PNAS

# Supporting Information for

## Diverse Genotype-by-Weather Interactions in Switchgrass

**Alice H. MacQueen, Li Zhang, Samuel A. Smith, Jason Bonnette, Arvid R. Boe, Phillip A. Fay, Felix B. Fritschi, David B. Lowry, Robert B. Mitchell, Francis M. Rouquette Jr, Yanqi Wu, Arbel Harpak, and Thomas E. Juenger**

**Thomas E. Juenger**
**E-mail: tjuenger@utexas.edu**

### This PDF file includes:



### Other supporting materials for this manuscript include the following:

Datasets S1 to S6

This PNAS template for Supporting Information (SI) may be used to organize your supporting material. The template is intended to provide a clearly organized PDF file that will ensure readers can easily navigate to sections or specific figures and tables. Movie files or large datasets can be presented as separate files. Further information is available in the PNAS Author Center.

**Using the template.** Specify the title, author list, and corresponding authors with the same YAML specification as in the corresponding PNAS Quarto template. The cover page will be automatically generated with the relevant description of the SI.

Figures should be placed on separate pages with legends set immediately below each figure. Table titles should be set immediately above each table. Note that tables extending beyond the width of the page can be included in the PDF or provided as separate dataset files. Oversized/nonstandard page sizes are accepted as part of your SI Appendix file.

References cited in the SI text should be included in a separate reference list at the end of this SI file: Varga and Edmonds (2016) and Olsen and Stensland (1992).

Supporting information for Brief Reports is limited to extended methods, essential supporting datasets, and videos (no additional tables or figures). Supporting figures and tables are not allowed for Brief Reports.

**Submitting SI.** Save your completed SI file as a PDF for submission. Further submission instructions are available here.

**Subhead.** Type or paste text here. This should be additional explanatory text such as an extended technical description of results, full details of mathematical models, etc.

## Supplementary Methods

**Section S1. Covariance matrices generated to jointly re-estimate SNP effects across eight sites.** Type or paste text here.

**Section S2. Greedy mash algorithm used to select covariance matrices that significantly improved the model log-likelihood.** Type or paste text here. Break this section up into subheads as needed

**Section S3. Narrow-sense heritability estimation.** In the diversity panel, we determined narrow-sense heritabilities ($h^2$) for green-up and flowering dates at single gardens using genomic relationship matrices calculated using the van Raden method (VanRaden 2008). Genomic relationship matrices were calculated within each subpopulation (Midwest and Gulf) and for both genetic subpopulations (Both). We used rrBLUP (Endelman 2011) to specify mixed models of the form:

$$\mathbf{y} = 1 + Zu + e$$

$$Var(u) = G\sigma_u^2$$

$$Var(e) = I\sigma_e^2$$

in which the vector $\mathbf{y}$ represents the flowering date or green-up date values for that garden, Z the design matrix for random effects, u the whole genome additive genetic effect, and e the residual. Matrix G is the whole genomic relationship matrix based on all SNPs retained for subpopulation-specific analyses. I is the rank-y identity matrix. Phenotypic variance $\sigma_p^2$ is $\sigma_u^2 + \sigma_e^2$. Narrow-sense heritability is then $h^2 = (\sigma_u^2/\sigma_p^2)$.

These models were run for each of the eight gardens, and across all gardens by adding an additional environmental effect of site without an interaction term. This resulted in 54 models: 3 sets of populations (the Gulf, Midwest, and Both subpopulations) for 9 garden sets (all eight gardens separately, and all eight gardens together) and two phenotypes (green-up date and flowering date).

**Section S4. Outbred pseudo-F2 mapping population and Quantitative Trait Locus mapping.** To confirm candidate genomic regions and patterns of allelic effects found in the diversity panel, we analyzed flowering in an outbred pseudo-F2 cross between four grandparents, two Midwest and two Gulf individuals. The formation of this mapping population has been described previously (Milano, Lowry, and Juenger 2016). The parents of this cross were DAC, an early flowering Midwest individual, VS16, a late flowering Midwest individual, AP13, an early flowering Gulf individual, and WBC, a late flowering Gulf individual. We made F1 crosses of the two early flowering genotypes, AP13xDAC, and the two late flowering genotypes, WBCxVS16. We then clonally propagated and planted the four parents, the two F1 genotypes (AP13xDAC, and VS16xWBC), and 801 F2 genotypes at eight field sites in May-July of 2015. To be directly comparable to the diversity panel data, only 2019 phenology data from the pseudo-F2 cross from the same eight common garden sites were used here.

Details on the genetic map construction, map polishing and fine-scale reordering can be accessed on DataDryad. QTL mapping was conducted with R/qtl2 (Broman et al. 2019). We performed a genome scan with a linear mixed model that accounts for the relationships among individuals and for environmental covariates (i.e., field sites). The full model can be expressed as:

$$phenotype = \mu + QTL + E + QTL * E + kinship + e$$

where $\mu$ is the population mean, QTL is the marker genetic effect, E is the environmental effects (here, common garden), QTL*E is the interaction between marker genetic and environmental effects, kinship corresponds to the background polygenic variation, and e is the error term. The genome scan was accomplished with the 'scan1' function. The statistical significance of the genome scan was established by performing a stratified (i.e., stratifying on common garden) permutation test (n=1000) using 'scan1perm' function. The estimated QTL effect was obtained using 'scan1coef' function in R/qtl2.

Figures should be cross-referenced like Figure S1. Each figure should be on its own page. You can control this by placing a pagebreak shortcode, with .

This can also be done with LaTeX's `\newpage` command. You may find `\FloatBarrier` useful if your figures or tables are running off to the wrong pages.

**Figure S1.** Figure

**Table S1. Simple table example**

| Species | CBS | CV | G3 |
|---|---|---|---|
| Acetaldehyde | 0.0 | 0.0 | 0.0 |
| Vinyl alcohol | 9.1 | 9.6 | 13.5 |
| Hydroxyethylidene | 50.8 | 51.2 | 54.0 |

**Datasets.** If your document relies on movies or datasets, please list them here with their captions. Use the `movie{your caption}` and `\dataset{file_name.ext}{your caption}` commands to do so.

Datasets 1 through 6 are csv files where the first two columns have the following definitions:

- Marker: The SNP marker in the format Chromosome_Position
- log10BF: log10(Bayes Factor) of the significance of the marker effect in the mash model

The remaining column names follow the pattern Effect_[Mean/StandardError/lfsr]_[Subpopulation]_[Phenotype]_[Garden], where Mean and Standard Error are estimates of the effect mean and standard error, lfsr is the local false sign rate statistic for the effect, and [Subpopulation], [Phenotype], and [Garden] follow the conventions of Figure 1.

**SI Dataset S1 (Dataset_1_Gulf_vegetative_growth_date.csv)**

SNP-associated effects for the start of vegetative growth jointly re-estimated in the Gulf genetic subpopulation.

**SI Dataset S2 (dataset_two.txt)**

SNP-associated effects and standard errors for the start of vegetative growth jointly re-estimated in the Midwest genetic subpopulation.

**SI Dataset S3 (dataset_three.txt)**

SNP-associated effects and standard errors for the start of vegetative growth jointly re-estimated in both Midwest and Gulf genetic subpopulations.

**SI Dataset S4 (dataset_four.txt)**

SNP-associated effects and standard errors for the start of reproductive growth jointly re-estimated in the Gulf genetic subpopulation.

**SI Dataset S5 (dataset_five.txt)**

SNP-associated effects and standard errors for the start of reproductive growth jointly re-estimated in the Midwest genetic subpopulation.

**SI Dataset S6 (dataset_six.txt)**

SNP-associated effects and standard errors for the start of reproductive growth jointly re-estimated in both the Midwest and Gulf genetic subpopulations.

## References


Broman, Karl W, Daniel M Gatti, Petr Simecek, Nicholas A Furlotte, Pjotr Prins, Śaunak Sen, Brian S Yandell, and Gary A Churchill. 2019. "R/Qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations." *Genetics* 211 (2): 495–502. https://doi.org/10.1534/genetics.118.301595.

Endelman, J. B. 2011. "Ridge Regression and Other Kernels for Genomic Selection with r Package rrBLUP" 4: 250–55.

Milano, Elizabeth R, David B Lowry, and Thomas E Juenger. 2016. "The Genetic Basis of Upland/Lowland Ecotype Divergence in Switchgrass (*Panicum Virgatum)*." *G3 Genes|Genomes|Genetics* 6 (11): 3561–70. https://doi.org/10.1534/g3.116.032763.

Olsen, Trond E, and Gunnar Stensland. 1992. "On Optimal Timing of Investment When Cost Components Are Additive and Follow Geometric Diffusions." *Journal of Economic Dynamics and Control* 16 (1): 39–51.

VanRaden, P. M. 2008. "Efficient Methods to Compute Genomic Predictions." *Journal of Dairy Science* 91 (11): 4414–23. https://doi.org/10.3168/jds.2007-0980.

Varga, Andrea, and Andrew N Edmonds. 2016. "Multilingual Extraction and Editing of Concept Strings for the Legal Domain." *Advances in Computer Science: An International Journal* 5 (4): 18–23.