-species variation in FFD, the choice of a bell-shaped distribution of flowering intensity during the flowering period of each species is agnostic about the prevalence of left- vs. right-skewed flowering among species, which is poorly documented across biomes.

Flowering season baselines prior to simulated warming— To generate a baseline against which to compare the structure of the flowering season after warming for every simulated community, I calculated: (1) the cumulative intensity of flowering across all species on each day throughout the growing season (defined below), (2) the length of the flowering season, (3) the amplitude of the season's flowering peak, and (4) the degree of flowering overlap between each pair of species in a community (Fig. 4h).

The cumulative intensity of flowering for each day throughout the season was calculated as the sum of the amplitudes of flowering curves across species. Because the amplitude of the flowering curve of a species ranges from 0 (no flowering) to 1 (its flowering peak), this community-level metric is bounded by 0 (no species flowering on that date) and a theoretical maximum equal to the number of species in the community (100 in all simulations), which would represent a scenario in which all species reach their flowering peaks on the same date. Accordingly, the more synchronous the flowering period of co-occurring species within a community, the greater its expected peak cumulative flowering intensity. Through this method, each species contributes to the cumulative intensity of the season in proportion to the intensity of its flowering on a given date. Consequently, this approach does not account for potential differences in the ecological importance of each species' flowering that may originate, for example, from variation among species in abundance or in the amount and quality of floral resources provided by individuals. Finally, the length of the season was calculated as the uninterrupted period during which at least 10 species were observed flowering, and the amplitude of the flowering peak as the maximum cumulative flowering intensity observed throughout the season.

Within each community, I measured flowering overlap among each pair of species (for 100 species, 4,950 unique pairs within each community). To do so, the flowering distribution of each species in a community was re-standardized so that it integrated to 1. Doing so ensured that the area under any segment of a population's flowering curve corresponded to the proportion of the total flowering effort that was observed during that period. Then, I measured the area under the intersection of the flowering curves of each pair of species. Therefore, the value of the flowering intersection between two species ranged from 0 (no overlap) to 1 (identical flowering curves).

Flowering season under warming— Prior to generating post-warming flowering distributions, each species was assigned temperature sensitivities for FFD (S_{FFD}) based on its mean flowering date and its community's $\beta_{S_{FFD}}$, and temperature sensitivities for duration based on its community's $\bar{S}_D$ and β_{S_D} parameters (as well as its mean flowering date) (Fig. 4i). S_{FFD} was assigned to each species using a normal distribution whose expected value for each species (i.e., its mean) was equal to the product between the slope parameter $\beta_{S_{FFD}}$ and the species' mean flowering date. Prior to computing these expected values, I centered mean flowering dates around the median species in each community, which resulted in an among-species distribution of S_{FFD} centered around 0 for all simulated communities. As previously noted, mean S_{FFD} was set to 0 because its value determined the extent to which the entire flowering season advanced or delayed without altering its internal structure (Fig. 2; Appendix 3—Fig. S1)

Additionally, by doing this, I ensured that pre- and post-warming flowering seasons for a community were centered around the same date, making it easier to isolate changes to the structure of the flowering season resulting from non-random variation in S_{FFD} among species. Once the expected S_{FFD} for each species was calculated, I obtained a randomized value around the predicted mean using a random error (i.e., standard deviation) of 3 d/°C. The magnitude of the random error was selected to approximate the level of scatter observed around the relationship between S_{FFD} and mean flowering date among species reported in empirical studies (e.g., Cook et al. 2012, Mazer et al. 2013, Wolkovich et al. 2012, Prevéy et al. 2019).

When $\beta_{S_D} = 0$ (Fig. 1a, b), S_D values were drawn randomly from a normal distribution with mean equal to $\bar{S}_D$ (ranging from -5 to 5 d/°C) and standard deviation equal to 3 d/°C. In turn, for scenarios where $\beta_{S_D} \neq 0$ (Fig. 1c-h), S_D values for each species were obtained from a normal distribution with mean equal to $\bar{S}_D$ plus the product of β_{S_D} and mean flowering date for each species, and a standard deviation of 3 d/°C.

Then, for each species in each community, post-warming FFDs and durations were generated based on (1) their initial FFDs and durations, and (2) their S_{FFD} and S_D (Fig. 2j) as:

$$(1) \qquad FFD_{post} = FFD_{pre} + S_{FFD} \times Warming$$

$$(2) \qquad Duration_{post} = Duration_{pre} + S_D \times Warming$$

I used a warming level of 2 °C for all communities because, as many regions in North America have already experienced temperature increases of about 1.5 °C, it reflects a scenario already occurring in warm years across the temperate zone (IPCC 2022, Chapter 14). Moreover, some species have documented decreases in phenological sensitivity to

temperature under warmer conditions (Fu et al. 2015), with declines often occurring beyond warming thresholds of 2 °C (Guo et al. 2023). Therefore, the assumption that phenology and temperature are linearly related is reasonable when simulating responses under a warming level of 2 °C. Nonetheless, I also ran simulations with temperature increases of 4 °C to determine whether qualitatively different flowering season responses may emerge with higher temperature increases.

From the post-warming FFDs and durations, I obtained flowering termination and median dates that were used to generate flowering periods for each species through the same procedure described for pre-warming communities (Fig. 4j). I calculated the same attributes of the flowering season for communities after warming as for each pre-warming community: cumulative flowering intensity throughout the season, season length, and peak amplitude, and pairwise flowering overlap among species (Fig. 4k). Then, I measured warming-induced changes in the composition of flowering overlaps among species in each simulated community using the Bray-Curtis Dissimilarity Index (henceforth 'BCI') (Bray & Curtis, 1957). The BCI is typically used to measure the degree of dissimilarity in species composition between communities or sites, accounting both for differences in the identity of species present and their abundance. More broadly, however, the BCI measures the compositional differences between two sets of observations of categorical entities, and therefore can be broadly used to measure compositional dissimilarity for data other than species surveys. In these data, the categorical entities corresponded to each pair of species within a community (4,950 unique pairs from among 100 unique species in each community) weighted by their degree of flowering overlap. Accordingly, the BCI measured compositional dissimilarity in flowering overlaps pre- and post-warming for each simulated community

accounting for changes in the identity and degree of overlap for each unique pair of species, with values of 0 indicating complete similarity pre- and post-warming for a community (i.e., same identity and degree of overlap among species pairs) to 1 (complete mismatch in the identity of overlapping pairs).

Finally, I summarized how variation in S_{FFD} and S_D among species affected community-level flowering season outcomes under warming by aggregating results from simulated communities generated using each combination of $\beta_{S_{FFD}}$, $\bar{S}_D$, and β_{S_D} values (Fig. 4l). Specifically, the range of outcomes—for each metric used—was calculated for the central 90% and the median of simulated communities for each pair of parameter values. To assess whether the influence of $\beta_{S_{FFD}}$, $\bar{S}_D$, and β_{S_D} differed depending on the distributions of FFDs among species in a community, changes in season length, peak amplitude, cumulative flowering intensity throughout the season, and turnover in species overlap were conducted separately for single-peak and bimodal simulated communities

D. Results

Scenario 1: independent variation in S_{FFD} and S_D among species

Non-random variation among species in S_{FFD} and S_D (represented by $\beta_{S_{FFD}}$, $\bar{S}_D$, and β_{S_D}) determined the magnitude and direction of changes to the flowering season under warming. As expected (Fig. 2), communities in which early species showed advances and late species delays in flowering ($\beta_{S_{FFD}} > 0$) tended to show season lengthening and less pronounced flowering peaks, whereas those showing the opposite patterns ($\beta_{S_{FFD}} < 0$) tended to show season contraction and amplified flowering peaks (Appendix 3—Fig. S2a, b). Also

expectedly, among communities in which S_D varied with species' mean flowering dates (i.e., $\beta_{S_D} \neq 0$), those in which early species tended to contract and late species to extend flowering showed season lengthening (or less shortening, depending on their $\bar{S}_D$) and less pronounced flowering peaks, whereas communities in which late species tended to contract and early species to extend flowering showed the opposite pattern (Fig. S2b, d). The range of variation among communities due to differences in β_{S_D}, however, was much narrower than that generated by differences in $\beta_{S_{FFD}}$ (e.g., Appendix 3—Fig. S2b vs S2d). Ultimately, changes to the amplitude and date of seasonal flowering peaks relative to the pre-warming baseline were determined by non-random variation in both S_{FFD} and S_D among species (Appendix 3—Fig. S3), with both sensitivity types either offsetting or exacerbating the effects of the other.

Changes to the distribution of flowering diversity and abundance throughout the season—which I measured as cumulative flowering intensity, interpretable also as community-wide flowering synchrony— were jointly influenced by $\bar{S}_D$, $\beta_{S_{FFD}}$, and β_{S_D}, but their precise effects differed among the early, mid, and late season (Fig. 5). For example, changes due to warming during the early flowering season (percentiles 0—33) differed only among communities with distinct values of $\beta_{S_{FFD}}$, with early-season flowering intensity increasing among communities in which FFDs advanced with warming among early species and delayed among later flowering species (note yellow line) and declining among communities exhibiting the opposite pattern (dark purple line). In contrast, cumulative changes during mid and late season varied widely with $\bar{S}_D$ and β_{S_D} (in addition to $\beta_{S_{FFD}}$) among communities. Communities showing average decreases in flowering duration among species under warming ($\bar{S}_D < 0$) tended to show decreases in cumulative flowering intensity

during the mid and late season, with variation in $\beta_{S_{FFD}}$ mediating their magnitude (Fig. 5a). In turn, communities whose species showed average increases in duration ($\bar{S}_D > 0$) showed increases in cumulative flowering intensity across much of the flowering season and whose magnitude was mediated by $\beta_{S_{FFD}}$ (Fig. 5c). Variation in S_D throughout the season (β_{S_D}; Fig. 5d-f) primarily impacted the timing of changes in cumulative flowering intensity mediated by $\bar{S}_D$. Among communities with $\bar{S}_D < 0$, those in which early species contracted flowering more than late species ($\beta_{S_D} > 0$) showed decreases in cumulative flowering intensity earlier in the season (yellow line) than communities where late species contracted more ($\beta_{S_D} < 0$, purple line; Fig. 5d). The converse was true among communities with $\bar{S}_D > 0$: those with $\beta_{S_D} > 0$ showed increases in cumulative flowering intensity later in the season (yellow line) than those with $\beta_{S_D} < 0$ (purple line; Fig. 5f). Communities with bimodal flowering distributions showed the same pattern of change in cumulative flowering intensity throughout the season as those with a single flowering peak (Appendix 3—Figs. S4, 5). Similarly, simulating flowering seasons under 4 °C resulted in intensified but qualitatively identical patterns of change (Appendix 3—Fig. S6).

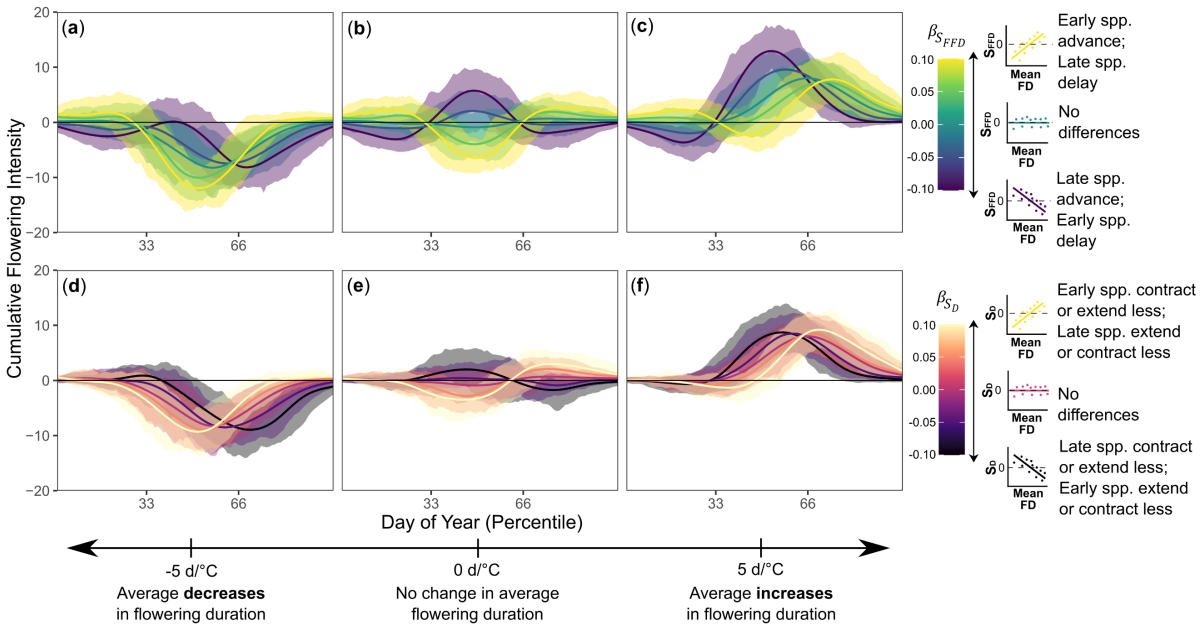

Figure 5—Changes in cumulative flowering intensity given uncorrelated variation between S_{FFD} and S_D within communities. Solid-colored lines in **a-c** correspond to the median change in cumulative intensity—across percentiles of the flowering season—due to warming compared to the pre-warming baseline among simulated communities with the same combination of $\beta_{S_{FFD}}$ and $\bar{S}_D$ values (and for which $\beta_{S_D} = 0$; see Figs. 1a,b). Solid-colored lines in **d-f** correspond to the median change in cumulative intensity due to warming compared to the pre-warming baseline among simulated communities with the same combination of β_{S_D} and $\bar{S}_D$ values (and for which $\beta_{S_D} = 0$; see Figs. 1c,d). Shaded regions indicate the 90% range of variation across communities grouped by each combination of parameter values. Scenarios of $\beta_{S_{FFD}} = 0$ or $\beta_{S_{FFD}} = 0$ respectively correspond to communities where S_{FFD} or S_D varied randomly (around the community mean) among species flowering successively throughout the season (middle curves in each panel).

Variation in S_{FFD} and S_D also altered the network of flowering overlaps among species (Fig. 6). Communities showing average decreases in flowering duration among species ($\bar{S}_D < 0$) tended to show the greatest changes in the composition of flowering overlaps, with the greatest differences observed among those also showing flowering advances among early-flowering species and delays among late-flowering species (Fig. 6a), or also showing flowering contractions among early species and flowering extension among late flowering species (Fig. 5b). Nonetheless, the composition of flowering overlaps changed substantially even among communities showing no average changes in flowering duration among species ($\bar{S}_D = 0$) and showing no average differences in S_D and S_{FFD} among early vs. late flowering species (i.e., those communities showing random variation in S_D and S_{FFD} among species, or $\beta_{S_{FFD}} = 0$ and $\beta_{S_D} = 0$; the middle group of Fig. 5a, b).

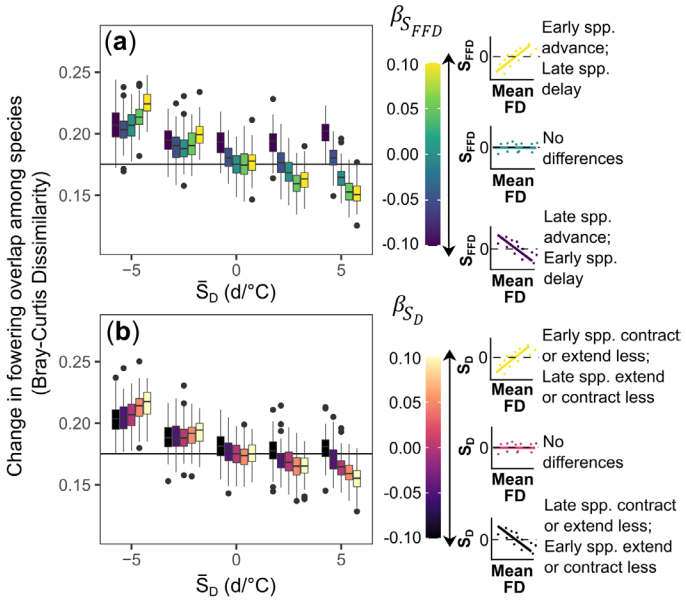

Figure 6—Changes in the composition of pairwise flowering overlaps among species due to warming across communities under scenarios of independent variation among species in S_{FFD} and S_D. In each community, change in the composition of flowering overlaps was measured using the Bray-Curtis

dissimilarity index, with values of 0 corresponding communities with the same identity and degree of pairwise species overlaps, and values of 1 indicating complete dissimilarity in the identity of its pairwise overlap pre- and post-warming communities. Colored boxplots in **a** and **b** depict the range of variation in turnover rates for communities varying in $\beta_{S_{FFD}}$ and β_{S_D}, respectively, with groups across the x-axis representing sets of communities with varying $\bar{S}_D$. In each panel, the horizontal solid black line indicates the degree of dissimilarity observed among communities showing no average changes in flowering duration (i.e., $\bar{S}_D = 0$), and random variation in both S_{FFD} and S_D among species throughout the season (i.e., $\beta_{S_{FFD}} = 0$ and $\beta_{S_D} = 0$).

The effects of non-random variation in S_{FFD} and S_D among species throughout the season (i.e., of $\beta_{S_{FFD}}$ and β_{S_D}) reversed for communities showing no average changes in flowering duration or showing average increases (i.e., $\bar{S}_D \geq 0$). For communities with $\bar{S}_D = 0$, flowering overlap dissimilarity was greatest in those showing flowering delays among early-flowering species and advances among late-flowering species, or showing flowering extension among early-flowering species and contraction among late-flowering species. Finally, communities showing averages increases in flowering duration ($\bar{S}_D > 0$) tended to show the least species compositional changes in flowering overlap. However, as was the case for communities with $\bar{S}_D = 0$, compositional changes were least among those showing advances among early flowering species and delays and among late flowering species (Fig. 6a) or showing flowering contractions among early-flowering species and extensions among late-flowering species (Fig. 6b). These patterns remained consistent when measuring the degree of change in pairwise flowering overlaps based on turnover of interacting pairs (i.e.,

sum of new overlaps gained or former overlaps lost relative to the total number of overlapping pairs pre- and post-warming; Appendix 3—Fig. S7).

Scenarios 2 and 3: Correlated variation in S_{FFD} and S_D among species throughout the season

Co-variation between S_{FFD} and S_D (mediated by β_{S_D} and $\beta_{S_{FFD}}$) either attenuated or amplified changes in flowering intensity during the mid and late season relative to scenarios of uncorrelated sensitivities (Fig. 7). Communities in which S_{FFD} and S_D were positively correlated throughout the season tended to show changes in cumulative flowering intensity of nearly twice the magnitude as those generated by equivalent scenarios (i.e., those with the same $\beta_{S_{FFD}}$ and $\bar{S}_D$) but in which S_{FFD} and S_D varied independently among species (Fig. 7a-e). In contrast, communities in which S_{FFD} and S_D were negatively correlated tended to show lesser changes in cumulative flowering intensity (of nearly half the magnitude) compared to those observed in scenarios in which S_{FFD} and S_D varied independently (Fig. 7f-j) (see Fig. S8 for results for all combinations of $\beta_{S_{FFD}}$ and β_{S_D}). Correlated S_{FFD} and S_D for communities with $\bar{S}_D \neq 0$ also generated amplified or attenuated differences in the degree of change between the mid or late season (for positive and negative correlations, respectively) (Appendix 3—Figs. S9, S10).

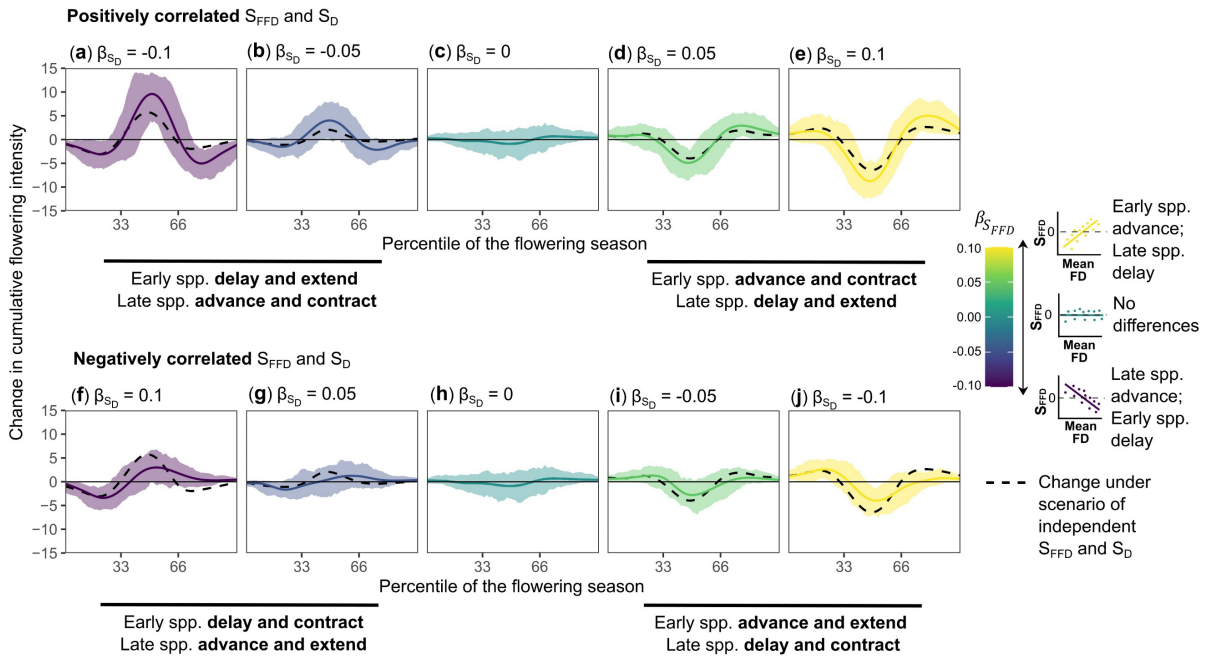

Figure 7—Warming-induced changes to cumulative flowering intensity due to correlated variation among species between S_{FFD} and S_D throughout the season (determined by $\beta_{S_{FFD}}$ and β_{S_D}) across communities with a single flowering peak. In each panel, solid-colored lines and shaded regions depict the median and 95% range of variation in cumulative flowering intensity change observed among communities sharing a combination of values $\beta_{S_{FFD}}$ and β_{S_D}. Black dashed lines in each panel depict a reference scenario of uncorrelated S_{FFD} and S_D, with $\beta_{S_{FFD}}$ varying within columns but β_{S_D} coefficient equal to 0 for all panels. Panels **a-e** show scenarios where S_{FFD} and S_D are positively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude and the same sign). In turn, panels **f-j** show a scenario in which S_{FFD} and S_D are negatively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude but opposite sign). $\bar{S}_D$ was set to 0 days per °C. Scenarios of $\beta_{S_{FFD}} = 0$ or $\beta_{S_{FFD}} = 0$ respectively correspond to communities where S_{FFD} or S_D varied randomly (around the community mean) among species flowering successively throughout the season.

A positive correlation between S_{FFD} and S_D (i.e., β_{S_D} and $\beta_{S_{FFD}}$ with the same sign) resulted in greater changes in season length and peak flowering intensity compared to equivalent scenarios in which S_{FFD} and S_D varied independently (i.e., those with the same $\beta_{S_{FFD}}$ and $\bar{S}_D$ but $\beta_{S_D} = 0$) (Appendix 3—Fig. S11a). In contrast, negatively correlated S_{FFD} and S_D (i.e., β_{S_D} and $\beta_{S_{FFD}}$ with different sign) resulted in lesser changes to season length and peak flowering intensity (Appendix 3—Fig. S11b). Correlated S_D and S_{FFD} also impacted the degree of change in the network of flowering overlaps among species within a community (Appendix 3—Fig. S12). Specifically, positively correlated S_{FFD} and S_D resulted in greater compositional changes in flowering overlap compared to a scenario of uncorrelated S_D and S_{FFD} (Appendix 3—Figs. S12a-e). In turn, negatively correlated S_D and S_{FFD} attenuated rates of composition change under warming relative to those expected under independent sensitivities (Appendix 3—Fig. S12f-j).

E. Discussion

Characterizing among-species variation in the sensitivity of flowering onset and duration is essential for resolving community-level responses to warming

Flowering sensitivity to temperature differs widely among co-occurring species; however, phenological datasets that include flowering duration and that sample sufficient species to characterize flowering dynamics of entire communities are rare (Willis et al. 2017), hindering generalizations of how species variation in temperature responses will scale to alter community-level flowering patterns.

Here, I illustrate the wide range of community-level flowering outcomes that could result from non-random variation among species in the temperature sensitivity of both first flowering dates (S_{FFD}) and flowering duration (S_D). Specifically, using simulations, this study shows that three empirically documented forms of among-species variation in sensitivity within a community—the magnitude and direction of change in S_{FFD} with the mean flowering date of a species ($\beta_{S_{FFD}}$), the mean S_D of a community ($\bar{S}_D$), and the magnitude and direction of change in S_D with the mean flowering date of a species (β_{S_D})—may have profound impacts on the post-warming structure of the flowering season both independently and interactively. While variation in S_{FFD} affected the structure of the entire flowering season, variation in S_D predominantly impacted its mid and late portions. This suggests that among-species variation in S_{FFD} and S_D will have distinct impacts on flowering-dependent ecological processes that occur at different times within the season. Moreover, the severity of simulated flowering reassembly depended on whether S_{FFD} and S_D were positively or negatively correlated among sequentially flowering species throughout the season, which respectively amplified or attenuated responses relative to scenarios of independent sensitivities. These results further demonstrate that models that solely evaluate changes in onset or median flowering dates of a population are insufficient to fully capture these dynamics, emphasizing the importance of also accounting for the temperature sensitivity of phenological terminations for understanding community-level flowering responses to climate change.

Changing flowering seasons will likely have profound but uncertain consequences from populations to communities

Redistribution of floral resources within a community—such as that observed in many of the simulation scenarios—might result in cascading ecological consequences within an ecosystem. For example, the availability and variety of floral resources throughout the season within a community strongly mediates the diversity and abundance of its pollinators (Potts et al. 2003, Fründ et al. 2010, Scheper et al. 2015). Studies of specialist bees have demonstrated that individuals typically travel longer distances for lesser pollen and nectar rewards during periods of floral scarcity (Minckley 1994, Pope and Jha 2018), and local scarcity of floral resources have been directly linked to decreases in brood provisioning and pollinator population declines in some systems (Williams and Kremen 2007, Schenk et al. 2018). Although measuring the broad resource base of generalist pollinators is more challenging than for specialists, the population sizes of the former are generally positively correlated over space and time with the local density of floral resources (Potts et al. 2003, Scheper et al. 2015). Therefore, if the flowering intensity of a community decreases, or if pollinators are phenologically mismatched with either their specialized mutualist plants (in the case of specialists) or with overall peaks in floral abundance (in the case of generalists), we might expect significant declines in pollinator abundance or species diversity.

In turn, net decreases in floral abundance or phenological synchrony between plants and their mutualists partners can negatively impact plant fitness through their effects on density-dependent processes such as pollinator attraction, fertilization and genetic recombination, seed dispersal, or predator attraction and satiation (Nilsson and Wastljung 1987, Elzinga et al. 2007, Carlo and Morales 2008, Jones and Comita 2010, Bergamo et al. 2020). Additionally, the pollen of many wind-pollinated tree and grass species across the temperate zone is allergenic to humans (Songnuan 2013, Garcia-Mozo 2017, Oh 2022).

Therefore, depending on the degree to which variation in S_{FFD} and S_D among allergenic species matches the broader community-level pattern, the substantial redistribution of flowering periods illustrated in many of the simulation scenarios could significantly alter the length and intensity of the allergy season.

The precise consequences of broad community-level changes, however, are not straightforward to predict. The simulations describe how the temporal distribution of floral resources might respond to warming, but not their overall abundance. Changes in net pollen and nectar production among species caused by warming might amplify or attenuate the ecological impacts of the changes in the temporal distribution of floral resources. For example, if among-species variation in S_{FFD} and S_D decrease the density of co-flowering species within a given season, then a lower production of floral resources among species that remain active during that period would exacerbate the consequence of phenologically-driven reductions in the richness of flowering species (e.g., where changes in cumulative flowering intensity are negative in Figs. 5, 7), whereas increases in floral resource production might attenuate such consequences.

In many taxa, fitness may be most strongly affected by a small number of species interactions. For example, specialist pollinators might rely on one or a few plant species for floral resources. Consequently, the persistence of specialist pollinator populations under warming depends on the degree to which they remain synchronized with a subset of plant species whose flowering shifts might not mirror the community-level pattern. Generalist florivores and pollinators might shift their phenology at different rates than the plant community (Memmot et al. 2007), and depending on initial patterns of synchrony, these changes could result in periods of activity overlapping with a greater or lesser diversity and

abundance of floral resources. Moreover, many plant communities are dominated by a handful of species (e.g., Sherry et al. 2007) whose flowering responses could deviate from the broader community pattern. Finally, while flowering synchrony among species can mediate density-dependent ecological processes, whether such processes increase or decrease plant reproductive success can vary among species and ecological contexts (Elzinga et al. 2007), further complicating simple predictions of the ecological consequences of changes to the flowering season.

Considering these complexities, evaluating the effects of warming on the quality and quantity of floral resource productions across species is essential for advancing our ability to predict the consequences of changes to the flowering season. Similarly, to illuminate the broader ecological consequences of their results, studies of flowering responses to climate should consider the local abundance and floral output (or other important attributes such as allergenic potential) of the species under examination.

Warming-induced changes to the flowering season will differ among temperate communities, but critical knowledge gaps remain

In North America, temperate and semi-arid plant communities consistently show greater advances in response to warming among species flowering early in the season compared to those flowering later (i.e., $\beta_{S_{FFD}} > 0$) (e.g., Ramirez-Parada et al. 2023). Therefore, the simulations presented here predict that temperate communities should typically show a longer flowering season under warming, with a greater proportion of floral resources becoming available during its early and late portions, resulting in decreased flowering peaks

(Fig. 5a-c). In contrast, patterns of variation in S_{FFD} among species throughout the season can vary in direction and magnitude among communities at high latitudes and elevations (Schmidt et al. 2016, Prevéy et al. 2019). Nonetheless, greater advances in first flowering date among late-flowering than early-flowering species appear to be common in such systems ($\beta_{S_{FFD}} < 0$) (e.g., Schmidt et al. 2016, Prevéy et al. 2019), which would result in a shorter flowering season with a higher proportion of floral resources becoming available around its median (Fig. 5a-c). Therefore, the results imply that such well-documented variation in $\beta_{S_{FFD}}$ among communities will generate distinct impacts on the structure of the flowering season in different regions.

The warming responses of the mid and late flowering season, however, were also highly sensitive to the structure of among-species variation in S_D within a community (i.e., $\bar{S}_D$ or β_{S_D}; Fig. 5d-f), for which geographic variation has not been extensively studied. Therefore, although recent research has established that communities often differ in both $\bar{S}_D$ and β_{S_D} (e.g., Nam and Kim 2020, Zhou et al. 2022), it is unclear whether and how these parameters vary systematically across plant assemblages that occupy different climate zones or that differ in the relative abundance of species across functional groups. Furthermore, flowering onset and termination can be mediated by different abiotic factors. For example, while flowering onset is often controlled by abiotic cues triggering the start or resumption of reproductive development (Amasino 2010), the termination of flowering may be mediated by constraints on the timing of subsequent phenophases, or by different cues, including the onset of physiologically stressful conditions, resource depletion, or seasonal cooling (Desclaux and Roumet 1996, Ettinger et al. 2019, Zohner et al. 2023). Therefore, it is likely that S_{FFD} and S_D will show different patterns of variation across regional floras, precluding the categorization

of many biomes into the scenarios explored in the simulations presented here, and consequently, general predictions of how the structure of the flowering season will change under warming across regions.

The degree of reassembly of flowering overlaps within a community was mediated interactively by non-random variation among species in the sensitivity of flowering onset and duration (Fig. 6). Consequently, as geographic differences in patterns of among-species variation in S_D within communities are not well characterized, it is hard to determine *a priori* which regional floras should show the greatest degree of flowering reassembly under ongoing warming. Moreover, the degree of change in community-level flowering patterns depended on the degree and direction of correlation between S_{FFD} and S_D among species (Fig. 7, Appendix 3—Fig. S12). As such, determining the ecological contexts in which S_{FFD} and S_D might be negatively correlated, independent, or positively correlated among successively flowering co-occurring species will be key to forecasting community-level phenological reassembly due to ongoing climatic change.

Conclusions

Despite complexities in predicting the ecological consequences of community-level flowering shifts and limited knowledge of among-species variation in S_D across communities, the simulations presented in this study demonstrate that $\beta_{S_{FFD}}$, $\bar{S}_D$, and β_{S_D} might have profound impacts in community-level flowering patterns. To date, among-species variation in S_{FFD} and S_D has been predominantly studied in temperate communities in North America and Europe (Piao et al. 2019), and a disproportionate attention to phenological onset

dates has limited assessments of flowering duration to comparatively few species. Expanding the biogeographic scope of empirical studies quantifying patterns of variation in sensitivity among species and a greater research focus on flowering duration are essential steps towards understanding how the structure of the flowering season might continue to change across geographically, ecologically, and climatically distinct plant assemblages. The worldwide abundance, vast temporal and taxonomic scope, and increasing digital availability of herbarium records provide rich data with which to estimate these parameters (Park et al. 2024). Additionally, determining the biomes in which advances or delays in flowering among early and late flowering species are typically associated with contractions or extensions in flowering duration would help determine where—and how frequently—each of the patterns demonstrated in the simulations should occur.

Finally, although I focused solely on temperature and flowering phenology, the simulation design can be extended to examine any phenophase (e.g., leaf out, fruiting), and to account for the independent and interactive effect of multiple climate variables for any hypothesized forms of structural variation in climate sensitivity among species in a community (e.g., non-linear variation in S_{FFD} and S_D among flowering species). Recent research suggests that differences in sensitivity among species between the early, mid, and late season might be common among taxa spanning several trophic levels (Roslin et al. 2021). Therefore, the effects of among-species variation in sensitivity in mediating community-level responses to warming described here might characterize the modes of phenological change due to warming not only for flowering, but for a much wider range of ecological phenomena.

F. Literature cited

Aide, T. M. 1988. "Herbivory as a selective agent on the timing of leaf production in a tropical understory community." *Nature*, *336*(6199), Article 6199. https://doi.org/10.1038/336574a0

Bergamo, P. J., Susin Streher, N., Traveset, A., Wolowski, M., & Sazima, M. 2020. "Pollination outcomes reveal negative density-dependence coupled with interspecific facilitation among plants." *Ecology Letters*, *23*(1), 129–139. https://doi.org/10.1111/ele.13415

CaraDonna, P. J., Iler, A. M., & Inouye, D. W. 2014. "Shifts in flowering phenology reshape a subalpine plant community." *Proceedings of the National Academy of Sciences*, *111*(13), 4916–4921. https://doi.org/10.1073/pnas.1323073111

Chen, B., Jin, Y., & Brown, P. 2019. "An enhanced bloom index for quantifying floral phenology using multi-scale remote sensing observations." *ISPRS Journal of Photogrammetry and Remote Sensing*, 156, 108–120. https://doi.org/10.1016/j.isprsjprs.2019.08.006

Chen, J., Luo, Y., Chen, Y., Felton, A. J., Hopping, K. A., Wang, R.-W., Niu, S., Cheng, X., Zhang, Y., Cao, J., Olesen, J. E., Andersen, M. N., & Jørgensen, U. 2020. "Plants with lengthened phenophases increase their dominance under warming in an alpine plant community." *Science of The Total Environment*, *728*, 138891. https://doi.org/10.1016/j.scitotenv.2020.138891

Chen, Y., Collins, S. L., Zhao, Y., Zhang, T., Yang, X., An, H., Hu, G., Xin, C., Zhou, J., Sheng, X., He, M., Zhang, P., Guo, Z., Zhang, H., Li, L., & Ma, M. 2022. "Warming reduced flowering synchrony and extended community flowering season in an alpine meadow on the Tibetan Plateau." *Ecology, 104(1)*, e3862. https://doi.org/10.1002/ecy.3862

Cook, B. I., Wolkovich, E. M., Davies, T. J., Ault, T. R., Betancourt, J. L., Allen, J. M., Bolmgren, K., et al. 2012. "Sensitivity of spring phenology to warming across temporal and spatial climate gradients in two independent databases." *Ecosystems*, *15*(8), 1283–1294. https://doi.org/10.1007/s10021-012-9584-5

Desclaux, D., & Roumet, P. 1996. "Impact of drought stress on the phenology of two soybean (*Glycine max* L. Merr) cultivars." *Field Crops Research*, *46*(1), 61–70. https://doi.org/10.1016/0378-4290(95)00086-0

Devaux, C., & Lande, R. 2009. "Displacement of flowering phenologies among plant species by competition for generalist pollinators." *Journal of Evolutionary Biology*, *22*(7), 1460–1470. https://doi.org/10.1111/j.1420-9101.2009.01762.x

Diez, J. M., Ibáñez, I., Miller-Rushing, A. J., Mazer, S. J., Crimmins, T. M., Crimmins, M. A., Bertelsen, C. D., & Inouye, D. W. 2012. "Forecasting phenology: From species variability to community patterns." *Ecology Letters*, *15*(6), 545–553. https://doi.org/10.1111/j.1461-0248.2012.01765.x

Dixon, D. J., Callow, J. N., Duncan, J. M. A., Setterfield, S. A., & Pauli, N. 2021. "Satellite prediction of forest flowering phenology." *Remote Sensing of Environment*, 255, 112197. https://doi.org/10.1016/j.rse.2020.112197

Elzinga, J. A., Atlan, A., Biere, A., Gigord, L., Weis, A. E., & Bernasconi, G. 2007. "Time after time: Flowering phenology and biotic interactions." *Trends in Ecology & Evolution*, *22*(8), 432–439. https://doi.org/10.1016/j.tree.2007.05.006

Ettinger, A. K., Gee, S., & Wolkovich, E. M. 2018. "Phenological sequences: How early-season events define those that follow." *American Journal of Botany*, *105*(10), 1771–1780. https://doi.org/10.1002/ajb2.1174

Fitter, A. H., & Fitter, R. S. R. 2002. "Rapid changes in flowering time in British plants." *Science*, *296*(5573), 1689–1691. https://doi.org/10.1126/science.1071617

Fründ, J., Linsenmair, K. E., & Blüthgen, N. (2010). Pollinator diversity and specialization in relation to flower diversity. *Oikos*, *119*(10), 1581–1590. https://doi.org/10.1111/j.1600-0706.2010.18450.x

Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., et al. 2015. "Declining global warming effects on the phenology of spring leaf unfolding". *Nature*, 526(7571), Article 7571. https://doi.org/10.1038/nature15402

Gavini, S. S., Sáez, A., Tur, C., & Aizen, M. A. (2021). Pollination success increases with plant diversity in high-Andean communities. Scientific Reports, 11(1), 22107. https://doi.org/10.1038/s41598-021-01611-w

González-Suárez, P., Walker, C. H., & Bennett, T. (2020). Bloom and bust: understanding the nature and regulation of the end of flowering. Current Opinion in Plant Biology, 57, 24-30. https://doi.org/10.1016/j.pbi.2020.05.009

Guo, J., Ma, Q., Xu, H., Luo, Y., He, D., Wang, F., Wu, J., et al. 2023. "Meta-analytic and experimental evidence that warmer climate leads to shift from advanced to delayed spring phenology". *Agricultural and Forest Meteorology*, 342, 109721. https://doi.org/10.1016/j.agrformet.2023.109721

Hu, X., Zhou, W., & Sun, S. 2020. "Responses of plant reproductive phenology to winter-biased warming in an alpine meadow." *Frontiers in Plant Science*, *11*, 1376. https://doi.org/10.3389/fpls.2020.534703

Huang, W., Dai, J., Wang, W., Li, J., Feng, C., & Du, J. 2020. "Phenological changes in herbaceous plants in China's grasslands and their responses to climate change: A meta-analysis." *International Journal of Biometeorology*. https://doi.org/10.1007/s00484-020-01974-1

Jabis, M. D., Winkler, D. E., & Kueppers, L. M. 2020. "Warming acts through earlier snowmelt to advance but not extend alpine community flowering." *Ecology*, *101*(9), e03108. https://doi.org/10.1002/ecy.3108

Jones, F. A., & Comita, L. S. 2010. "Density-dependent pre-dispersal seed predation and fruit set in a tropical tree." *Oikos*, *119*(11), 1841–1847. https://doi.org/10.1111/j.1600-0706.2010.18547.x

Li, D., Barve, N., Brenskelle, L., Earl, K., Barve, V., Belitz, M. W., Doby, J., et al. 2021. "Climate, urbanization, and species traits interactively drive flowering duration." *Global Change Biology*, *27*(4), 892–903. https://doi.org/10.1111/gcb.15461

Mazer, S. J., Travers, S. E., Cook, B. I., Davies, T. J., Bolmgren, K., Kraft, N. J. B., Salamin, N., & Inouye, D. W. 2013. "Flowering date of taxonomic families predicts phenological sensitivity to temperature: Implications for forecasting the effects of climate change on unstudied taxa." *American Journal of Botany*, *100*(7), 1381–1397. https://doi.org/10.3732/ajb.1200455

Miller-Rushing, A. J., & Primack, R. B. 2008. "Global warming and flowering times in Thoreau's Concord: a community perspective." Ecology, 89(2), 332–341. https://doi.org/10.1890/07-0068.1

Morandin, L. A., & Kremen, C. 2013. "Hedgerow restoration promotes pollinator populations and exports native bees to adjacent fields." *Ecological Applications*, 23(4), 829–839. https://doi.org/10.1890/12-1051.1

Nagahama, A., Kubota, Y., & Satake, A. 2018. "Climate warming shortens flowering duration: A comprehensive assessment of plant phenological responses based on gene expression analyses and mathematical modeling." *Ecological Research*, *33*(5), 1059–1068. https://doi.org/10.1007/s11284-018-1625-x

Nam, B. E., & Kim, J. G. 2020. "Flowering season of vernal herbs is shortened at elevated temperatures with reduced precipitation in early spring." *Scientific Reports*, *10*(1), 17494. https://doi.org/10.1038/s41598-020-74566-z

Oh, J.-W. 2022. "Pollen Allergy in a Changing Planetary Environment." *Allergy, Asthma & Immunology Research*, 14(2), 168–181. https://doi.org/10.4168/aair.2022.14.2.168

Park, D. S., Breckheimer, I., Williams, A. C., Law, E., Ellison, A. M., & Davis, C. C. (2019). Herbarium specimens reveal substantial and unexpected variation in phenological sensitivity across the eastern United States. Philosophical Transactions of the Royal Society B, 374(1763), 20170394. https://doi.org/10.1098/rstb.2017.0394

Park, I. W., Ramirez-Parada, T., Record, S., Davis, C., Ellison, A. M., & Mazer, S. J. (2024). Herbarium data accurately predict the timing and duration of population-level flowering displays. Ecography, e06961. https://doi.org/10.1111/ecog.06961

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. 2019. "Plant phenology and global climate change: Current progresses and challenges." *Global Change Biology*, *25*(6), 1922–1940. https://doi.org/10.1111/gcb.14619

Pope, N. S., & Jha, S. 2018. "Seasonal food scarcity prompts long-distance foraging by a wild social bee." *The American Naturalist*, *191*(1), 45–57. https://doi.org/10.1086/694843

Potts, S. G., Vulliamy, B., Dafni, A., Ne'eman, G., & Willmer, P. 2003. "Linking bees and flowers: How do floral communities structure pollinator communities?". *Ecology*, *84*(10), 2628–2642. https://doi.org/10.1890/02-0136

Prevéy, J. S., Rixen, C., Rüger, N., Høye, T. T., Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., et al. 2019. "Warming shortens flowering seasons of tundra plant communities." *Nature Ecology & Evolution*, *3*(1), 45–52. https://doi.org/10.1038/s41559-018-0745-6

Reilly, J. R., Artz, D. R., Biddinger, D., Bobiwash, K., Boyle, N. K., Brittain, C., Brokaw, J., et al. 2020. "Crop production in the USA is frequently limited by a lack of pollinators." *Proceedings of the Royal Society B: Biological Sciences*, 287(1931), 20200922. https://doi.org/10.1098/rspb.2020.0922

Ramirez-Parada, T. H., Park, I. W., Sydne R., Davis, C. C., Ellison, A. M., Mazer, S. J. (2023). "Plasticity and not adaptation is the primary source of temperature-mediated variation in flowering phenology in North America", 07 July 2023, PREPRINT (Version 1) available at *Research Square*. https://doi.org/10.21203/rs.3.rs-3131821/v1

Renner, S. S., & Zohner, C. M. 2018. "Climate change and phenological mismatch in trophic interactions among plants, insects, and vertebrates." *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 165–182. https://doi.org/10.1146/annurev-ecolsys-110617-062535Roulston, T. H., & Goodell, K. 2011. "The Role of Resources and Risks in Regulating Wild Bee Populations." *Annual Review of Entomology*, *56*(1), 293–312. https://doi.org/10.1146/annurev-ento-120709-144802

Scheper, J., Bommarco, R., Holzschuh, A., Potts, S. G., Riedinger, V., Roberts, S. P. M., Rundlöf, M. et al. (2015). "Local and landscape-level floral resources explain effects of wildflower strips on wild bees across four European countries." *Journal of Applied Ecology*, *52*(5), 1165–1175. https://doi.org/10.1111/1365-2664.12479

Sherry, R. A., Zhou, X., Gu, S., Arnone, J. A., Schimel, D. S., Verburg, P. S., Wallace, L. L., & Luo, Y. (2007). "Divergence of reproductive phenology under climate warming." *Proceedings of the National Academy of Sciences*, 104(1), 198–202. https://doi.org/10.1073/pnas.0605642104

Songnuan, W. 2013. "Wind-pollination and the roles of pollen allergenic proteins." *Asian Pacific journal of allergy and immunology*, 31(4), 261.

Stinson, K. A., Wheeler, J. A., Record, S., & Jennings, J. L. 2018. "Regional variation in timing, duration, and production of flowers by allergenic ragweed." *Plant Ecology*, 219(9), 1081–1092. https://doi.org/10.1007/s11258-018-0860-0

Theobald, E. J., Breckheimer, I., & HilleRisLambers, J. 2017. "Climate drives phenological reassembly of a mountain wildflower meadow community." *Ecology*, 98(11), 2799–2812. https://doi.org/10.1002/ecy.1996

Williams, N. M., & Kremen, C. 2007. "Resource distributions among habitats determine solitary bee offspring production in a mosaic landscape." *Ecological Applications*, *17*(3), 910–921. https://doi.org/10.1890/06-0269

Willis, C. G., Ellwood, E. R., Primack, R. B., Davis, C. C., Pearson, K. D., Gallinat, A. S., Yost, et al. 2017. "Old Plants, New Tricks: Phenological Research Using Herbarium Specimens." *Trends in Ecology & Evolution*, *32*(7), 531–546. https://doi.org/10.1016/j.tree.2017.03.015

Wolkovich, E. M., Cook, B. I., Allen, J. M., Crimmins, T. M., Betancourt, J. L., Travers, S. E., Pau, S. 2012. “Warming experiments underpredict plant phenological responses to climate change.” *Nature*, *485*(7399), 494–497. https://doi.org/10.1038/nature11014

Zhang, H., Yuan, W., Liu, S., Dong, W., & Fu, Y. 2015. “Sensitivity of flowering phenology to changing temperature in China.” *Journal of Geophysical Research: Biogeosciences*, 120(8), 1658–1665. https://doi.org/10.1002/2015JG003112

Zhou, Z., Zhang, K., Sun, Z., Liu, Y., Zhang, Y., Lei, L., ... & Miao, Y. (2022). Lengthened flowering season under climate warming: evidence from manipulative experiments. Agricultural and Forest Meteorology, 312, 108713. https://doi.org/10.1016/j.agrformet.2021.108713

Zohner, C. M., Mirzagholi, L., Renner, S. S., Mo, L., Rebindaine, D., Bucher, R., Palouš, D. et al. 2023. “Effect of climate warming on the timing of autumn leaf senescence reverses after the summer solstice.” *Science*, 381(6653), eadf5098. https://doi.org/10.1126/science.adf5098

Zurbuchen, A., Cheesman, S., Klaiber, J., Müller, A., Hein, S., & Dorn, S. 2010. “Long foraging distances impose high costs on offspring production in solitary bees.” *Journal of Animal Ecology*, *79*(3), 674–681. https://doi.org/10.1111/j.1365-2656.2010.01675.x

V. Climate change restructures the flowering season across North America but effects vary among ecoregions

A. Abstract

Climate change impacts the structure of the flowering season within communities by shifting species distributions and phenology, causing cascading ecological impacts. However, data constraints and incomplete knowledge of population-level flowering responses to climate limit our understanding of how environmental change may reassemble the flowering season across biomes. Using millions of herbarium and occurrence records, I modeled the distribution and flowering phenology of 2,837 species across the United States under historical, current, and projected climate and land cover conditions, scaling onset, duration, and termination responses from species to communities, and from local to continental levels. Within species, the onset, duration, and termination of flowering responded differently to climate, with substantial variation in sensitivity among species both within and between communities. At the community level, climate change altered species composition and the timing and duration of flowering seasons, with ecoregion-specific changes in the seasonal distribution of co-flowering species diversity and in the network of flowering overlaps between species that are projected to intensify with ongoing climate and land use changes. This study reveals broad macroecological patterns of change and identifies biomes in which the ecological consequences of altered flowering seasons may be most severe.

B. Introduction

The flowering season is crucial for plant fitness and the survival and reproduction of organisms that depend on floral resources. Its structure determines the identity, diversity, and abundance of floral resources throughout the year, which can mediate pollinator population growth rates (Roulston and Goodell 2011) and various density-dependent ecological outcomes in plants—such as pollinator success and seed predation—that can shape the evolution of their life-history strategies (Elzinga et al. 2007). Recent climate change has led to widespread shifts in flowering phenology and plant distributions (Cleland et al. 2007, Kelly and Goulden 2008), potentially disrupting ecological interactions through altered spatial and seasonal synchrony between species (Renner and Zohner 2018). However, our understanding of how these shifts affect community-level flowering patterns is limited due to a lack of long-term datasets that represent enough species to characterize a community's flowering season or that measure floral resources throughout the year (e.g., Caradonna et al. 2014). Additionally, while remote sensing can capture the growing season's start, peak, and end, current methods cannot detect the weak spectral signals associated with community-level flowering patterns at continental scales.

Inferring community-level flowering patterns from the responses of co-occurring species is typically not possible due to an incomplete understanding of how climate affects flowering. Most research has concentrated on the onset of the flowering period within populations or sites (e.g., Fitter and Fitter 2002, Wolkovich et al. 2012, Prevéy et al. 2019), with less focus on its termination and duration. However, at the individual level, flowering termination can be influenced by different factors than those affecting onset (Nagahama et al. 2018, Gonzalez-Suarez et al. 2020). Consequently, responses to climate of flowering

termination or duration—which have been measured in few species (e.g., Li et al. 2021)—may differ from those of flowering onset. These knowledge gaps are compounded by a lack of long-term datasets on flowering duration, limiting our ability to predict how species-level responses will scale to alter community-level flowering patterns (Park et al. 2024).

The growing availability of herbarium specimens and research-grade records from community science initiatives (e.g., iNaturalist) provides new opportunities to overcome these limitations (Willis et al. 2017, Belitz et al. 2020). These data record the date, location, and often the flowering status of individuals, offering snapshots of plant phenology across different times and places that have been increasingly used in phenoclimatic research (Calinger et al. 2013, Park et al. 2019, Li et al. 2021, Ramirez-Parada 2022, 2024). These records are often opportunistically collected from local populations, and collectively encompass a combination of individuals flowering early, near the median, or late relative to their source populations. When sample sizes are sufficient, these records can reflect the distribution of flowering dates under a set of environmental conditions, allowing for estimation of the onset, termination, and duration of a species' flowering period (Belitz et al. 2023, Austin et al. 2024, Park et al. 2024).

In this study, I analyzed a dataset of over 2.7 million herbarium and community-science records for 2,837 species to examine how climate change affects the structure of the flowering season in the contiguous United States (CONUS). I modeled each species' geographic distribution under historical, current, and future climate conditions, accounting for land cover and land use changes. Additionally, I assessed how interannual variation in temperature and precipitation—independently and interactively—influence the onset, termination, and duration of the flowering period for each species. By concurrently

predicting species occurrences and their flowering periods across sites, these analyses allowed me to scale changes from individual species to communities and from local to continental levels. This approach enabled me to address knowledge gaps in the biogeographic patterns of flowering sensitivity and community-level responses to climate change. For example, previous studies have shown that flowering onset sensitivity varies among communities in different climates and among species blooming at different times of the year (e.g., Wang et al. 2015, Prevéy et al. 2017, Ramirez-Parada et al. 2024). However, it is unclear how the sensitivity of flowering termination and duration varies among biomes and among species flowering on different dates within the same community. By modeling distributions and flowering responses for thousands of species, I assessed how sensitivities to temperature and precipitation—henceforth 'S_{TMEAN}' and 'S_{PPT}'—for flowering onset, duration, and termination differ i) among communities distributed across ecoregions with varying climate regimes and species compositions, and ii) within communities among species that flower at different times of the year.

Community-level changes in the flowering season have been studied in only a few systems (e.g., Jabis et al. 2020, Li et al. 2020, Zhou et al. 2022), with studies primarily focusing on simple directional changes in season length. Consequently, studies collectively cover only a few regional floras, and more detailed seasonal attributes—such as the temporal distribution of flowering diversity—remain unexamined except in select cases (e.g., Caradonna et al. 2014, Austin et al. 2024). Similarly, the extent to which changes in geographic distributions and flowering responses will result in novel flowering assemblages remains unclear for most biomes (Theobald et al. 2017, Austin et al. 2024). To address these gaps, I scaled species-level responses to evaluate how recent and future climate change

across CONUS affects i) species composition, ii) the start of the flowering season (SOS), iii) the end of the flowering season (EOS), iv) the duration of the flowering season (DOS), v) the distribution of co-flowering species diversity throughout the year, and vi) the network of flowering synchronies among co-occurring species, which determines the potential for flowering-mediated interactions between species.

These analyses offer the most comprehensive assessment of how recent and projected climate change impacts flowering patterns across North American floras. This study demonstrates that i) recent climate change has significantly affected all characteristics of the flowering season, ii) the extent and nature of these changes vary widely among ecoregions, and iii) future climate trends could lead to even more severe shifts in flowering patterns than currently observed, with uneven effects across CONUS.

C. Methods

Phenological and Occurrence data

I compiled specimen records from 220 herbaria, accessed digitally through 16 consortia from Mexico, the U.S., and Canada (in July and August 2022; Park et al. 2023). Only specimens explicitly recorded as bearing flowers were retained, identified by summarizing unique entries in the DarwinCore 'reproductiveCondition' column that clearly indicated presence of flowers. Specimens missing geographic coordinates, collection dates, or species-level identification were excluded. To avoid pseudoreplication, conspecific specimens collected within 1 km of each other on the same day were removed. Since over 92% of the remaining specimens were collected within the United States, and to match the spatial extent of land

use/land cover (LULC) data used in species distribution models (SDMs, I excluded specimens collected outside CONUS. Specimens collected before 1958 were also removed to align with the temporal range of TerraClimate climatic data used in these analyses. After harmonizing species names using the Global Biodiversity Information Facility (GBIF) taxonomic backbone, the data were filtered further to include only species represented by at least 100 specimens. The day of year (DOY) of collection was used as a proxy for flowering date, with an azimuthal correction applied to address the discontinuity between 31 December and 1 January, converting prior year DOYs into negative values.

I obtained an additional 13.2 million research-grade occurrence records from GBIF for species well-represented in the flowering phenology dataset (accessed July 11, 2024; https://www.gbif.org/occurrence/download/0021084-240626123714530). These records, primarily from iNaturalist and herbarium sources, were combined with those from Park et al. (2023). I removed duplicates using the 'occurrenceID' column in DarwinCore. To match the temporal and spatial extent of LULC data for SDMs, I retained only occurrences collected between 1999 and 2023, and limited the dataset to occurrences within CONUS.

Preliminary analyses of GBIF occurrences revealed significant spatial biases towards urban areas and major roads. To address this, I identified occurrences within urban areas as defined by the US Census Bureau (2012) using the 'tigris' package v2.1. I thinned the data using the 'spThin' package v0.2.0 (Aiello-Lammens et al., 2015), keeping only occurrences of the same species recorded at least 20 km apart within urban areas. Additionally, I removed occurrences within 2 km of 'primary roads' mapped by the US Census Bureau (2012). To further reduce spatial bias, another thinning step was applied, keeping only conspecifics recorded at least 5 km apart regardless of urban or road proximity. After cleaning using

BONAP records (detailed in the next section), I retained only species with at least 50 occurrences to ensure adequate data for species distribution modeling.

Final cleaning of specimens and occurrences using BONAP

Species misidentifications or geolocation errors in herbarium and occurrence databases can bias SDMs or phenoclimatic models by distorting the climate space or flowering period represented among by observations. To mitigate this, I removed implausible records using expertly curated data from the Biota of North America Program's (BONAP; Kartesz, 2015) North American Plant Atlas (NAPA), which documents 19,039 taxa from 227 families across 3,067 counties across CONUS. BONAP compiles species presence/absence from herbarium records, museums, and bibliographic reviews, most of which is verified by taxonomic and floristic experts. Species names were harmonized across the specimen, occurrence, and BONAP datasets using BONAP's taxonomic backbone. I then excluded observations from counties where BONAP did not report occurrences for the species.

After taxonomic harmonization and BONAP cleaning, the final phenology dataset included 1,042,939 specimens (collected from 1958 to 2022) representing 2,837 species in 1,042 genera and 139 families. The final occurrence dataset contained 1,673,454 records (collected from 1999 to 2023), comprising the same species, genera, and families. Of these, 1,369,657 were human observations from iNaturalist, and 303,797 were herbarium specimens not included in Park et al. (2023).

Climate data

I obtained historical monthly climatic rasters from TerraClimate (Abatzoglou et al. 2018) available from January 1958 to December of 2023 at a 4km × 4km resolution. These data consisted of monthly time series for minimum temperature (TMIN), mean temperature (TMEAN), maximum temperature (TMAX), and cumulative precipitation (PPT), as well as modeled water balance metrics including actual evapotranspiration (AET), climate water deficit (DEF), soil moisture (SOIL), and snow water equivalent (SWE).

Climate variables for species distribution modelling

I used monthly climate data to calculate annual bioclimatic variables known to influence plant distributions. For each year and location across CONUS, I computed annual means (or sums for precipitation), minimum and maximum monthly values (e.g., mean minimum temperature of the coldest month, mean maximum of the warmest month), annual ranges (difference between maximum and minimum mean monthly values), and seasonality (standard deviation of the 12 monthly values). For temperature, I also calculated the approximate mean diurnal temperature range (mean difference between TMAX and TMIN across months) and approximate isothermality (mean approximate diurnal range divided by the annual range). For precipitation, I seasonality was calculated relative to cumulative annual precipitation within each site. Minimum monthly SWE was removed from the analyses, as it was 0 across CONUS. This resulted in 31 climate variables: 7 for temperature, 5 related to PPT, AET, DEF, SOIL, and 4 for SWE, calculated annually across all CONUS locations. For each occurrence record, I computed long-term averages of these variables over the 20 years preceding its collection date. Additionally, I obtained elevation data from USGS, calculating mean elevation and elevational heterogeneity within 800m × 800m grid cells. The

coarser resolution for elevation was used to account for uncertainties in georeferencing of herbarium specimens, which may lead to merger of incorrect values especially in steep areas where altitude changes rapidly (Gamble & Mazer, 2022).

Since many of the climate variables were highly collinear (Appendix 4—Fig. S1) and are causally related, I performed a principal component analysis (PCA) to reduce the dimensionality of the climate space. The PCA used 20-year averages of all variables for the most recent period available (2004-2023) across all 4 km × 4 km grid cells in CONUS. I retained the five principal components (PCs) with eigenvalues ≥ 1, which collectively explained 88.2% of the variance in the climate data (Table S1). PC1 represented a gradient of increasing aridity, PC2 a gradient of decreasing temperature and increasing temperature seasonality, and PC3 a gradient of increasing elevational heterogeneity and mean elevation with decreasing temperature seasonality. PC4 primarily captured increasing soil moisture, while PC5 reflected increasing actual evapotranspiration and elevation (Appendix 4—Fig. S2). I then projected the 20-year average climate conditions associated with each occurrence record onto these PCA axes, reducing the number of climatic predictors from 33 variables to 5.

To predict species distribution across different periods, I calculated 20-year averages for each of the 33 climate variables for the historical period (1961-1980) and the present period (2001-2020). I also obtained projected climate conditions from TerraClimate for a scenario where global temperatures rise by 2°C above pre-industrial levels. This scenario is not tied to a specific time frame or emissions pathway; instead, TerraClimate interpolates climate normals from 1985-2015, adjusting for the changes in means and seasonality

expected under 2°C of warming. I then projected historical, present, and future climatic conditions onto the 5 principal components derived from the 2001-2020 data.

Climate variables for phenoclimatic modelling

Variation in TMEAN and PPT among sites and years of specimen collection was partitioned into spatial and temporal components by calculating long-term means (reflecting geographic differences in chronic climatic conditions) and year-specific deviations (reflecting interannual differences). For each species at each site and year, I obtained data for the climatic conditions during the 3-month periods leading up to its average flowering onset, peak, and termination. To estimate conditions approximately before flowering onset, I used the 10th percentile collection date of each species and calculated the mean TMEAN and cumulative PPT for the 3 months leading up to that month. The same approach was applied for the 50th percentile (flowering median) and 90th percentile (flowering termination) collection dates. For each specimen, I characterized its site's long-term TMEAN and PPT (normals) by averaging the observed conditions across all years between 1961 and 1990 for each 3-month period approximating that species' flowering onset, median, and termination. I then calculated climatic deviations (anomalies) from the 1961-1990 normals in the year of each specimen's collection for these 3-month periods.

Assuming phenological changes are driven by interannual variation in TMEAN and PPT rather than spatial phenology-climate relationships, I calculated deviations from 1961-1990 normals for all 3-month windows. This was done for the historical period (1961-1980), the present period (2001-2020), and the future 2°C warming scenario. These TMEAN and

PPT deviations were then used to predict changes in flowering onset and termination between reference periods at each species' occurrence site.

Land use and land cover data

I obtained land use and land cover (LULC) data from the National Land Cover Database (NLCD) (USGS, Dewitz 2019), available for 2001, 2004, 2006, 2008, 2011, 2013, 2016, 2019, and 2021. The NLCD uses Landsat spectral data to classify 30m resolution grid cells into land cover and land use classes, providing a consistent, high-resolution dataset across CONUS. I separated each year's multiclass raster into layers representing the presence or absence of each LULC type. I retained all cover classes except those not present in CONUS (e.g., lichen, moss, sedge classes from Alaska) or those that were rare (e.g., barren land). For land cover, I kept forest classes (deciduous, evergreen, mixed), scrubland (shrub/scrub), herbaceous grasslands, and wetlands (herbaceous and woody). For land use, I included four urban categories (open, low, mid, high) and two agricultural classes (cultivated crops, pasture/hay). To match the format of the LULC data available for forecasting and backcasting (see next paragraph), I aggregated all urban classes into a single category. To account for uncertainty in occurrence coordinates and because plant occurrence can be influenced by landscape context at broader scales than 30m (Mazerolle & Villard, 1999), I measured the proportion of each class cover within 750m × 750m grid cells (625 30m cells) around each occurrence. LULC class proportions were sourced from the NLCD layer closest to the year of collection for each record. These class proportions were then used as predictors in species distribution modeling.

Because NLCD data was available only from 2001 to 2021, I obtained historical (1961-1980) and future (2061-2080) LULC projections from the Earth Resources Observation and Science Center (EROS) (USGS; Sohl et al. 2014, Sohl et al. 2016) at a 250m resolution. EROS' projections use the same modeling framework as NLCD, integrating land use trends with spatial explicit allocation based on regional suitability for each LULC class. Though EROS projections were based on the Special Reports Emissions Scenarios (SRES) from the IPCC (2000)—replaced later by Representative Concentration Pathways (RCP; IPCC 2013) and Shared Socioeconomic Pathways (SSP; IPCC 2021)—they align closely with RCP and SSP scenarios (Rogelj et al. 2012, Riahi et al. 2016). I chose the B1 scenario for forecasting, as it is the closest match to RCP4.5 and SSP2-3, representing 'middle-of-the-road' emissions and development scenarios. As with NLCD data, I calculated the proportion of each land cover class in 750m resolution blocks (containing 9 grid cells) to generate historical and future predictions used in SDMs.

Analyses

Training SDMs

Species distributions were modeled using presence-background random forest classifier models implemented in the ‘randomForest’ package v4.7-1.1 (Liaw & Wiener 2002). Random forests are a supervised machine learning technique that uses an ensemble of decision trees to identify relationships between a response (presence/background data) and predictors (climatic and LULC variables). By combining multiple decision trees, the ensemble often performs better than any single model, leveraging the "wisdom of the

crowds." This approach does not require predefined model structures (e.g., linear relationships) and its non-parametric nature allows for discovering complex relationships and interactions (Cutler et al. 2007). This flexibility was crucial for these analyses of thousands of species with diverse distributional responses to climate and LULC. Additionally, random forests are computationally efficient and have been demonstrated to be among the most accurate SDM methods available (Valavi et al. 2022).

SDMs for each species were trained using occurrence data from 1999-2021. This period ensured availability of high-quality LULC data from NLCD within two years of each collection date. All models included the 5 bioclimatic PCs and the proportion of each LULC class around collection sites as predictors. To address the challenge of selecting pseudo-absences—where it is often unclear if unoccupied regions are due to sampling bias or true distributional patterns—I used curated BONAP records to identify counties where each species was not documented. I drew 10,000 random locations per species from these counties (excluding areas within 2km of major roads). For these locations, I obtained 20-year climatic averages (2001-2020) projected onto the 5 climatic PCs and LULC variables from a randomly selected year between 2001 and 2021. To address class imbalance, I downsampled pseudo-absences to match the number of occurrences in each initial tree. Each species-specific model used 500 trees with a maximum of 5 predictors at each split.

Decision trees are built using bootstrap samples of the data. Typically, these samples contain about 2/3 of the original data, with the remaining third (out-of-bag or ‘OOB’ data) used to calculate each tree's error rate (Cutler et al. 2007). I evaluated model performance by averaging the error rates across all trees, which provides an unbiased estimate of the model’s generalization error. Specifically, I calculated the true skill statistic (TSS), which is the sum

of the true positive rate (TPR) and true negative rate (TNR) minus one. The median TSS among species was 0.91 (range: 0.51 to 0.99).

The SDMs output a probability of occurrence under specific environmental conditions, derived from the proportion of trees predicting the positive class. Given that SDMs used presence-only data, these probabilities are interpreted as habitat suitability rather than actual probabilities of occurrence. To set a suitability threshold for considering a species present at a site, I calculated the receiving operating curve (ROC) for each model and determined the threshold (0-1) that maximized the true positive rate (minimizing false negatives) while keeping the false positive rate below 0.05. While this criterion maximized true positive detection at the expense of higher false negative rates, this tradeoff is justified since implausible occurrence predictions could be identified and removed using BONAP county records.

Training phenoclimatic models

For each species, I modeled how flowering onset, termination, and duration varied with long-term climatic conditions and interannual climatic variation. I used quantile regression (via the 'quantreg' package v5.97, Koenker et al. 2019) to assess how collection date distributions among conspecifics responded to geographic and interannual variations in TMEAN and PPT (i.e., normal and anomalies, respectively). I used the 10th percentile of the distribution to represent population-level flowering onset, the 90th percentile to represent flowering termination, and the interquartile distance between them to represent flowering duration. I chose the 10th and 90th percentiles and focused on well-sampled species because

estimation of extreme quantiles can be biased with small samples. Moreover, recent simulations show that quantile regression accurately estimates 10^{th} and 90^{th} percentiles of opportunistically sampled data for sample sizes similar to ours (Park et al. 2024), and this approach has been effective in studying phenological distributions in both plants and insects (Belitz et al. 2023, Austin et al. 2024).

In each species-specific model, predictors included TMEAN normal, PPT normal, and their interaction for the 3-month period before the average date of flowering onset or termination, as well as TMEAN anomaly, PPT anomaly, and their interaction during the same period (6 predictors total). The coefficients for the main terms in these quantile regressions indicate how the 10th and 90th percentiles of flowering are affected by geographic or interannual variation in TMEAN and PPT, assuming average values for interacting variables. Interaction coefficients between normals represent the degree to which long-term precipitation affects the magnitude of phenological changes due to variation in long-term TMEAN across sites (or vice versa), whereas the interaction coefficients between anomalies indicate how the effects of interannual variation in TMEAN varies among drier- or wetter-than-average years (and viceversa).

This approach models phenological variation as a response to: i) geographic variation in chronic TMEAN and PPT conditions across sites, using temporally invariant normals from 1961-1990, and ii) TMEAN and PPT anomalies reflecting temporal variation within sites, which primarily capture plastic phenological responses (Ramirez-Parada et al. 2024). Thus, I assumed that any temporal changes in a species' flowering season within sites are driven by deviations from their 1961-1990 TMEAN and PPT normals.

Species-level predictions of distributions and phenology

Each species' SDM was used to generate habitat suitability maps for historical (1961-1980), present (2001-2020), and future (2 °C warming, B1 LULC scenario for 2080) conditions. Predictors were resampled to a 12km resolution for computational ease. Suitability estimates were then converted to binary occurrence maps by applying a threshold that maximized the true positive rate (see '*Analyses—Training SDMs*' subsection). Presence-only SDMs can predict unsuitable areas outside a species' range or beyond its dispersal capacity. To address this, predictions were constrained to within 40km of counties where BONAP confirmed each species' presence, which allowed for moderate range expansion to areas adjacent to currently occupied regions between periods. These SDMs predicted substantial variation in species richness across CONUS, from 56 to 1,445 species (from a total 2,837) for the historical period (Appendix 3—Fig. S4). Species richness was generally lowest in arid regions of the Great Plains and higher in the West compared to the East, consistent with more comprehensive assessments of plant diversity in North America (e.g., Daru et al. 2024).

Each species' phenoclimatic model were used to predict flowering onset, termination, and duration for each location where the species was projected to occur during historical, present, and future periods. This was done by applying deviations of average TMEAN and PPT conditions from the 1961-1990 normals for each period. Climate rasters were resampled to a 12km resolution before estimating phenological onset, termination, and duration for each site and period.

Geographic variation in flowering sensitivity

I assessed how sensitivity of population-level flowering onset, termination, and duration to interannual variation in TMEAN and PPT varied among communities across CONUS. First, I calculated the mean sensitivity to TMEAN and PPT anomalies for species in each 12km grid cell using the coefficients from their main effects in the quantile regressions (with interacting anomalies set to their average value of 0). Species with different flowering times often show significant variation in climate sensitivity, affecting community-level responses to warming (Ramirez-Parada et al. 2024). Accordingly, to measured flowering sensitivity varied among species throughout the flowering season, sensitivities to TMEAN and PPT anomalies were regressed against the median flowering date for each species within a location. The regression coefficients were then multiplied by the range of DOYs from the 5th to the 95th percentile median flowering dates within each location. This provided the average change in sensitivity between the 5th and 95th percentile median flowering dates in the community (in d $°C^{-1}$ for TMEAN and d $100mm^{-1}$ for PPT).

Changes in community composition and flowering structure

The SDM and phenoclimatic modeling provided predictions for species presence, flowering onset, and termination under historical, present, and future conditions. I used these predictions to measure changes in species composition between historical and present periods, and between present and future scenarios. Specifically, I assessed species gains and losses proportionally to local species richness in the previous period and calculated overall species turnover (gains + losses divided by species present in both periods).

Next, I examined changes in the start, end, and duration of the flowering season across these periods. The onset of the flowering season was defined as the DOY when 5% of species had started flowering (5th percentile), and the end as the DOY when 95% of species had ceased flowering, with duration as the span between these dates. For each location, I calculated the difference in days for the season's start, end, and duration between historical and present conditions, and between present and future conditions. I also measured changes in the richness of flowering species each month by calculating the percent difference in species numbers under historical versus present and present versus future conditions, relative to local species richness in the preceding period.

Finally, I assessed how patterns of flowering synchrony among species change in response to environmental trends. For each location, I first calculated the overlap in flowering periods between each pair of species in each period, calculating changes in overlap between historical and present conditions, or present and future conditions. Flowering overlap ranged from 0 (no overlap) to 1 (complete overlap in flowering dates). If a species was present in one period but absent in the other, the overlap for that pair was considered 0 in both periods. Using these pairwise overlaps, I measured changes in flowering synchrony within each community using the Bray-Curtis Dissimilarity Index (BCI) (Bray and Curtis, 1957). While BCI is typically used to assess species composition dissimilarity between communities using abundance data, it is also applicable to other categorical data. In this context, BCI measured compositional differences in flowering overlaps between periods, reflecting changes in both the identity and degree of overlap among species pairs. BCI values range from 0 (complete similarity) to 1 (complete dissimilarity), indicating how flowering synchrony shifts between periods for each community.

D. Results

Variation in onset, duration, and termination sensitivities

Across CONUS, mean S_{TMEAN} for both flowering onset and termination was predominantly negative, indicating that species generally advance both the start and end of their flowering periods in warmer-than-average years (Fig. 1a, b). Mean S_{TMEAN} was consistently greater for flowering onset than for termination, leading to average increases in flowering duration under higher temperatures among species in most communities (Fig. 1c). There was significant variation in mean S_{TMEAN} of onset and termination among ecoregions, with the highest sensitivities observed in the Western Cordillera and the lowest in the semi-arid prairies of the Great Plains.

Within communities, the mean S_{TMEAN} for flowering onset and termination varied significantly throughout the season, with increasingly positive values (i.e., smaller advances or greater delays) among successively flowering species (Fig. 1d, e). The S_{TMEAN} of flowering termination generally increased more than that of flowering onset as the season progressed. This caused the S_{TMEAN} of flowering duration to increase throughout the season in most communities (Fig. 1c). Onset and duration sensitivities typically did not change at the same rate, leading to significant variation in seasonal patterns of S_{TMEAN} for flowering duration across CONUS (e.g., Western Cordillera vs. Warm Deserts).

The mean S_{PPT} for flowering onset and termination were generally positive across CONUS, indicating that species usually started and ended flowering later under wetter conditions (Fig. 1g, h). However, the mean S_{PPT} of flowering termination was more variable

than that of flowering onset, with extensive regions showing negative values (i.e., earlier flowering under wetter conditions). This caused substantial regional variation in the mean S_{PPT} of flowering duration, with a notable lengthening of the flowering period among species in Warm Deserts and semi-arid prairies, and widespread shortening in Cold Deserts, the Western Cordillera, and most Eastern ecoregions (Fig. 1i).

Throughout the season, the mean S_{PPT} for flowering onset and termination generally decreased (Fig. 1j, k), suggesting that, under wetter conditions, late-flowering species either delayed flowering onset and termination less or advanced more compared to early-flowering species. As with S_{TMEAN}, the seasonal variation in mean S_{PPT} was greater for flowering termination than for flowering onset, with regional differences leading to variable patterns of seasonal change in the S_{PPT} of flowering duration (Fig. 1l).

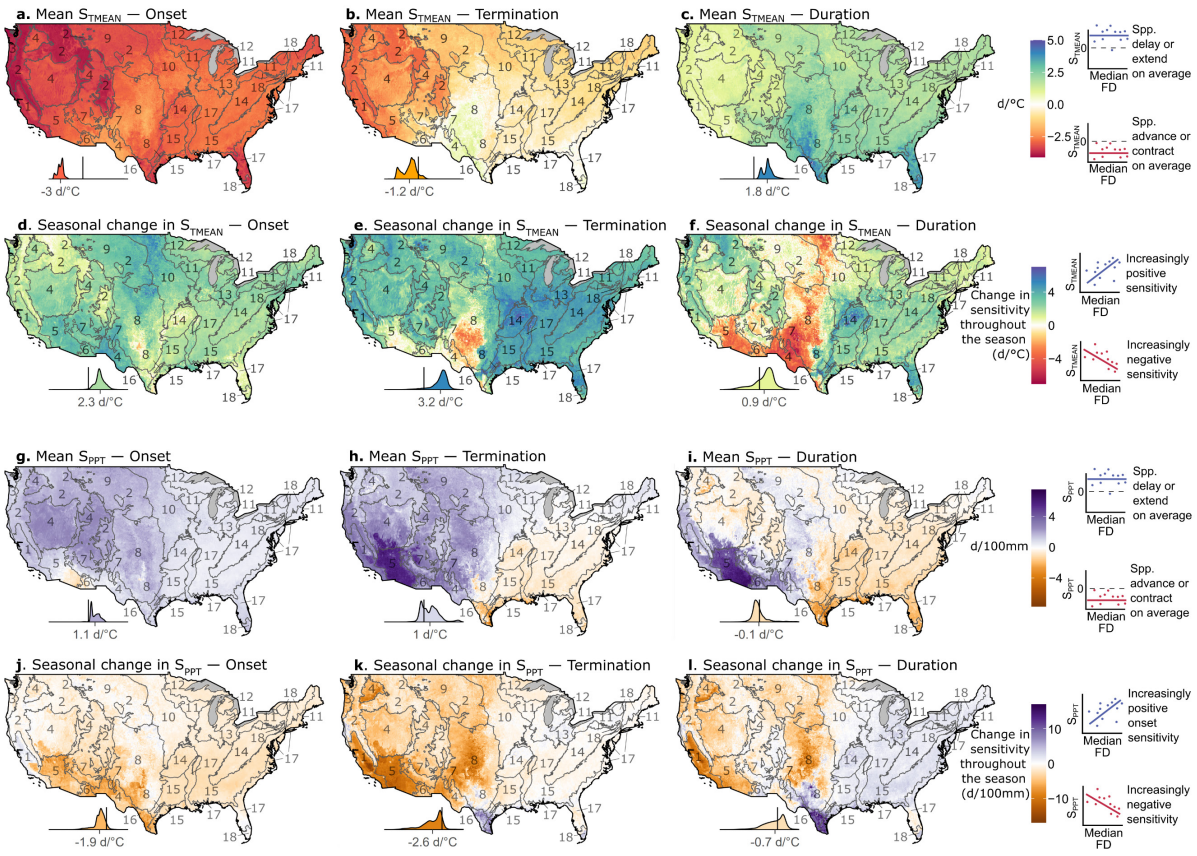

Figure 1—Geographic variation in flowering sensitivity to temperature (S_{TMEAN}) and precipitation (S_{PPT}) within and among communities across the conterminous United States. Panels **a**, **b**, and **c** show the mean S_{TMEAN} for flowering onset, termination, and duration (in days per °C) for species predicted to co-occur in each 12 km grid cell in 2001-2020. Panels **d**, **e**, and **f** illustrate the average change in S_{TMEAN} between species flowering early (5th percentile median flowering dates) and late (95th percentile median flowering dates). Positive values indicate greater delays or lesser advances in flowering with warmer temperatures as the season progresses. Panels **g**, **h**, and **i** display variation in mean S_{PPT} among co-occurring species in each grid cell. Panels **j**, **k**, and **l** show the change in mean S_{PPT} among species throughout the season. For graphing, the color scale was capped to the central 99% of the data to avoid distortion of the range from extreme values. Insets in each panel depict the distribution of the focal parameter, with vertical bars indicating a value of 0 and the x-axis value representing the mean. Subdivisions labeled 1-18 represent level II ecoregions. *1) Mediterranean*

California, 2) Western Cordillera, 3) Marine West Coast Forest, 4) Cold Deserts, 5) Warm Deserts, 6) Western Sierra Madre Piedmont, 7) Upper Gila Mountains, 8) South-Central Semi-arid Prairies, 9) West-Central Semi-arid Prairies, 10) Temperate Prairies, 11) Mixed Wood Plains, 12) Mixed Wood Shield, 13) Central USA Plains, 14) Ozark, Ouachita-Appalachian Forests, 15) Southeastern USA Plains, 16) Tamaulipas-Texas Semi-arid Plains, 17) Mississippi Alluvial and Southeast USA Coastal Plain, 18) Atlantic Highlands.

Changes in species composition

The SDMs predicted moderate but regionally variable species gains and losses across CONUS due to climate and land cover changes from the historical (1961-1980) to the present period (2001-2020) (Fig. 2a, b). Gains and losses were most pronounced in the Great Plains, with the greatest gains in the Temperate Prairies and Central USA plains, and the largest losses in the South-central Semi-arid Prairies. Eastern and Western ecoregions saw comparatively lower changes, resulting in the highest species turnover in the central U.S. and the lowest in the Western Cordilleras (e.g., Rocky Mountains) and Southeastern USA Plains (Fig. 2c). Within the Great Plains, the areas experiencing the greatest species gains and losses occurred respectively East and West of the 100th meridian West, a longitudinal boundary separating the humid Eastern and arid Western climates (Seager et al. 2018). Climate change differentially impacted both sides of the 100th meridian between the historical and present period (1961-1980, and 2001-2020, respectively), with intensifying water deficit in the West and wetter conditions in the East (Appendix 4—Fig. S2).

Future climate and land cover changes are expected to cause species gains of similar magnitude to those from historical to present periods, with the most pronounced gains also

expected in the Great Plains (but not in the Central USA Plains) (Fig. 2d). Projected species losses were much greater than in recent decades, especially in the Great Plains and Central USA Plains (Fig. 2e). Western ecoregions are expected to see more severe losses than during recent decades, while Eastern ecoregions remain relatively stable. Overall, the greatest species turnover is predicted for the Great Plains, primarily driven by losses, with moderate turnover in Western ecoregions and mild changes in Eastern ecoregions (Fig. 2f). The areas of greatest projected species losses and turnover in the central United States are also predicted to experience the greatest aridification, agricultural intensification, and loss of grasslands in coming decades (Appendix 3—Figs. S3, S4).

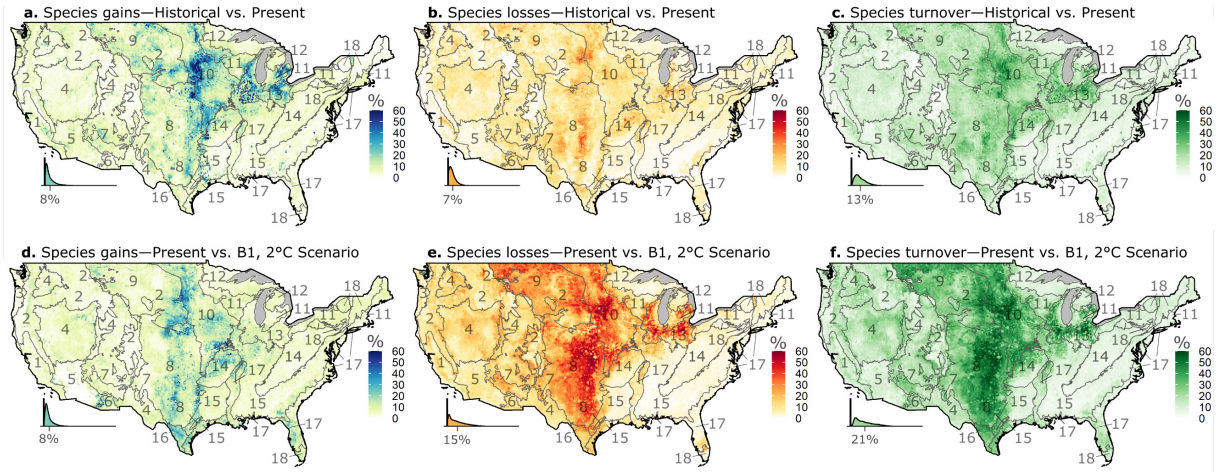

Figure 2—Predicted species gains, losses, and turnover between historical (1961-1980) and present (2001-2020) environmental conditions, and between present and future conditions. Percentages were calculated using species distributions from the preceding period as a baseline (1961-1980 for panels **a-c**, 2001-2020 for panels **d-f**). Future conditions are based on 2 °C warming and 2080 land use scenarios. The color scales are capped at the 99th percentile of the data to prevent distortion from extreme values. Insets show the distribution of each metric across locations, with vertical black lines indicating 0% and the x-axis value representing the mean. Subdivisions labeled 1-18 represent level II ecoregions. *1) Mediterranean California, 2) Western Cordillera, 3) Marine West Coast Forest, 4)*

Cold Deserts, 5) Warm Deserts, 6) Western Sierra Madre Piedmont, 7) Upper Gila Mountains, 8) South-Central Semi-arid Prairies, 9) West-Central Semi-arid Prairies, 10) Temperate Prairies, 11) Mixed Wood Plains, 12) Mixed Wood Shield, 13) Central USA Plains, 14) Ozark, Ouachita-Appalachian Forests, 15) Southeastern USA Plains, 16) Tamaulipas-Texas Semi-arid Plains, 17) Mississippi Alluvial and Southeast USA Coastal Plain, 18) Atlantic Highlands.

Changes to the start, duration, and end of the flowering season

Climate change-induced shifts in species distributions and flowering phenology altered the start, end, and duration of the flowering season across ecoregions. SOS advanced between the historical and present periods, with the largest shifts in the Sierra Nevada and Klamath Mountains (Fig. 3a). EOS was delayed in Eastern ecoregions, while responses in the West were more variable (e.g., some areas of the Western Cordillera saw delays, others did not) (Fig. 3b). SOS and EOS generally moved in opposite directions, with SOS showing larger shifts where they moved in the same direction. Consequently, the flowering season duration (DOS) increased across most of CONUS, with the greatest increases in the Klamath Mountains and Western Cordillera, and decreases mainly in the Warm Desert and some areas within the semi-arid prairies of the Great Plains (Fig. 3c).

Future environmental conditions were projected to generate more drastic changes than those observed in recent decades. SOS advances were projected to be larger, with fewer areas showing SOS delays (e.g., Warm Deserts) (Fig. 3d). EOS delays were projected to be greater and more consistent across regions than in recent decades (e.g., few EOS advances in Western ecoregions) (Fig. 3e). As a result, DOS was predicted to increase significantly across most of CONUS (Fig. 3f).

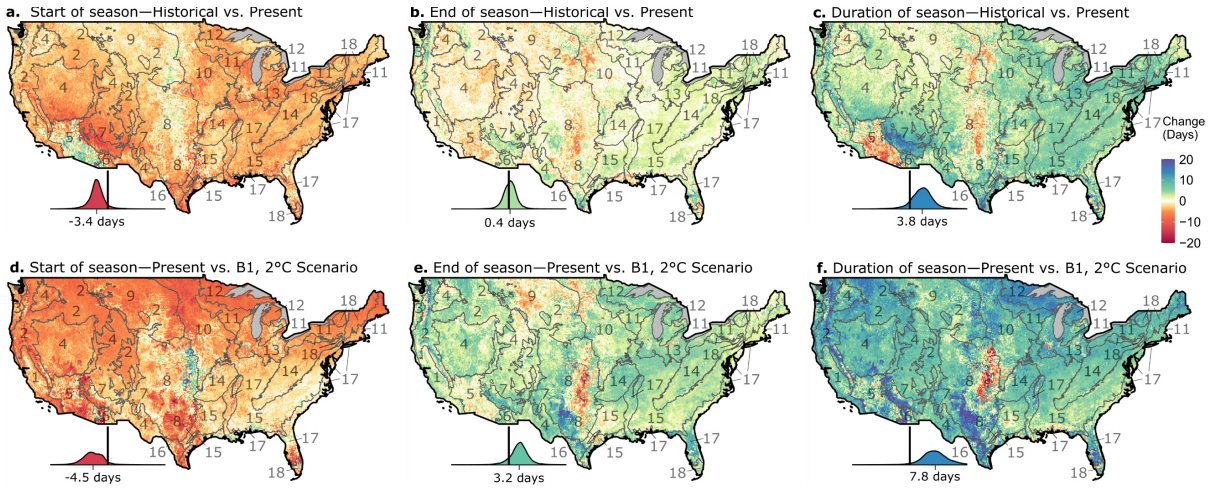

Figure 3—Change in the start, end, and duration of the flowering season under recent and future climatic and land cover change. **a**, **b** and **c** depict changes in the start, end, and duration of the season predicted between the historical (1961-1980) and present (2001-2020) periods. **d**, **e**, and **f** show predicted changes between recent conditions and future conditions expected under 2 °C of warming and land cover patterns for the year 2080 under the Special Report on Emissions Scenario (SRES) B1. For graphing, the color scale was capped to the central 99% of the data to avoid distortion of the range from extreme values. Insets in each panel show the distribution of the focal metric across all locations within the conterminous United States, with vertical black lines representing values of 0 days and the marked value in the x-axis representing the mean. Subdivisions in each panel—labeled 1-18—represent level II ecoregions. *1) Mediterranean California, 2) Western Cordillera, 3) Marine West Coast Forest, 4) Cold Deserts, 5) Warm Deserts, 6) Western Sierra Madre Piedmont, 7) Upper Gila Mountains, 8) South-Central Semi-arid Prairies, 9) West-Central Semi-arid Prairies, 10) Temperate Prairies, 11) Mixed Wood Plains, 12) Mixed Wood Shield, 13) Central USA Plains, 14) Ozark, Ouachita-Appalachian Forests, 15) Southeastern USA Plains, 16) Tamaulipas-Texas Semi-arid Plains, 17) Mississippi Alluvial and Southeast USA Coastal Plain, 18) Atlantic Highlands.*

Change in the seasonal distribution of flowering species diversity

Flowering species diversity remained mostly unchanged during the winter months before the flowering season in most regions (Fig. 4a, b), but increased by February in low-latitude ecoregions with early spring conditions (e.g., Warm Deserts) (Fig. 4c). More significant changes occurred in spring, with widespread increases due to the earlier onset of spring across CONUS (Fig. 4d-f). Decreases first occurred in April and May in regions experiencing early onset of summer drought (e.g., Warm Deserts) or where sharp species losses between historical and present periods occurred (e.g., South-Central Semi-arid Prairies; Fig. 2b). More ecoregions experienced declines in flowering diversity during summer (Fig. 4g-i), with the onset of declines occurring earlier in the year in arid ecoregions (e.g., May to June for Cold Deserts vs. June to July for the Western Cordillera). Fall changes were smaller across CONUS (Fig. 4j-l), except in regions that experienced sharp species gains or losses (e.g., South-Central Semi-arid Prairies, Southeastern USA Plains; Fig. 2a, b).

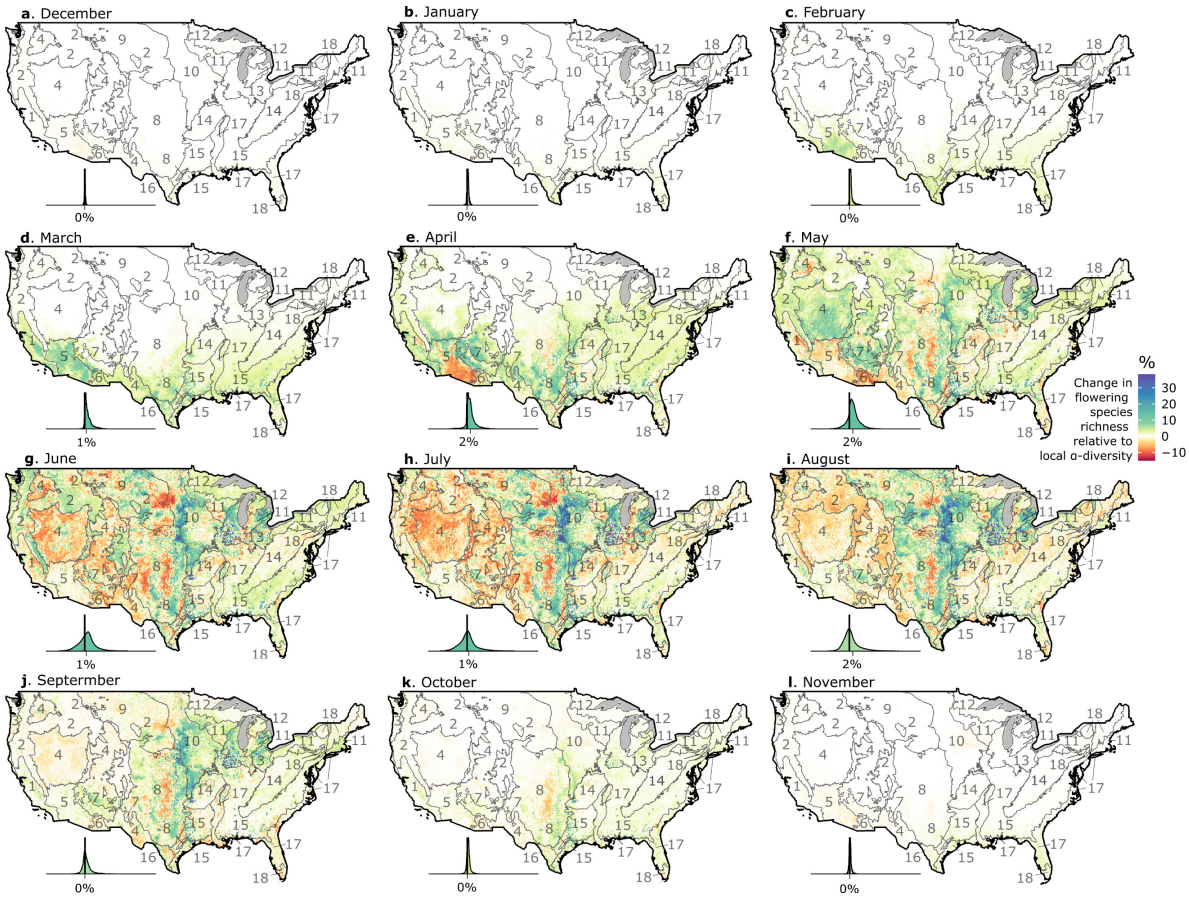

Figure 4—Changes between historical and recent periods in the richness of species flowering each month across the conterminous United States. The color scale in each panel shows the change in the number of species observed in flowering in that month for each location relative to the total species diversity of that site during the historical period (1961-1990). For graphing, the color scale was capped to the central 99% of the data to avoid distortion of the range from extreme values. Insets in each panel show the distribution of changes across all locations within the conterminous United States for that month, with vertical black lines representing values of 0% and the marked value in the x-axis representing the mean. Subdivisions in each panel—labeled 1-18—represent level II ecoregions. *1) Mediterranean California, 2) Western Cordillera, 3) Marine West Coast Forest, 4) Cold Deserts, 5) Warm Deserts, 6) Western Sierra Madre Piedmont, 7) Upper Gila Mountains, 8) South-Central Semi-arid Prairies, 9) West-Central Semi-arid Prairies, 10) Temperate Prairies, 11) Mixed Wood Plains, 12) Mixed Wood Shield, 13) Central USA Plains, 14) Ozark, Ouachita-*

Appalachian Forests, 15) Southeastern USA Plains, 16) Tamaulipas-Texas Semi-arid Plains, 17) Mississippi Alluvial and Southeast USA Coastal Plain, 18) Atlantic Highlands.

Under projected climate conditions, decreases in flowering species diversity were larger and more widespread across ecoregions compared to recent decades. Changes during winter were modest (Fig. 5a-c), and increases in spring flowering diversity were more modest than in recent decades, with declines occurring earlier in many ecoregions (e.g., Warm Deserts, Mediterranean California) (Fig. 5d-f). Summer and fall changes in flowering diversity largely reflected projected species gains and losses (Fig. 2), with sharp decreases in regions facing severe species losses (e.g., Great Plains) (Fig. 5g-l). In contrast, flowering diversity was projected to increase during summer and fall in many Eastern ecoregions (e.g., Southeast USA Plains, Southeast USA Coastal Plain) (Fig. 5g-l), which were predicted to have mild species losses, moderate gains, and delayed EOS (Figs. 2d-f, 3d-f).

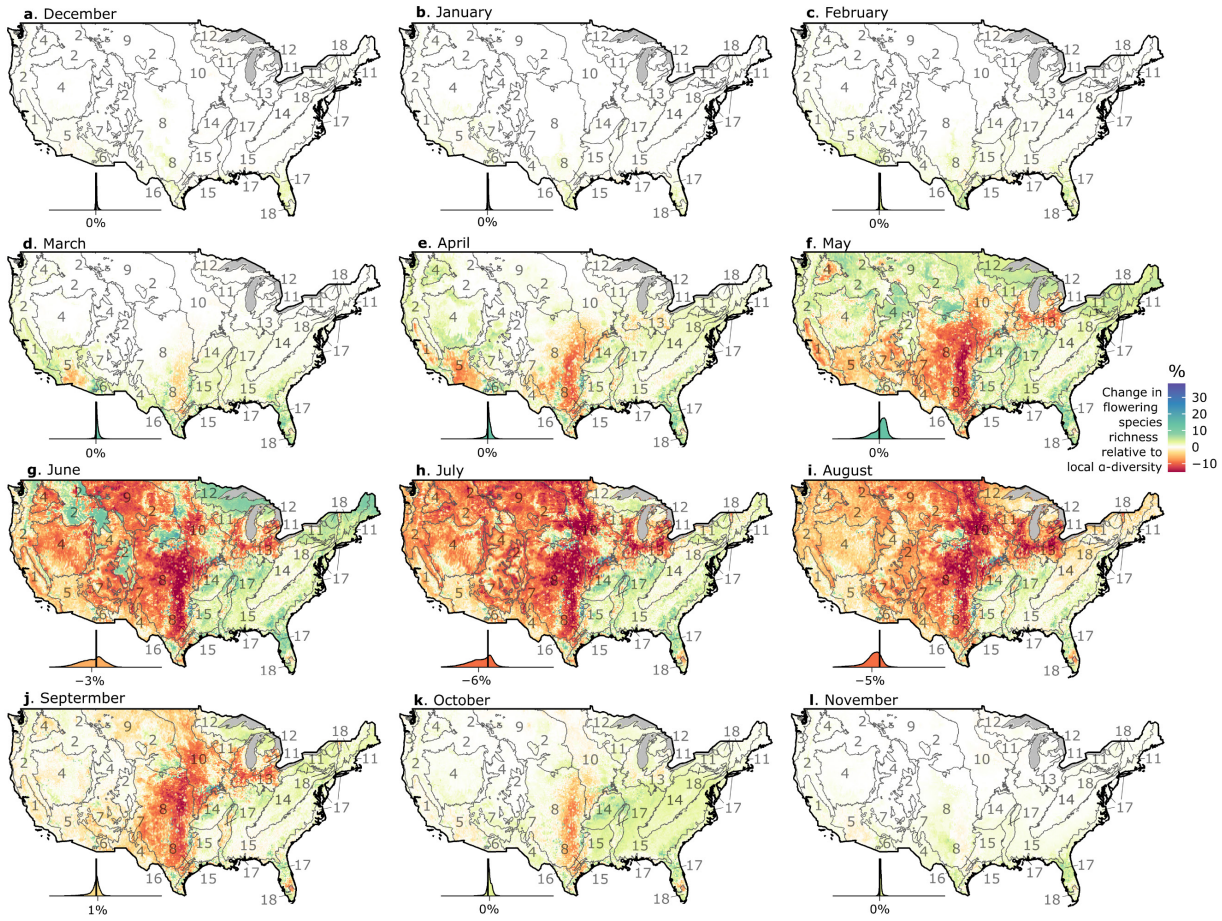

Figure 5—Changes between present and future environmental conditions in the richness of species flowering each month across the conterminous United States. The color scale in each panel shows the projected change—under a 2 °C warming scenario—in the number of species observed flowering each month in each location relative to the total species diversity of that site during the present period (2001-2020). For graphing, the color scale was capped to the central 99% of the data to avoid distortion of the range from extreme values. Insets in each panel show the distribution of changes across all locations within the conterminous United States for that month, with vertical black lines representing values of 0% and the marked value in the x-axis representing the mean. Subdivisions in each panel—labeled 1-18—represent level II ecoregions. *1) Mediterranean California, 2) Western Cordillera, 3) Marine West Coast Forest, 4) Cold Deserts, 5) Warm Deserts, 6) Western Sierra Madre Piedmont, 7) Upper Gila Mountains, 8) South-Central Semi-arid Prairies, 9) West-Central Semi-arid*

Prairies, 10) Temperate Prairies, 11) Mixed Wood Plains, 12) Mixed Wood Shield, 13) Central USA Plains, 14) Ozark, Ouachita-Appalachian Forests, 15) Southeastern USA Plains, 16) Tamaulipas-Texas Semi-arid Plains, 17) Mississippi Alluvial and Southeast USA Coastal Plain, 18) Atlantic Highlands.

Changes in the composition of flowering synchronies

Changes in patterns of flowering synchrony among species were overwhelmingly determined by shifts in species composition across CONUS, so the degree of flowering dissimilarity between periods largely resembled geographic patterns in species turnover among sites (Fig. 6). Flowering dissimilarity between periods was more severe under projected environmental conditions than observed in recent decades, with the greatest reassembly observed in ecoregions across the Great Plains and in the Central USA Plains.

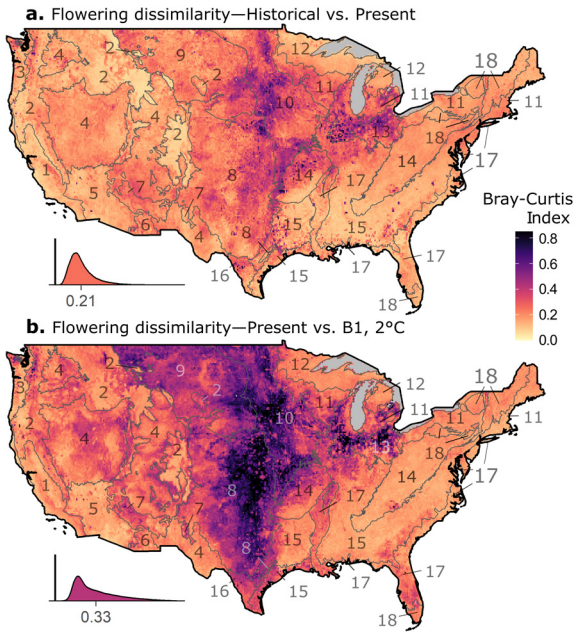

Figure 6—Dissimilarity in the composition of flowering synchronies among sympatric species. The color scale represents the Bray-Curtis Dissimilarity index, with values of 0 indicating no changes in species composition and degree of flowering overlap, and values of 1 indicating all pairs of overlapping species were gained or lost relative to the preceding period. For graphing, the color scale was capped to the central 99% of the data to avoid distortion of the range from extreme values. Insets in each panel show the distribution of changes across all locations within the conterminous United States for that month, with vertical black lines representing a BCI value of 0 and the marked value in the x-axis representing the mean. Subdivisions in each panel—labeled 1-18—represent level II ecoregions. *1) Mediterranean California, 2) Western Cordillera, 3) Marine West Coast Forest, 4) Cold Deserts, 5) Warm Deserts, 6) Western Sierra Madre Piedmont, 7) Upper Gila Mountains, 8) South-Central Semi-arid Prairies, 9) West-Central Semi-arid Prairies, 10) Temperate Prairies, 11) Mixed Wood Plains, 12) Mixed Wood Shield, 13) Central USA Plains, 14) Ozark, Ouachita-Appalachian Forests, 15) Southeastern USA Plains, 16) Tamaulipas-Texas Semi-arid Plains, 17) Mississippi Alluvial and Southeast USA Coastal Plain, 18) Atlantic Highlands*.

E. Discussion

Restructuring of the flowering season under climate change could have cascading consequences across trophic levels, but data constraints and incomplete knowledge of species' flowering responses to climate limit our understanding of community-level responses. Modeling the distribution and flowering phenology of thousands of species across the conterminous United States shows that—within species—the onset, duration, and end of flowering respond differently to temperature and precipitation, with broad variation in mean sensitivity among and within communities. Climate change has resulted in widespread shifts in species composition and the timing and length of the flowering season across ecoregions, leading to region-specific changes in the seasonal distribution of flowering diversity and patterns of flowering overlap that are projected to intensify under ongoing climate trends. These findings outline broad macroecological changes, revealing uneven impacts of climate change on the identity, diversity, and flowering patterns of co-occurring species across biomes.

Onset and termination respond differently to temperature and precipitation

Previous studies indicate that the mean S_{TMEAN} of flowering onset varies across communities (e.g., Cook et al. 2012, Zhang et al. 2015, Panchen and Gorelick 2017, Prevéy et al. 2017, Park et al. 2019, Miller et al. 2023) and that greater mean sensitivity of onset than termination often extends flowering among species (Zhang et al. 2015, Augspurger and Zaya 2020, Bucher and Romermann 2020, Li et al. 2021, Chen et al. 2023, Austin et al. 2024). By assessing sensitivities for onset, termination, and duration within a spatially explicit

framework, these results corroborate these findings throughout CONUS while also uncovering novel patterns of variation within and between communities. For example, S_{TMEAN} and S_{PPT} of flowering termination and duration showed more regional variability than those of flowering onset. Consequently, as most research has focused on flowering onset, these results caution against extrapolating patterns of variation in phenological firsts to make inferences about variation in termination and duration among species and regions.

Within communities, this study's findings support studies showing decreasing phenological sensitivity to temperature throughout the season in temperate biomes (e.g., Cook et al. 2012, Mazer et al. 2013, Park et al. 2019) and others showing advances in spring but delays in autumn phenology under warming (Delgado et al. 2020, Beil et al. 2021, Li et al. 2021, Roslin et al. 2021). These results are consistent with the hypothesis that temperature advances phenology during spring due to accelerated developmental rates, while phenophases occurring during autumn are cued directly by seasonal cooling, which may be lessened or delayed in warmer-than-average years (Fu et al. 2018, Zohner et al. 2023). However, termination sensitivity to temperature decreased but did not reverse to delays late in the season in some regions (e.g., Western Cordillera), and remained stable or even transitioned from delays to advances in others (e.g., arid ecoregions of the South and Southwest), highlighting that the primary cuing mechanisms for flowering termination (e.g., photoperiod, drought) vary among floras (Gill et al. 2015). Overall, seasonal changes in onset and termination S_{TMEAN} generated stark differences among regions in the direction and degree to which flowering duration sensitivity varies throughout the season (e.g., South-Central Semi-arid Prairies versus the Western Cordillera), suggesting that the impacts of

warming on floral resource availability early and late in the season will vary among ecoregions.

Mean S_{PPT} also varied markedly throughout the season, with different geographic patterns among communities from those of S_{TMEAN}. Consequently, the joint impacts on flowering of correlated temperature and precipitation variation (e.g., responses to warmer- and drier-than-average conditions) are likely to be highly idiosyncratic across regions. If so, dissimilar geographic variation in S_{TMEAN} and S_{PPT} among and within communities might explain why the effects of precipitation on phenology can be highly heterogeneous at large spatial scales (Peñuelas et al. 2004, Wang et al. 2022).

Flowering reassembly is widespread but regionally variable

This study reveals major changes in the structure of the flowering season, driven by shifts in i) species composition, ii) the season's timing and duration, iii) the seasonal distribution of co-flowering species diversity, and iv) flowering synchronies between species. These changes could have cascading ecological impacts. For instance, the seasonal availability and diversity of floral resources strongly influences pollinator diversity and abundance (Potts et al. 2003, Fründ et al. 2010, Scheper et al. 2015). Specialist bees often travel further for fewer resources under floral scarcity, reducing brood provisioning (Minckley 1994, Williams and Kremen 2007, Pope and Jha 2018, Schenk et al. 2018), and populations sizes of generalist pollinators tend to correlate with local floral density (Potts et al. 2003, Scheper et al. 2015). Thus, decreases in flowering diversity or phenological mismatches can result in pollinator declines, potentially disrupting ecosystem services like crop pollination (Woodcock et al.

2018). In turn, decreases in phenological synchrony between plant species can affect their fitness by impacting density-dependent processes like pollinator attraction, fertilization, seed dispersal, or flower and seed predation (Nilsson and Wastljung 1987, Elzinga et al. 2007, Carlo and Morales 2008, Jones and Comita 2010, Bergamo et al. 2020). Additionally, pollen from many wind-pollinated trees and grasses is allergenic to humans, so longer population- and community-level flowering might intensify the allergy season (Anderegg et al. 2021).

The wide regional variation in flowering reassembly reported here shows that potential ecological impacts likely differ across floras. For example, the diversity of co-flowering species sharply increased early and decreased late in the season in many ecoregions (e.g., Western Cordillera, Cold Deserts, Mediterranean California), which could result in opposing effects on density-dependent processes such as pollinator attraction or insect foraging success during spring and summer. Some communities experienced consistent declines in flowering diversity throughout the year due to substantial species losses (e.g., areas of the South-Central Semi-arid Prairies), which could decrease the diversity of organisms reliant on flowers throughout the year. Other regions experienced consistent increases in flowering diversity throughout the year due to net species gains (e.g., Southeastern ecoregions), which may not affect different seasons disproportionately but could alter ecological processes through novel species interactions. Regardless of specific patterns, these alterations to flowering diversity and flowering synchrony within communities have the potential to substantially alter the selective environments encountered by plants and interaction taxa across CONUS (Elzinga et al. 2007).

Although the nature and degree of ecological impacts may vary across ecoregions, these analyses identify biomes that might be particularly vulnerable to climate change.

Flowering reassembly due to phenological shifts and species turnover was most pronounced within the Great Plains, where it was primarily driven by severe species losses. These areas, located around the 100th meridian West—a bioclimatic boundary separating the humid East from the arid West (Sieger et al. 2018)—saw increasingly arid conditions in the West and more humid conditions in the East between the historical and present periods (Appendix 4—Fig. S2). This aligns with studies showing that drier conditions reduce biomass, shift species dominance, and cause rapid species losses in grasslands (Chase et al. 2000, Cleland et al. 2013, McDougall et al. 2024). Future projections indicate that aridification will intensify throughout the South-Central U.S. (a region harboring most of the United State's agricultural land) generating species losses that may be worsened by ongoing grassland conversion to agriculture (Appendix 4—Fig. S4). To the extent that losses in flowering species diversity reduce pollinator diversity and abundance, these trends might have proximate impacts on human welfare through widespread disruptions to crop pollination services.

Interestingly, regions like the Sierra Nevada and Cascades within the Western Cordillera (facing strong aridification) or many Southeastern ecoregions (facing substantial land cover changes) showed much lower historical and projected species turnover than the Great Plains. While I cannot identify the ultimate causes of these trends from these analyses, the greater changes observed across the Great Plains might be caused by potential limits to species ranges imposed by the arid-humid bioclimatic boundary, with aridification trends leading to local extirpation of humidity-adapted species at the boundaries of their ranges (Barnes and Harrison 1982, Epstein et al. 1996, Anderegg and HilleRisLambers 2015, Berdugo et al. 2020).

Limitations and future directions

This study provides a broad macroecological outline of changes to the flowering season due to the combined effects of flowering responses and shifts in species distributions across North American ecoregions. However, multiple factors make it difficult to predict specific ecological consequences from these changes. First, these results show shifts in the timing and diversity of flowering across scales, but not in the overall abundance of floral resources, which could amplify or reduce the ecological impacts of altered flowering times. For example, if warming decreases the density of co-flowering species, reduced flower production per species would worsen the effects of lower flowering richness, while increased production could mitigate these effects. Additionally, many plant communities are dominated by a few species whose flowering responses may deviate from the wider community. Similarly, ecological outcomes often depend on a handful of species interactions (e.g., fitness of specialist pollinators and their mutualist plants) whose responses might not match those of the wider community. Given these and other complexities, assessing the effects of climate change on floral resource production is crucial to determine whether the spatiotemporal redistribution of co-flowering diversity will lead to concordant changes in floral resource availability. In turn, forecasting more precise ecological outcomes will require focusing analyses on key species based on local abundance, floral output, functional traits, or other relevant attributes.

F. Literature cited

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific data*, *5*(1), 1-12. https://doi.org/10.1038/sdata.2017.191

Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B. and Anderson, R.P. (2015), spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, 38: 541-545. https://doi.org/10.1111/ecog.01132

Anderegg, L. D. L., & HilleRisLambers, J. (2016). Drought stress limits the geographic ranges of two tree species via different physiological mechanisms. *Global Change Biology*, *22*(3), 1029–1045. https://doi.org/10.1111/gcb.13148

Anderegg, W. R. L., Abatzoglou, J. T., Anderegg, L. D. L., Bielory, L., Kinney, P. L., & Ziska, L. (2021). Anthropogenic climate change is worsening North American pollen seasons. *Proceedings of the National Academy of Sciences*, *118*(7), e2013284118. https://doi.org/10.1073/pnas.2013284118

Augspurger, C. K., & Zaya, D. N. (2020). Concordance of long-term shifts with climate warming varies among phenological events and herbaceous species. *Ecological Monographs*, *90*(4), e01421. https://doi.org/10.1002/ecm.1421

Austin, M. W., Smith, A. B., Olsen, K. M., Hoch, P. C., Krakos, K. N., Schmocker, S. P., & Miller-Struttmann, N. E. (2024). Climate change increases flowering duration, driving phenological reassembly and elevated co-flowering richness. *New Phytologist*, *243*(6), 2486–2500. https://doi.org/10.1111/nph.19994

Barnes, P. W., & Harrison, A. T. (1982). Species distribution and community organization in a Nebraska Sandhills mixed prairie as influenced by plant/soil-water relationships. *Oecologia*, *52*(2), 192–201. https://doi.org/10.1007/BF00363836

Beil, I., Kreyling, J., Meyer, C., Lemcke, N., & Malyshev, A. V. (2021). Late to bed, late to rise—Warmer autumn temperatures delay spring phenology by delaying dormancy. *Global Change Biology*, *27*(22), 5806–5817. https://doi.org/10.1111/gcb.15858

Belitz, M. W., Larsen, E. A., Ries, L., & Guralnick, R. P. (2020). The accuracy of phenology estimators for use with sparsely sampled presence-only observations. *Methods in Ecology and Evolution*, *11*(10), 1273–1285. https://doi.org/10.1111/2041-210X.13448

Belitz, M. W., Larsen, E. A., Shirey, V., Li, D., & Guralnick, R. P. (2023). Phenological research based on natural history collections: Practical guidelines and a lepidopteran case study. *Functional Ecology*, *37*(2), 234–247. https://doi.org/10.1111/1365-2435.14173

Berdugo, M., Delgado-Baquerizo, M., Soliveres, S., Hernández-Clemente, R., Zhao, Y., Gaitán, J. J., Gross, N., Saiz, H., Maire, V., Lehmann, A., Rillig, M. C., Solé, R. V., & Maestre, F. T. (2020). Global ecosystem thresholds driven by aridity. *Science*, *367*(6479), 787–790. https://doi.org/10.1126/science.aay5958

Bergamo, P. J., Susin Streher, N., Traveset, A., Wolowski, M., & Sazima, M. (2020). Pollination outcomes reveal negative density-dependence coupled with interspecific facilitation among plants. *Ecology Letters*, *23*(1), 129–139. https://doi.org/10.1111/ele.13415

Bucher, S. F., & Römermann, C. (2020). Flowering patterns change along elevational gradients and relate to life-history strategies in 29 herbaceous species. *Alpine Botany*, *130*(1), 41–58. https://doi.org/10.1007/s00035-020-00231-w

Calinger, K. M., Queenborough, S., & Curtis, P. S. (2013). Herbarium specimens reveal the footprint of climate change on flowering trends across north-central North America. *Ecology Letters*, *16*(8), 1037–1044. https://doi.org/10.1111/ele.12135

CaraDonna, P. J., Iler, A. M., & Inouye, D. W. (2014). Shifts in flowering phenology reshape a subalpine plant community. *Proceedings of the National Academy of Sciences*, *111*(13), 4916–4921. https://doi.org/10.1073/pnas.1323073111

Carlo, T. A., & Morales, J. M. (2008). Inequalities in fruit-removal and seed dispersal: Consequences of bird behaviour, neighbourhood density and landscape aggregation. *Journal of Ecology*, *96*(4), 609–618. https://doi.org/10.1111/j.1365-2745.2008.01379.x

Chase, J. M., Leibold, M. A., Downing, A. L., & Shurin, J. B. (2000). The Effects of Productivity, Herbivory, and Plant Species Turnover in Grassland Food Webs. *Ecology*, *81*(9), 2485–2497. https://doi.org/10.1890/0012-9658(2000)081[2485:TEOPHA]2.0.CO;2

Chen, Y., Collins, S. L., Zhao, Y., Zhang, T., Yang, X., An, H., Hu, G., Xin, C., Zhou, J., Sheng, X., He, M., Zhang, P., Guo, Z., Zhang, H., Li, L., & Ma, M. (2023). Warming reduced flowering synchrony and extended community flowering season in an alpine meadow on the Tibetan Plateau. *Ecology*, *104*(1), e3862. https://doi.org/10.1002/ecy.3862

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*(7), 357–365. https://doi.org/10.1016/j.tree.2007.04.003

Cleland, E. E., Collins, S. L., Dickson, T. L., Farrer, E. C., Gross, K. L., Gherardi, L. A., Hallett, L. M., Hobbs, R. J., Hsu, J. S., Turnbull, L., & Suding, K. N. (2013). Sensitivity of grassland plant community composition to spatial vs. Temporal variation in precipitation. *Ecology*, *94*(8), 1687–1696. https://doi.org/10.1890/12-1006.1

Cook, B. I., Wolkovich, E. M., Davies, T. J., Ault, T. R., Betancourt, J. L., Allen, J. M., Bolmgren, K., Cleland, E. E., Crimmins, T. M., Kraft, N. J. B., Lancaster, L. T., Mazer, S. J., McCabe, G. J., McGill, B. J., Parmesan, C., Pau, S., Regetz, J., Salamin, N., Schwartz, M. D., & Travers, S. E. (2012). Sensitivity of Spring Phenology to Warming Across Temporal and Spatial Climate Gradients in Two Independent Databases. *Ecosystems*, *15*(8), 1283–1294. https://doi.org/10.1007/s10021-012-9584-5

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, *88*(11), 2783-2792. https://doi.org/10.1890/07-0539.1

Daru, B. H. (2024). Predicting undetected native vascular plant diversity at a global scale. *Proceedings of the National Academy of Sciences*, *121*(34), e2319989121. https://doi.org/10.1073/pnas.2319989121

Delgado, M. del M., Roslin, T., Tikhonov, G., Meyke, E., Lo, C., Gurarie, E., Abadonova, M., Abduraimov, O., Adrianova, O., Akimova, T., Akkiev, M., Ananin, A., Andreeva, E., Andriychuk, N., Antipin, M., Arzamascev, K., Babina, S., Babushkin, M., Bakin, O., … Ovaskainen, O. (2020). Differences in spatial versus temporal reaction norms for spring and autumn phenological events. *Proceedings of the National Academy of Sciences*, *117*(49), 31249–31258. https://doi.org/10.1073/pnas.2002713117

Elzinga, J. A., Atlan, A., Biere, A., Gigord, L., Weis, A. E., & Bernasconi, G. (2007). Time after time: Flowering phenology and biotic interactions. *Trends in Ecology & Evolution*, *22*(8), 432–439. https://doi.org/10.1016/j.tree.2007.05.006

Epstein, H. e., Lauenroth, W. k., Burke, I. c., & Coffin, D. p. (1996). Ecological responses of dominant grasses along two climatic gradients in the Great Plains of the United States. *Journal of Vegetation Science*, *7*(6), 777–788. https://doi.org/10.2307/3236456

Fitter, A. H., & Fitter, R. S. R. (2002). Rapid Changes in Flowering Time in British Plants. *Science*, *296*(5573), 1689–1691. https://doi.org/10.1126/science.1071617

Fründ, J., Linsenmair, K. E., & Blüthgen, N. (2010). Pollinator diversity and specialization in relation to flower diversity. *Oikos*, *119*(10), 1581–1590. https://doi.org/10.1111/j.1600-0706.2010.18450.x

Fu, Y. H., Piao, S., Delpierre, N., Hao, F., Hänninen, H., Liu, Y., Sun, W., Janssens, I. A., & Campioli, M. (2018). Larger temperature response of autumn leaf senescence than spring leaf-out phenology. *Global Change Biology*, *24*(5), 2159–2168. https://doi.org/10.1111/gcb.14021

Gamble, D. E., & Mazer, S. J. (2022). Spatial uncertainty in herbarium data: simulated displacement but not error distance alters estimates of phenological sensitivity to

climate in a widespread California wildflower. *Ecography*, 2022(10), e06107. https://doi.org/10.1111/ecog.06107

Gill, A. L., Gallinat, A. S., Sanders-DeMott, R., Rigden, A. J., Short Gianotti, D. J., Mantooth, J. A., & Templer, P. H. (2015). Changes in autumn senescence in northern hemisphere deciduous trees: A meta-analysis of autumn phenology studies. *Annals of Botany*, *116*(6), 875–888. https://doi.org/10.1093/aob/mcv055

González-Suárez, P., Walker, C. H., & Bennett, T. (2020). Bloom and bust: understanding the nature and regulation of the end of flowering. Current Opinion in Plant Biology, 57, 24-30. https://doi.org/10.1016/j.pbi.2020.05.009

Jabis, M. D., Winkler, D. E., & Kueppers, L. M. (2020). Warming acts through earlier snowmelt to advance but not extend alpine community flowering. *Ecology*, *101*(9), e03108. https://doi.org/10.1002/ecy.3108

Jones, F. A., & Comita, L. S. (2010). Density-dependent pre-dispersal seed predation and fruit set in a tropical tree. *Oikos*, *119*(11), 1841–1847. https://doi.org/10.1111/j.1600-0706.2010.18547.x

Kelly, A. E., & Goulden, M. L. (2008). Rapid shifts in plant distribution with recent climate change. *Proceedings of the National Academy of Sciences*, *105*(33), 11823–11826. https://doi.org/10.1073/pnas.0802891105

Koenker, R., Portnoy, S., Ng, P. T., Zeileis, A., Grosjean, P., & Ripley, B. D. (2018). Package 'quantreg'. *Reference manual available at R-CRAN:* https://cran.rproject.org/web/packages/quantreg/quantreg.pdf

Li, D., Barve, N., Brenskelle, L., Earl, K., Barve, V., Belitz, M. W., Doby, J., Hantak, M. M., Oswald, J. A., Stucky, B. J., Walters, M., & Guralnick, R. P. (2021). Climate, urbanization, and species traits interactively drive flowering duration. *Global Change Biology*, *27*(4), 892–903. https://doi.org/10.1111/gcb.15461

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, *2*(3), 18-22. https://CRAN.R-project.org/doc/Rnews/

MacDougall, A. S., Esch, E., Chen, Q., Carroll, O., Bonner, C., Ohlert, T., Siewert, M., Sulik, J., Schweiger, A. K., Borer, E. T., Naidu, D., Bagchi, S., Hautier, Y., Wilfahrt, P., Larson, K., Olofsson, J., Cleland, E., Muthukrishnan, R., O'Halloran, L., … Seabloom, E. W. (2024). Widening global variability in grassland biomass since the 1980s. *Nature Ecology & Evolution*, 1–12. https://doi.org/10.1038/s41559-024-02500-x

Mazer, S. J., Travers, S. E., Cook, B. I., Davies, T. J., Bolmgren, K., Kraft, N. J. B., Salamin, N., & Inouye, D. W. (2013). Flowering date of taxonomic families predicts phenological sensitivity to temperature: Implications for forecasting the effects of climate change on unstudied taxa. *American Journal of Botany*, *100*(7), 1381–1397. https://doi.org/10.3732/ajb.1200455

Mazerolle, M. J., & Villard, M. A. (1999). Patch characteristics and landscape context as predictors of species presence and abundance: A review[1]. *Écoscience*, *6*(1), 117–124. https://doi.org/10.1080/11956860.1999.11952204

Miller, T. K., Heberling, J. M., Kuebbing, S. E., & Primack, R. B. (2023). Warmer temperatures are linked to widespread phenological mismatch among native and non-native forest plants. *Journal of Ecology*, *111*(2), 356–371. https://doi.org/10.1111/1365-2745.14021

Minckley, R. L., Wcislo, W. T., Yanega, D., & Buchmann, S. L. (1994). Behavior and Phenology of a Specialist Bee (Dieunomia) and Sunflower (Helianthus) Pollen Availability. *Ecology*, *75*(5), 1406–1419. https://doi.org/10.2307/1937464

Nagahama, A., Kubota, Y., & Satake, A. (2018). Climate warming shortens flowering duration: A comprehensive assessment of plant phenological responses based on gene expression analyses and mathematical modeling. *Ecological Research*, *33*(5), 1059–1068. https://doi.org/10.1007/s11284-018-1625-x

Nilsson, S. G., & Wastljung, U. (1987). Seed predation and cross-pollination in mast-seeding beech (Fagus sylvatica) patches. *Ecology*, *68*(2), 260-265. https://doi.org/10.2307/1939256

Panchen, Z. A., & Gorelick, R. (2017). Prediction of Arctic plant phenological sensitivity to climate change from historical records. *Ecology and Evolution*, *7*(5), 1325–1338. https://doi.org/10.1002/ece3.2702

Peñuelas, J., Filella, I., Zhang, X., Llorens, L., Ogaya, R., Lloret, F., Comas, P., Estiarte, M., & Terradas, J. (2004). Complex spatiotemporal phenological shifts as a response to rainfall changes. *New Phytologist*, *161*(3), 837–846. https://doi.org/10.1111/j.1469-8137.2004.01003.x

Park, D. S., Breckheimer, I., Williams, A. C., Law, E., Ellison, A. M., & Davis, C. C. (2019). Herbarium specimens reveal substantial and unexpected variation in phenological sensitivity across the eastern United States. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1763), 20170394. https://doi.org/10.1098/rstb.2017.0394

Park, Isaac et al. (2023). Herbarium-Derived Phenological Data in North America [Dataset]. *Dryad*. https://doi.org/10.25349/D9WP6S

Park, I. W., Ramirez-Parada, T., Record, S., Davis, C., Ellison, A. M., & Mazer, S. J. (2024). Herbarium data accurately predict the timing and duration of population-level flowering displays. *Ecography*, e06961. https://doi.org/10.1111/ecog.06961

Pope, N. S., & Jha, S. (2018). Seasonal Food Scarcity Prompts Long-Distance Foraging by a Wild Social Bee. *The American Naturalist*, *191*(1), 45–57. https://doi.org/10.1086/694843

Potts, S. G., Vulliamy, B., Dafni, A., Ne'eman, G., & Willmer, P. (2003). Linking bees and flowers: How do floral communities structure pollinator communities? *Ecology*, *84*(10), 2628–2642. https://doi.org/10.1890/02-0136

Prevéy, J., Vellend, M., Rüger, N., Hollister, R. D., Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., Clark, K., Cooper, E. J., Elberling, B., Fosaa, A. M., Henry, G. H. R., Høye, T. T., Jónsdóttir, I. S., Klanderud, K., Lévesque, E., Mauritz, M., Molau, U., Natali, S. M., … Rixen, C. (2017). Greater temperature sensitivity of plant phenology

at colder sites: Implications for convergence across northern latitudes. *Global Change Biology*, *23*(7), 2660–2671. https://doi.org/10.1111/gcb.13619

Prevéy, J. S., Rixen, C., Rüger, N., Høye, T. T., Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., Ashton, I. W., Cannone, N., Chisholm, C. L., Clark, K., Cooper, E. J., Elberling, B., Fosaa, A. M., Henry, G. H. R., Hollister, R. D., Jónsdóttir, I. S., Klanderud, K., Kopp, C. W., … Wipf, S. (2019). Warming shortens flowering seasons of tundra plant communities. *Nature Ecology & Evolution*, *3*(1), 45–52. https://doi.org/10.1038/s41559-018-0745-6

Ramirez-Parada, T. H., Park, I. W., & Mazer, S. J. (2022). Herbarium specimens provide reliable estimates of phenological responses to climate at unparalleled taxonomic and spatiotemporal scales. *Ecography*, *2022*(10), e06173. https://doi.org/10.1111/ecog.06173

Ramirez-Parada, T. H., Park, I. W., Record, S., Davis, C. C., Ellison, A. M., & Mazer, S. J. (2024). Plasticity and not adaptation is the primary source of temperature-mediated variation in flowering phenology in North America. *Nature Ecology & Evolution*, *8*(3), 467–476. https://doi.org/10.1038/s41559-023-02304-5

Renner, S. S., & Zohner, C. M. (2018). Climate change and phenological mismatch in trophic interactions among plants, insects, and vertebrates. *Annual review of ecology, evolution, and systematics*, *49*(1), 165-182. https://doi.org/10.1146/annurev-ecolsys-110617-062535

Roslin, T., Antão, L., Hällfors, M., Meyke, E., Lo, C., Tikhonov, G., Delgado, M. del M., Gurarie, E., Abadonova, M., Abduraimov, O., Adrianova, O., Akimova, T., Akkiev, M., Ananin, A., Andreeva, E., Andriychuk, N., Antipin, M., Arzamascev, K., Babina, S., … Ovaskainen, O. (2021). Phenological shifts of abiotic events, producers and consumers across a continent. *Nature Climate Change*, *11*(3), 241–248. https://doi.org/10.1038/s41558-020-00967-7

Roulston, T. A. H., & Goodell, K. (2011). The role of resources and risks in regulating wild bee populations. *Annual review of entomology*, *56*(1), 293-312. https://doi.org/10.1146/annurev-ento-120709-144802

Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'neill, B. C., Fujimori, S., ... & Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global environmental change*, *42*, 153-168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Schenk, M., Krauss, J., & Holzschuh, A. (2018). Desynchronizations in bee–plant interactions cause severe fitness losses in solitary bees. *Journal of Animal Ecology*, *87*(1), 139–149. https://doi.org/10.1111/1365-2656.12694

Scheper, J., Bommarco, R., Holzschuh, A., Potts, S. G., Riedinger, V., Roberts, S. P. M., Rundlöf, M., Smith, H. G., Steffan-Dewenter, I., Wickens, J. B., Wickens, V. J., & Kleijn, D. (2015). Local and landscape-level floral resources explain effects of wildflower strips on wild bees across four European countries. *Journal of Applied Ecology*, *52*(5), 1165–1175. https://doi.org/10.1111/1365-2664.12479

Sohl, T. L., Sayler, K. L., Bouchard, M. A., Reker, R. R., Friesz, A. M., Bennett, S. L., Sleeter, B. M., Sleeter, R. R., Wilson, T., Soulard, C., Knuppe, M., & Van Hofwegen, T. (2014). Spatially explicit modeling of 1992–2100 land cover and forest stand age for the conterminous United States. *Ecological Applications*, *24*(5), 1015–1036. https://doi.org/10.1890/13-1245.1

Sohl, T., Reker, R., Bouchard, M., Sayler, K., Dornbierer, J., Wika, S., Quenzer, R., & Friesz, A. (2016). Modeled historical land use and land cover for the conterminous United States. *Journal of Land Use Science*, *11*(4), 476–499. https://doi.org/10.1080/1747423X.2016.1147619

Theobald, E. J., Breckheimer, I., & HilleRisLambers, J. (2017). Climate drives phenological reassembly of a mountain wildflower meadow community. *Ecology*, *98*(11), 2799–2812. https://doi.org/10.1002/ecy.1996

Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2022). Predictive performance of presence-only species distribution models: A benchmark study with reproducible code. *Ecological Monographs*, *92*(1), e01486. https://doi.org/10.1002/ecm.1486

Wang, H., Ge, Q., Rutishauser, T., Dai, Y., & Dai, J. (2015). Parameterization of temperature sensitivity of spring phenology and its application in explaining diverse phenological responses to temperature change. *Scientific Reports*, *5*(1), 8833. https://doi.org/10.1038/srep08833

Wang, J., Liu, D., Ciais, P., & Peñuelas, J. (2022). Decreasing rainfall frequency contributes to earlier leaf onset in northern ecosystems. *Nature Climate Change*, *12*(4), 386–392. https://doi.org/10.1038/s41558-022-01285-w

Williams, N. M., & Kremen, C. (2007). Resource distributions among habitats determine solitary bee offspring production in a mosaic landscape. *Ecological Applications*, *17*(3), 910–921. https://doi.org/10.1890/06-0269

Willis, C. G., Ellwood, E. R., Primack, R. B., Davis, C. C., Pearson, K. D., Gallinat, A. S., Yost, J. M., Nelson, G., Mazer, S. J., Rossington, N. L., Sparks, T. H., & Soltis, P. S. (2017). Old Plants, New Tricks: Phenological Research Using Herbarium Specimens. *Trends in Ecology & Evolution*, *32*(7), 531–546. https://doi.org/10.1016/j.tree.2017.03.015

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. (2012). Warming experiments underpredict plant phenological responses to climate change. *Nature*, *485*(7399), 494–497. https://doi.org/10.1038/nature11014

Woodcock, B. A., Garratt, M. P. D., Powney, G. D., Shaw, R. F., Osborne, J. L., Soroka, J., Lindström, S. a. M., Stanley, D., Ouvrard, P., Edwards, M. E., Jauker, F., McCracken, M. E., Zou, Y., Potts, S. G., Rundlöf, M., Noriega, J. A., Greenop, A., Smith, H. G., Bommarco, R., … Pywell, R. F. (2019). Meta-analysis reveals that pollinator functional diversity and abundance enhance crop pollination and yield. *Nature Communications*, *10*(1), 1481. https://doi.org/10.1038/s41467-019-09393-6

Zhang, H., Yuan, W., Liu, S., Dong, W., & Fu, Y. (2015). Sensitivity of flowering phenology to changing temperature in China. *Journal of Geophysical Research: Biogeosciences*, *120*(8), 1658–1665. https://doi.org/10.1002/2015JG003112

Zhou, Z., Zhang, K., Sun, Z., Liu, Y., Zhang, Y., Lei, L., Li, Y., Wang, D., Hu, M., Wang, S., Lu, Q., Cui, Y., Zhong, M., Han, S., & Miao, Y. (2022). Lengthened flowering season under climate warming: Evidence from manipulative experiments. *Agricultural and Forest Meteorology*, *312*, 108713. https://doi.org/10.1016/j.agrformet.2021.108713

Zohner, C. M., Mirzagholi, L., Renner, S. S., Mo, L., Rebindaine, D., Bucher, R., Palouš, D., Vitasse, Y., Fu, Y. H., Stocker, B. D., & Crowther, T. W. (2023). Effect of climate warming on the timing of autumn leaf senescence reverses after the summer solstice. *Science*, *381*(6653), eadf5098. https://doi.org/10.1126/science.adf5098

VI. Appendix 1—Supplementary materials for **II**

Note S1—Herbarium data sources

Data used in this study was contributed by the Yale Peabody Museum of Natural History, the George Safford Torrey Herbarium at the University of Connecticut, the Acadia University Herbarium, the Chrysler Herbarium at Rutgers University, the University of Montreal Herbarium, the Harvard University Herbarium, the Albion Hodgdon Herbarium at the University of New Hampshire, the Academy of Natural Sciences of Drexel University, the Jepson Herbarium at the University of California-Berkeley, the University of California-Berkeley Sagehen Creek Field Station Herbarium, the California Polytechnic State University Herbarium, the University of Santa Cruz Herbarium, the Black Hills State University Herbarium, the Luther College Herbarium, the Minot State University Herbarium, the Tarleton State University Herbarium, the South Dakota State University Herbarium, the Pittsburg State University Herbarium, the Montana State University-Billings Herbarium, the Sul Ross University Herbarium, the Fort Hays State University Herbarium, the Utah State University Herbarium, the Brigham Young University Herbarium, the Eastern Nevada Landscape Coalition Herbarium, the University of Nevada Herbarium, the Natural History Museum of Utah, the Western Illinois University Herbarium, the Eastern Illinois University Herbarium, the Northern Illinois University Herbarium, the Morton Arboretum Herbarium, the Chicago Botanic Garden Herbarium, the Field Museum of Natural History, the University of Wisconsin-Madison Herbarium, the University of Michigan Herbarium, the Indiana University Herbarium, the Universidad de Sonora Herbarium, the Centro de Investigaciones Biológicas del Noroeste, S. C., the Instituto Politécnico Nacional, CIIDIR Unidad Durango, the University of California-Riverside Herbarium, the San Diego State University Herbarium, the Granite Mountains Desert Research Center, the University of South Carolina Herbarium, the Auburn University Museum of Natural History, the Clemson University Herbarium, the Eastern Kentucky University Herbarium, the College of William and Mary Herbarium, the Appalachian State University Herbarium, the University of North Carolina Herbarium, the University of Memphis Herbarium, the Mississippi State University Herbarium, the University of Mississippi Herbarium, the University of Southern Mississippi Herbarium, the Mississippi Museum of Natural Science, the Marshall University Herbarium, the Longwood University Herbarium, the Herbarium of Western Carolina University, the Northern Kentucky University Herbarium, the Salem College Herbarium, the Troy University Herbarium, the Arizona State University Herbarium, the University of Arizona Herbarium, the Desert Botanical Garden, the Deaver Herbarium, the Navajo Nation Department of Fish and Wildlife, the Grand Canyon National Park Herbarium, the University of New Mexico Herbarium, the Western New Mexico University Herbarium, the Museum of Northern Arizona, the Gil National Forest Herbarium, the Arizona Western College Herbarium, and the Natural History Institute.

Note S2—Stan code for varying-intercept and varying-slopes Bayesian Model

```
stan_mod1 <- "data{
  int<lower=1> N;
  int<lower=1> N_ID;
  real AzimuthDOY[N];
  int type[N];
  real Tmin_AnnDev_max[N];
  real Tmin_Normal_max[N];
  int ID[N];
}
parameters{
  matrix[6,N_ID] z_N_ID;
  cholesky_factor_corr[6] L_rho;
  vector<lower=0>[6] sigma_sp_i;
  vector[N_ID] sig_sp;
}
transformed parameters{
  vector<lower=0>[N] sigma = exp(sig_sp[ID]);
  matrix[N_ID,6] v_N_ID;
  vector[N_ID] a;
  vector[N_ID] a2;
  vector[N_ID] b1;
  vector[N_ID] b2;
  vector[N_ID] b3;
  vector[N_ID] b4;
  matrix[6,6] rho;
  v_N_ID = (diag_pre_multiply(sigma_sp_i,L_rho)*z_N_ID)';
    a = col(v_N_ID,1);
    a2 = col(v_N_ID,2);
    b1 = col(v_N_ID,3);
    b2 = col(v_N_ID,4);
    b3 = col(v_N_ID,5);
    b4 = col(v_N_ID,6);
    rho = L_rho * L_rho';

}
model{
  vector[N] mu;
  L_rho ~ lkj_corr_cholesky( 2 );
  sigma_sp_i ~ exponential( 1 );
  to_vector(z_N_ID) ~ normal( 0 , 1 );
  sig_sp[ID] ~ normal(0,5);

  for ( i in 1:N ) {
   AzimuthDOY[i] ~ normal( a[ID[i]] * (type[i] == 0) +
                           a2[ID[i]] * (type[i] == 1) +
                           b1[ID[i]] * Tmin_Normal_max[i] *(type[i] == 0) +
                           b2[ID[i]] * Tmin_AnnDev_max[i] * (type[i] == 0) +
                           b3[ID[i]] * Tmin_Normal_max[i] * (type[i] == 1) +
                           b4[ID[i]] * Tmin_AnnDev_max[i] * (type[i] == 1),
                           sigma[i]); }}
```

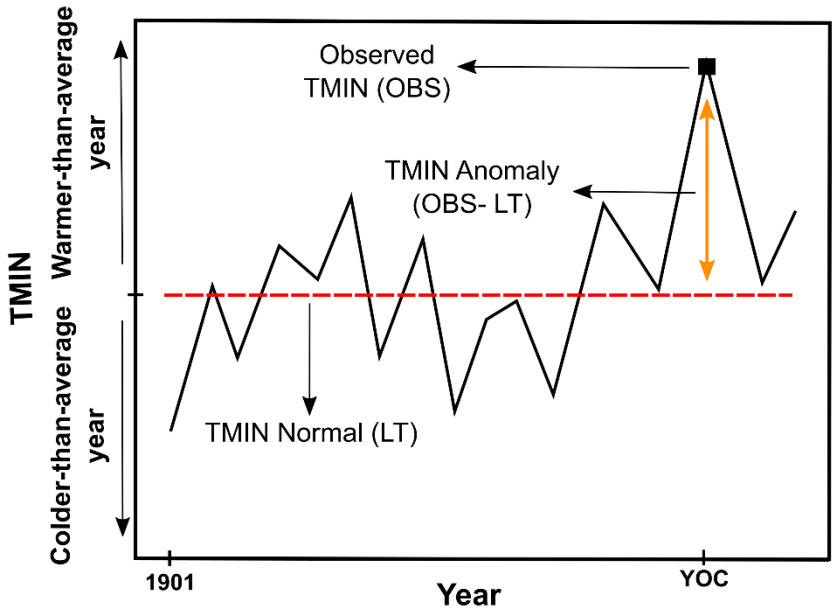

Figure S1—Hypothetical time series (solid black line), long-term mean (i.e., normal; LT; dashed red line), and anomaly in the year of collection or observation (YOC; orange arrows) of mean minimum temperature (TMIN) observed at a site of phenological monitoring or herbarium specimen collection. For a given YOC (OBS; solid black square), the observed climate condition and climate anomaly (orange double-headed arrow) are displayed. Long-term, mean climate conditions at the site are the average observed climate conditions between 1901 and 2016 for the season of interest (3 months leading to and including mean flowering dates for each species). Values higher than the long-term mean line represent years in which TMIN was higher-than-average for that location, whereas those below the long-term mean represent years in which TMIN was lower than average. TMIN anomalies represent the difference between observed climate conditions in a year of collection and long-term, mean conditions (i.e., normals) for that site.

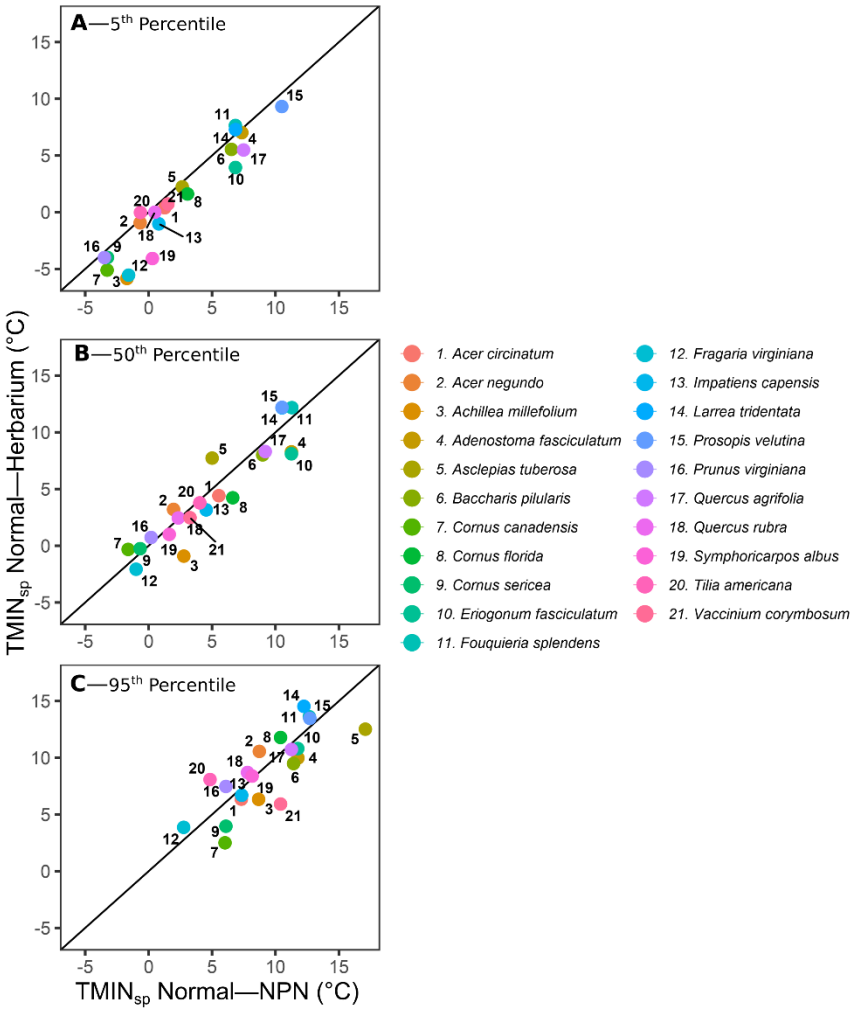

Figure S2 – Correspondence between quantiles of the TMIN normal envelope spanned by herbarium collections and field observations in the NPN for 21 species in the continental United States.

Species	Standardized DOY				Sens. to TMIN anomaly (d/°C)				Sens. to TMIN normal (d/°C)			
	NPN		Herb.		NPN		Herb.		NPN		Herb.	
	2.5%	*97.5%*	*2.5%*	*97.5%*	*2.5%*	*97.5%*	*2.5%*	*97.5%*	*2.5%*	*97.5%*	*2.5%*	*97.5%*
Acer circinatum	107	134	138	168	-6.8	3.1	-6.1	1.1	-10.5	3.8	-7.6	2.7
Acer negundo	92	108	118	133	-5.3	-2.1	-3.4	-0.5	-8.6	-1.3	-8.1	-1.7
Achillea millefolium	179	206	191	194	-3.8	1.2	-4.5	-3.6	-13.3	0.9	-5.0	-2.1
Adenostoma fasciculatum	150	181	147	162	-11.8	-3.8	-9.3	-0.2	-11.7	2.8	-9.2	-0.3
Asclepias tuberosa	165	208	208	238	-1.5	2.5	-5.4	-2.5	-9.0	5.2	-4.7	3.6
Baccharis pilularis	221	253	237	282	2.5	7.8	0.1	8.6	-1.7	8.1	-2.9	7.4
Cornus canadensis	152	164	154	166	-6.2	-2.4	-5.9	-3.2	-7.2	1.4	-6.7	-0.8
Cornus florida	84	96	88	105	-5.3	-3.1	-5.3	-3.3	-5.3	-1.4	-6.8	-0.6
Cornus sericea	153	165	160	170	-5.0	-0.7	-4.2	-2.0	-9.9	0.8	-4.4	0.6
Eriogonum fasciculatum	157	195	166	182	-8.1	-1.0	-9.7	-5.1	-6.6	8.0	-5.1	1.6
Fouquieria splendens	98	124	102	120	-8.2	0.8	-9.4	-0.9	-16.7	-1.2	-8.5	0.6
Fragaria virginiana	136	145	141	149	-5.9	-1.6	-5.3	-3.9	-5.1	3.2	-6.2	-1.8
Impatiens capensis	189	239	190	228	-1.4	4.3	0.1	4.7	-2.5	8.9	0.5	8.2
Larrea tridentata	112	130	109	120	-13.6	-6.8	-12.9	-8.1	-7.4	3.2	-5.1	1.7
Prosopis velutina	146	187	132	175	-15.8	-3.0	-12.5	-2.8	-12.4	0.7	-8.2	-0.1
Prunus virginiana	132	145	147	154	-7.9	-3.6	-3.3	-1.7	-10.3	-0.4	-6.5	-1.9
Quercus agrifolia	88	107	87	105	-5.1	3.3	-7.3	2.0	-11.1	-0.4	-7.7	1.8
Quercus rubra	113	125	173	193	-4.2	-1.5	-4.3	-0.2	-5.8	0.4	-4.4	3.8
Symphoricarpos albus	182	198	189	197	-4.1	0.3	-3.0	-0.4	-7.3	2.5	-5.6	0.9
Tilia americana	157	218	185	222	-6.4	2.9	-4.4	0.8	-7.2	6.3	-4.0	4.9
Vaccinium corymbosum	104	112	122	141	-6.9	-4.7	-4.0	0.3	-7.7	-1.0	-7.1	-2.4

Table S1 – 95% credible intervals (CRIs) for estimates for standardized flowering dates, sensitivity to TMIN normal, and sensitivity to TMIN anomaly derived from USA-NPN and herbarium data for 21 species in the continental United States.

VII. Appendix 2—Supplementary materials for III

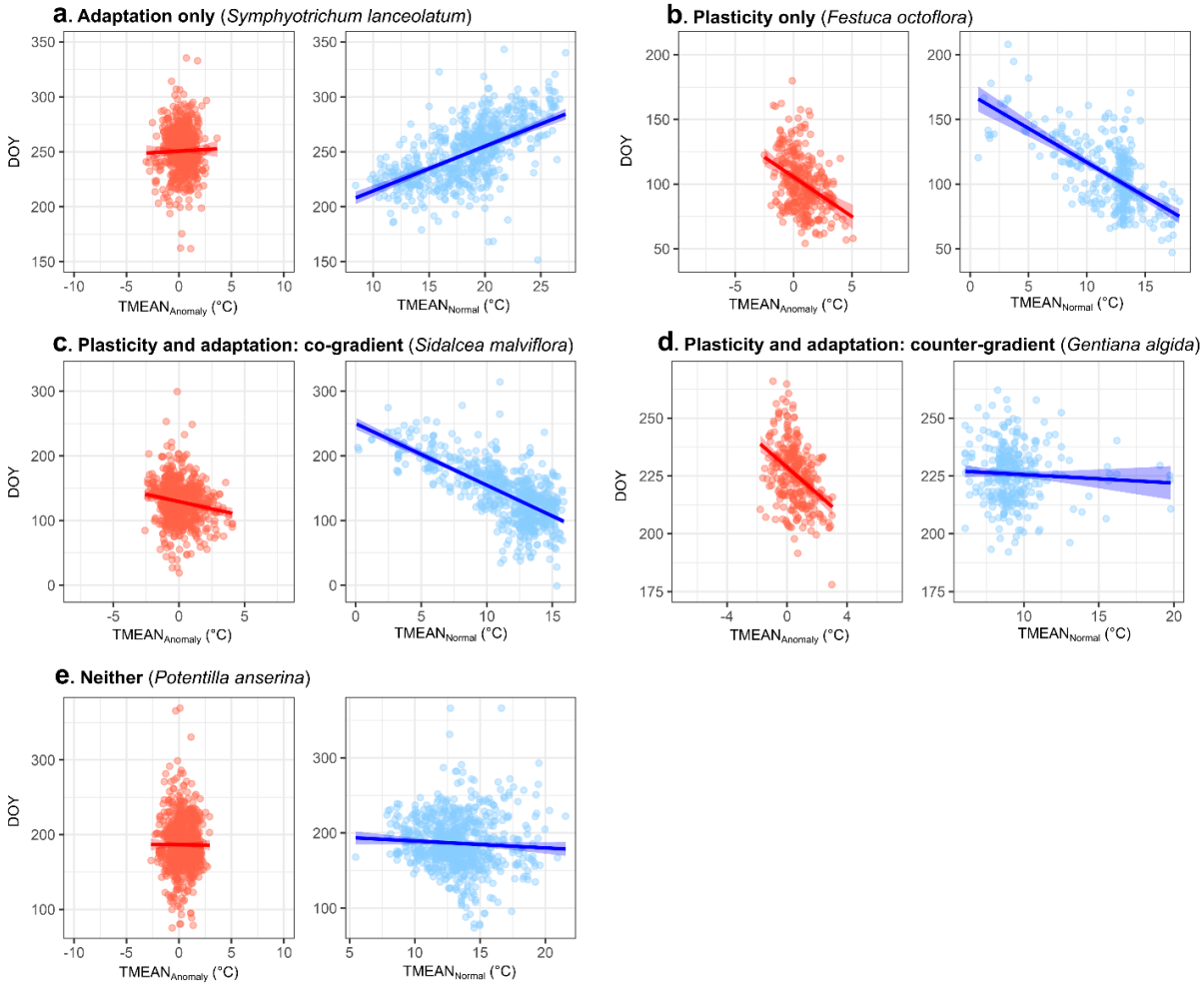

Figure S1—Phenological sensitivities to interannual and spatial temperature variation for five species exemplifying various relative contributions of plasticity and adaptation. (a) Sensitivity patterns of the panicle aster (*Symphyotrichum lanceolatum,* Asteraceae), a widespread perennial herb primarily occupying moist habitats across the United States. Its lack of responsiveness to $TMEAN_{Anomaly}$ (i.e., non-significant S_{time}) and significantly greater responsiveness to $TMEAN_{Normal}$ (i.e., significant $S_{space} - S_{time}$) are consistent with a scenario in which adaptation is the sole driver of flowering time variation along the temperature gradient (see Fig. 1a,b in the main text). (b) Sensitivity patterns of the sixweeks fescue (*Festuca octoflora,* Poaceae), a widespread annual grass most abundant in arid prairies and deserts. *F. octoflora* exhibits significant S_{time} that does not substantially differ from S_{space} (i.e., non-significant $S_{space} - S_{time}$), a pattern consistent with plasticity as the sole cause of variation in flowering time along the temperature gradient (Fig. 1c,d in the main text). (c) Sensitivity patterns of the checker bloom (*Sidalcea malviflora,* Malvaceae), a perennial herb widespread in the West Coast of the United States. *S. malviflora* exhibits significant responsiveness to

interannual temperature variation (i.e., significant S_{time}) and significantly greater sensitivity to spatial variation (i.e., S_{space} has greater magnitude and the same direction as S_{time}; $S_{space} - S_{time}$ is significant), a scenario consistent with the joint effects of plasticity and adaptation in a co-gradient pattern (Fig. 1e,f in the main text). (d) Sensitivity patterns of the whitish gentian (*Gentiana algida*¸Gentianaceae), a herbaceous perennial occupying mid-to-high elevations in the southern, mid, and northern Rocky Mountains. Its significant responsiveness to interannual TMEAN variation (i.e., significant S_{time}) and significantly lesser responsiveness to geographic temperature variation (i.e., S_{space} has lesser magnitude or opposite direction than S_{time}; $S_{space} - S_{time}$ is significant) is consistent with the joint effects of plasticity and local adaptation along the temperature gradient in a counter-gradient pattern (Fig. 1g,h in the main text). (e) Sensitivity patterns of silverweed (*Potentilla anserina*, Rosaceae), a widespread perennial herb occurring primarily in moist soils near bodies of water. Its apparent lack of sensitivity to interannual and geographic variation in temperature (and its non-significant $S_{space} - S_{time}$) make it inconsistent with the apparent roles of either plasticity or adaptation. Solid lines in each panel correspond to the regression lines of DOY vs. $TMEAN_{Anomaly}$ or $TMEAN_{Normal}$, with colored ribbons indicating the standard error of the predicted value of the response at each value of the predictor.

Note S1—Hosting herbaria and databases used to access data

Herbaria hosting the specimen data used in this study (Darwin Core Institution Codes):

ACAD, AMES, BALT, BLMRD, BSCA, CalBG, SFV, UCR, SD, SDSU, SBBG, UCJEPS, IRVC, LOB, CDA, OBI, RSA, CSUSB, CSLA, GMDRC, FSC, LA, JOTR, POM, DAV, OBS, HSC, UCSC, UCSB, MACF, PUA, JROH, SCFS, SOC, UNLV, OSC, CHRB, CM, NY, University of Connecticut, Université de Montréal Biodiversity Centre, ANSP, Rutgers University, UConn, University of New Hampshire, NEBC, GH, A, YPM, Yale Peabody Museum of Natural History, Woods Hole Oceanographic Institution, ECON, Harvard University, University of Maine, Memorial University of Newfoundland, VT, ELH, Harvard, HUDC, Utah State University, ENLC, RENO, USU, BRY, SUU, RM, UT, NTS, EDOBLM, NDOA, BLM, SEINet, WSCO, NPS, SLCTNL, MARY, UNISON, HCIB, CIAD, BCMEX, IBUG, UJED, URUZA, UAEH, UADY, DEK, EIU, IND, MOR, MSC, WIS, MU, BUT, CINC, MICH, MWI, NC, F, LUC, CHIC, KE, MIN, ILL, PH, TAWES, PAC, APSC, USFWS, BHSC, USFS/BHSC, LCDI, MISU, CSCN, KSP, MSUB, FHKSC, MSUNH, PPWD, ASU, ARIZ, ASC, NAVA, DES, RHNM, USFS, UNM, SNM, MNA, SJNM, JEMEZ, KAIB, AWC, NHI, BTA, TAF, MUR, USCH, TROY, NCU, USF, MISSA, FLAS, DUKE, SWSL, SEL, WILLI, NCSM, BOON, MMNS, ANHC, AUA, LSU, CLEMS, NCSC, VDB, MISS, UNCC, USMS, TENN, PEMB, MUHW, GA, UARK, NLU, KNK, WVA, STAR, ODU, WWC, HBSH, NO, ENO, SC, ECUH, UAM, NCZP, CATU, HTTU, CAU, WEWO, WCUH, GMUF, URV, UOS, APCR, EKY, UCHT, MEM, SBAC, VPI, FARM, SFSU, dtnm, gtnp, WYAC, FBNM, USNH, YELLO, GRTE, dtnp, DTNM, RMBL, DBG, FLD, WSC, MESA, PUSC, COLO, ALAM, CIBO, SAT, SRSC, BRIT, TAC, PAUH, TTC, TLU, JWC, OKL.

Herbarium specimen data were accessed and downloaded from the following platforms:

SEINET, the Consortium of Pacific Herbaria, the Consortium of California Herbaria, the Consortium of Northeastern Herbaria, the Consortium of Pacific Northwest Herbaria, the consortium of MidAtlantic Herbaria, the Consortium of Canadian Herbaria, the Consortium of Midwest Herbaria, the North American Network of Small Herbaria, The Consortium of Northeastern Herbaria, the Red de Herbarios de Noroeste de México, the SouthEast Regional Network of Expertise and Collections (SERNEC), and the Texas Oklahoma Regional Consortium of Herbaria (TORCH).

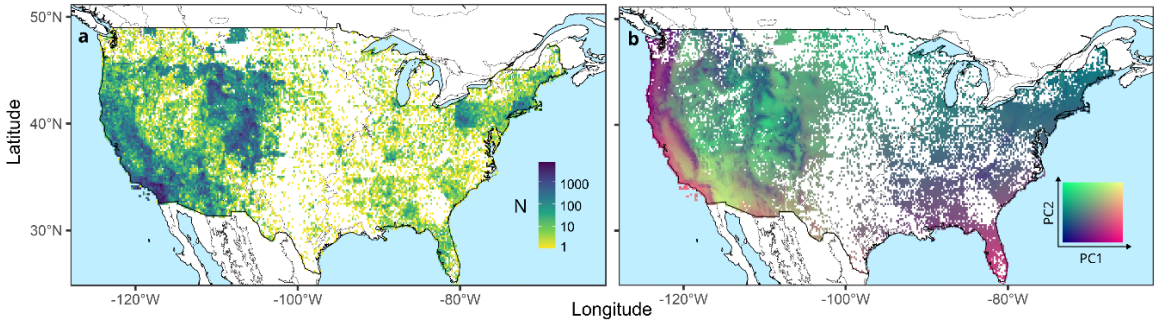

Figure S2—Sampling intensity and long-term climatic conditions across collection sites in the continental United States. Pixels correspond to 20 × 20-km grid cells, with their color representing (a) the total number of specimens collected and (b) their mean PC1 and PC2 values. PC1 represents a gradient of increasing precipitation seasonality, decreasing temperature seasonality and increasing long-term mean annual temperature. In turn, PC2 represents a gradient of decreasing long-term mean annual precipitation and increasing temperature seasonality (see 'Climatic data').

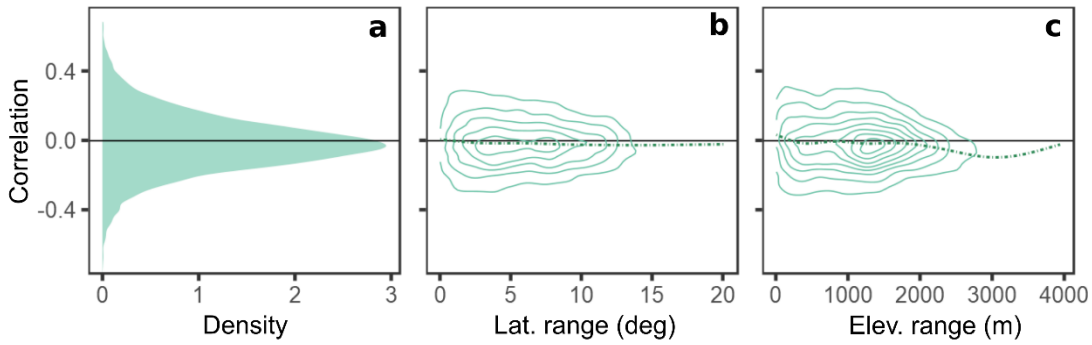

Figure S3—Correlation between $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ among species. (a) Distribution of correlation coefficients among 1,605 angiosperms in the continental United States. (b) Relationship between normal vs. anomaly correlations and the latitudinal range of a species. (c) Relationship between normal vs. anomaly correlations and the elevational range of a species.. Contour lines in (b) and (c) correspond to bivariate density plots for each of the correlations between $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ and the predictor (latitudinal or elevational range).

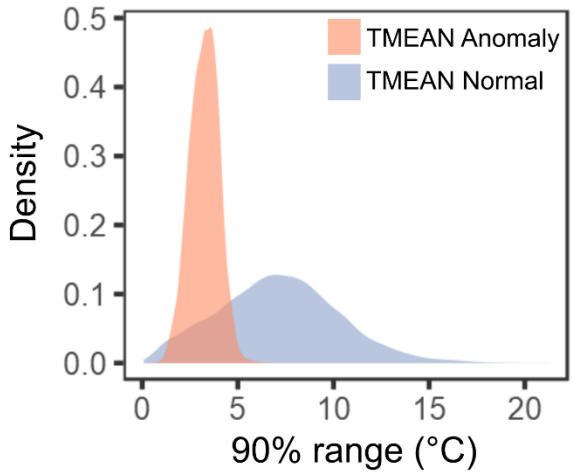

Figure S4—Distribution of temperature ranges among species. Temperature ranges were calculated using the central 90% of specimens of each of 1,605 species across the continental United States for both $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$.

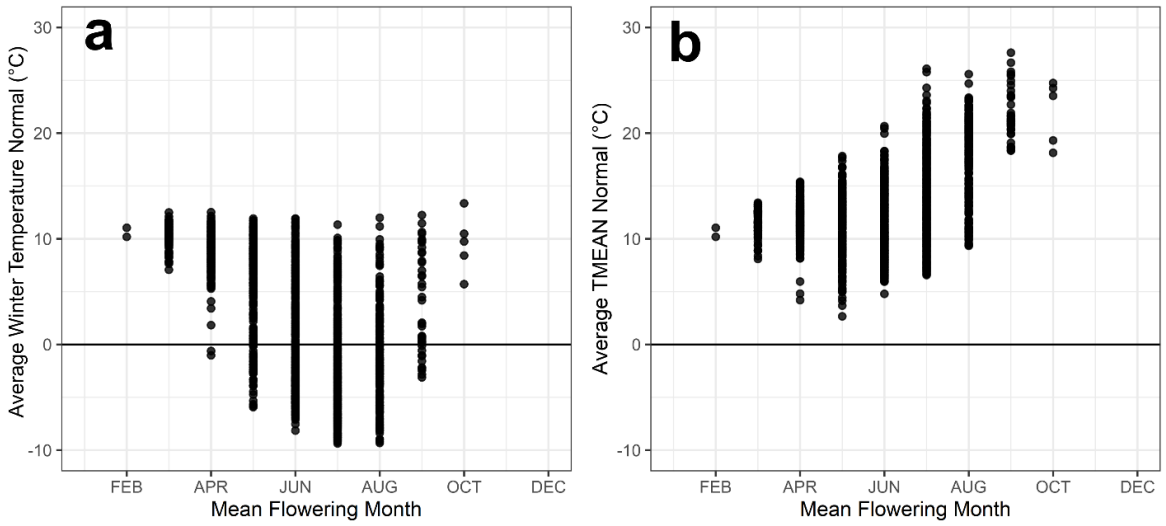

Figure S5—Normal temperatures during winter and the periods preceding flowetring among species. (a) Average winter temperatures and (b) average $TMEAN_{Normal}$ (for the 3-month period leading up to mean flowering dates) experienced by 1,605 angiosperm species with varying mean months of flowering.

Note S2—Model performance on simulated data: estimation accuracy and precision

Our estimate of apparent adaptation is derived from the difference of two random variables (i.e., $S_{space} - S_{time}$). As such, apparent adaptation may show greater estimation uncertainty than apparent plasticity (S_{time}), which could result in lower probability of direction for apparent adaptation estimates compared to those of the same magnitude for apparent plasticity. If so, our criteria for classifying sensitivity patterns as consistent with plasticity or adaptation (see Table 1 of the main text) would be biased towards plasticity. Similarly, any directional biases in estimating S_{space} and S_{time} may be compounded when estimating their difference, making our estimates of apparent adaptation less reliable.

To assess whether our model (see **Methods**: ***Analyses*** in main text; and Supplementary Note 2) could reliably recover estimates of apparent plasticity (S_{time}) and apparent adaptation ($S_{space} - S_{time}$), we simulated data and assessed how estimation accuracy and precision varied depending on 1) the average sample size among species, 2) the magnitude of the true value of $S_{space} - S_{time}$, and 3) whether $S_{space} - S_{time}$ and S_{time} differed in a co- or counter-gradient manner.

Simulation outline

In each iteration of the simulation, we generated batches of 30 species, with 5 species within each batch sharing an assigned true value of $S_{space} - S_{time}$ (i.e., apparent adaptation) varying from 0 to 5 d/°C in increments of 1 d/°C. In each batch, every species shared a sample size of 100, 200, 300, 400, or 500 observations. After repeating the process for 30 iterations and pooling results across them, we obtained a total of 150 simulated species for each combination of true $S_{space} - S_{time}$ and sample size. Each batch of 30 species was eventually analyzed using the model described in the 'Methods: Analyses" section of the main text (Stan code provided in Supplementary Note 2); accordingly, we chose to simulate 30 species per iteration to balance 1) computational constraints, given that multi-species models with thousands of species, as those in the main text, may take days/weeks to run, 2) the need to analyze results in a multi-species context, and 3) the need to cover a sufficient range of sample and effect sizes (i.e., the magnitude of $S_{space} - S_{time}$ in d/°C)

We assigned each simulated specimen of a species long-term TMEAN conditions in their site of collection (i.e., $TMEAN_{Normal}$) and deviations from the long-term average in the year of collection (i.e., $TMEAN_{Anomaly}$) by sampling values from actual TMEAN conditions within species in our data. Specifically, we subset the actual herbarium specimen dataset to include only those species meeting the most stringent sample size requirement (i.e., 776 species with 500+ specimens), and sampled random pairs of $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ values from a single species (calculated during the 3-month period preceding the mean flowering date of the sampled species) equal to the focal sample size. Sampling $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ from real observations ensured that the simulated data were generated using a realistic range of temperature conditions that helped replicate the signal-to-noise ratio found in our data for the relationships between DOY and each TMEAN variable.

Next, for each species, we sampled a "true" S_{time} parameter from a normal distribution with mean = -4 d/°C and SD = 3 d/°C, values consistent with much of the literature estimating phenological sensitivity to temperature across species (e.g., Wolkovich

et al. 2012, Calinger et al. 2013). Next, we generated a "true" S_{space} parameter by randomly adding (or subtracting) the true $S_{space} - S_{time}$ (0 to 5 d/°C) to (or from) S_{time}. Based on S_{space} and S_{time}, we then generated predicted DOYs for each specimen of the focal species using its assigned $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ in the following linear equation:

$$\text{Predicted DOY} = S_{space} \times TMEAN_{Normal} + S_{time} \times TMEAN_{Anomaly}$$

Since we were interested in modelling parameters that measure *variation* in DOY due to temperature (S_{space}, S_{time}, and $S_{space} - S_{time}$), the numerical values of the response are immaterial in these analyses. Therefore, for simplicity, we assigned an intercept of 0 to all simulated species.

To generate realistic DOYs, we added normally distributed noise to each predicted DOY obtained in the previous step. For each species, we did this by randomly drawing observation errors from a normal distribution with mean equal to the predicted DOY of each specimen. In turn, to generate a realistic degree of variance around the predicted DOY, we used the specimen data to first generate an among-species distribution of standard deviations in the residual variation in DOY after fitting a linear regression of DOY vs. $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ for each species represented by a minimum of 100 specimens in our data. Next, we defined the standard deviation of the normal distribution used to generate noise around the predicted DOYs by sampling a standard deviation from the distribution generated in the previous step. As we did when sampling $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$, setting the dispersion parameter based on the actual data helped preserve the signal-to-noise ratio for DOY vs. $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ that we are likely to observe in real herbarium datasets. Once the DOYs for a batch of 30 simulated species were generated, we fitted the model presented in the main text (See **Methods: Analyses** in the main text; Stan code for the model is provided in Supplementary Note 2) using one MCMC sampling chain, 300 iterations for warmup, and 2000 for sampling. As in the analyses in the main text, $S_{space} - S_{time}$ was estimated within the "generated quantities" block in STAN as the difference between S_{space} and S_{time}, fully propagating their estimation uncertainty when generating $S_{space} - S_{time}$. All R-hat values for S_{space}, S_{time}, and $S_{space} - S_{time}$ estimates obtained across iterations were less than 1.01.

After model fitting, we calculated estimation error as the difference between the maximum a posteriori (MAP) estimate of $S_{space} - S_{time}$ for each simulated species and its true $S_{space} - S_{time}$. In turn, for each simulated species, we measured estimation uncertainty in S_{time} and $S_{space} - S_{time}$, respectively, as the standard deviation of their posterior distributions. We then grouped results by sample size and true value of $S_{space} - S_{time}$ (150 species per combination) to assess how sample and effect sizes affected 1) the ability of the models to recover the parameters and 2) the uncertainty of the estimates (i.e., their precision). We then further grouped results based on whether the true S_{time} and $S_{space} - S_{time}$ values for species represented a co- or counter-gradient apparent adaptation pattern (Table 1), eliminating species for which true $S_{space} - S_{time}$ was set to equal 0 (i.e., no apparent adaptation, precluding an assignment of co- or counter-gradient adaptation). As the direction of $S_{space} - S_{time}$ in our model was random, grouping by true $S_{space} - S_{time}$, sample size, and co- vs. counter-gradient patterns resulted in a variable number of simulated species per group (min = 62, median = 75, max = 88).

Simulation Results

Overall, we found that our model produced accurate estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$ and that estimates of apparent plasticity and adaptation showed very similar degrees of precision. The distribution of estimation errors for S_{space} and S_{time} among simulated species was centered on or very close to 0 across all sample size categories, with decreasing variability in accuracy for increasing sample sizes (Supplementary Figure 5). S_{space} estimates tended to show less variability in error than S_{time} estimates, a result consistent with the greater signal-to-noise ratio (relative to DOY) observed for $TMEAN_{Normal}$ vs. $TMEAN_{Anomaly}$ (Supplementary Figure 3).

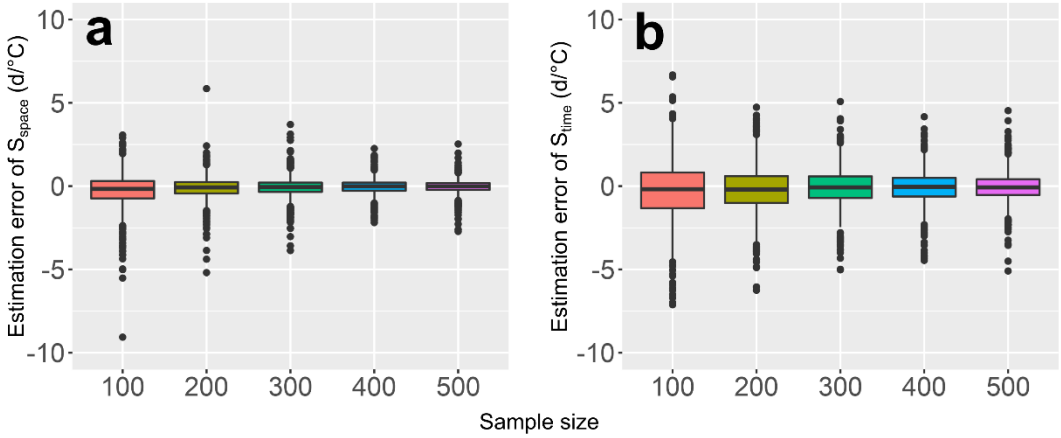

Figure S6—Estimation error of simulated (a) S_{space} and (b) S_{time}. Boxes bound values between the 25th and 75th percentiles, whereas solid lines range from the 5th to the 95th percentiles. The solid, horizontal lines in each box indicate the medians.

Accordingly, we found that the distribution of estimates of $S_{space} - S_{time}$ was 0-centered among species for all combinations of effect size and sample size (Supplementary Figure 6), indicating that the models did not show systemic biases when estimating differences between S_{space} and S_{time}. The magnitude of the estimation error was consistent across true $S_{space} - S_{time}$ values, but the variability in accuracy moderately increased with higher values of true $S_{space} - S_{time}$. The variability in estimation error decreased with sample size for all values of true $S_{space} - S_{time}$.

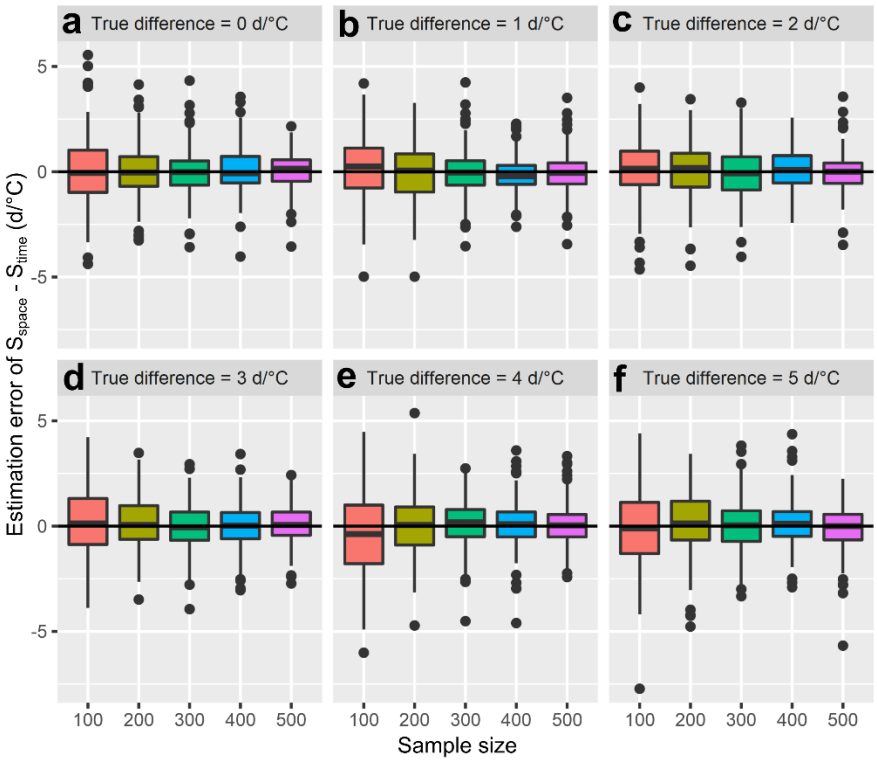

Figure S7—Error among simulated species for $S_{space} - S_{time}$ estimates relative to true values of $S_{space} - S_{time}$ for different sample and effect sizes. Boxes bound values between the 25th and 75th percentiles, whereas solid lines range from the 5th to the 95th percentiles. The solid, horizontal lines in each box indicate the medians.

Moreover, the absolute error of apparent adaptation estimates did not differ substantially among species showing apparent co- or counter-gradient adaptation patterns regardless of sample size or the true value of $S_{space} - S_{time}$ (Supplementary Figure 7). The magnitude of estimation error decreased with sample size but did so only marginally for sample sizes >300 observations.

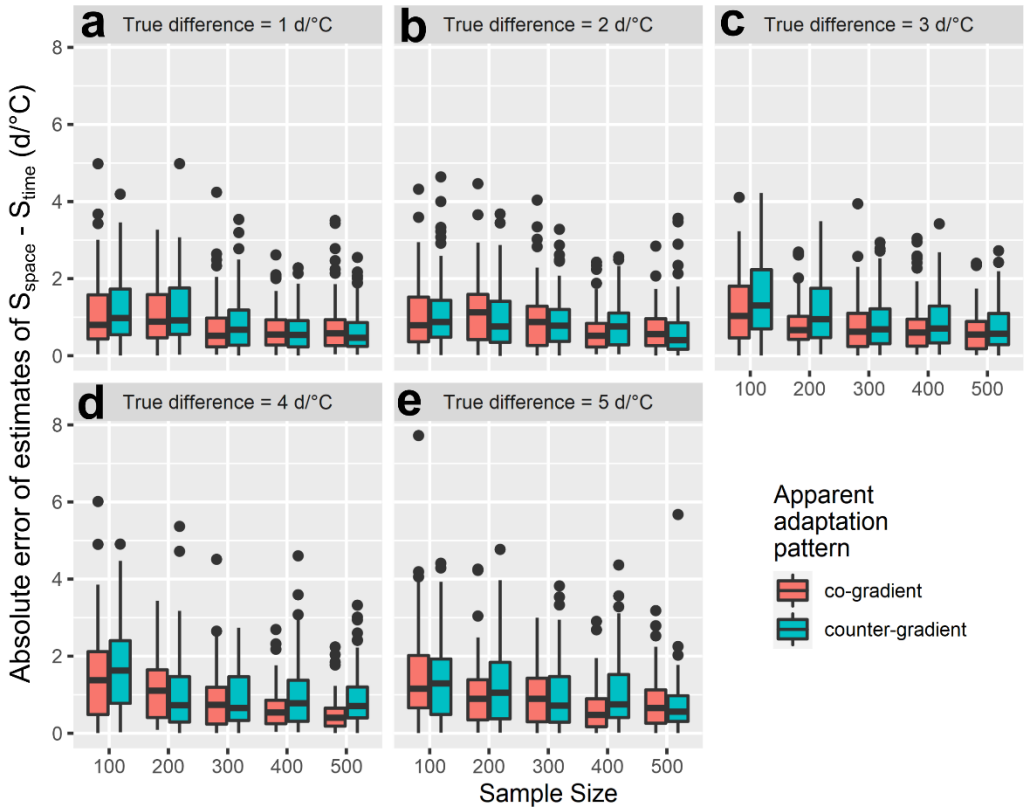

Figure S8—Variation in absolute error of $S_{space} - S_{time}$ estimates for simulated species showing co- or counter-gradient apparent adaptation. Species differed in both sample size and the true magnitude of $S_{space} - S_{time}$. Boxes bound values between the 25th and 75th percentiles, whereas solid lines range from the 5th to the 95th percentiles. The solid, horizontal lines in each box indicate the medians.

Apparent adaptation and plasticity showed similar degrees of estimation uncertainty for all sample size categories and true values of $S_{space} - S_{time}$, with only marginally greater uncertainty for $S_{space} - S_{time}$ across groups (mean difference in posterior SD = 0.035 ± 0.013 d/°C; Supplementary Figure 8). Estimation uncertainty for both parameters decreased substantially with sample size for all values of true $S_{space} - S_{time}$.

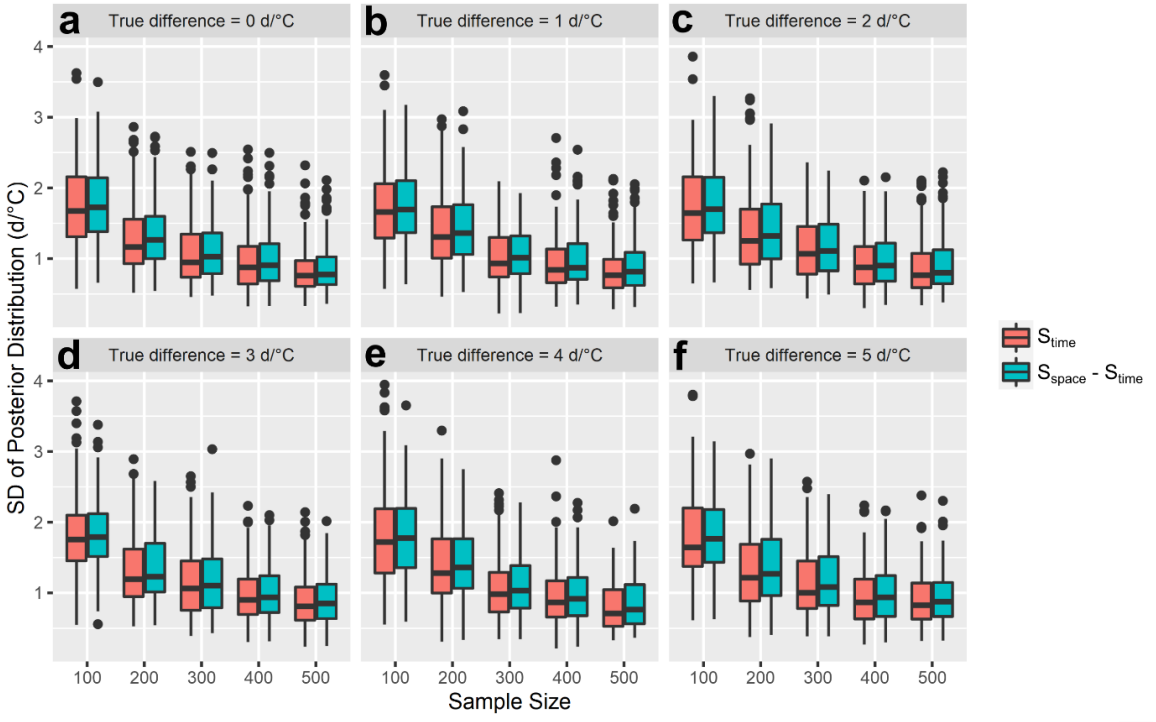

Figure S9—Estimation uncertainty of apparent plasticity (S_{time}) and apparent adaptation ($S_{space} - S_{time}$). Estimates were derived from simulations of species differing in their true value of $S_{space} - S_{time}$ and in the number of observations per species. Boxes bound values between the 25th and 75th percentiles, whereas solid lines range from the 5th to the 95th percentiles. The solid, horizontal lines in each box indicate the medians.

Conclusions

Together, our results demonstrate that the model presented in the main text does not exhibit systemic biases when estimating apparent plasticity and adaptation. Crucially, the estimation uncertainties of S_{time} and $S_{space} - S_{time}$ were similar despite the fact that $S_{space} - S_{time}$ is the difference between two random variables, suggesting that our classification method (Table 1 in the main text) should not result in a substantially larger likelihood of classifying a species' sensitivity pattern as consistent with apparent plasticity (i.e., $P(S_{time} \neq 0) \geq 0.95$) than with apparent adaptation (i.e., $P(S_{space} - S_{time} \neq 0) \geq 0.95$). Finally, while the simulation showed increases in estimation precision with sample size, the magnitude of such increases declined as sample size increased. This observation—and the fact that the analyses presented in the main text included species represented by a median of 491 specimens (min = 300, max = 6,192)—suggests that estimation uncertainty would decrease marginally by using higher sample size cutoffs for inclusion of species in our models, which would come at the cost of substantial taxonomic breadth (e.g., requiring $\geq$ 500 specimens would retain only 776 spp. while requiring $\geq$ 300 specimens retained 1,605 species).

Although these results demonstrate that our model can recover the parameters of interest with sufficient accuracy and precision, our simulations assumed neither phylogenetic nor spatial

structure in DOY, S_{space}, S_{time}, or $S_{space} - S_{time}$ among species. However, the model presented in the main text was fitted under the same set of assumptions, and we confirmed that accounting for phylogeny or spatial autocorrelation did not substantially affect our estimates (Extended Data Fig. 7). Similarly, the model presented in the main text included many more species and larger sample sizes (specimens per species) than these simulations. The larger scale of the model in the main text, however, should improve model performance, as a larger number of groups should improve estimation of hyper-parameters—such as among-species means or standard deviations—thus improving partial pooling of species-specific effects.

Accordingly, we conclude that these simulations—in combination with other Supplementary analyses verifying assumptions such as the lack of substantial effects of phylogeny or spatial autocorrelation on our estimates—demonstrate the statistical reliability of the model used to derive the results presented in the main text.

Note S3—Does our model account for temporal trends in temperature and phenology observed over recent decades? Demonstration through simulations and empirical analysis.

Our model (see **Methods:** ***Analyses*** in main text; and Supplementary Note 2) did not include an explicitly temporal term (i.e., year), and as such, it may appear to omit the potential influence of widespread warming and temporal shifts in phenology observed in recent decades. However, because $TMEAN_{Anomaly}$ represents an interannually variable predictor, values for recent decades do indeed show temporal trends due to recent global warming. Therefore, to the extent that shifts in phenology are caused by TMEAN increases, our model should not only account for temporal trends in temperature, but also in phenology. Moreover, the magnitude of such phenological trends should depend on a species' sensitivity to $TMEAN_{Anomaly}$ (i.e., on S_{time}), as we would expect temperature increases to generate flowering shifts only among species that are responsive to interannual temperature variation.

To demonstrate this, we used simulations—much like those described in Supplementary Note 3—to assess whether temporal trends in temperature generated concordant trends in phenology, and whether the magnitude of those phenological trends was mediated by S_{time}. Moreover, to evaluate the assumption that temporal trends in phenology are generated by increases in $TMEAN_{Anomaly}$ in our data, we used the herbarium dataset to determine whether observed temporal trends in $TMEAN_{Anomaly}$ and a species' S_{time} explain observed trends in DOY.

Simulation outline

To assess whether our model would generate temporal trends in phenology, we simulated data for 1,605 species (as many as in our main analyses) following a similar procedure to the simulation described in our previous response: we generated DOYs by 1) randomly assigning sensitivities to $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ (i.e., S_{space} and S_{time}) to a simulated species, 2) randomly selecting temperature conditions from real observations within species in our dataset, 3) generating predicted DOYs based on S_{time}, S_{space}, and $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$, and 4) adding randomly distributed noise to the predicted DOYs.

In each iteration of our simulation, we generated data for a single species, first assigning it a random S_{time} parameter obtained from a normal distribution with a mean of -4 d/°C and standard deviation of 3 d/°C, and a randomly selected $S_{space} - S_{time}$ parameter of 0, 1, 2, 3, 4, or 5 d/°C. We then generated an S_{space} parameter by randomly adding (or subtracting) $S_{space} - S_{time}$ to (or from) S_{time}. Next, the focal species was randomly assigned an initial sample size of 100, 200, 300, 400, or 500 specimens, and for each sample, we obtained a pair of $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ values obtained from observations of species represented by at least 500 specimens in our data (776 species). As the onset of rapid warming began in the late 20th century, in each iteration, we retained only simulated specimens assigned TMEAN conditions from the year 1980 and onwards. If the resulting sample consisted of less than 30 observations, we skipped to the next iteration, discarding the current simulated species. In turn, if a simulated sample met this criterion, we generated predicted DOYs for each specimen of the focal species using the following linear equation:

$$\text{Predicted DOY} = S_{space} \times TMEAN_{Normal} + S_{time} \times TMEAN_{Anomaly}$$

As in Supplementary Note 3, we were interested in modelling parameters measuring *variation* in DOY due to temperature (S_{space}, S_{time}, and $S_{space} - S_{time}$), so the numerical values of the response are immaterial in these analyses. Therefore, for simplicity, we assigned an intercept of 0 to all simulated species.

This approach generated simulations with sample sizes ranging from 30 to 479 observations (median = 167). To generate simulated DOYs, we then sampled values from a normal distribution with mean equal to the predicted DOY of each specimen. In turn, the standard deviation was drawn from an among-species distribution of standard deviations for the residual variation in DOY obtained after fitting a linear regression of DOY vs. $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ (as described in Supplementary Note 3).

For each dataset simulated this way, we quantified temperature and phenological trends by fitting two simple regressions that included $TMEAN_{Anomaly}$ or DOY as a response and year as a predictor. The coefficients from these regressions (indicating the rate of change of $TMEAN_{Anomaly}$ or DOY per year) were stored in each iteration, as well as the S_{time} value used to simulate the data. Once the target threshold of 1,605 species was reached, we stopped the simulation and assessed whether temporal trends in sampled $TMEAN_{Anomaly}$ and S_{time} explained simulated trends in DOY using the following linear regression:

$$\text{Phenological Trend} = \beta_1 \times TMEAN_{Anomaly}\text{ Trend} + \beta_2 \times TMEAN_{Anomaly}\text{ Trend} \times S_{time}$$

Empirical assessment

We complemented these simulations by assessing whether—among species in our data—temporal trends in $TMEAN_{Anomaly}$ and a species' sensitivity to them (S_{time}) indeed mediated observed trends in phenology between 1980 and 2020. First, we subset the dataset to include only specimens collected from the year 1980 onwards. While all 1,605 species were retained after this step, subsetting resulted in sample sizes ranging from 34 to 4600 observations (mean = 410). As in the simulation, we calculated temporal trends in phenology and temperature by fitting two linear regressions, including $TMEAN_{Anomaly}$ or DOY as a response and year as the only predictor. Then, we quantified the degree of responsiveness to interannual temperature variation by estimating S_{time} from a regression of DOY vs. $TMEAN_{Normal}$ (as a control) and $TMEAN_{Anomaly}$. Finally, as in the simulated dataset, we assessed whether phenological trends were jointly explained by $TMEAN_{Anomaly}$ trends and S_{time} by fitting the same regression model as for the simulation.

Results

Simulated species showed average trends in phenology of -1.4 days/decade, generated from sampled $TMEAN_{Anomaly}$ conditions showing average trends of 0.24 °C/decade. As predicted, the magnitude of phenological trends was mediated by the magnitude of both a species' sensitivity to $TMEAN_{Anomaly}$ and of the temporal trend in $TMEAN_{Anomaly}$ (Supplementary Figure 9a). Accordingly, species for which S_{time} was close to 0 tended to show no temporal trends in DOY regardless of the magnitude of the temporal trends in TMEAN anomaly; in turn, species responsive to $TMEAN_{Anomaly}$ tended to show phenological trends proportional to the magnitude of the temperature trend. Accordingly, species that were more sensitive tended

to show greater increases in the magnitude of phenological trends for equivalent increases in the magnitude of the temperature trends.

We detected the same patterns in the specimen dataset: phenological trends were mediated by the degree of warming and the sensitivity to temperature of a species (Supplementary Figure 9b). However, the magnitude of phenological trends tended to be moderately greater than that of the trends obtained from the simulated data. Specifically, species in the dataset showed average trends in phenology of -1.9 d/decade for average trends in $TMEAN_{Anomaly}$ of the same magnitude as those in the simulate data (0.24 °C/decade).

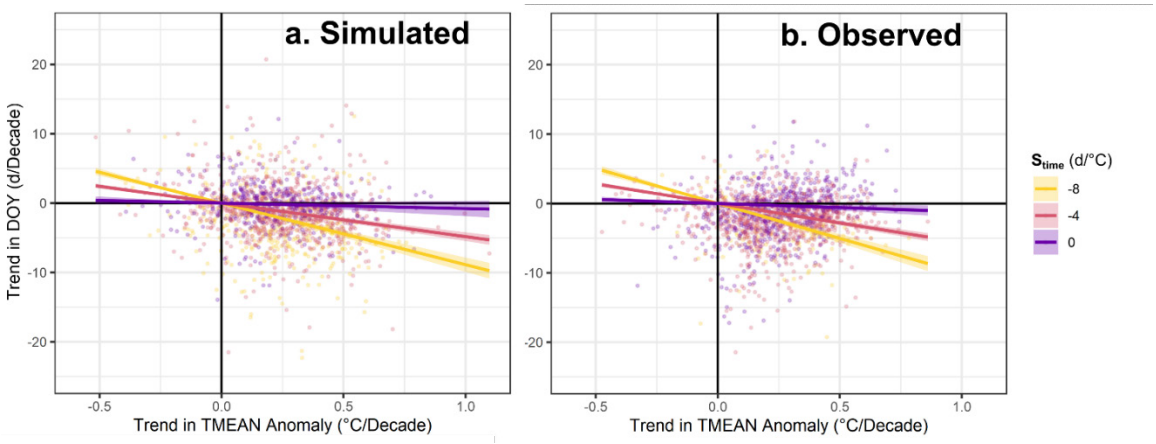

Figure S10—Relationship between temporal trends in flowering phenology (DOY) and the interaction between temporal trends in TMEAN anomaly and species sensitivity to TMEAN anomaly. This relationship was evaluated for (a) simulated data, and (b) 1,605 species represented by 658,778 specimens collected from 1980 onwards (i.e., after the approximate onset of rapid global atmospheric warming). Solid lines correspond to the regression lines obtained from the model described in Supplementary Note 4, with colored ribbons showing the standard error of the mean trend in DOY predicted by each value of the trend in TMEAN anomaly.

The results from our simulation demonstrate that—despite not including an explicitly temporal term—the model presented in the main text does account for trends through time in phenology due to trends in the $TMEAN_{Anomaly}$ data caused by climate warming. However, this model assumes that temporal trends in phenology result exclusively from species responses to increasing $TMEAN_{Anomaly}$. If temporal trends in DOY and $TMEAN_{Anomaly}$ were not causally related in our data, omitting their shared temporal trend in the model (e.g., by excluding time as a covariate) could lead to spurious results (Iler et al. 2017). Reassuringly, we recovered the same patterns using real data: the magnitude of a species' phenological trends was mediated by its sensitivity to $TMEAN_{Anomaly}$ and by the magnitude of temporal trends in $TMEAN_{Anomaly}$. These results suggest that observed phenological trends are indeed causally related to increases in TMEAN over time. Moreover, detrending DOY and $TMEAN_{Anomaly}$ prior to fitting the model did not affect estimates of S_{time} (Extended Data Fig. 3), further indicating that DOY and $TMEAN_{Anomaly}$ are mechanistically correlated, and that omitting temporal trends in the model is unlikely to bias our results.

Conclusions

The analyses outlined above demonstrate that 1) our results are robust to selection of different reference periods over which to calculate $TMEANN_{ormal}$, 2) that the model presented in the main text does account for temporal trends in temperature and phenology, and 3) that its assumptions (i.e., phenological shifts are caused by temperature shifts) are reasonable for our data.

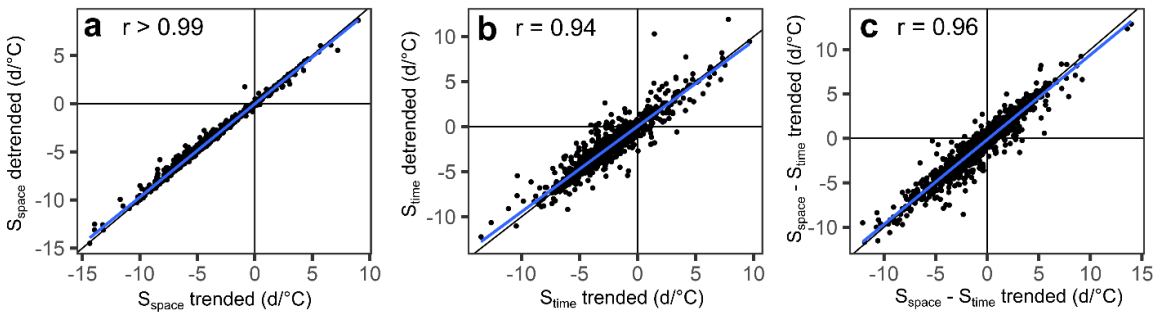

Figure S11—Estimates with and without detrending DOY and $TMEAN_{Anomaly}$. Iler et al. (2017) showed that shared temporal trends between DOY and temperature can generate spurious relationships between these variables that often disappear when the phenological and temperature time series are detrended prior to estimating their relationship. Alternatively, a non-spurious but trended relationship between DOY and temperature might reflect the effects of adaptation to directional changes in temperature, at least in short-lived species. Therefore, relationships between phenology and temperature that persist following detrending are more likely to reflect phenological plasticity. Accordingly, we assessed whether estimates of sensitivity to $TMEAN_{Anomaly}$ (S_{time}) presented in the main text could be confounded by temporal trends in DOY and $TMEAN_{Anomaly}$. To do so, we first ran single-species linear regressions using DOY or $TMEAN_{Anomaly}$ as responses and year as a single predictor, storing the resulting residuals as detrended versions of both responses. Then, for each species, we ran two linear models of DOY against $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$: one with observed DOY and $TMEAN_{Anomaly}$ and another with detrended DOY and $TMEAN_{Anomaly}$. Trended and detrended estimates of sensitivity to $TMEAN_{Anomaly}$ were very highly correlated among species, suggesting that TMEAN sensitivity estimates presented in the main text do not reflect the confounding effect of shared temporal trends. Similarly, detrending DOY and $TMEAN_{Anomaly}$ did not substantially alter estimates of S_{space} and $S_{space} - S_{time}$. a–c, In each panel, points represent the combinations of trended or detrended estimates of S_{space}, S_{time}, or $S_{space} - S_{time}$ for each species in the data, whereas diagonal black lines correspond to 1:1 to relationships denoting perfect agreement between trended and detrended estimates. Solid blue lines in each panel indicate the observed relationship between trended and detrended estimates, with the shaded region around the trend line (nearly imperceptible due to the large sample size) indicating the standard error of the predicted value.

Note S4—Rationale for using $TMEAN_{Normal}$ calculated between 1901 and 2020, and impacts of using $TMEAN_{Normal}$ for different reference periods.

Because there is a long-term trend in temperature due to climate warming across most sites of specimen collection, it is possible that the 120-year window used to calculate $TMEAN_{Normal}$ will result in biases relative to the conditions that individuals or their source populations experienced preceding the collection of a specimen. Specifically, long-term temperatures (i.e., $TMEAN_{Normal}$) used in our models could be much too warm for specimens collected in the early 1900s and much too cool for specimens collected in the 2000s.

We used $TMEAN_{Normal}$ with the objective of characterizing long-term temperature conditions of each site of specimen collection, whose variation therefore represents spatial temperature gradients for each species irrespective of the years in which the specimens were collected. In doing this, the main motivation is *not* to capture the conditions that a plant or its source population experienced prior to a specimen's collection, but to differentiate locations within species' ranges based on their chronic temperature conditions. Here, we aim to describe *geographic temperature gradients* (i.e., *variation* across locations), not the numerical temperature average that most closely matches conditions in the decades prior to the collection of a specimen. If alternative ways of measuring such geographic temperature gradients (i.e., $TMEAN_{Normal}$ estimated over different time periods) do not substantially impact *differences* in long-term conditions observed among sites, then they should reveal the same associations between phenology and geographic temperature variation (i.e., they should result in equivalent estimates of S_{space}).

Nonetheless, to assess possible pitfalls of calculating $TMEAN_{Normal}$ over the entire study period, we evaluated the degree to which our choice of reference period affected estimates of S_{space} and S_{time}. To do so, we measured the relationship between $TMEAN_{Normal}$ for winter (Dec, Jan, Feb), spring (Mar, Apr, May), summer (Jun, Jul, Aug), and fall (Sep, Oct, Nov) calculated between 1901 and 2020 vs. those obtained using four 30-year reference periods (1901-1930, 1931-1960, 1961-1990, or 1991-2020) among all unique collection locations in our data. Before calculating $TMEAN_{Normal}$ for all these periods, we downscaled the spatial resolution of the data to 111m (4 decimal points of a degree) to make the analyses computationally feasible, which yielded a total of 588,650 unique locations. We then calculated winter, spring, summer, and Fall $TMEAN_{Anomaly}$ between 1901 and 2020 for each unique site of collection (approximately 72 million site-by-year combinations) relative to $TMEAN_{Normal}$ calculated using each 30-year reference period, comparing their values to those obtained using $TMEAN_{Normal}$ from 1901 to 2020. Finally, we evaluated how the choice of reference period for calculating normals and anomalies affected estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$. To do this, we fitted single-species models (for each of 1,605 spp. included in the main analyses) of DOY vs. $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ relative to the 1901-1930, 1931-1960, 1961-1990, 1991-2020, or 1901-2020 periods. As in the main analyses, $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ were averaged over the 3-month period preceding the mean flowering date of each species. We then compared estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$ based on $TMEAN_{Normal}$ estimated from 1901-2020 to those obtained using TMEAN variables for each alternative 30-year period.

Results

The values of $TMEAN_{Normal}$ calculated between 1901 and 2020 (hereafter '1901-2020 normals') were nearly identical to those obtained using alternative periods. For every season, 1901-2020 normals showed correlation coefficients greater than 0.99 and 1:1 in relationship to normals for other reference periods (Supplementary Figure 10). This is likely because the magnitude of temperature increases in recent decades—and the magnitude of variation in temperature trends among sites—is negligible compared to the range of geographic variation in long-term conditions typically spanned by a species. For example, locations in our data showed average increases of 0.23 °C/Decade in mean spring temperature between 1980 and 2020, whereas the average species spanned a 99% range of spring $TMEAN_{Normal}$ conditions of 13.1 °C throughout its range.

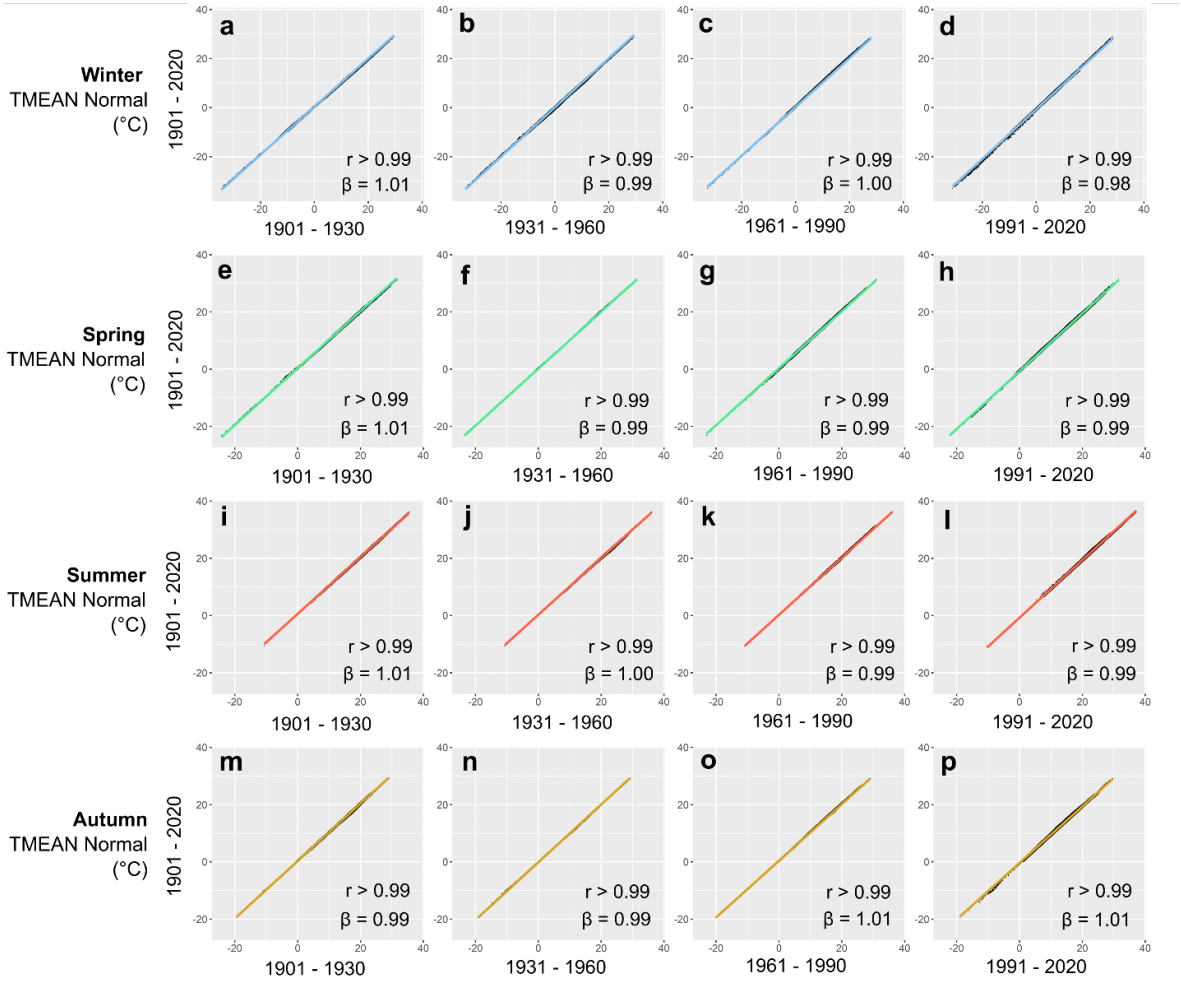

Figure S12—Relationship between winter, spring, summer, and autumn $TMEAN_{Normal}$ calculated between 1901 and 2020 vs. those obtained using different reference periods.

Similarly, seasonal $TMEAN_{Anomaly}$ calculated using 1901-2020 $TMEAN_{Normal}$ (hereafter "1901-2020 anomalies") were highly correlated with those obtained using normals for different reference periods (mean correlation = 0.97; Supplementary Figure 11). The relationship between 1901-2020 anomalies and those calculated relative to normals for other

periods had a mean slope of 0.95, indicating that 1901-2020 anomalies tended to span, on average, a marginally narrower range of variation.

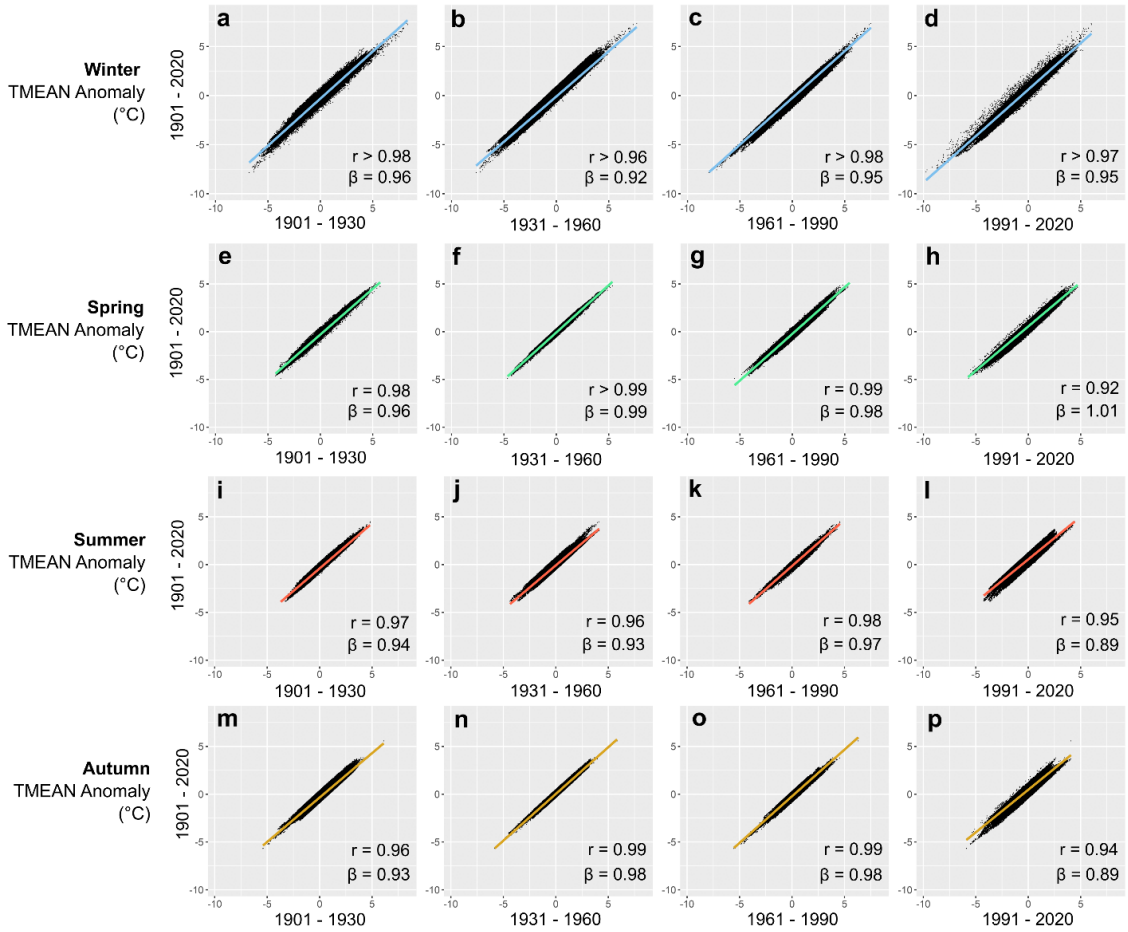

Figure S13—Relationship between $TMEAN_{Anomaly}$ calculated using 1901—2020 $TMEAN_{Normal}$ relative to those obtained using $TMEAN_{Normal}$ for different periods.

Estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$ obtained using normals for 1901-2020 (hereafter '1901-2020 estimates') were highly correlated to those obtained using other reference periods (Supplementary Figure 12). Estimates of S_{space} were least sensitive to period selection, with estimates for all periods showing nearly identical values and 1:1 relationships with those obtained using 1901-2020 normals and anomalies (r > 99; Supplementary Figure 12a-d). Similarly, 1901-2020 estimates of S_{time} and $S_{space} - S_{time}$ were highly correlated to those obtained using different periods ($r \geq 0.89$, mean r = 0.93; Supplementary Figures 12d-l). The relationships between 1901-2020 S_{time} or $S_{space} - S_{time}$ and those obtained using alternative reference periods tended to deviate slightly from a 1:1 line, with estimates of slightly lesser magnitude than those obtained using different reference periods (i.e., slopes < 1 reflecting estimates closer to 0 in both the positive and the negative range). However, the magnitude of the deviation from the 1:1 relationship was nearly identical for S_{time} and $S_{space} - S_{time}$ (mean S_{time} β= 0.84; mean $S_{space} - S_{time}$ β= 0.84), indicating that selecting alternative

periods for calculating $TMEAN_{Normal}$ would have very similar effects on estimates of apparent plasticity and apparent adaptation of the same magnitude.

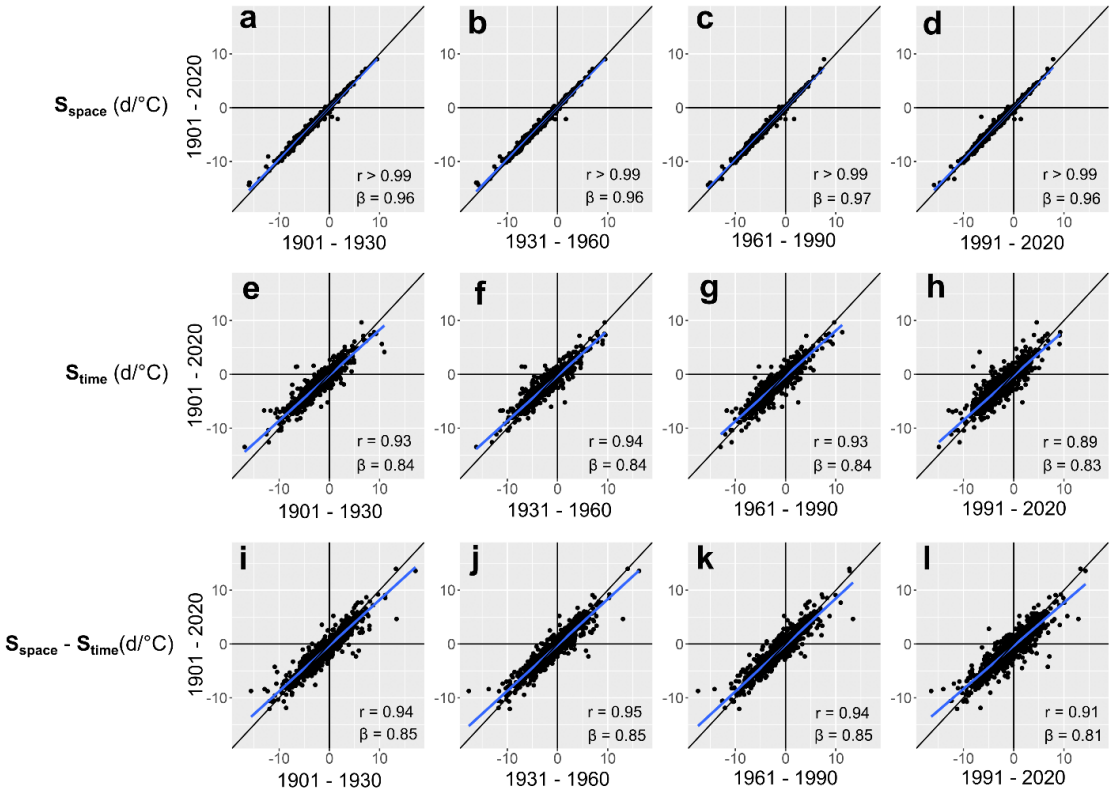

Figure S14—Relationship between estimates of S_{space}, S_{time}, and S_{space} - S_{time} obtained using TMEAN conditions calculated over different reference periods.

Together, these results demonstrate that the choice of period over which to calculate $TMEAN_{Normal}$ does not substantially affect resulting patterns of geographic and interannual variation in temperature. Estimates of S_{space}, S_{time}, and S_{space} - S_{time} calculated using 1901-2020 normals and anomalies were largely equivalent to those obtained using alternative reference periods. While 1901-2020 estimates of apparent plasticity and adaptation (S_{time}, and S_{space} - S_{time}) tended to have slightly lower magnitude than those for other reference periods, the magnitude of these differences was nearly identical for estimates of apparent plasticity and of apparent adaptation (i.e., slopes in Supplementary Figure 12e-l). In other words, estimates of S_{time} and $S_{space} - S_{time}$ of similar sign and magnitude would change by nearly the same amount when calculating $TMEAN_{Normal}$ over a different period. Therefore, choice of $TMEAN_{Normal}$ should not influence the likelihood that species are classified as showing sensitivity patterns consistent with apparent plasticity or apparent adaptation.

Note S5—Motivation for using a statistical design different from that in Phillimore et al. (2010)

Phillimore et al. (2010) generated spatial and temporal slopes by modeling the effects on frog spawning phenology of long-term, mean temperature among spatial grid cells (assumed to represent populations) and of the anomalies in the year of phenological observations from mean temperature conditions in each population. They used a generalized linear mixed-effects modelling approach (GLMM), treating phenology and temperature as a bivariate response and population and year as random effects, and estimated spatial and temporal slopes from population- and year-specific terms in the variance-covariance matrix of the model. Although statistically robust and largely derived from quantitative genetics theory, their method requires intensive sampling across space and time so that sufficient sampling within spatial units (at spatial scales fine enough that could be plausibly interpreted as populations) would be achieved across multiple years. Indeed, Phillimore et al. (2010) piloted this approach using a single-species dataset of 55,602 frog spawning dates from across Great Britain, a sample size unattainable for even the best sampled species in our extensive dataset.

Delgado et al. (2020) successfully applied this approach to a dataset consisting of occurrence dates for various phenological events across multiple taxa (including plants, saprotrophs, and primary and secondary consumers) using smaller sample sizes per species than those in Phillimore et al. (2010). However, phenological observations in Delgado et al.'s dataset were made regularly at individual sites over time, enabling extensive sampling within each location and relatively precise estimation of occurrence dates. In contrast, although they are abundant and spatiotemporally extensive, herbarium specimens of a given species rarely are collected repeatedly at the same location over many years. Because of this, even among the well-sampled species in the data set analyzed here, we found few areas with repeated observations across years and at spatial scales small enough to plausibly represent single populations. Additionally, specimens could have been collected at any time between the onset and termination of a phenophase (Panchen et al. 2019), which introduces noise to the phenological signal and increases the sample size required to estimate phenology-climate relationships (Ramirez-Parada et al. 2022).

Thus, we estimated spatial and temporal TMEAN sensitivities using different statistical methods from those in Phillimore et al. (2010) and Delgado et al. (2020), partitioning observed TMEAN conditions in sites and years of collection into climate normals and anomalies, and measuring sensitivities in a varying-slopes and varying intercepts Bayesian GLMM framework instead (see Methods: *Analyses* in the main text). Our methods relied on the same set of biological assumptions, whose plausibility we evaluated in the analyses referenced within the 'Methods: *Exploring assumptions*' subsection.

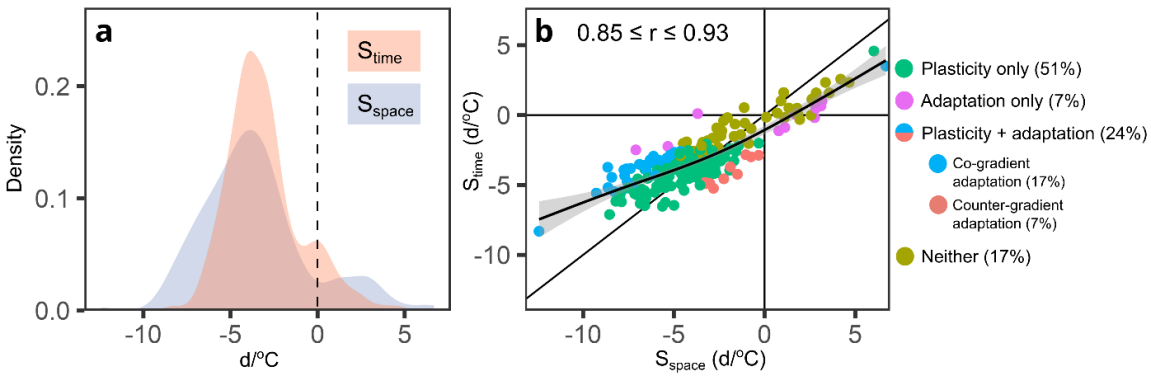

Figure S15—Distributions of and relationship between S_{space} and S_{time} among 201 long-lived species in the continental United States. The solid black line in (b) indicates a 1:1 relationship corresponding to perfect agreement between sensitivity types. The solid curved line indicates the line of best fit obtained from a Generalized Additive Model (GAM) of S_{time} vs. S_{space}, with the shaded area around it denoting the standard error of the predicted mean value. Each point in (b) represents a species whose *x, y* coordinates are given by the maximum a posteriori (MAP) estimates for S_{space} and S_{time}, respectively. Point shapes and colors in (b) indicate whether sensitivity patterns were consistent with plasticity or adaptation as the sole drivers of flowering time variation along the temperature gradient, with both plasticity and adaptation having significant effects in a co- or counter-gradient adaptation pattern, or not showing statistically significant adaptation nor plasticity. The percent of species showing each pattern is shown in the legend in parenthesis. The 95% credible interval for the correlation between S_{space} and S_{time} is provided as a text inset in (b).

The subset of 201 species was selected based on growth form data from the United States Department of Agriculture Plant Database (USDA Plant Database, https://plants.usda.gov). We downloaded all records of growth habit information available through the search tool, and subset the resulting dataset to contain only species represented among the 1,605 species included in the analyses presented in the main text. We then retained only flowering specimens from species classified as "Tree" (n = 5), "Shrub" (n = 164), "Subshrub" (n = 27) or "Vine" (n = 5), which yielded a dataset of 201 species. Using this subset dataset, we ran the same varying intercepts, varying slopes model, obtaining estimates of sensitivity to $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ and of their difference for each species, as well as an estimate of their correlation accounting for parameter uncertainty. The resulting patterns closely mirrored those of the larger dataset, with a high correlation and agreement in magnitude between S_{space} and S_{time}, and similar relative frequencies among species for each sensitivity pattern.

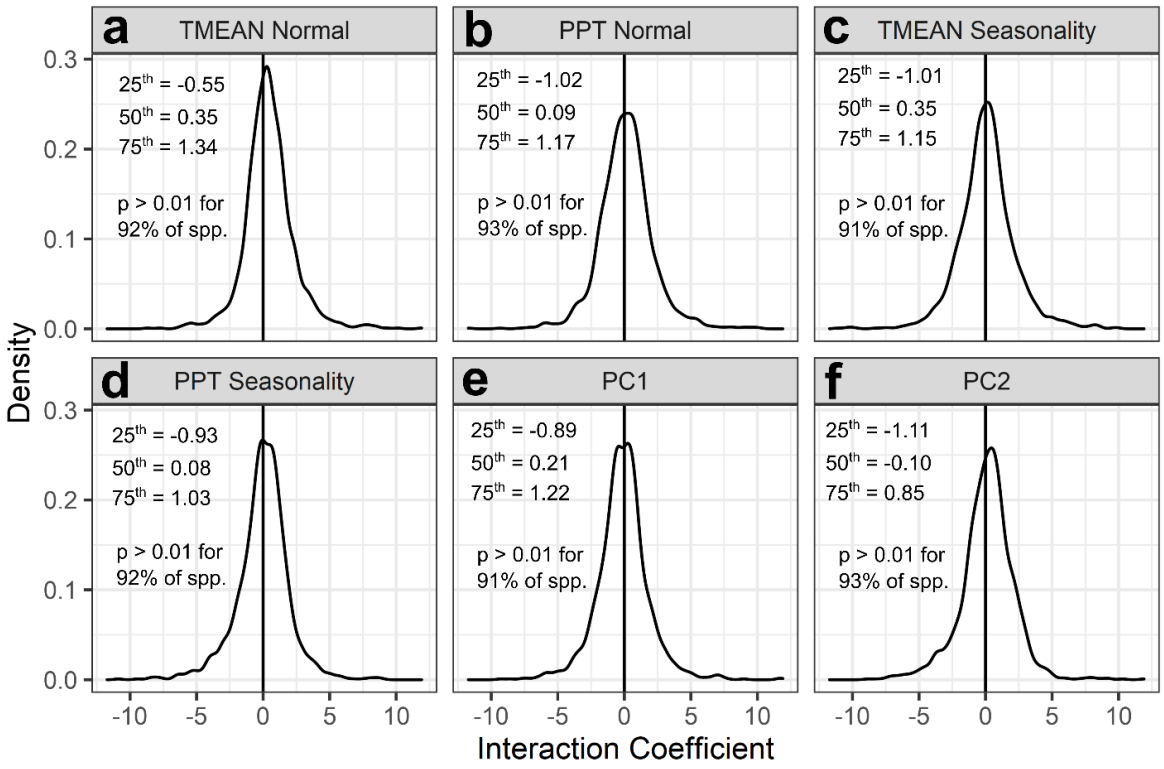

Figure S16—Variation in S_{time} across geographic climatic gradients for 1,605 across the coterminous United States. (a) Distribution of interaction terms between $TMEAN_{Anomaly}$ and long-term climatic conditions within sites of specimen collection, including (a) $TMEAN_{Normal}$, (b) PPT Normal, (c) TMEAN Seasonality, (d) PPT Seasonality, (e) the gradient of increasing temperature and precipitation seasonality described by PC1, and (f) the gradient of decreasing precipitation and increasing temperature seasonality described by PC2. The interaction coefficients for all variables were obtained from single-species models including flowering DOY as a response, the focal long-term climatic variable and $TMEAN_{Anomaly}$ as a predictor, and an interaction term between them. Long-term climatic variables were standardized (mean = 0, SD = 1) before fitting the models. Accordingly, the interaction terms quantify the change in the slope of $TMEAN_{Anomaly}$ vs. DOY (S_{time}, in days/°C) for an increase of 1 SD in the long-term climatic variable. The values for the 25th, 50th and 75th percentiles of each distribution is indicated in each panel as a text inset, as well as the proportion of species for which the interaction coefficient had a p value greater than 0.01.

Within a species, phenological sensitivity to temperature can vary among portions of the range with different long-term climatic conditions. Therefore, differences between S_{space} and S_{time} presented in the main text may result from variation in the DOY–$TMEAN_{Anomaly}$ slope across the geographic climatic gradients and not from the effects of adaptation across the gradient. Despite this, found no evidence of pervasive variation in S_{time} along various geographic climatic gradients, suggesting intraspecific variation in phenological sensitivity is unlikely to generate the patterns reported in the main text.

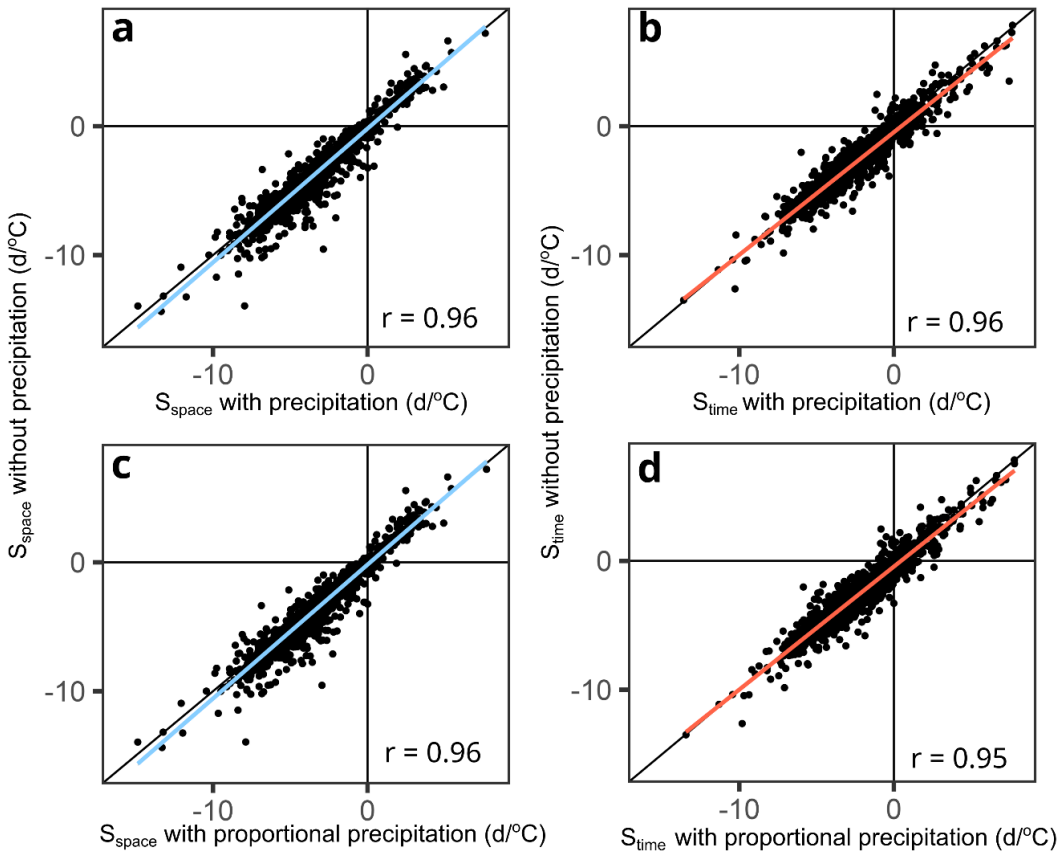

Figure S17—Consequence of including precipitation in models estimating S_{space} and S_{time}. Relationship between estimates of flowering sensitivity to $TMEAN_{Normal}$ (S_{space}) (a, c) and $TMEAN_{Anomaly}$ (S_{time}) (b, d) with or without accounting for the effects of cumulative precipitation normal and anomaly (x-axis and y-axis, respectively) during the same 3-month periods used to calculate $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ for each species (see Methods in main text). Panels (a) and (b) show the relationships between estimates from temperature-only models with those obtained from models including PPT normal and net PPT anomaly for the focal 3-month period in the year of collection. In turn, panels (c) and (d) show the same relationship but with estimates from a model including PPT normal and PPT anomaly *proportional* to the long-term average for that period (i.e., divided by the PPT normal). Proportional anomalies were included to account for differences in the biological significance that the same amount of precipitation might have in chronically dry compared to chronically wet locations.

The method developed by Phillimore et al. (2010) assumed that the variables causing phenological variation along spatial temperature gradients were correctly identified and included in the model. Although temperature has been found to be a predominant environmental cue inducing flowering in temperate biomes, other variables such as precipitation, or those that emerge from the interaction between temperature and precipitation, such as snow cover or water stress, routinely have been implicated in phenological variation in many North American species. Therefore, it is possible that

differences in spatial vs. temporal patterns of temperature-related phenological variation might stem from the confounding effects of phenologically important variables not included in our models. Estimates of phenologically important variables such as the timing of snowmelt or the onset of drought conditions in xeric environments are not available at the temporal and spatial scales spanned by our data. However, most of these variables are highly correlated with precipitation and temperature over space and time, and including both of these variables in phenoclimatic models might account for the effects of predictors other than temperature and precipitation. Accordingly, we assessed whether estimated S_{space} and S_{time} changed when accounting for the effects of long-term cumulative precipitation (PPT normal) and PPT anomalies in the year of collection, separately assessing the effects of both net PPT anomalies and of anomalies scaled proportionally to long-term means (PPT normal) for the focal 3-month period. Estimates of phenology-temperature relationships in space and time did not change substantially when including precipitation variables, resulting in a very high correlation between estimates from temperature-only models and those from models including precipitation ($r = 0.95$ or 0.96). Therefore, the estimates presented in the main text are unlikely to be biased by the omission of precipitation during the months leading up to flowering.

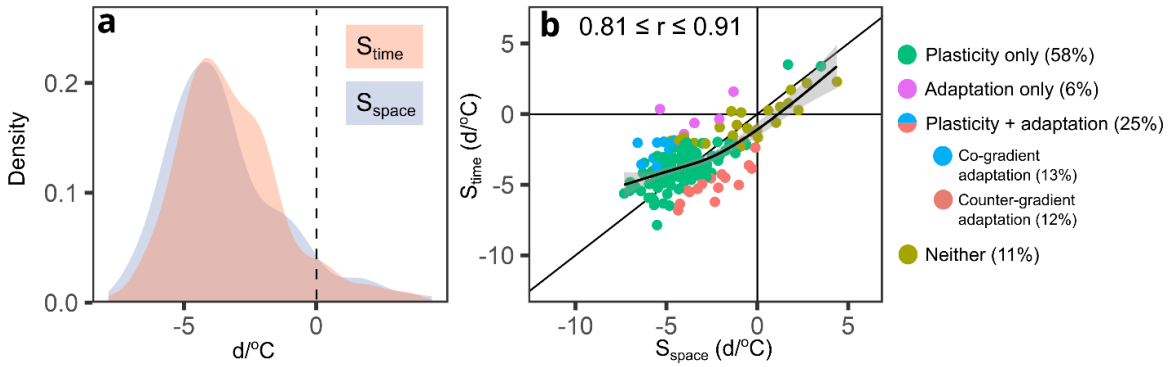

Figure S18—Distribution of and relationship between S_{space} and S_{time} among species collected within narrow latitudinal bands. These analyses included 157 species with 200 or more specimens collected within a latitudinal band of 1° (~111km) in the continental United States (analogous to Figure 3 of the main text). The solid black line in (b) indicates a 1:1 relationship corresponding to perfect agreement between the two types of sensitivity. The solid curved line indicates the line of best fit obtained from a Generalized Additive Model (GAM) of S_{time} vs. S_{space}, with the shaded area around it denoting the standard error of the predicted mean value. Each point in (b) represents a species whose *x, y* coordinates are given by the maximum a posteriori (MAP) estimates for S_{space} and S_{time}, respectively. Point shapes and colors in (b) indicate whether sensitivity patterns were consistent with plasticity or adaptation as the sole drivers of flowering time variation along the temperature gradient, with both plasticity and adaptation having significant effects in a co- or counter-gradient adaptation pattern, or not showing statistically significant adaptation nor plasticity. The percent of species showing each pattern is shown in the legend in parenthesis. The 95% credible interval for the correlation between S_{space} and S_{time} is provided as a text inset in (b).

Both temperature and photoperiod are known to be the predominant environmental cues controlling both vegetative and reproductive phenology among plants in temperature regions. Therefore, across latitudinal ranges such as those spanned by most species in our data (median 90% latitudinal range = 5.7°), it is possible that differences in S_{time} and S_{space} (e.g., geographic temperature gradients) might reflect the confounding influence of latitudinal shifts in photoperiod on our estimates of sensitivity to $TMEAN_{Normal}$. To account for this possibility, we identified 157 species in our data that were well sampled (200 or more specimens) within narrow latitudinal bands (≤ 1°). Using this subset of species and including only specimens from such 1° bands, we ran the same varying intercepts, varying slopes model, obtaining estimates of S_{time} and S_{space} and their difference for each species, and an estimate of their correlation accounting for parameter uncertainty. The results did not qualitatively differ from those presented in the main text, with a high correlation between S_{space} and S_{time}, and similar relative frequencies of each sensitivity pattern among species..

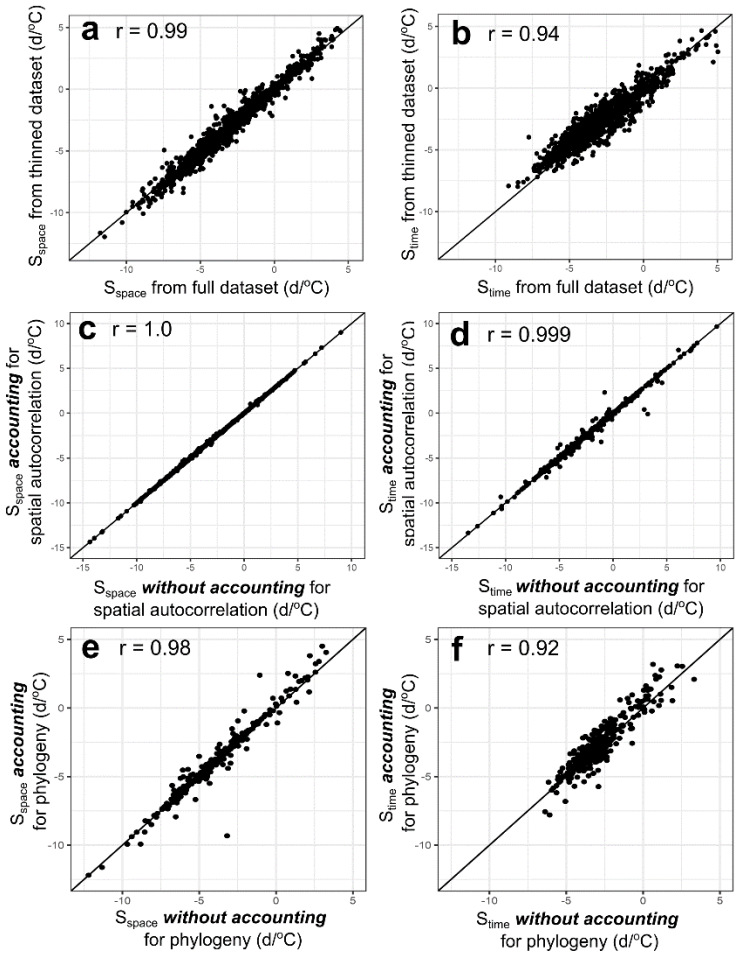

Figure S19—Effects of sample size differences, spatial autocorrelation and phylogeny on estimates of S_{space} and S_{time}. Comparison of S_{space} and S_{time} estimates obtained by (a) homogenizing sample sizes among species, (b) accounting for spatial autocorrelation among observations and (c) accounting for phylogenetic relationships among species against estimates generated ignoring these factors (as those presented in the main text). In (a, b), we fit the model presented in the main text using a thinned dataset were each species was represented by 300 specimens, comparing its output to that of the model in the main text. In (c, d), we compared the results of models omitting or accounting for phylogenetic relationships. We selected a random subset of 300 species from which to generate a phylogeny, thinning these data to include only 300 specimens for each species (to make the model computationally tractable). S_{space} and S_{time} estimates that did not account for phylogeny were obtained using the model described in the main text. In turn, the model accounting for phylogeny included a prior for the covariance structure of species-specific parameters consisting of the evolutionary distance between
each pair of species as estimated from a phylogenetic hypothesis and a model of trait divergence among species. The phylogenetic tree (or hypothesis) was generated using the R package 'v.PhyloMaker' version 0.1.071 and generated a phylogeny resolved to the genus level. Using this tree, we then calculated the
variance–covariance phylogenetic matrix predicted by a Brownian model of trait evolution using the R package 'ape' version 5.6-2. Finally, both models were implemented using the 'brms' package version 2.18.073. Finally, in (e, f) we compared estimates obtained from models ignoring or

accounting for spatial autocorrelation of the residuals. All S_{space} and S_{time} estimates were obtained using single-species models, but those accounting for spatial autocorrelation included a covariance structure for the residuals determined by the geographic distance between each pair of points. All models were fitted using the 'nlme' package version 3.1 in R. Estimates of S_{space} and S_{time} obtained accounting for or ignoring spatial autocorrelation were nearly identical across species. Across panels, the x-axes show the estimates obtained when omitting the focal factor (sample size, phylogeny, or spatial autocorrelation), whereas the y-axes show estimates obtained when accounting for it. Solid black lines represent a 1:1 line, representing perfect agreement in magnitude and direction between estimates. S_{space} and S_{time} estimates obtained ignoring sample size differences, phylogeny and spatial autocorrelation were highly correlated to estimates obtained from models accounting for these factors. Accordingly, we consider it unlikely that omitting these factors could have biased our results.

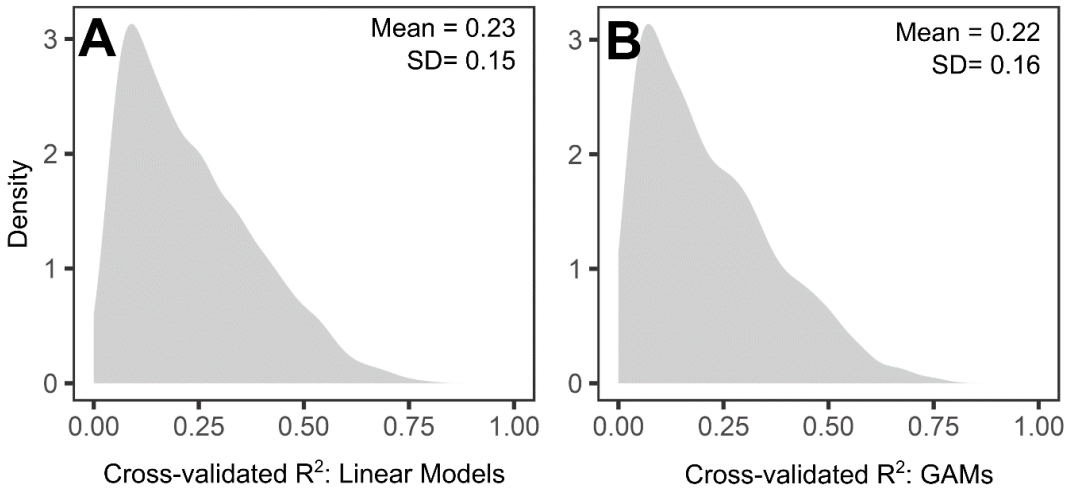

Figure S20—Assessing evidence for non-linear phenology-temperature relationships. Comparison of R^2 values obtained using 10-fold cross-validation of models of flowering DOY versus $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ obtained from (A) linear regressions assuming linear relationships between phenology and temperature or (B) generalized additive models (GAMs) accounting for potential non-linear relationships. The shaded region in each panel represents the among-species kernel distribution of cross-validated R^2 values obtained using each model type (linear regression or GAM). The mean and SD of R^2 values each is presented as a text inset in each panel.

The model that generated the sensitivity estimates presented in the main text assumed linear relationships between flowering dates and $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$. To verify whether such an assumption was warranted for our data, we compared the predictive ability of single-species models assuming linear relationships between phenology and temperature (fitted using linear regression) and models accounting for possible non-linear relationships (fitted using Generalized Additive Models). We reasoned that if omitted non-linear relationships between flowering time and temperature were pervasive in our data and potentially biased our results, then models accounting for non-linear relationships would tend to perform better than linear regressions among species in our data. We used 10-fold cross-validation to compare the out-of-sample performance (quantified through R^2 values) of linear regressions and GAMs. For each model type (linear regression or GAM), this procedure randomly split the observations for each species into 10 groups, each of which was omitted from a model estimated from the remaining 9 groups. The performance of each of these models was then assessed against the observations omitted in fitting the model, generating 10 out-of-sample R^2 values for each model type (linear or GAM) per species. We then compared the distribution of mean cross-validated R^2 values obtained from linear models and GAMs to assess whether non-linear models explained additional variance.

GAMs did not perform better than linear models among species in our data; in fact, linear models exhibited marginally higher R^2 values among species than GAMs (0.23 vs. 0.22). Therefore, we conclude that the assumption of linearity is warranted in the analyses presented in the main text.

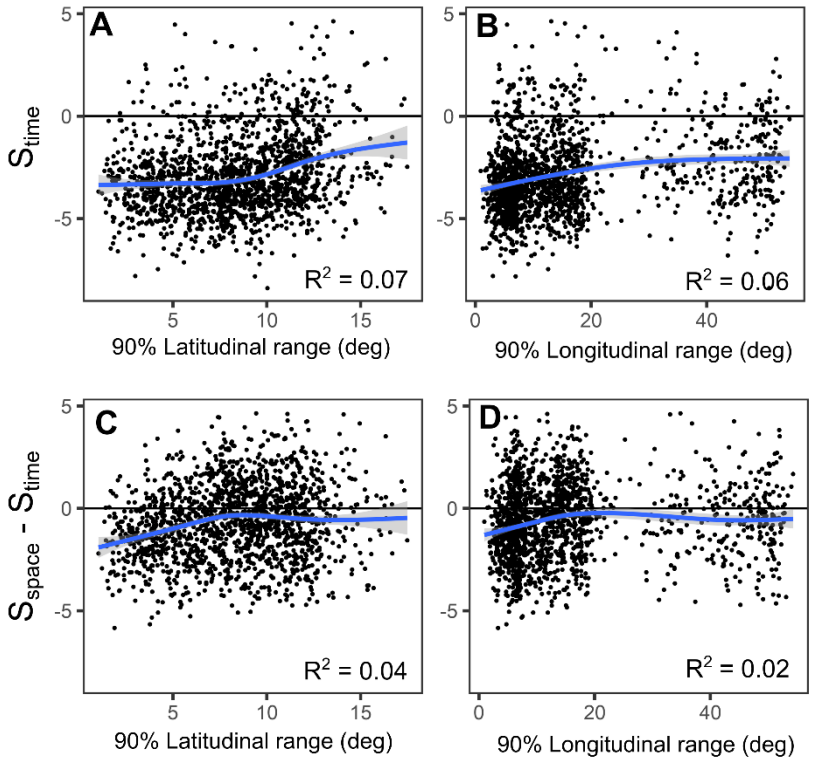

Figure S21—Effects of geographic range on apparent plasticity and adaptation. Relationships between the latitudinal and longitudinal range spanned by the central 90% of specimens of a species and estimates of apparent plasticity (S_{time}) and apparent adaptation ($S_{space} - S_{time}$). Blue lines in each panel correspond to best-fit lines obtained using generalized additive models (GAMs). R^2 are provided as text insets in each panel.

Although apparent plasticity and adaptation showed marginally greater magnitude among species with narrower latitudinal and longitudinal range, these relationships explained a very small proportion of the variance. Therefore, we conclude that it is unlikely that differences in latitudinal or longitudinal range size could confound the results presented in the main text. GAMs using apparent plasticity or apparent adaptation as a response and including both latitudinal and longitudinal range as predictors also explained a marginal proportion of the variance ($R^2 = 0.08$ and $R^2 = 0.05$, respectively).

Note S6—Assessing collection patterns relative to major roads and population density and their effects on estimation of apparent plasticity and adaptation.

Daru et al. (2018) showed severe spatial collection biases for collections within Australia, South Africa, and New England (USA), chief among them the concentration of specimen collection near major roads and urban areas. Such environments might show different climatic conditions from undisturbed areas. To the extent that such microenvironments might be severely different and not captured by available climate products, estimates of phenology-climate relationships obtained without accounting for these factors could be inaccurate.

However, Daru et al. (2018) also showed that these regions differed in the degree of specimen clustering around major roads or urban areas, suggesting the severity of these biases might not be universal. As our analyses relied on an overwhelmingly non-overlapping set of collections from those analyzed by Daru et al. (2018), the severity of these biases in our dataset was unknown. Accordingly, we conducted analyses describing the extent of spatial bias in our data with respect to population density (a common proxy for urbanization; Park et al. 2023) and proximity to major roads (i.e., what Daru et al. [2018] call "infrastructure bias"). Then, we assessed whether including proximity to roads and population density into our models in various ways affected our estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$.

Population density

As urbanization typically encompasses large areas, climatic data resolved at the 4km level (e.g., that used in our study) are likely to capture its effects on local climatic conditions. For example, Park et al. (2023) used 4km-by-4km climate data from PRISM and population density information from the US Census Bureau to show that geographic and temporal climatic variation and urbanization interactively mediate variation in phenology across the Eastern coastal plain of the United States. If urbanization affects phenology by modifying climate, such effects should be captured by the PRISM data used in our study. However, if urbanization affects the slopes of phenology-climate relationships (e.g., Park et al. 2023), our estimates might be biased, as we omitted interactions between urbanization and climate in our models.

To evaluate these possibilities, we obtained decadal population density estimates for each county of the United States between 2000 and 2020 from the US Census Bureau using the *tidycensus* package v.1.4.4 in R (Walker 2023). To match the temporal extent of the specimen and census data, we subset the data to include only specimens collected from 1990 onwards. This yielded a dataset of 562,631 specimens (54% of the full dataset), with 663 species represented by 300+ specimens. Then, we linked each collection to the population density of its county from the census following its decade of collection. To examine whether specimens were disproportionately collected in highly populated areas, we compared the population density of collection locations in our data to that of an equally sized random sample of points across the contiguous United States. Random points were assigned to decades in the same proportion as observed in our dataset (39% in the 90s, 43% in the 2000s, and 18% in the 2020s) and linked to the corresponding US Census Bureau decadal population density data. Next, for each species, we assessed whether accounting for

population density affected our estimates by including population density (log-transformed) into a model of DOY vs. $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$, entering the model either as a predictor, or both as a predictor and in interaction terms with both temperature variables. Estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$ obtained from both types of model were compared to those obtained omitting population density.

Results

The distribution of population densities among specimen collection locations showed marked peaks compared to the random sample of locations (Supplementary Figure 13), reflecting a concentration of specimen collections within areas of the United States of differing population density (e.g., Southern California, Rocky Mountains, New England, etc.; Extended Data Fig. 2). Overall, specimens were more likely to be collected in relatively densely populated areas, with 19% of specimens occurring in areas with 100 or more residents per km^2 compared to 8% for the random sample of points. However, specimens were also disproportionately collected in areas of low population density, with the modal specimen being collected in counties with approximately 4 residents per km^2. Moreover, <1% of species were collected within counties with 994 residents per km^2 or more (i.e., the median population density of urban areas in the United States as of 2020).

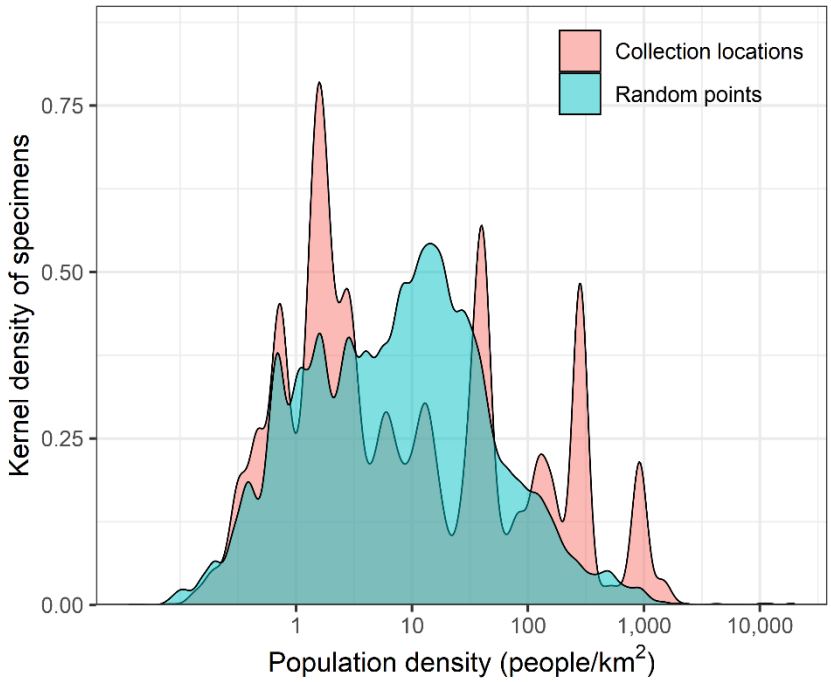

Figure S22—Population density of collection locations vs. random locations. Distribution of county-level population density among collection locations of 562,631 specimens collected since 1990 and of an equally sized random sample of points across the contiguous United States. Decadal county-level population densities were obtained from the US Census Bureau for the years 2000, 2010, and 2020).

Including population density in our models did not substantially alter estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$. Estimates generated from models including population density (with or without interaction terms) were highly correlated to those obtained when excluding population density, and tended to agree in both direction and magnitude (i.e., they were sorted along a 1:1 line) (Supplementary Figure 14).

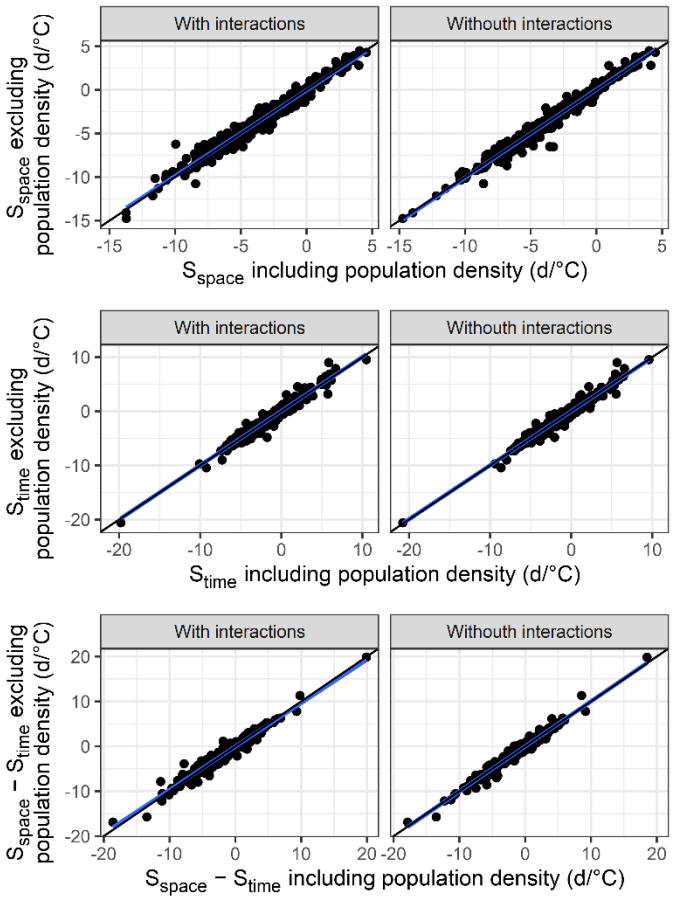

Figure S23—Effects of accounting for population density in our models. Comparison of estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$ obtained from models omitting urbanization (y-axes) and those from models including urbanization either as a predictor, or both as a predictor and in interaction terms with $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ (x-axes). Models were fit separately for each of 662 species represented by 300 or more specimens collected since 1990.

Summary and conclusions

Specimens were disproportionately collected in areas of high population density compared to a random set of locations in the United States, but we also observed higher-than-random collection intensity in counties with low population density. Nonetheless, including population density (both as a linear or interaction term) yielded nearly identical estimates of

S_{space}, S_{time}, and $S_{space} - S_{time}$. This suggests that population density—and by proxy, urbanization—do not result in substantial climatic effects that are not captured by PRISM data, and that the interactive effects of urbanization and temperature on flowering time might not be (on average) large enough to result in biased estimates when omitting interactions.

Many limitations of this analysis are of note. For example, we could only access population density data at the county level, but such density might not be homogenous, and it is possible that conditions in the site of collection of a specimen do not match the county mean (especially among western counties, which are typically larger). Similarly, we had to use collections from recent decades to match the temporal extent of the census data, excluding 46% of observations. While the excluded data were collected throughout a period of lower overall population density (1896-1989), collection biases might have changed over time, and it is conceivable that the urbanization bias among early collection could be greater despite lower national population density. Additionally, it is possible that omitting urbanization might bias estimates of climate variables other than temperature. Nonetheless, these results suggest that—at least for temperature and flowering time—omission of urbanization does not lead to biased estimates of S_{space} and S_{time}.

Proximity to major roads

Following Daru et al. (2018), we measured infrastructure bias as the proximity of each collection location to a major road within the US. We obtained shapefiles of primary and secondary roads for the year 2020 from the US Census Bureau using the *tigris* package v.2.0.3 in R (Walker 2016), and calculated distance to the nearest one for each specimen. As with our analyses of population density (Supplementary Figures 13, 14), we used specimens collected from 1990 to ensure that most roads in the census maps were present at the time of collection. To assess whether collection locations clustered around major roads, we generated a sample of 562,631 random points (same sample size as our post 1990 specimen data) and compared its distribution of to that of the collection locations in our data. Proximity to major roads can only bias our results if, for most species, a sizeable proportion of specimens occurs close enough to roads for them to experience altered microclimatic conditions. Because the exact spatial scale at which roads might affect microclimates is unclear, we measured the proportion of specimens of each species occurring within various distance thresholds of a major road: 50m, 100m, 500m, and 1km.

Finally, we assessed whether including distance to road in models measuring S_{space} and S_{time} affected our estimates. As the scale at which distance to road will affect microclimatic conditions is unclear, we included it in our models in a variety of ways. First—in addition to $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$—we included distance to road as a predictor (without interaction terms), either as a continuous variable (log-transformed), or as a one of three dummy variables indicating whether a specimen occurred within 50m, 100m, or 500m of a road. Then, we fitted versions of these models including interaction terms between the temperature variables and the version of distance to road that entered the model. Each model was only fit for a species if at least 30 specimens occurred within the distance threshold used in the model. Accordingly, model type captured a different number of species: all 663 species for continuous distance, 561 species for the 500m threshold, 205 species for the 100m threshold, and 70 species for the 50m threshold. This approach generated 8 sets of S_{space} and

S_{time} estimates (2 per each distance to road variable) that we compared to those obtained from models excluding distance to major roads.

Results

Specimens were more likely to be collected near major roads than randomized locations across the US: 4.4% of specimens were collected within 100m of a major road compared to 2% among random locations. However, overall, specimens tended to be collected further from major roads than expected at random. The median specimen and random location were 5.8km vs. 3.8km away from a major road (42% greater distance among specimens), and the modal specimen and random location were 9.1km vs. 4.7km away from a major road (94% greater distance among specimens) (Supplementary Figure 15).

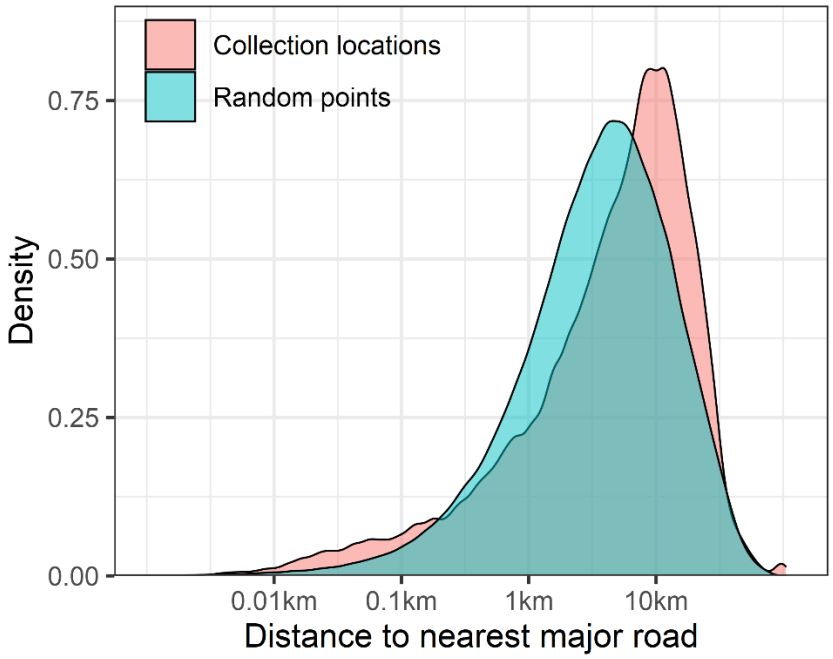

Figure S24—Distance to roads of specimen collection location vs. random locations. Distribution of distance to major roads for 562,631 collection locations of specimens and for an equal number of randomly generated locations within the continental United States. Geospatial data for major roads (i.e., primary and secondary) was obtained from the US Census Bureau for the year 2020.

Additionally, we found that, for most species, only a small proportion of specimens tended to occur very close to major roads (Supplementary Figure 16). For the median species, only 2% of specimens were collected within 50m of a major road, and only 4% within 100m. While the median proportion of specimens for greater distance thresholds was

much greater for the 500m and 1km thresholds (11% and 17%, respectively), the likelihood that roads affect microclimates decreases substantially with distance.

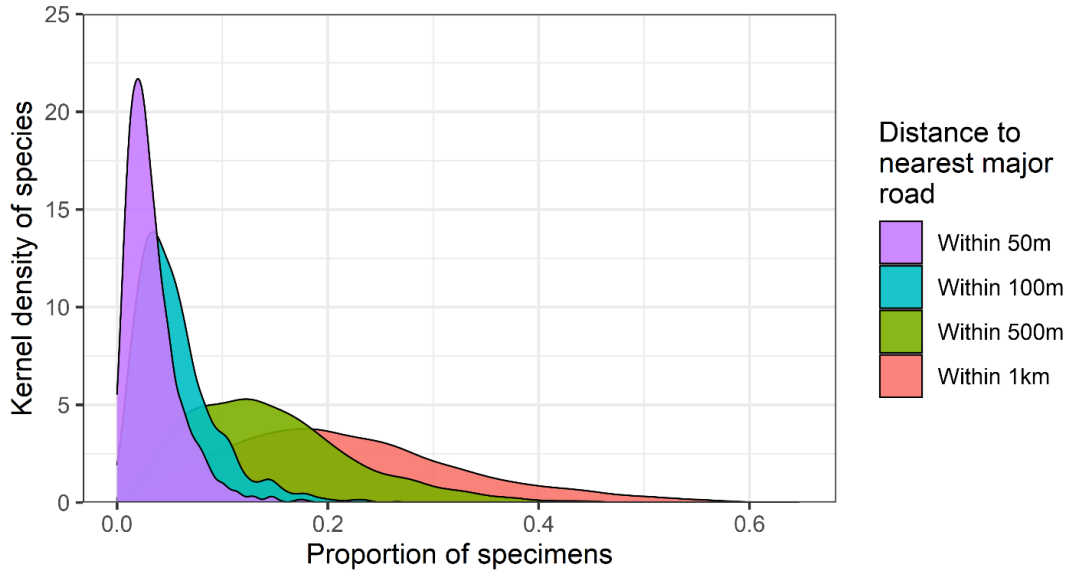

Figure S25—Distribution of the proportion of specimens of a species occurring within 50m, 100m, 500m, or 1000m of a major road. Each colored distribution is generated from the proportion of specimens within a distance threshold across 1,605 species included in our analyses. Geospatial data for major roads (i.e., primary and secondary) was obtained from the US Census Bureau for the year 2020.

Inclusion of distance to major roads—with or without interaction terms—did not substantial impact our estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$. Models including distance to road with or without interaction terms were very highly correlated to those that only included $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ ($r > 0.96$ for all estimates), and their magnitude and direction tended to correspond in a 1:1 ratio (Supplementary Figures 17, 18).

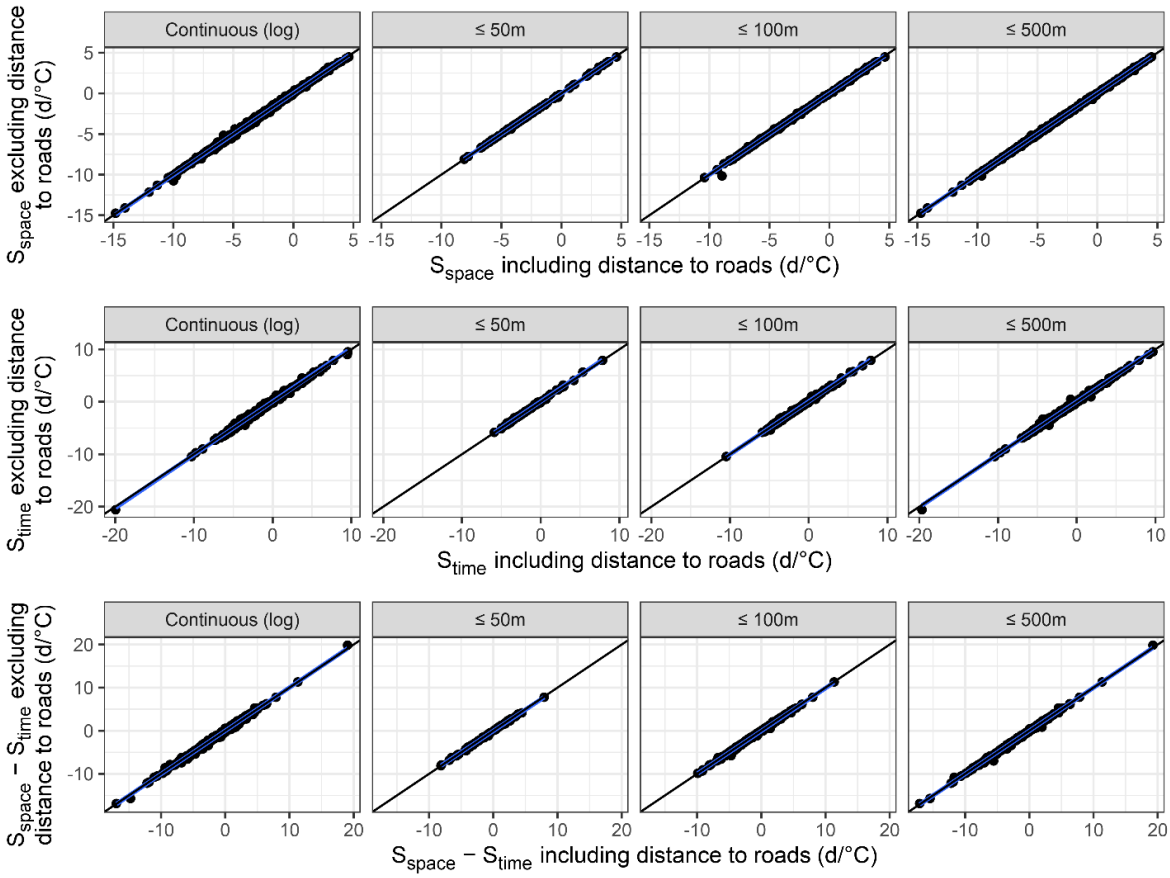

Figure S26—Accounting for distance to roads in our models without interactions. Comparison of estimates of S_{space}, S_{time} and $S_{space} - S_{time}$ obtained including distance to major roads as a predictor (***without*** interaction terms; x-axes) to those obtained omitting distance to roads (y-axes). Distance to road was included in four different ways across models: as a continuous variable (log-transformed) or as one of three dummy variables indicating whether a specimens occurred within 50m, 100m, or 500m of a road.

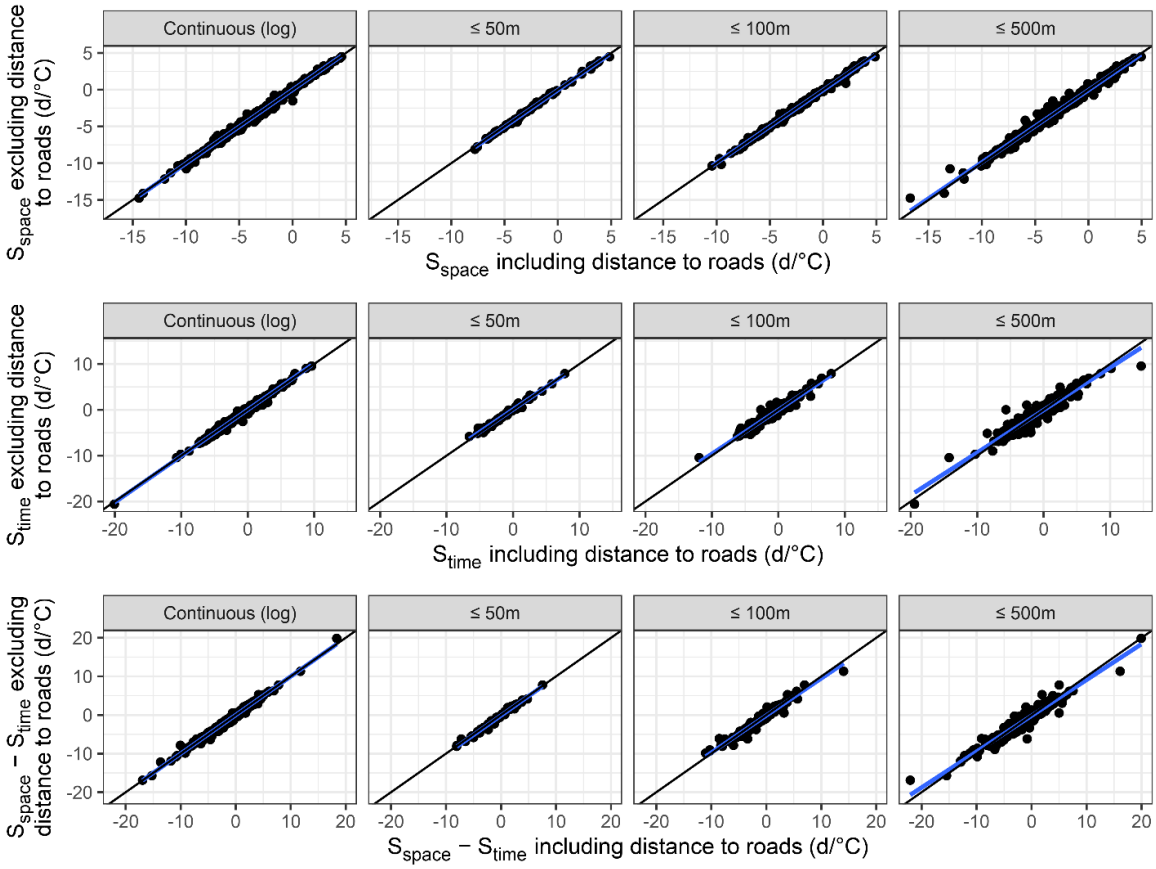

Figure S27—Accounting for distance to roads in our models with interactions. Comparison of estimates of S_{space}, S_{time} and $S_{space} - S_{time}$ obtained including distance to major roads as a predictor (*with* interaction terms between distance to roads and $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$) to those obtained omitting distance to roads. Distance to road was included in four different ways across models: as a continuous variable (log-transformed) or as one of three binary variables indicating whether a specimens occurred within 50m, 100m, or 500m of a road.

Summary and conclusions

While a slightly greater proportion of specimen collections occurred near major roads relative to randomized locations across the United States, specimens tended to occur further from major road than expected at random sample of locations (Supplementary Figure 17). Moreover, only a small proportion of specimens per species was collected within relatively short distances from roads (i.e., within 50m or 100m) (Supplementary Figure 18). These patterns suggest that, despite some degree of opportunistic sampling near highly accessible areas, botanists contributing to the datasets that we obtained might have preferentially collected away from major infrastructure (at least since the 1990s). However, it is conceivable that these patterns might have differed for collections earlier in the 20th century for which we do not have temporally resolved road data.

Ultimately, estimates of the effects of $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ were robust to inclusion of distance to road in the models (Supplementary Figures 17, 18). This lack of apparent effects of distance to roads on our estimates might be because they do not tend to substantially bias the climate space captured by the data (Kadmon et al. 2004). Alternatively, this might be caused by the fact that the microclimatic effects of roads often operate at spatial scales much lower than 50m (Delgado et al. 2007). If so, such fine-scale effects would be unlikely to alter our results, as species tended to show a very low proportion of specimens within 50m of a road, and such proportion would be substantially lower for smaller distance thresholds.

Note S7—Assessing the temporal distribution of specimens and the potential for temporal non-independence among observations

Herbarium data are unevenly distributed over time, and collections might often sparsely sample recent decades during which most global warming has occurred (Daru et al. 2018). Collections may be clustered within individual years, or the date of collection of conspecific individuals in a phenophase might not be independent among years, as trends in climate over time might also cause trends in phenology.

To explore these potential issues, we characterized the temporal distribution of our data, and assessed the proportion of specimens within each species collected before the approximate onset of rapid warming in North America (1980), and during the 20th century. While a rigorous analysis of temporal autocorrelation and its potential impact on estimates of S_{space} and S_{time} would require auto-regressive (AR) models—or at least analysis of residuals—relying on complete time series during the study period (or also on imputation methods for incomplete time series provided missing observations are few and occur randomly). Unfortunately, most species have several multi-year gaps, making our data unsuitable for these approaches.

Accordingly, we evaluated whether models accounting for various forms of temporal non-independence generated different estimates from model assuming temporal independence among observations (such as that in the main text). First, to account for potential temporal non-independence resulting from temporal trends in both temperature and phenology in each species, we included year as a predictor in a model of DOY vs. $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$. Second, as rapid warming trends only started towards the end of the 20th century (ca. 1980), and as temperatures have fluctuated non-linearly over time, we included year as a non-linear predictor using a spline smooth within a Generalized Additive Model (GAM; mgcv package v.1.8.40, Wood 2017). Finally, uneven sampling across years within species might result in an overrepresentation of years and climatic conditions; accordingly, we ran a third type of model including year as a random effect (through the lme4 package v.1.1-30, Bates et al. 2015), allowing for variation in the intercept of specimens clustered within the same year. Estimates of S_{space} and S_{time} (and their difference) for each species obtained using each type of model were compared to those obtained from models omitting year (as did the model in our main analyses).

Results

Specimens in our data were collected between 1896 and 2020, with collection intensity fluctuating but tending to increase over time (Supplementary Figure 19). The number of specimens collected each year decreased gradually throughout the 2010s, likely reflecting lags in digitization for more recent collections. Overall, 62% of our data was collected following the onset of rapid warming trends (ca. 1980), with 35% collected during the 20th century. This temporal distribution was mirrored at the species level. The median species had 61% of its specimens collected on or after 1980, and 34% of its specimens collected in the 21st century.

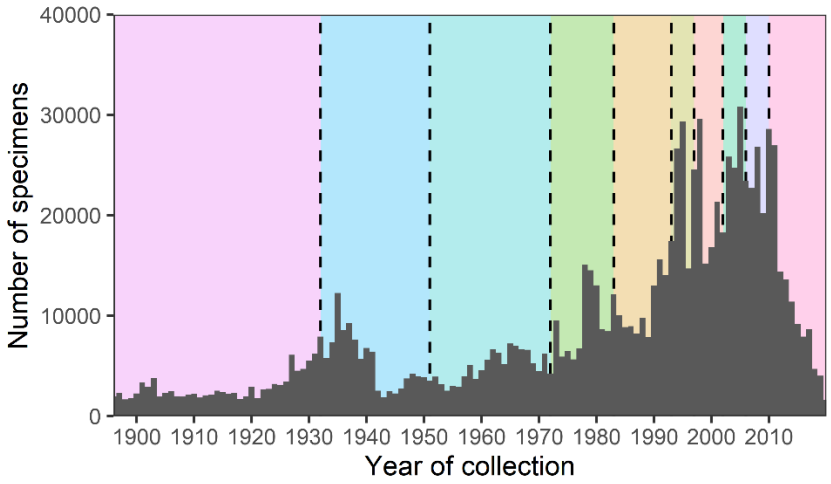

Figure S28—Temporal distribution of the 1,038,027 specimens included in the analyses within the main text. Colored segments divided by dashed lines indicate year of collection deciles, and therefore divide the data in 10 bins of 103,802 or 103,803 specimens. The width of each segment indicates the period over which each successive batch of ~103,802 specimens was collected. Bars show the number of specimens collected within a single year.

Including year in our models did not generate substantially different estimates from those obtained from models omitting year (Supplementary Figure 20). Regardless of the way in which year entered the model, estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$ were highly correlated to those models including only $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ ($r \geq 0.92$), and their magnitude and direction tended to correspond in a 1:1 ratio.

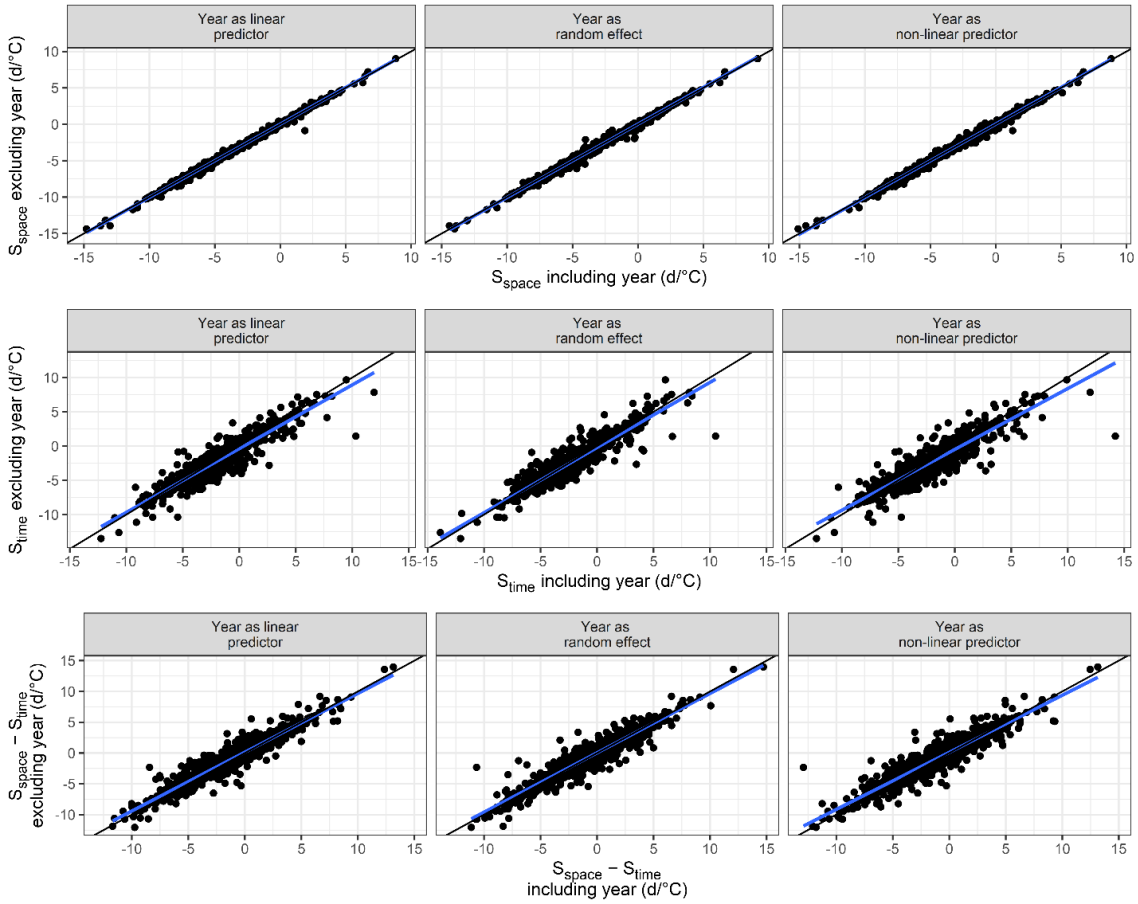

Figure S29—Assessing potential for temporal non-independence. Comparison of estimates of S_{space}, S_{time}, and $S_{space} - S_{time}$ obtained omitting year vs. those obtained by including year as a linear predictor, a random effect, or a non-linear predictor. Models were fit separately for each of the 1,605 species included in our main analyses.

Conclusions

Species in our data were typically represented by sizeable proportions of specimens collected before and after the onset of rapid warming and during the much warmer conditions of the 21st century. Accordingly, we conclude that our data likely provide sufficient observations before, during, and after the onset of warming trends to properly capture variation across the spectrum of conditions spanning the 20th and early 21st centuries.

While time-series analysis methods for assessing temporal autocorrelation where not applicable to our data, including year in our models in ways that account for various possible scenarios of non-independence generated estimates highly consistent with those from models including only $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$. Together with the analyses showing extensive sampling before and after the onset of rapid warming for most species (Supplementary Figure 19), as well as those demonstrating no evidence of non-linear phenology-temperature relationships (Appendix 2—Fig. S20), these results suggest that the

models presented in the main text are likely to estimate S_{space}, S_{time}, and $S_{space} - S_{time}$ accurately.

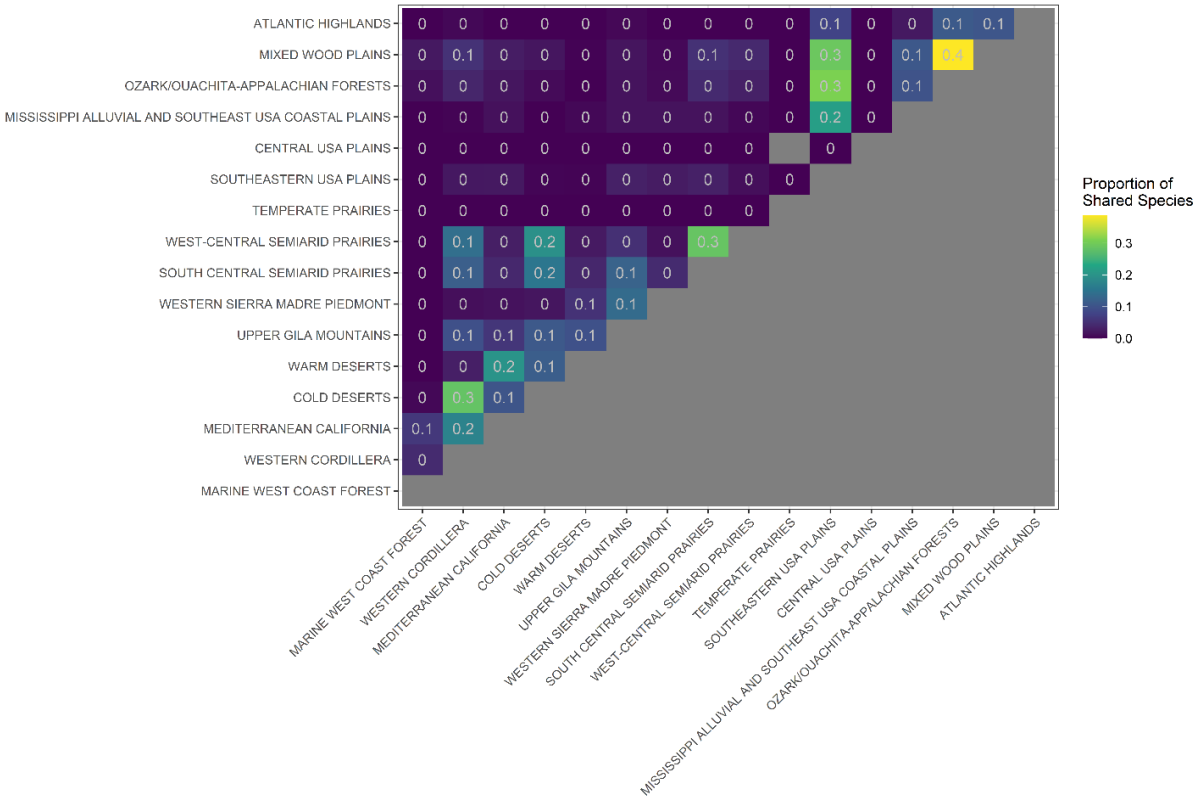

Figure S30—Compositional similarity among ecoregions of the United States. Proportion of species (n = 1,605) shared between pairs of Level II Ecoregions within the contiguous United States. The proportion of species overlap for each pair was calculated as the number of species occurring in both ecoregions (i.e., those for which at least 10% of the specimen were collected within that ecoregion) divided by the total number of unique species present in the two ecoregions combined.

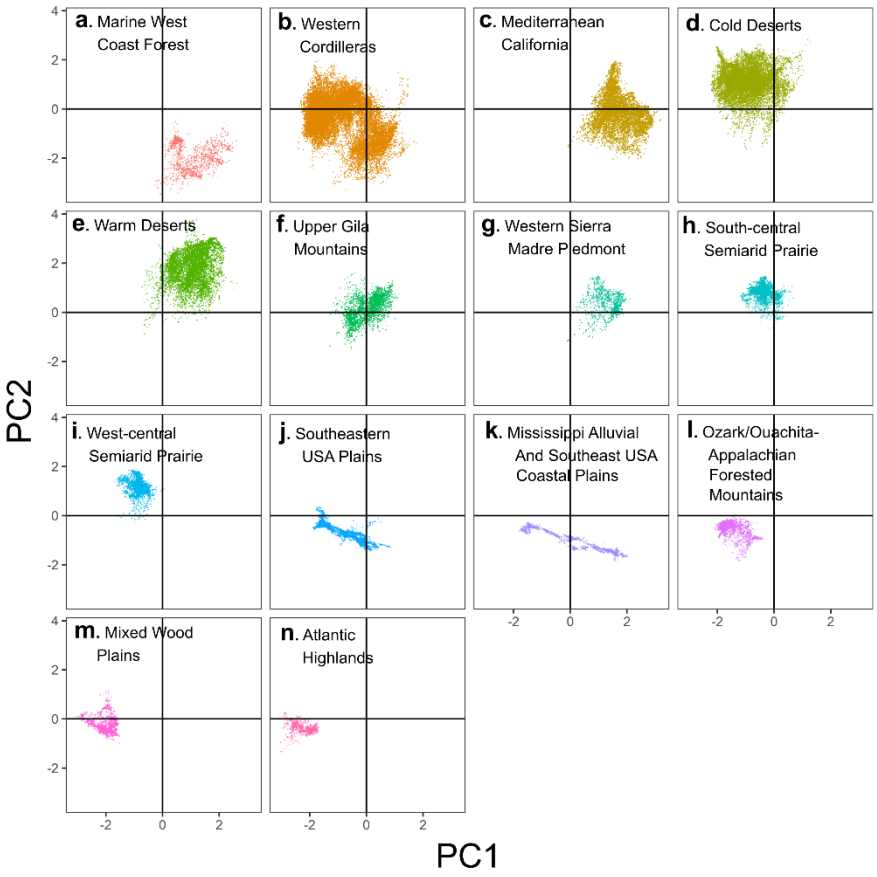

Figure S31—Climatic space captured among specimen collection locations across ecoregions. Variation in long-term climatic conditions among sites of specimen collection occurring within different Level II ecoregions throughout the contiguous United States. Variation in long-term conditions was calculated using principal components (PCs). PC1 represents a gradient of increasing precipitation seasonality (PPT seas. in the legend), decreasing MAT seasonality (T° seas.), and increasing mean annual temperature (MAT) normal (Mean T°). In turn, PC2 represents a gradient of decreasing mean annual precipitation (Mean PPT), and increasing MAT seasonality.

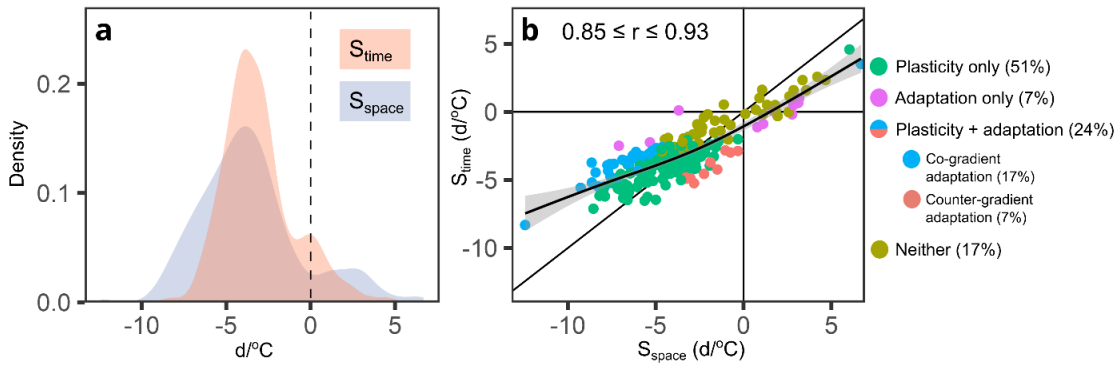

Figure S32—Distributions of and relationship between S_{space} and S_{time} among 201 long-lived species in the continental United States. The solid black line in (b) indicates a 1:1 relationship corresponding to perfect agreement between sensitivity types. The solid curved line indicates the line of best fit obtained from a Generalized Additive Model (GAM) of S_{time} vs. S_{space}, with the shaded area around it denoting the standard error of the predicted mean value. Each point in (b) represents a species whose *x, y* coordinates are given by the maximum a posteriori (MAP) estimates for S_{space} and S_{time}, respectively. Point shapes and colors in (b) indicate whether sensitivity patterns were consistent with plasticity or adaptation as the sole drivers of flowering time variation along the temperature gradient, with both plasticity and adaptation having significant effects in a co- or counter-gradient adaptation pattern, or not showing statistically significant adaptation nor plasticity. The percent of species showing each pattern is shown in the legend in parenthesis. The 95% credible interval for the correlation between S_{space} and S_{time} is provided as a text inset in (b).

The subset of 201 species was selected based on growth form data from the United States Department of Agriculture Plant Database (USDA Plant Database, https://plants.usda.gov). We downloaded all records of growth habit information available through the search tool, and subset the resulting dataset to contain only species represented among the 1,605 species included in the analyses presented in the main text. We then retained only flowering specimens from species classified as "Tree" (n = 5), "Shrub" (n = 164), "Subshrub" (n = 27) or "Vine" (n = 5), which yielded a dataset of 201 species. Using this subset dataset, we ran the same varying intercepts, varying slopes model, obtaining estimates of sensitivity to $TMEAN_{Normal}$ and $TMEAN_{Anomaly}$ and of their difference for each species, as well as an estimate of their correlation accounting for parameter uncertainty. The resulting patterns closely mirrored those of the larger dataset, with a high correlation and agreement in magnitude between S_{space} and S_{time}, and similar relative frequencies among species for each sensitivity pattern.

VIII. Appendix 3—Supplementary materials for IV

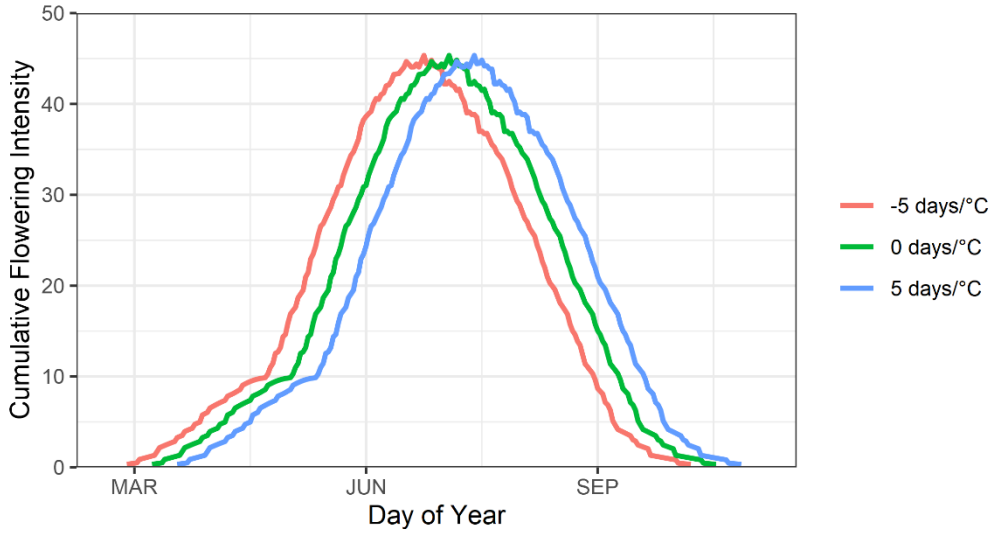

Figure S1—Time series of cumulative flowering intensity under 2 °C of warming for three communities differing only in the mean temperature sensitivity of first flowering date (i.e., in mean S_{FFD}) among species. These communities were generated from the same initial set of 100 simulated species, each randomly assigned first flowering dates, flowering durations, flowering termination dates, and mean flowering dates as described in the "Methods" section of the main text. Species were then assigned random duration sensitivities (i.e., S_D), with the mean S_D among species set at 0 days/°C. In turn, each species was assigned three S_{FFD} values. An initial S_{FFD} was obtained as described in the "Methods" section of the main text, resulting in a distribution of S_{FFD} values among species centered at 0 days/°C. Then, we assigned two additional S_{FFD} values for each species: one generated by adding and another by subtracting 5 days/°C to the initial S_{FFD} assigned to each species. This resulted in 3 different sets of S_{FFD}, each showing identical variation among species but differing in their mean. Using each of these three S_{FFD} distribution, we simulated a time series of flowering intensity for each species under 2 °C of warming, and calculated the cumulative flowering intensity of the community as described in the "Methods" section of the main text. This resulted in three post-warming time series of cumulative flowering intensity differing only in the among-species mean S_{FFD} used to generate them.

Generating identical communities that differed only in their mean S_{FFD} among species merely shifted the relative position of an otherwise identical flowering season under 2 °C of warming. As we were interested in assessing changes to the temporal distribution of flowering species at the community level—not the precise position of the flowering season within the year—we set mean S_{FFD} to be 0 in all simulations for simplicity.

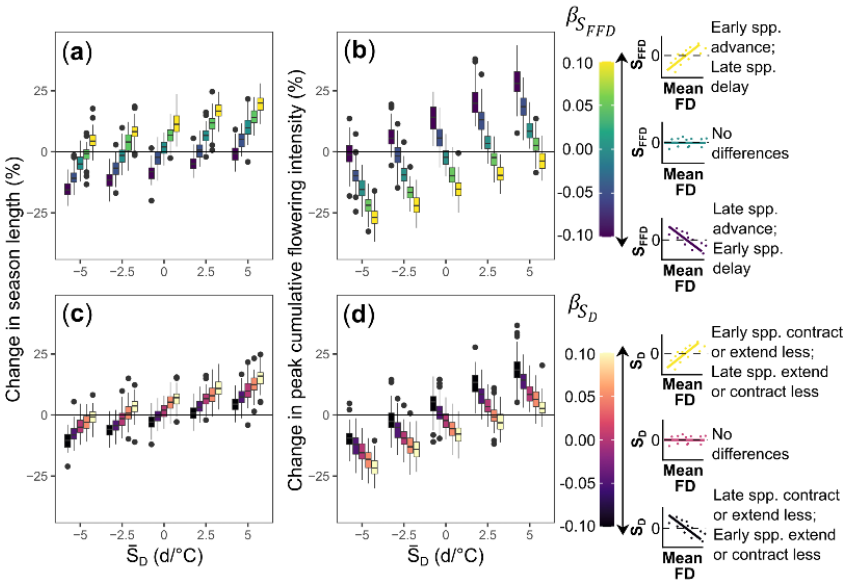

Figure S2—Percent change in flowering season length and seasonal peaks in cumulative flowering intensity relative to the pre-warming baseline due to independent variation among species in S_{FFD} and S_D for communities with a single flowering peak. Colored boxplots in **a** and **b** depict the range of variation in season length change for communities varying in $\beta_{S_{FFD}}$, with groups across the x-axis representing sets of communities with varying $\bar{S}_D$. In turn, colored boxplots in **a** and **b** depict the range of variation in season length change for communities varying in β_{S_D}.

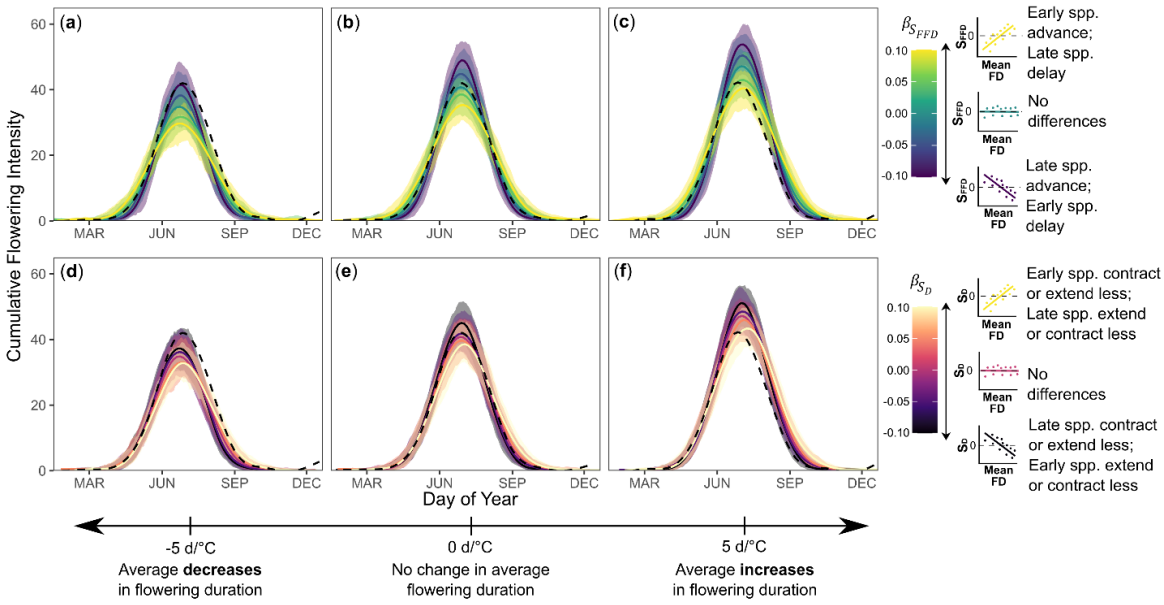

Figure S3—Variation in cumulative flowering intensity—the sum of the amplitude of species-level flowering curves on each day—throughout the season under 2 °C warming given independent variation among species in S_{FFD} and S_D. The black dashed line in all panels corresponds to the pre-warming time series of median cumulative flowering intensity across all simulated communities with the same $\underline{S_D}$ value. Dashed black lines in **a-c** correspond to the median post-warming cumulative flowering intensity among simulated communities with the same combination of $\beta_{S_{FFD}}$ and $\underline{S_D}$ values (and for which $\beta_{S_D} = 0$; see Figs. 1a,b). In turn, solid-colored lines in **d-f** correspond to the median post-warming cumulative flowering intensity among simulated communities with the same combination of β_{S_D} and $\underline{S_D}$ values (for which $\beta_{S_{FFD}} = 0$; see Figs. 1c,d). Shaded regions across panels indicate the 90% range of variation across communities grouped for each parameter combination.

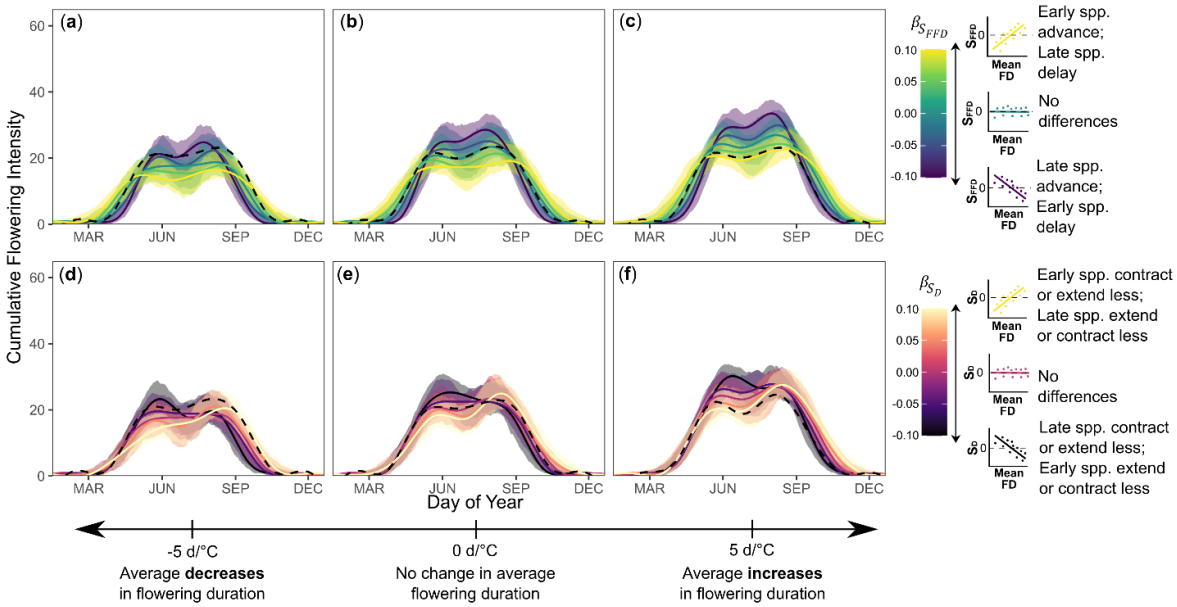

Figure S4—Variation in cumulative flowering intensity—the sum of the amplitude of species-level flowering curves on each day—throughout the season under 2 °C warming given independent variation among species in S_{FFD} and S_D within communities with two flowering peaks. The black dashed line in all panels corresponds to the pre-warming time series of median cumulative flowering intensity across all simulated communities with the same $\underline{S_D}$ value. Dashed black lines in **a-c** correspond to the median post-warming cumulative flowering intensity among simulated communities with the same combination of $\beta_{S_{FFD}}$ and $\underline{S_D}$ values (and for which $\beta_{S_D} = 0$; see Figs. 1a,b). In turn, solid-colored lines in **d-f** correspond to the median post-warming cumulative flowering intensity among simulated communities with the same combination of β_{S_D} and $\underline{S_D}$ values (for which $\beta_{S_{FFD}} = 0$; see Figs. 1c,d). Shaded regions across panels indicate the 90% range of variation across communities grouped for each parameter combination.

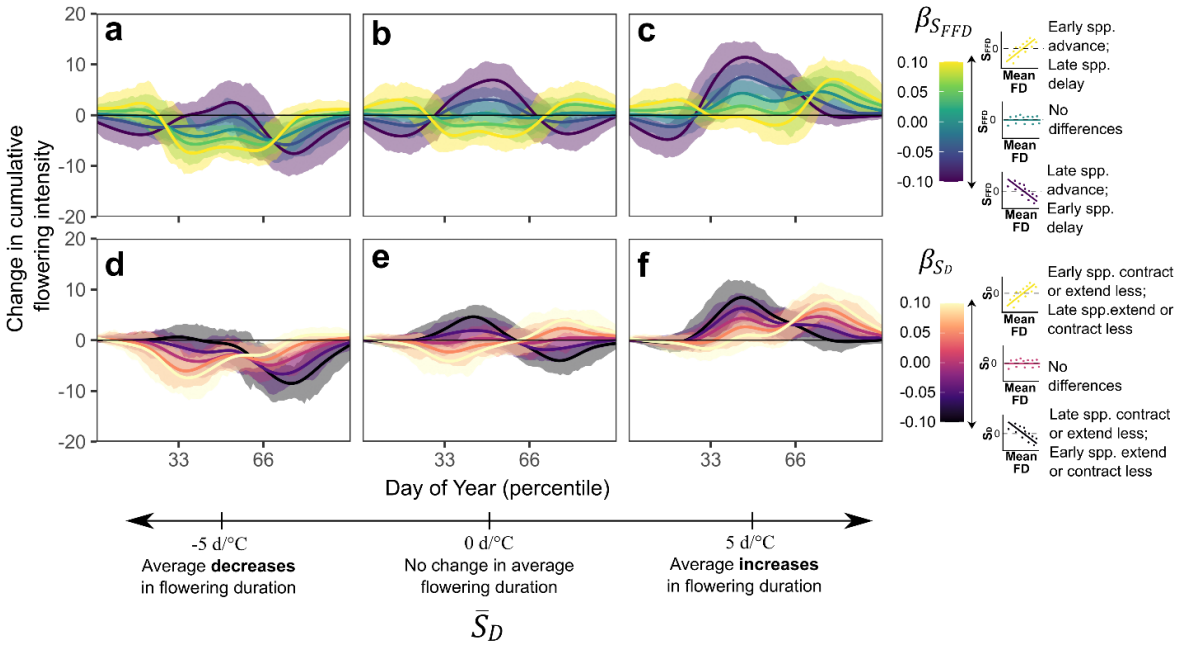

Figure S5— Changes in cumulative flowering intensity due to uncorrelated variation between S_{FFD} and S_D within communities with two flowering peaks. Solid-colored lines in **a-c** correspond to the median change in cumulative intensity—across percentiles of the flowering season—due to warming compared to the pre-warming baseline among simulated communities with the same combination of $\beta_{S_{FFD}}$ and $\bar{S}_D$ values (and for which $\beta_{S_D} = 0$; see Figs. 1a,b). Solid-colored lines in **d-f** correspond to the median change in cumulative intensity due to warming compared to the pre-warming baseline among simulated communities with the same combination of β_{S_D} and $\bar{S}_D$ values (and for which $\beta_{S_D} = 0$; see Figs. 1c,d). Shaded regions indicate the 90% range of variation across communities grouped by each combination of parameter values.

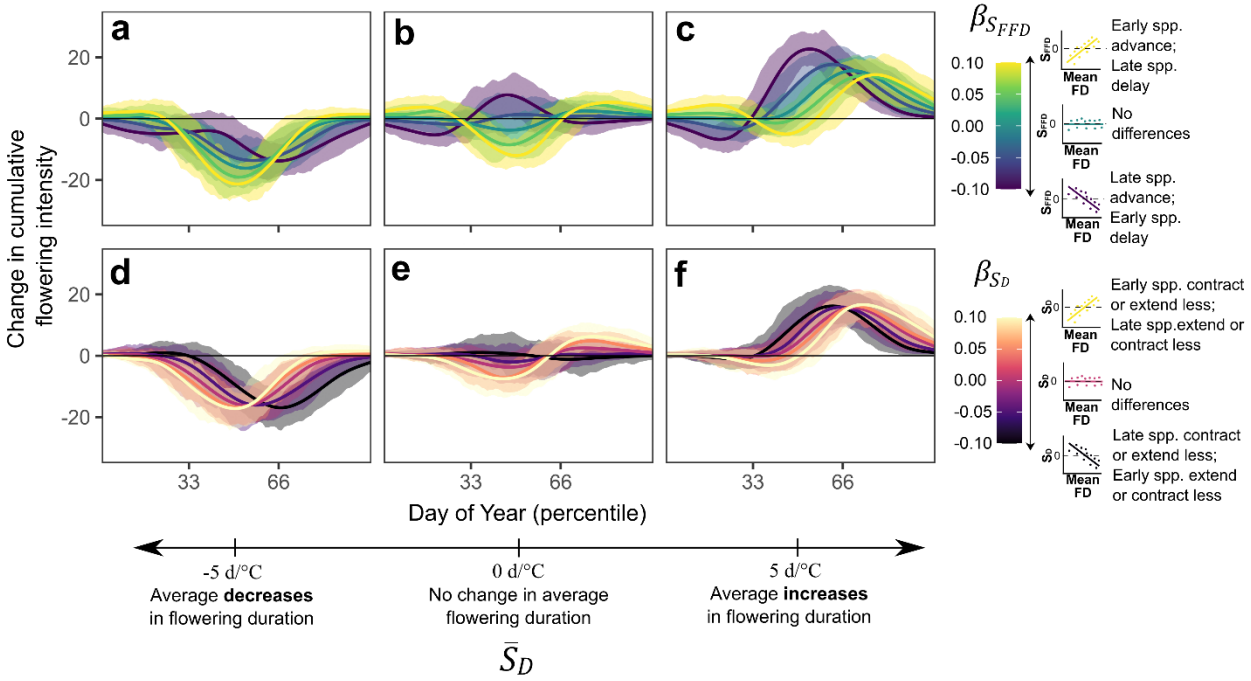

Figure S6— Changes in cumulative flowering intensity due to uncorrelated variation between S_{FFD} and S_D within communities with a single flowering peak under 4 °C of warming. Solid-colored lines in **a-c** correspond to the median change in cumulative intensity—across percentiles of the flowering season—due to warming compared to the pre-warming baseline among simulated communities with the same combination of $\beta_{S_{FFD}}$ and $\bar{S}_D$ values (and for which $\beta_{S_D} = 0$; see Figs. 1a,b). Solid-colored lines in **d-f** correspond to the median change in cumulative intensity due to warming compared to the pre-warming baseline among simulated communities with the same combination of β_{S_D} and $\bar{S}_D$ values (and for which $\beta_{S_D} = 0$; see Figs. 1c,d). Shaded regions indicate the 90% range of variation across communities grouped by each combination of parameter values.

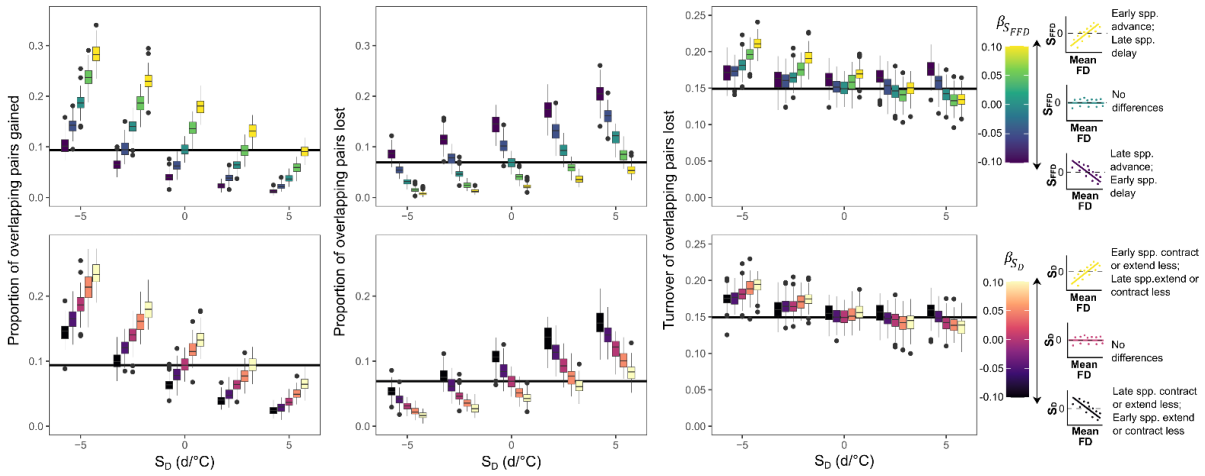

Figure S7—Change in the composition of flowering overlap among species within communities based on the number of new overlapping pairs (a, b), formerly overlapping pairs lost (c, d), and the turnover of overlapping species pairs based on gains and losses (e, f). In each community, the proportion of overlaps gained was measured as the number of post-warming overlapping pairs that did not overlap before warming, divided by the total number of overlapping pairs post warming. Species were considered to overlap if the intersection of their population-level flowering curves equaled 10% or more of the area under the curve for either species. Overlap losses were measured as the number of overlapping pairs that overlapped pre warming but did not post warming, divided by the total number of overlapping pairs pre warming. Overlap turnover was measured as the sum of the gains and losses in overlapping pairs after warming, divided by the total number of overlapping pairs observed pre and post warming. Colored boxplots in **a** and **b** depict the range of variation in turnover rates for communities varying in $\beta_{S_{FFD}}$ and β_{S_D}, respectively, with groups across the x-axis representing sets of communities with varying $\underline{S_D}$. In each panel, the horizontal solid black line indicates the degree of dissimilarity observed among communities showing no average changes in flowering duration (i.e., $\bar{S}_D = 0$), and random variation in both S_{FFD} and S_D among species throughout the season (i.e., $\beta_{S_{FFD}} = 0$ and $\beta_{S_D} = 0$)

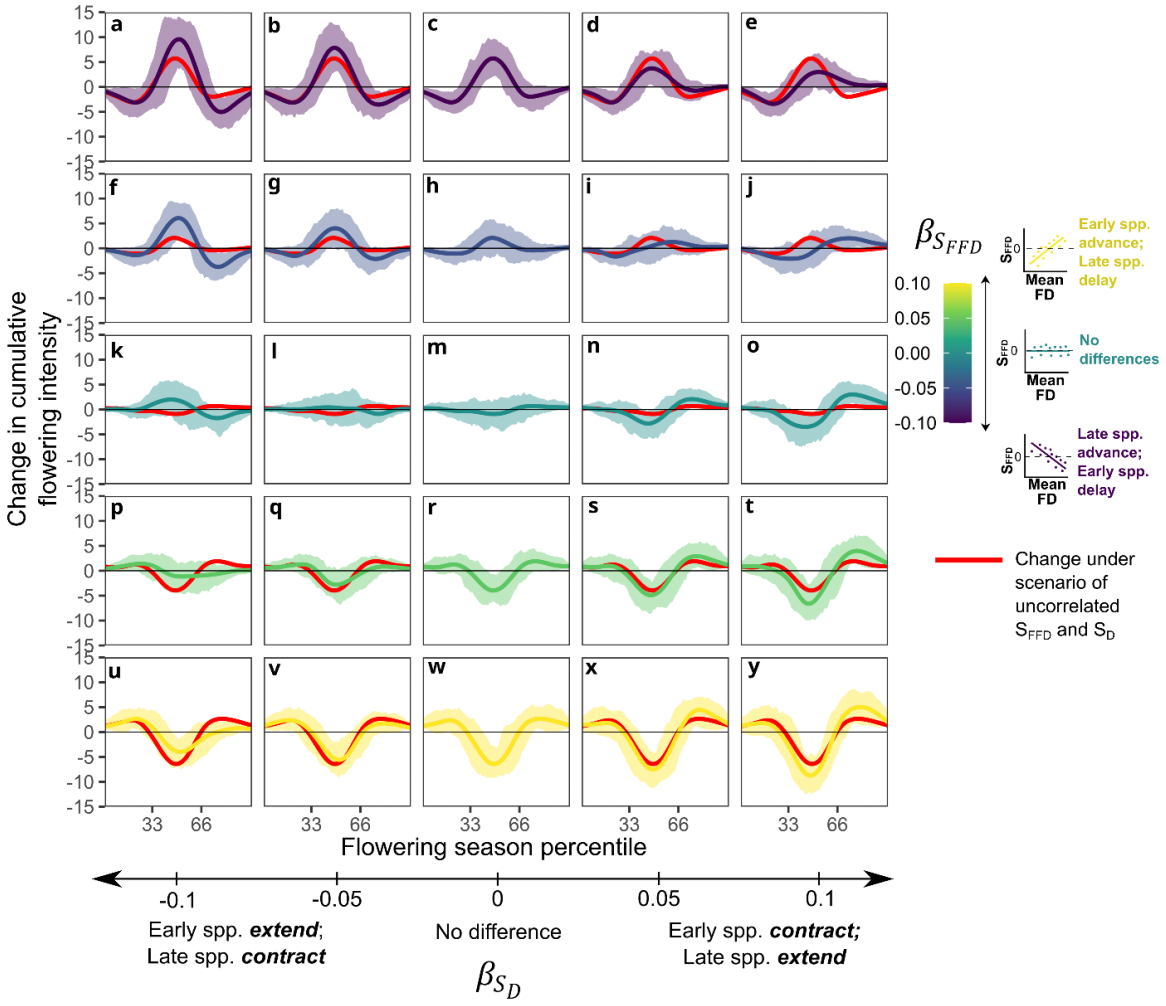

Figure S8—Warming-induced changes to cumulative flowering intensity for all combinations of $\beta_{S_{FFD}}$ and β_{S_D} across communities with a single flowering peak (with $\bar{S}_D = 0$). In each panel, colored lines (other than red) and shaded regions depict the median and 95% range of variation in cumulative flowering intensity change observed among communities sharing a combination of values $\beta_{S_{FFD}}$ and β_{S_D}. Red lines in each panel depict a reference scenario of uncorrelated S_{FFD} and S_D, with $\beta_{S_{FFD}}$ varying within columns but β_{S_D} coefficient equal to 0 for all panels. Panels **a**, **g**, **m**, **s**, and **y** show scenarios where S_{FFD} and S_D are positively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude and the same sign), corresponding to the panels in Figs. 7a-e. In turn, panels **e, i, m, q** and **u** show a scenario in which S_{FFD} and S_D are negatively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude but opposite sign), corresponding to Figs. 7f-j.

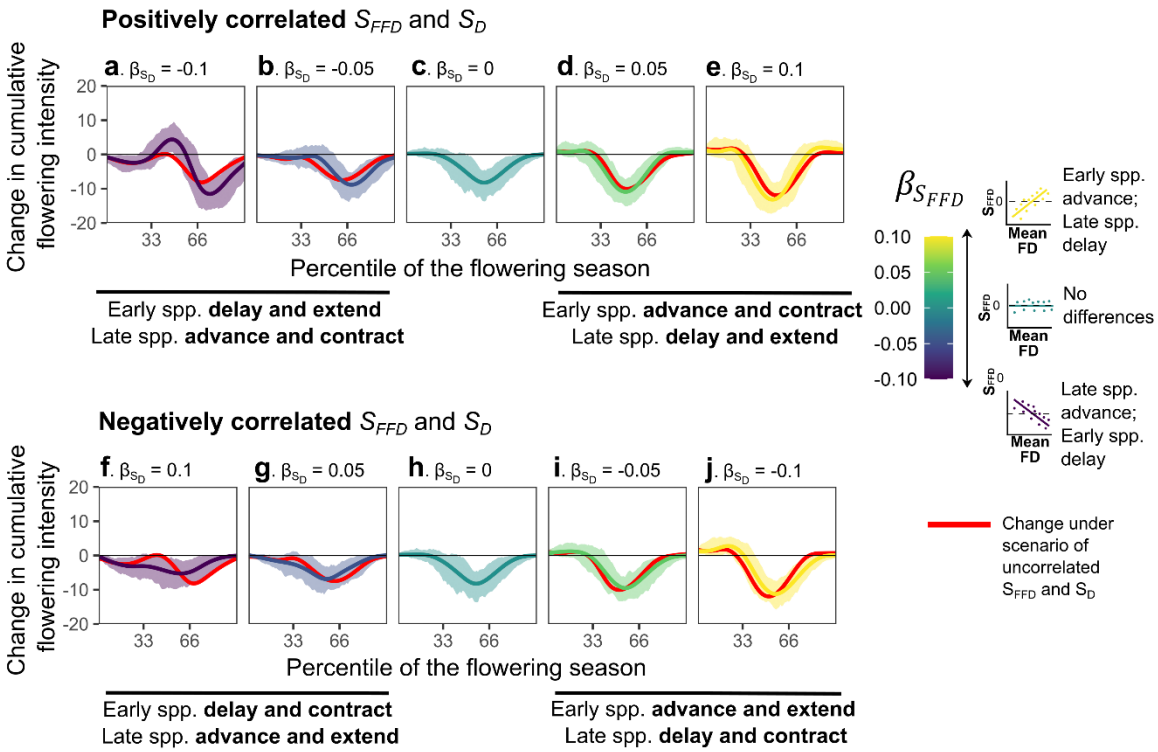

Figure S9—Warming-induced changes to cumulative flowering intensity due to correlated variation among species between S_{FFD} and S_D throughout the season (determined by $\beta_{S_{FFD}}$ and β_{S_D}) across communities with a single flowering peak ($\bar{S}_D = -5$ days per °C). In each panel, colored lines (other than red) and shaded regions depict the median and 90% range of variation in cumulative flowering intensity change observed among communities sharing a combination of values $\beta_{S_{FFD}}$ and β_{S_D}. Red lines in each panel depict a reference scenario of uncorrelated S_{FFD} and S_D, with $\beta_{S_{FFD}}$ varying within columns but β_{S_D} coefficient equal to 0 for all panels. Panels **a-e** show scenarios where S_{FFD} and S_D are positively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude and the same sign). In turn, panels **f-j** show a scenario in which S_{FFD} and S_D are negatively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude but opposite sign). $\bar{S}_D$ was set to 0 days per °C.

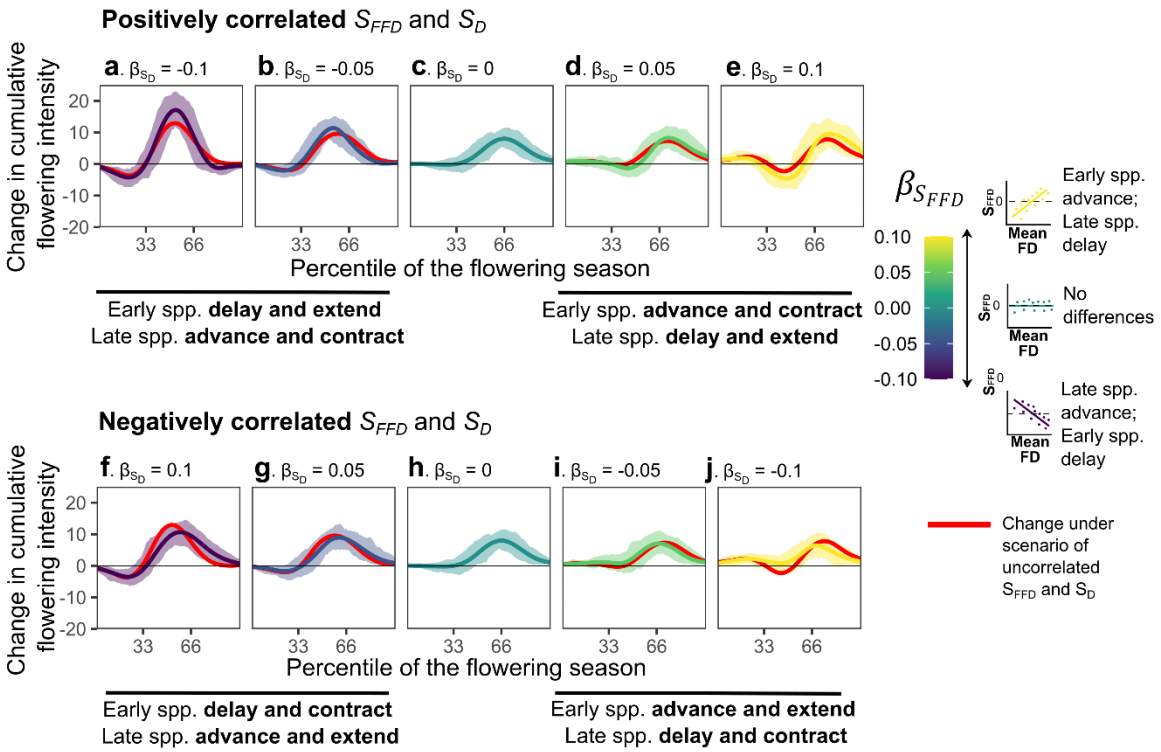

Figure S10—Warming-induced changes to cumulative flowering intensity due to correlated variation among species between S_{FFD} and S_D throughout the season (determined by $\beta_{S_{FFD}}$ and β_{S_D}) across communities with a single flowering peak ($\bar{S}_D = 5$ days per °C). In each panel, colored lines (other than red) and shaded regions depict the median and 95% range of variation in cumulative flowering intensity change observed among communities sharing a combination of values $\beta_{S_{FFD}}$ and β_{S_D}. Red lines in each panel depict a reference scenario of uncorrelated S_{FFD} and S_D, with $\beta_{S_{FFD}}$ varying within columns but β_{S_D} coefficient equal to 0 for all panels. Panels **a-e** show scenarios where S_{FFD} and S_D are positively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude and the same sign). In turn, panels **f-j** show a scenario in which S_{FFD} and S_D are negatively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude but opposite sign). $\bar{S}_D$ was set to 0 days per °C.

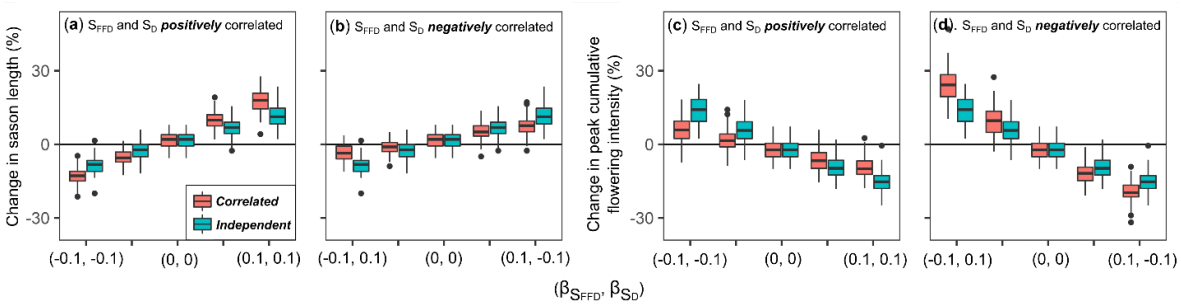

Figure S11— Change in flowering season length (**a**, **b**) and peak flowering intensity (**c**, **d**) for scenarios of independent vs. correlated S_{FFD} and S_D (blue boxes and red boxes, respectively) among sequentially flowering species throughout the season. For communities with correlated S_{FFD} and S_D, results are presented separately for those showing (**a, c**) positively correlated S_{FFD} and S_D throughout the season (Scenario 2) and (**b, d**) negatively correlated S_{FFD} and S_D throughout the season (Scenario 3). For communities with correlated S_{FFD} and S_D in (**a, c**), β_{S_D} increases from -0.1 to 0.1 along the x-axis, generating communities for whom $\beta_{S_{FFD}}$ and β_{S_D} had the same direction and magnitude. In (**b, d**), β_{S_D} decreases from 0.1 to -0.1 along the x-axis, resulting in communities for which $\beta_{S_{FFD}}$ and β_{S_D} had the same magnitude but opposite direction. In contrast, blue boxes depict reference scenarios of uncorrelated S_{FFD} and S_D, with $\beta_{S_{FFD}}$ varying along the x-axis, β_{S_D} coefficient equal to 0, and $\bar{S}_D = 0$. Results are shown only for communities showing a single flowering peak.

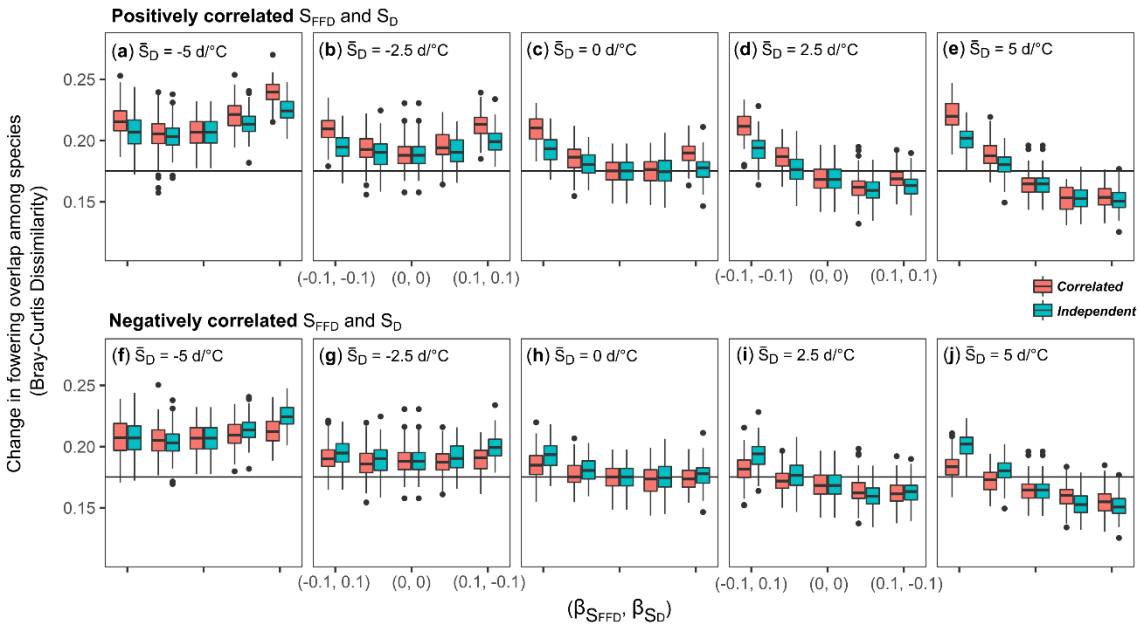

Figure S12— Changes in the composition of pairwise flowering overlaps among species under warming for scenarios of independent vs. correlated S_{FFD} and S_D (blue boxes and red boxes, respectively) among sequentially flowering species throughout the season. In each community, change in the composition of flowering overlaps was measured using the Bray-Curtis dissimilarity index, with values of 0 corresponding communities with the same identity and degree of pairwise species overlaps, and values of 1 indicating complete dissimilarity in the identity of its pairwise overlap pre- and post-warming communities. Panels **a-e** show scenarios where S_{FFD} and S_D are positively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude and the same sign). In turn, panels **f-j** show a scenario in which S_{FFD} and S_D are negatively correlated among species throughout the season (i.e., $\beta_{S_{FFD}}$ and β_{S_D} have the same magnitude but opposite sign). Panels from left to write (i.e., **a** to **e** and **f** to **j**) represent increasing values of the mean sensitivity of flowering duration in a community ($\bar{S}_D$), from average decreases to average increases in flowering duration among species. In each panel, the horizontal solid black line indicates the degree of dissimilarity observed among communities showing no average changes in flowering duration (i.e., $\bar{S}_D = 0$), and random variation in both S_{FFD} and S_D among species throughout the season (i.e., $\beta_{S_{FFD}} = 0$ and $\beta_{S_D} = 0$).

IX. Appendix 4—Supplementary materials for V

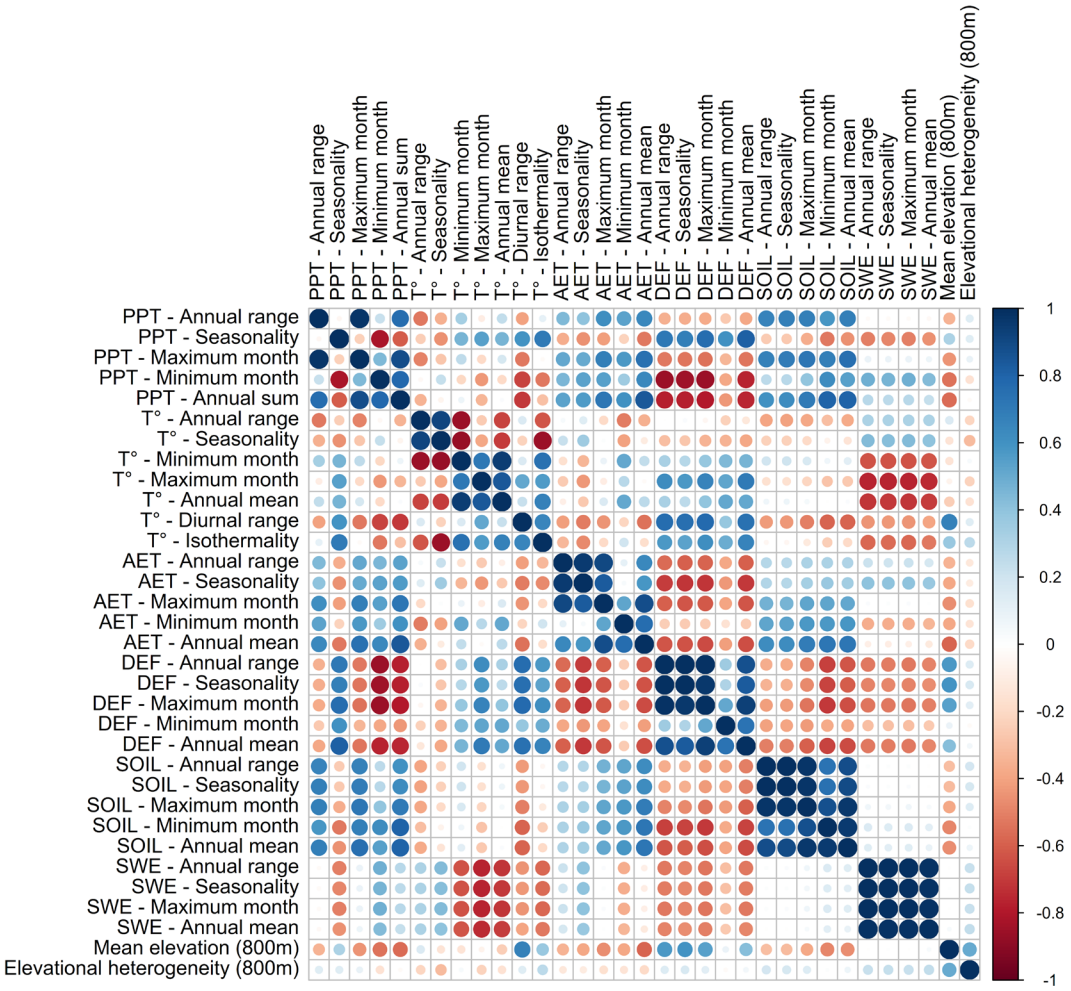

Figure S1—Correlations among 2004-2023 normals for 31 climatic variables, mean elevation, and elevation heterogeneity across 4km resolution grid cells throughout the conterminous United States. Climate variables include annual sums or means, maximum monthly values, minimum monthly values, annual monthly range, and seasonality for precipitation (PPT), temperature (T°), actual evapotransporation (AET), climate water deficit (DEF), soil moisture (SOIL), and snow-water equivalent (SWE). Variables for T° also include approximate mean daily range, and isothermality. Seasonality for precipitation was calculated proportionally to the mean cumulative annual precipitation in each site.

	PC1 (1.9, 43%)	PC2 (1.7, 43%)	PC3 (1.3, 9%)	PC4 (1.2, 6%)	PC5 (1.1, 5%)
PPT - Annual range	-0.15	-0.21	0.14	-0.05	**0.22**
PPT - Seasonality	0.20	-0.10	0.10	-0.11	0.18
PPT - Monthly maximum	-0.19	-0.18	0.10	-0.09	0.16
PPT - Monthly minimum	-0.22	0.02	-0.11	-0.17	-0.20
PPT - Annual sum	**-0.23**	-0.12	0.01	-0.10	-0.03
T° - Annual range	0.01	**0.27**	-0.20	**0.29**	0.08
T° - Seasonality	-0.07	**0.26**	**-0.25**	0.23	0.00
T° - Monthly minimum	0.08	**-0.30**	0.04	-0.22	-0.07
T° - Monthly maximum	0.16	-0.20	-0.20	-0.02	-0.03
T° - Annual mean	0.09	**-0.29**	-0.12	-0.22	-0.05
T° - Diurnal range	0.21	-0.01	0.04	0.10	0.19
T° - Isothermality	0.16	-0.21	0.18	-0.16	0.12
AET - Annual range	-0.17	-0.02	-0.19	-0.11	**0.49**
AET - Seasonality	-0.20	0.04	-0.15	-0.12	**0.42**
AET - Monthly maximum	-0.19	-0.13	-0.19	-0.07	**0.35**
AET- Monthly minimum	-0.09	**-0.26**	-0.06	0.06	-0.15
AET - Annual mean	-0.20	-0.19	-0.16	0.02	0.07
DEF - Annual range	**0.24**	-0.04	0.08	0.15	0.11
DEF - Seasonality	**0.24**	-0.03	0.10	0.19	0.09
DEF - Monthly maximum	**0.25**	-0.05	0.07	0.06	0.07
DEF- Monthly minimum	0.16	-0.06	-0.02	**-0.38**	-0.16
DEF - Annual mean	**0.25**	-0.06	0.03	-0.11	-0.02
SOIL - Annual range	-0.16	-0.18	0.21	**0.27**	0.01
SOIL - Seasonality	-0.16	-0.18	**0.21**	**0.26**	-0.01
SOIL - Monthly maximum	-0.19	-0.17	0.18	**0.23**	-0.06
SOIL - Monthly minimum	-0.21	-0.13	0.08	0.11	-0.19
SOIL - Annual mean	-0.21	-0.16	0.14	0.16	-0.15
SWE - Annual range	-0.14	0.24	0.21	-0.19	-0.02
SWE - Seasonality	-0.14	0.24	0.21	-0.20	-0.02

SWE - Monthly maximum	-0.14	0.24	0.21	-0.19	-0.02
SWE – Annual mean	-0.13	0.24	**0.22**	-0.21	-0.02
Mean Elevation (800m)	0.14	0.10	**0.28**	0.13	**0.25**
Elevational heterogeneity	0.01	0.01	**0.46**	-0.02	0.19

Table S1—Loadings from a principal component analysis (PCA) of 2004-2023 normals for 31 climatic variables, mean elevation, and elevation heterogeneity across 4km resolution grid cells throughout the conterminous United States. Loadings are reported for the 5 principal components (PCs) explaining more variance than any input variable in the data. Highlighted values in each column correspond to the 5 input variables with the highest loadings on each PC. Values next to each PC's name indicate its eigenvalue and variance explained

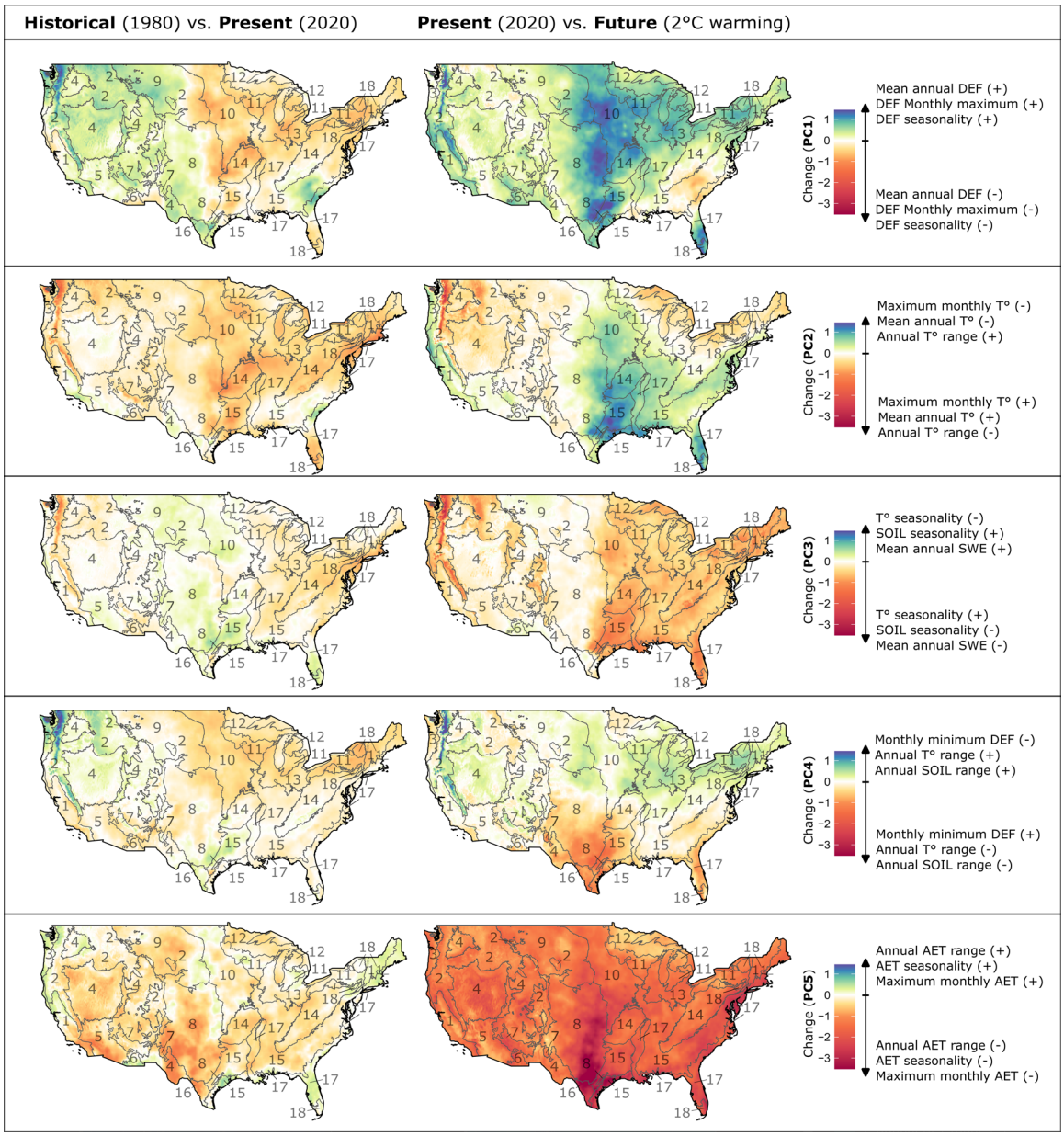

Figure S2—Climate change between the historical period (1960-1980) and the present period (2001-2020), and between the present period and projected conditions under a scenario of 2°C warming above pre-industrial levels. Climate change is shown as the difference in the 5 principal components summarizing 31 climatic variables, as well as mean elevation and elevational heterogeneity within 4km grid cells throughout the conterminous United States (see '*Methods*' section of the main text) between periods. The variables listed in each legend correspond to those with the greatest loadings for each PC. Positive and negative signs next to each variable indicate whether positive or negative values in the color scale are associated to increases or decreases between periods. Subdivisions labeled 1-18 represent level II ecoregions. *1) Mediterranean California, 2) Western Cordillera, 3) Marine West Coast Forest, 4) Cold Deserts, 5) Warm Deserts, 6) Western Sierra Madre Piedmont, 7) Upper Gila Mountains, 8) South-Central Semi-arid Prairies, 9) West-Central Semi-arid Prairies, 10) Temperate Prairies, 11) Mixed Wood Plains, 12) Mixed Wood Shield, 13)*

Central USA Plains, 14) Ozark, Ouachita-Appalachian Forests, 15) Southeastern USA Plains, 16) Tamaulipas-Texas Semi-arid Plains, 17) Mississippi Alluvial and Southeast USA Coastal Plain, 18) Atlantic Highlands.

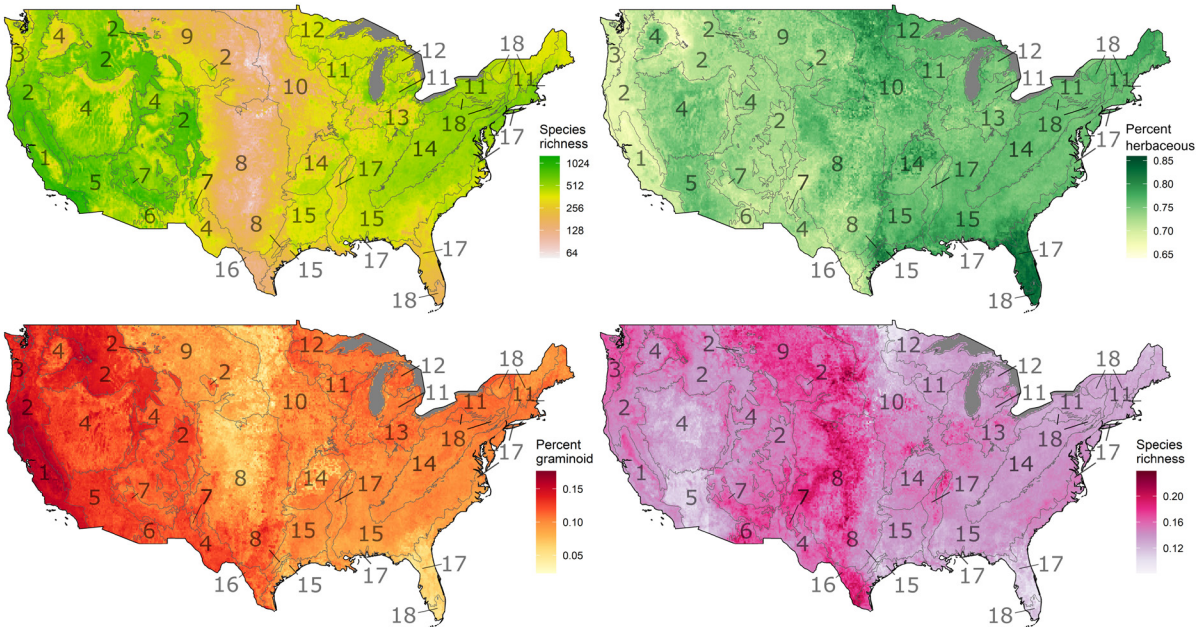

Figure S3—Predicted species richness (from a total of 2,837 species) within 12km resolution grid cells across the conterminous United States estimated for the 2001-2020 period, and the percent that have herbaceous, graminoid, or woody growth habit. Subdivisions labeled 1-18 represent level II ecoregions. *1) Mediterranean California, 2) Western Cordillera, 3) Marine West Coast Forest, 4) Cold Deserts, 5) Warm Deserts, 6) Western Sierra Madre Piedmont, 7) Upper Gila Mountains, 8) South-Central Semi-arid Prairies, 9) West-Central Semi-arid Prairies, 10) Temperate Prairies, 11) Mixed Wood Plains, 12) Mixed Wood Shield, 13) Central USA Plains, 14) Ozark, Ouachita-Appalachian Forests, 15) Southeastern USA Plains, 16) Tamaulipas-Texas Semi-arid Plains, 17) Mississippi Alluvial and Southeast USA Coastal Plain, 18) Atlantic Highlands*.

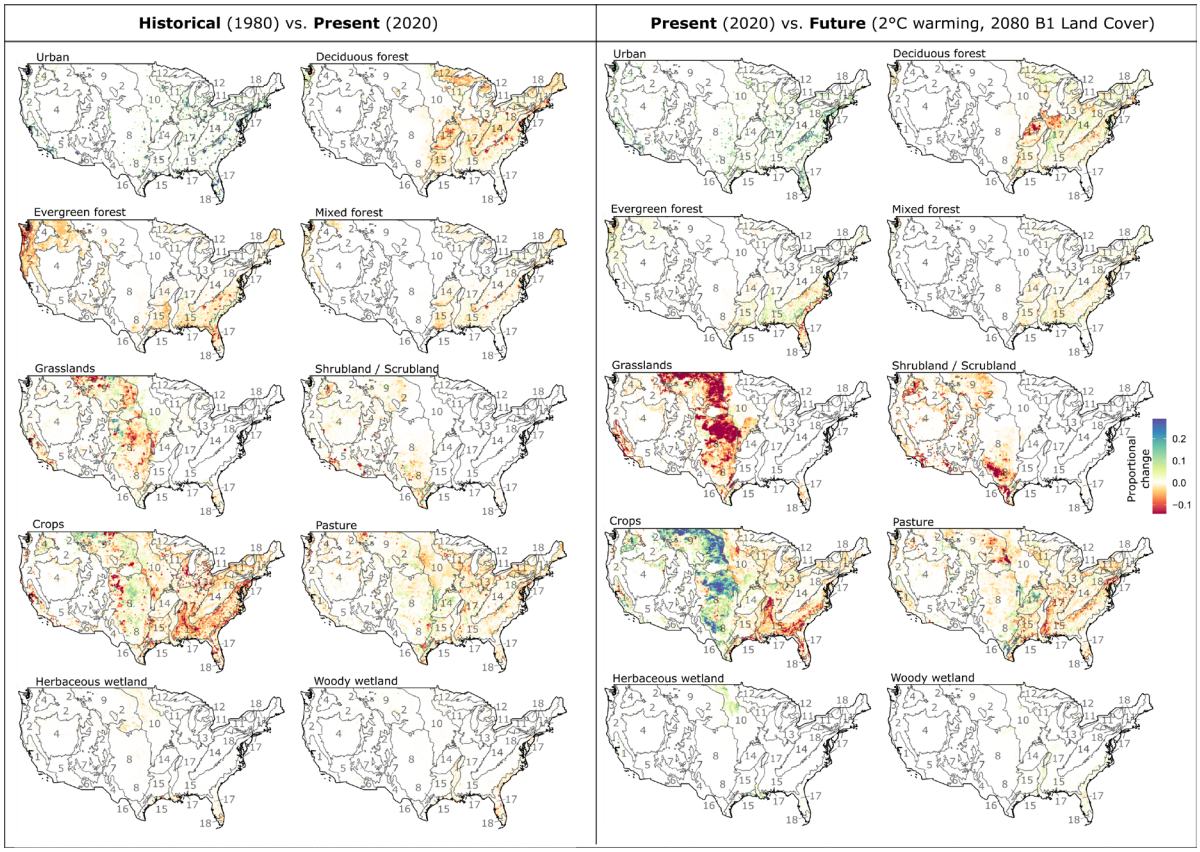

Figure S4—Change in land cover between 1980 and 2020, and between 2020 and 2080 under SRES B1 scenario of land use and land cover change. Each panel shows changes in the proportion of 250m cells of the focal class found within each 750m resolution grid cell across the conterminous United States (CONUS). Subdivisions labeled 1-18 represent level II ecoregions. *1) Mediterranean California, 2) Western Cordillera, 3) Marine West Coast Forest, 4) Cold Deserts, 5) Warm Deserts, 6) Western Sierra Madre Piedmont, 7) Upper Gila Mountains, 8) South-Central Semi-arid Prairies, 9) West-Central Semi-arid Prairies, 10) Temperate Prairies, 11) Mixed Wood Plains, 12) Mixed Wood Shield, 13) Central USA Plains, 14) Ozark, Ouachita-Appalachian Forests, 15) Southeastern USA Plains, 16) Tamaulipas-Texas Semi-arid Plains, 17) Mississippi Alluvial and Southeast USA Coastal Plain, 18) Atlantic Highlands*.