Supplement: Reconciling competing hypotheses regarding flower-leaf sequences in temperate forests for fundamental and global change biology

February 24, 2020

Methods

Climate Change and FLS:

To evaluate how FLS patterns have changed over time in association with climate change we obtained phenological data for four European woody plant species with long term phenology records of both flower (BBCH 60) and leafout phenology (BBCH 11) from the Pan European Phenological Database (Templ *et al.*, 2018). We restricted the data set to include only stations with more than 50 years worth of data. Following conventions for modeling effects of climate change, we modeled the number of days between flowering and leafing as a function of time for each species, using a hinge model with 1980 as a break point (Stocker *et al.*, 2013; Kharouba *et al.*, 2018). For each species, we display the pre-1980 mean and 95% credible intervals of the time between flowering and leafing and the post-1980 change in mean time between phenophases that can be driven by climate change.

Defining FLS with MTSV and USFS data

For these two, categorical, species level-case studies, we converted verbal descriptions of flower-leaf sequences into a binary response variable. For our more inclusive "functional" definition of hysteranthy, which allows for some overlap between floral and vegetative phenophases, we included species entries with descriptions *"flowers before the leaves"*, *"flowers before or with leaves"* and *"flowers with leaves"* as hysteranthous. Our more restrictive "physiological" hysteranthy definition only included species described as *"flowers before the leaves"* as hysteranthous.

For modeling trait associations with FLS, we chose three predictors to represent the three major FLS hypotheses; pollination syndrome, average flowering time and minimum precipitation levels across the species range. We obtained pollination syndrome and average flowering time information directly from the respective data sources and estimates of minimum precipitation across range from the USDA/NRCS Conservation Plants Characteristics database (USDA/NRCS, 2020). We coded pollination syndrome as biotic- or wind-pollinated, and assigned known ambophilous species in the genus *Salix* as biotic-pollinated. We re-coded flowering time as the average of the range of months of flowering reported in each data source.

For these case studies, we modeled associations between hysteranthy and the trait predictors with all two-way interactions with logistical regressions in phylogenetic generalized linear modeling framework (Ives & Garland, 2010) using the R package "phylolm" (Ho & Ane, 2014). Our models incorporated a published angiosperm phylogenetic tree (Zanne *et al.*, 2013) pruned to match the species list for each case study. Species found in the trait data set but not in the original phylogenetic tree were added to the pruned tree at the generic root. In total 32 species were added to the generic roots for the MTSV data

set and eight for the USFS data set. We visualize phylogenetic patterning of FLS across the tree of each case study (Fig. S4).

We ran the models with 599 bootstrapped re-sampling iterations for each data set (Wilcox, 2010). We standardized all predictors by subtracting the mean and dividing by two standard deviations to allow for a reasonable comparison of effect sizes between the binary and continuous predictors in this model (Gellman & Hill, 2997).

Harvard Forest models

For each individual tree per year in the data, we calculated the time between flowers opening and leaves reaching 75% of their final size. Positive FLS values indicate flowering-first and negative values leafing-first. To compare the inference between categorical and continuous measure of FLS, we re-coded the continuous FLS measures as binary responses with positive values coded as hysteranthous and negative values as seranthous. These models used the same predictors as the MTSV and USFS datasets (flowering time, pollination syndrome, minimum precipitation across species' range and all two way interactions between predictors), except that we estimated flowering time directly from the HF data. For the categorical model, we used a logistic regression with a binary outcome variable. The phylogeny and non-phylogenetic species effects were each modeled as separate, random intercepts in the model. For the continuous model, we used a linear mixed model. Here we included an effect of individual tree as random intercept nested within species in addition to the random intercept of phylogeny. We ran all models with the R package "brms" (Bürkner, 2018) to estimate the relationship between FLS and the predictors with a phylogenetic mixed model in a Bayesian framework (de Villemeruil P. Nakagawa, 2014). We ran 4 chains with 4000 iterations and a warm-up of 3000 iterations each, resulting in 4000 total sampling iterations. All models used weakly informative priors on the intercept and error terms. Increasing priors three-fold did not impact the model estimates. As our primary goal was to directly compare the effects each predictor, we standardized these variables to allow for a reasonable comparison between them (Gellman & Hill, 2997).

We calculated the marginal effects for the Harvard Forest continuous model using the R-package "ggeffects" (Daniel Lüdecke, 2018). Figure 3 shows the water dynamics effect of FLS give a flowering day of May 1. This relationship trends between pollination syndrome and minimum precipitation remained across flowering dates from mid April-June.

Though we make broad comparisons between the HF and MTSV/USFS case studies, differences in data structure between the datasets required us to use alternative modeling frameworks. The MTSV and USFS data provide one response variable for each species while the HF data contains intra-specific differences in FLS, providing several different response values per species. The current phylogenetic generalized linear model framework can only fit models with one response value per species, while the phylogenetic mixed model in brms over-fits models with this kind of data structure (Paul Burkner, personal communication) and performs better on multi-response per species datasets like HF. We ran both model types on each case study and while they do yield different absolute estimates, the patterns we found were consistent across each framework, and we report results from the most accurate model for each dataset.

Phylogenetic signals

For all categorical specifications of FLS (MTSV, USFS and HF), we assessed the phylogenetic structure of hysteranthous flowering in all with Fritz's D-statistic (Fritz & Purvis, 2010) using the "Caper" package (Orme *et al.*, 2013) in R. Fritz's D calculates the sum of changes in estimated node values of a binary trait along edges in a phylogeny and compares this observed value to both a model of Phylogenetic randomness

and Brownian threshold model. The means of the two data simulations scale values of D to set points of 0 (as phylogenetically conserved) and 1 (random) (Orme *et al.*, 2013).

For the quantitative Harvard Forest model, we estimated the phylogenetic signal for FLS (lambda) based covarience of residuals among species. Estimated phylogenetic signals from all case studies are reported below (Fig. S3). Add citation for ggtree.
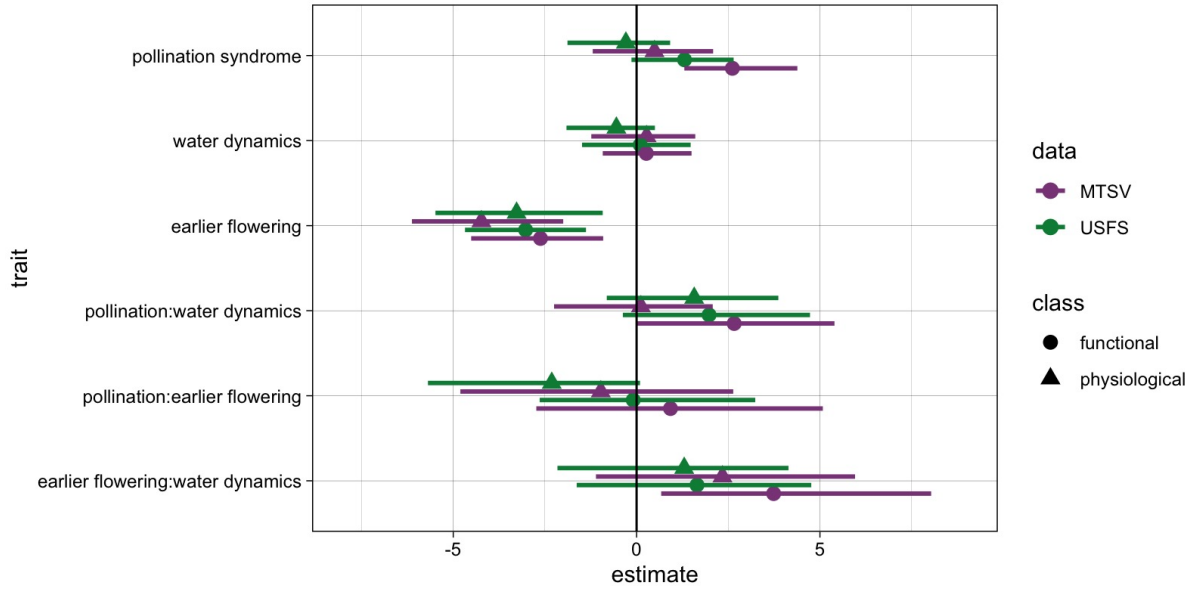
Supplemental Tables and Figures

Figure S1: **Mean estimates of the effects of FLS predictors on the likelihood a species is hysteranthous vary across datasets and definitions of FLS.** We used phylogenetic adjustments and standardized units to make a basic comparison of two datasets and classes (physiological= no overlap between flowering and leafing, functional= moderate overlap) of FLS. While there is some agreement across models (strong effects of flowering time, no consistent effect interactions between predictors), the effect of other predictors (pollination syndrome, water dynamics) were highly sensitive to how data were defined, biasing any inference from models and compromising the ability to validate the existing FLS hypotheses. Lines represent 95% bootstrap intervals.
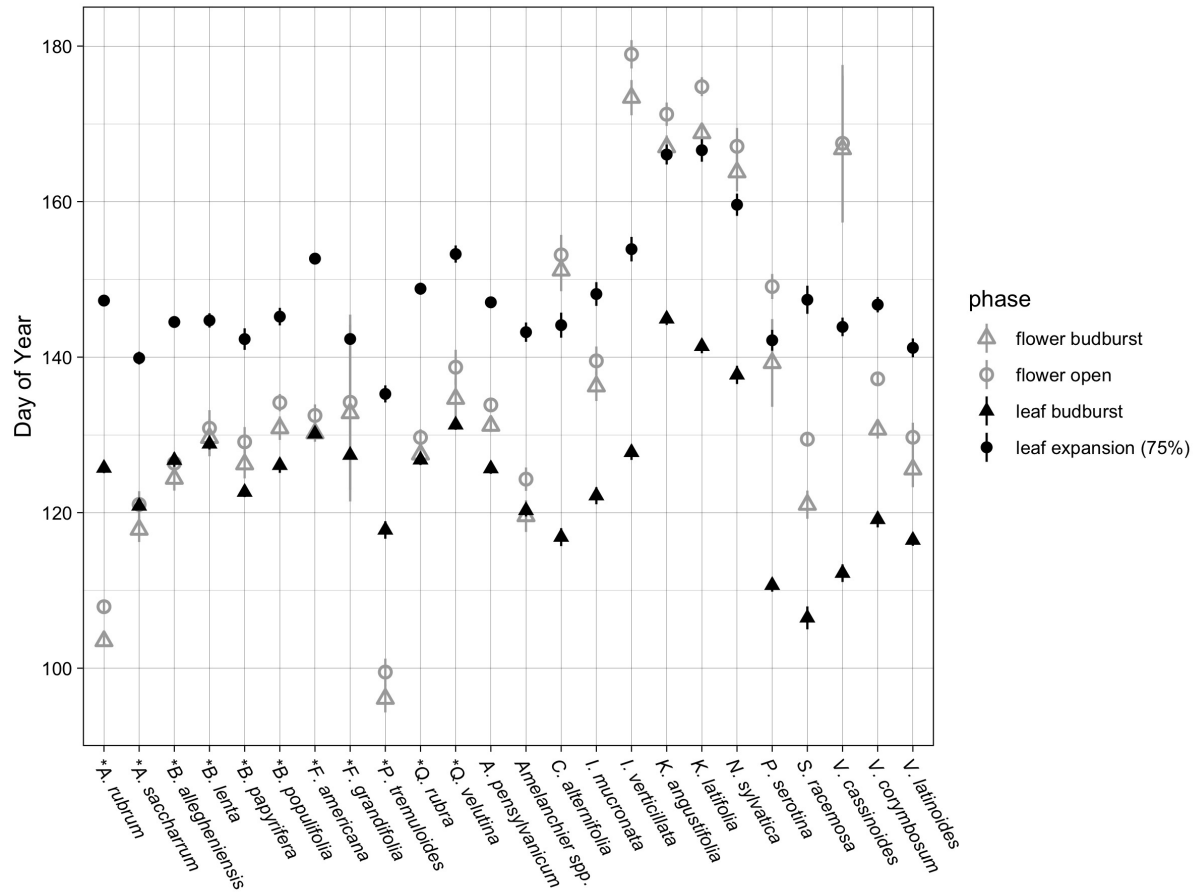
Figure S2: **Quantitative FLS patterns for woody plants at Harvard Forest in Pertersham, MA.** Because phenological sequences consist of several sub-stages if is difficult to unambiguously categorize many species into the currently FLS categories.
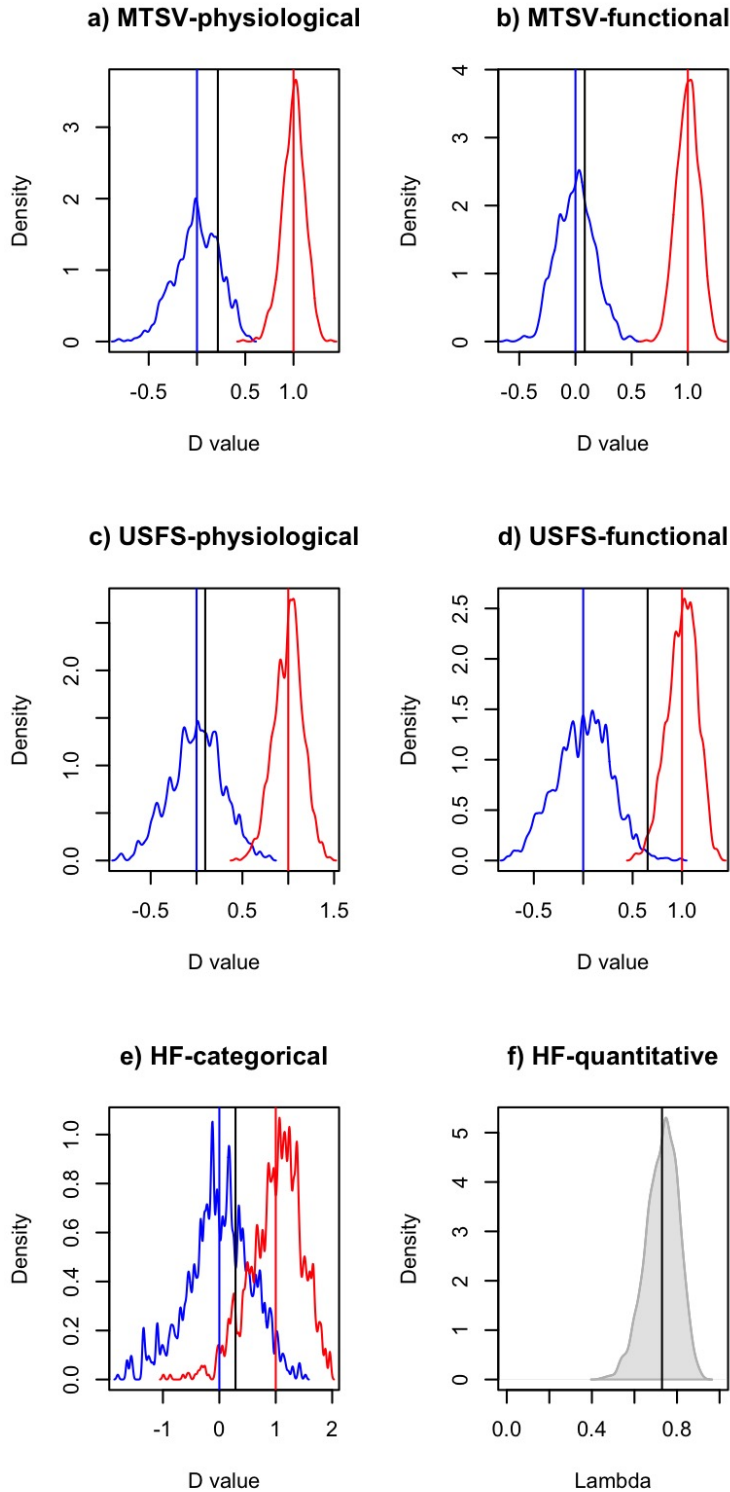
Figure S3: **The phylogenetic signal for FLS varies between datasets, and is sensitive to how FLS patterns are categorized.** In a)-e),the black vertical line show the the Fritz's D statistic for binary classifications of FLS estimated from the data, with blue and red lines representing expected D values based on simulations under Brownian threshold model and random model respectively. Panel f) shows the the estimated λ values of FLS from the the continuous modeling framework. Higher values indicate stronger phylogenetic structure.

5
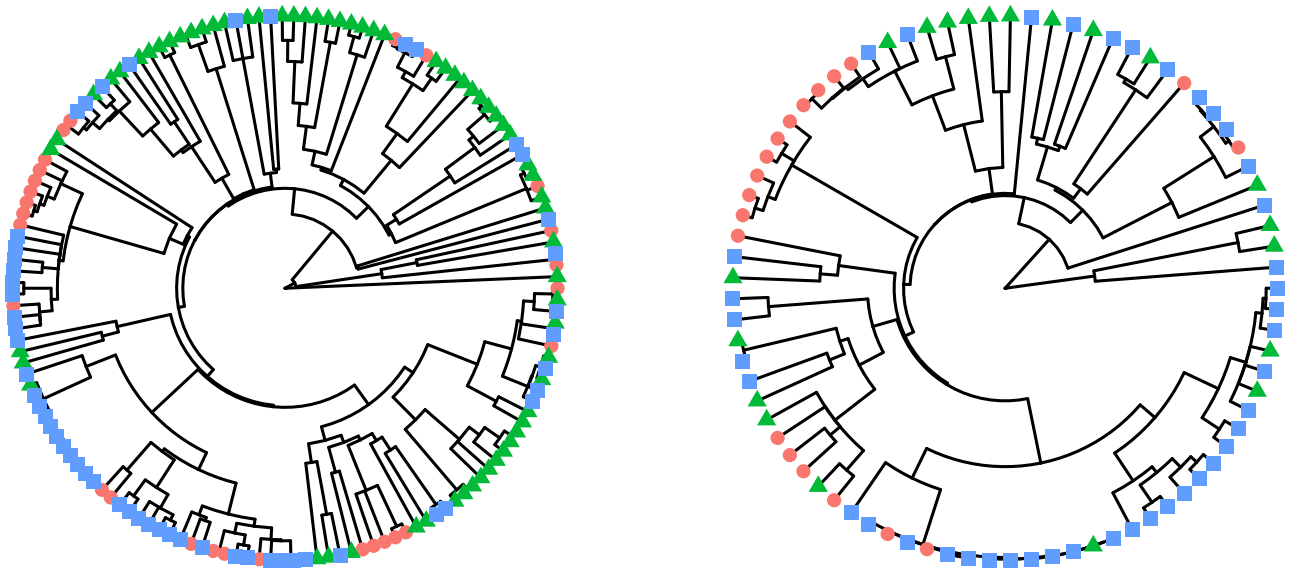
hysteranthy ● always hysteranthous ▲ always seranthous ■ transitional

Figure S4: **Phylogenetic structure of FLS in MTSV and USFS varies significantly depending on how FLSs are defined.** Many species get re-assigned to either hysteranthy or seranthy depending on whether FLS is defined functionally (partial overlap between flowering and leafing allowed) or physiologically (no overlap between flowering and leafing allowed) (blue squares). This modeling choice dramatically alters FLS patterning across the tree, resulting in an unstable phylogentic signal for this trait.

References

Boose E, G.E. (2004) Shaler meteorological station at harvard forest 1964-2002. Tech. rep., Harvard Forest Data Archive: HF000.

Bürkner, P.C. (2018) Advanced bayesian multilevel modeling with the r package brms. *R Journal* **10**, 395–411.

Daniel Lüdecke (2018) ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software* **3**, 772.

de Villemeruil P. Nakagawa, S. (2014) *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*, chap. General quantitative genetic methods for comparative biology, pp. pp. 287–303. Springer, New York.

Fritz, S.A. & Purvis, A. (2010) Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits **24**, 1042–1051.

Gellman, A. & Hill, J. (2997) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Ho, L.S.T. & Ane, C. (2014) A linear-time algorithm for gaussian and non-gaussian trait evolution models. *Systematic Biology* **63**, 397–408.

Ives, A.R. & Garland, Jr., T. (2010) Phylogenetic logistic regression for binary dependent variables. *Systematic Biology* **59**, 9–26.

Kharouba, H.M., Ehrlén, J., Gelman, A., Bolmgren, K., Allen, J.M., Travers, S.E. & Wolkovich, E.M. (2018) Global shifts in the phenological synchrony of species interactions over recent decades. *Proceedings of the National Academy of Sciences* **115**, 5211.

Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N. & Pearse, W. (2013) *caper: Comparative Analyses of Phylogenetics and Evolution in R*.

Stocker, T., Qin, D., Plattner, G.K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V. & Midgley, P. (2013) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Cambridge, United Kingdom and New York, NY.

Templ, B., Koch, E., K.Bolmgren, Ungersböck, M., Paul, A., Scheifinger, H. & et al. (2018) Pan european phenological database (pep725): a single point of access for european data. *Int. J. Biometeorology* .

USDA/NRCS (2020) Conservation plants characteristics.

Wilcox, R.R. (2010) *Fundamentals of modern statistical methods: Substantially improving power and accuracy.* Springer.

Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., FitzJohn, R.G., McGlinn, D.J., O'Meara, B.C., Moles, A.T., Reich, P.B. & et al. (2013) Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92.