**Working Title:** The genetic basis of multiple distinct flowering time cues in switchgrass (*Panicum virgatum*)

**Authors:** Alice MacQueen\*, Li Zhang\*, Jason Bonette, …who else? …, Tom Juenger

**Abstract**

Switchgrass (*Panicum virgatum*) is a perennial, warm-season species native to North America which has evolved into multiple divergent populations that vary in ploidy, morphology, and phenological timing. Its potential uses have expanded in the past few decades to include biofuels. Flowering time is a key life-history trait for biofuel production, as plants cease biomass accumulation upon completion of floral development. Here, we present evidence that photoperiodic sensitivity in switchgrass flowering time varies between genetic subpopulations using a diversity panel grown at eight field sites that cover 17 degrees of latitude in the central United States. We then map the genetic basis of flowering time in this population and in an independent four-way cross created from two highly divergent southern lowland and northern upland populations. We describe eight additive QTLs across these seven field sites with moderate effects on flowering, five of which had overlapping significant associations in the diversity panel. \_sentence about major gene candidates found – prevalence of genes involved in \_\_/expressed in \_\_.

**Introduction**

*# Flowering and phenology importance to the plant*

*…this will be the first paragraph, as outlined above.*

*# Walkthrough of what is known about mechanisms controlling greenup and flowering in switchgrass*

Over the course of one season, switchgrass behaves as a determinate plant. It typically produces a single flush of tillers, which all become reproductive after a period of vegetative (leaf) development, and – critically – cease biomass accumulation upon completion of floral development (Van Esbroeck et al 2003). (Parrish and Fike 2005).

Spring growth is initiated by “adequate temperature” according to McMillan and Weiler (1959), with adequacy thought to be dependent on the cultivar (Parrish and Fike 2005). During vegetative growth, switchgrass phenology is closely correlated with growing degree days (GDD), the cumulative mean daily temperature less a base temperature (Madakadze et al 1998c; Sanderson and Wolf, 1995a, 1995b). A base temperature of 12 C for vegetative and reproductive development is commonly used for growth models (Kiniry et al 2005, Kinery et al 2008a, Berhman et al 2013). However, base temperatures vary by cultivar (Madakadze et al 2003) and there may be a photoperiod or vernalization mechanism rather than a temperature threshold *per se* that triggers spring growth (Parish and Fike 2005).

Switchgrass is considered a short-day plant that flowers when exposed to shortening days of a specific length (Benedict, 1940) and reproductive development is strongly linked to day-of-the year (Cornelius and Johnston, 1941; Eberhart and Newell, 1959; Hopkins et al., 1995a; Sanderson and Wolf, 1995a). However, the nature of switchgrass photoperiodicity may be genotype dependent – both upland and lowland cultivars flowered under both 12 and 16 hour photoperiods, but flowering was delayed in the upland cultivar (Van Esbroeck et al 2003). Photoperiodicity likely differs with plant latitude of origin (Parish and Fike 2005). Moving plants from southern populations northward is thought to delay flowering, increasing leaf number and yields, while moving northern populations to southern latitudes is thought to hasten the transition to reproductive development, reducing vegetative growth and biomass yield (Sanderson et al 1996).

*# Management of switchgrass for biofuels is informed by an understanding of the biology underpinning plant responses to the environment.*

Genetic differences in flowering date and photoperiod response can be a point of exploitation in biomass crop breeding. Breeding for photoperiodic sensitivity in plants with earlier greenup dates may allow growers in the northern United States to take advantage of longer growing seasons, as photoperiod-sensitive strains will accumulate more biomass before flowering, contributing to higher biomass yields (cite?). Alternatively, breeding for cultivars with larger cumulative GDD requirements and reduced photoperiod sensitivity could increase the heat requirement needed for switchgrass development and stabilize switchgrass biomass yields at higher levels, which could help offset the negative impacts of climate warming (cite). However, despite substantial study of the mechanisms controlling switchgrass development, genetics of flowering and prediction equations with broad application remain elusive.

*# Here, we do xxx, yyy, zzz.*

Here, we grow and phenotype a diversity panel of 978 distinct switchgrass genotypes, clonal replicates of which were planted at eight common garden sites across 17 degrees of latitude. We use this panel to interrogate environmental mechanisms controlling greenup and flowering. We then use a F2 cross between individuals from the most distinctive subpopulations for flowering response to dissect the genetics of flowering in these groups. Finally, we combine the results from this cross with genome-wide association results from the diversity panel to narrow in on candidate genes affecting flowering date. Taken together, our results allow us to describe the distinct genes and alleles controlling flowering across multiple distinct switchgrass populations.

**Results**

*Diversity panel captures the range of phenological responses across the common gardens*

In 2019, we grew and phenotyped a diversity panel of 978 distinct, clonally propagated switchgrass genotypes at eight common gardens. These common gardens cover the majority of the latitudinal and climatic range of switchgrass and therefore capture the most comprehensive picture to date of genotype-specific environmental plasticity, or genotype-by-environment interactions. We scored plant green up and flowering across these locations every three days. We divided these switchgrass genotypes into five categories: tetraploid individuals in the Atlantic, Midwest, and Gulf genetic subpopulations (Lovell et al 202X), admixed/uncategorized tetraploid individuals, and octoploid individuals. We then explored the overall patterns of greenup and flowering within these five subpopulations.

We first evaluated the evidence that switchgrass photoperiodicity was genotype dependent – specifically, that switchgrass genotypes flowered when exposed to shortening days of a specific length. Many genotypes flowered while days were lengthening at the three Texas sites (Figure 1). At these three sites, only individuals from the Gulf subpopulation consistently flowered when days were shortening (Figure 1; 85.3%, 94%, 99.5%). Our data thus only supports a consistent flowering time photoperiodicity cue of shortening days within the Gulf subpopulation, while response to this cue may be segregating within the Atlantic subpopulation and within 8X individuals.

For plants with sequenced genomes, we observed little evidence for a specific day length that triggered flowering during shortening days. To test this, we modelled daylength at flowering as a function of common garden and the kinship matrix as random effects and removed daylength values for genotypes that flowered during lengthening days. The heritability for daylength at flowering during shortening days was 12%, while the heritability for flowering as a Julian date for the same set of individuals was 23.7% (Supplementary Table: Variance Components analysis).

Next, we evaluated the evidence that photoperiodicity, defined simply as exposure to shortening days, differed by plant latitude of origin (Parish and Fike 2005). We observed a strong signal of latitude of origin on whether plants grown in Texas common gardens flowered in lengthening or shortening days (glm, *Pr* < 2x10-16, Supp. Figure 1). The majority of plants from latitudes of origin below 35 degrees did not flower until days were shortening at the Texas sites, while the majority of plants from latitudes of origin above 38 degrees flowered while days were lengthening. Interestingly, the Oklahoma (OK) site, our first common garden at which most genotypes flowered during shortening days, was at ~36 degrees of latitude; in contrast, the three sites where a photoperiod cue was evident were below 32 degrees of latitude. At our common gardens in 2019, plants grown at latitudes >35 degrees did not have sufficient vegetative growth to flower before the summer solstice, and thus were not competent to respond to a photoperiod cue. As this is a common feature of growing seasons at these latitudes, we suggest that plants from more northern latitudes have evolved a flowering time response to a separate, non-photoperiod based environmental cue.

A screenshot of a cell phone

Description automatically generated**Figure 1.** *Many genotypes do not use a shortening day photoperiodicity cue at the three Texas (TX) common garden sites. Bars represent the number of distinct genotypes that had 50% of tillers flowering before the summer solstice (when days were lengthening) or after the summer solstice (when days were shortening). Colors represent the five categories we grouped genotypes into: tetraploid individuals in the Atlantic, Midwest, and Gulf genetic subpopulations, admixed/uncategorized tetraploid individuals, and octoploid individuals.*

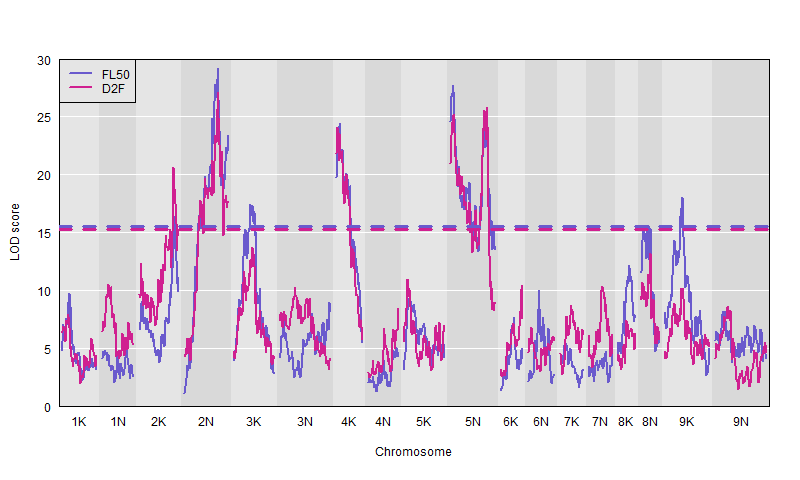
A screenshot of a cell phone

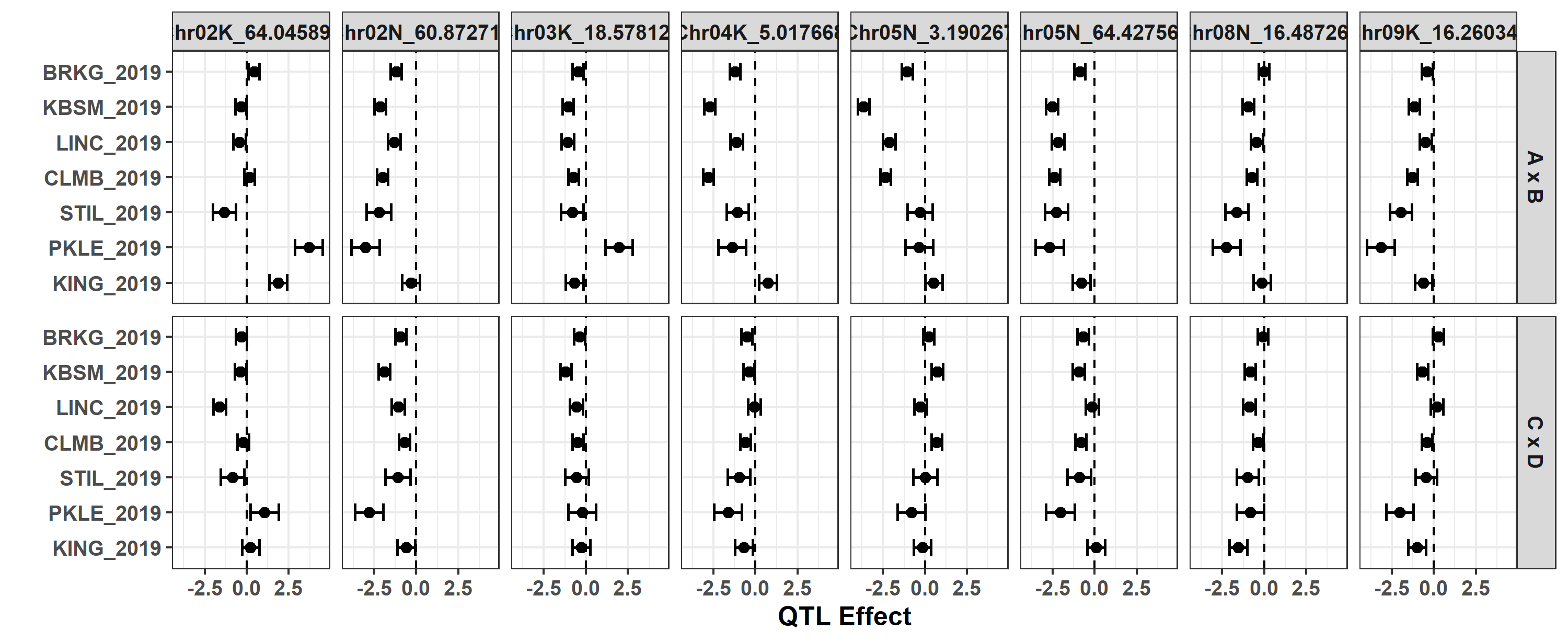
Description automatically generated**Supplementary Figure 1.** *Latitude of origin correlates with flowering response to an environmental cue of shortening day length. Bars represent the number of distinct genotypes that had 50% of tillers flowering before the summer solstice (when days were lengthening) or after the summer solstice (when days were shortening). Colors represent the five categories we grouped genotypes into: tetraploid individuals in the Atlantic, Midwest, and Gulf genetic subpopulations, admixed/uncategorized tetraploid individuals, and octoploid individuals.*

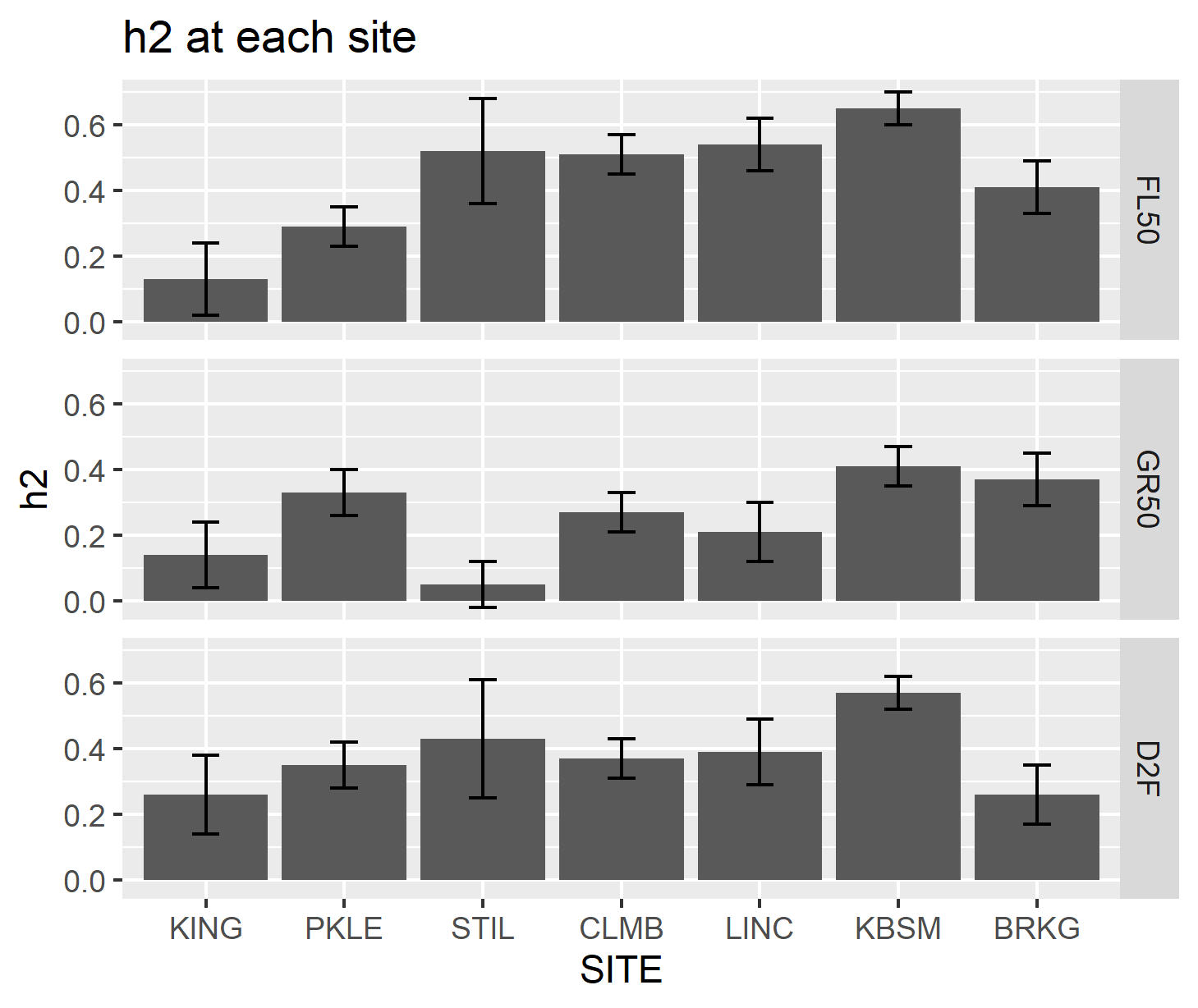
To test for a temperature-based environmental cue for flowering, we evaluated flowering date as a function of cumulative GDD between plant green up and flowering using a base temperature of 12 C (Kiniry et al 2005, Behrman 2013). We then modelled GDD as a function of subpopulation and the interaction between subpopulation and latitude of origin, both as random effects. To avoid confounding with a photoperiod-based environmental cue, we removed GDD values for genotypes predicted to have a photoperiod cue at common gardens south of 38 degrees latitude. Including these values substantially increased the residual variance for photoperiod sensitive individuals at these sites (data not shown//in supplement). Subpopulation explained most of the variation in flowering as a function of GDD (% Varsubpop = 85.9%). Predicted values of GDD necessary for flowering were largest for the Gulf subpopulation (1056; 95% CI 973-1099), smallest for the Midwest (667; 95% CI 517 – 814) and intermediate for the other three groups (741 (599-1019); 776 (483-1195); 747 (573 – 1043); Supplementary table of effects). The Midwest subpopulation response to GDD varied significantly by latitude of origin, with plants from the northernmost 20% of the range flowering at a GDD of 568, and plants from the southernmost 20% of the range flowering at a GDD of 779. In contrast, the southernmost and northernmost 20% of the Gulf plants differed only slightly in GDD (1008 vs 1090). The heritability for GDD at flowering for photoperiod insensitive, sequenced individuals was 30%, while the heritability for flowering as a Julian date for the same set of individuals was 10.8% (Supplementary Table: Variance Components analysis). Taken together, these data support latitude-of-origin based, low GDD cue for flowering in the Midwest subpopulation, and a consistent, high GDD cue for flowering in the Gulf subpopulation that is superseded by a photoperiod cue.

*Fourway cross breaks up genetics of flowering for the two most distinctive subpopulations*

The Midwest and Gulf subpopulations had the most distinct phenological responses out of all subpopulations across our common gardens and appeared to flower in response to two distinct environmental cues. To analyze the genes and alleles underlying the phenological responses in the most distinct two subpopulations, we used an F2 cross between four individuals, two Midwest and two Gulf individuals. 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 individuals, AP13xDAC, and the two late flowering individuals, WBCxVS16. We then planted the four parents, the two F1 individuals (AP13xDAC, and VS16xWBC), and XXX F2 individuals at eight field sites, and recorded greenup and flowering date for the 2016-2019 seasons. Though there was a X day gap in flowering date between the two F1 crosses, F1 individuals flowered at similar dates as the Midwest parents and Midwest subpopulation individuals in general, indicating that the flowering time genetic response to the Midwestern GDD levels was dominant to the photoperiod response and the higher GDD levels required by the Gulf subpopulation. At the northern five sites and at Kingsville, F2 individuals had similar greenup and flowering dates as individuals from the Midwest subpopulation. However, F2 individuals at PKLE and TMPL had flowering dates similar to late flowering Midwest individuals to A close up of a map

Description automatically generatedearly- to mid-range flowering Gulf individuals. Thus, at the northern five sites, we expected to find QTL controlling flowering date from the Midwest subpopulation, potentially by affecting the cumulative GDD requirements for the plant before flowering. In contrast, at TMPL and PKLE, there was the potential to isolate QTL that affected the ability of the plant to flower in response to a photoperiod cue.

We analyzed greenup and flowering in 2019 for seven fourway sites. We anticipated finding QTL at the northern common gardens that distinguished between early and late-flowering Midwestern alleles, while Pickle offered the opportunity to examine some effects of Gulf alleles. There were no significant QTL for greenup. There were eight QTL for flowering date in the fourway cross; the largest QTL were on Chr02N, Chr04K, and Chr05N. All eight QTL exhibited significant GxE between common garden sites. Due to our cross design, we could estimate allelic effects of alleles from both F1 individuals. In the early flowering alleles, AxB, at PKLE, five of eight lowland alleles delayed flowering date. In the CxD cross at PKLE, four of eight lowland alleles delayed flowering date – four of the same five regions that delayed flowering date in the AxB cross. These effects were consistent with the observed order of flowering in the F0 and F1. We therefore hypothesized that these regions: Chr02N, Chr04K, Chr05N, (Chr08N), and Chr09K were important regions affecting flowering in the Gulf subpopulation. At the northern sites, because of the dominance of Midwest phenotypes and alleles that we observed, we compared effects of the Midwest allele from the early and late F1 cross. For 28 site by QTL combinations for the AxB early flowering alleles, the Midwest allele was accelerating flowering. For 16 site by QTL combinations for the CxD late flowering alleles, the Midwest allele was accelerating flowering, and for 2 site by QTL the upland allele was delaying flowering. These effects were also consistent with the observed order of flowering in the F0 and F1. All eight QTL affected flowering date for at least one northern site for both the early and late flowering allele sets. Thus, we hypothesized that all eight regions were important regions affecting flowering in the Midwest subpopulation. 

Kingsville represents a marked departure from the temperate growth habits at the other nine sites. In general, heritability for flowering was lowest at Kingsville in both the diversity panel and the fourway cross, indicating a larger effect of environment on phenotypic variance at this site. At this site, 2 AxB Midwest QTL accelerated flowering, and 2 delayed flowering. In CxD, alleles at four Midwest QTL accelerated flowering.

*Genome wide association identifies candidates affecting flowering in fourway QTL intervals*

Though QTLs detected using the fourway cross are too coarse to be informative about gene identity by themselves, combining linkage and association mapping can outperform each method used in isolation (cite Brachi, others?). We therefore conducted GWAS on greenup and flowering date at seven common garden sites for individuals from both the Midwest and Gulf subpopulation. We then analyzed the allelic effects across common garden sites for the top XK SNPs using mash. Using this data, we focused on identifying associations in or near the important QTL regions for flowering date from the fourway cross, as well as the population that allele was found within and the effect that that SNP had across sites.

Five of the eight QTL – the four largest QTL, and all five QTL that we hypothesized affected flowering in the Gulf subpopulation – colocalized with one or more of the top 25 mash hits. Discussion of candidate genes, where they are in the QTL interval, what the effects look like in mash, what subpops they are segregating in, for:

* Candidate on Chr02N
* Candidate on Chr04K
* Candidate on Chr05N 4.3 Mb
* Candidate on Chr05N 64.4 Mb
* Candidate on Chr08N

*Possible Genomic Prediction Section*

* I don’t know how to tie genomic prediction in here. Ideally we’d perhaps have information on which parts of the genome being Midwestern vs Gulf would make you sensitive to photoperiod cues, and which parts affect how much cumulative GDD you need, if you’re not sensitive to photoperiod cues. That would be a nice way to bring this full circle.

**Discussion**

* Possible reasons why we didn’t find candidates in GWAS for other three QTL intervals: false negatives, type of allele shifts we are testing here vs the fourway, other things
* And what are other types of region we can detect with GWAS that we couldn’t detect in our fourway? Particularly differences between Gulf alleles… maybe offer up Gulf specific GWAS/mash here. Or GWAS on any individual that is tetraploid and looks like it’s photoperiod sensitive – including some admixed and Atlantic individuals. As an idea.
* Want to bring back discussion of environmental cues… may be able to lead in to this by talking about which subpopulations are segregating for these things.

**References**

Parrish DJ, Fike JH. 2005. The Biology and Agronomy of Switchgrass for Biofuels. Critical Reviews in Plant Sciences 24:423-459.

Tables & Figures, Maybe Supplementary

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Site | Latitude | Genotypes that flowered while: | 4X | 8X | Atlantic | Gulf | Midwest |
| Kingsville, TX, USDA/PMC | 27.55 | days lengthening | 9 | 8 | 44 | 23 | 23 |
| Kingsville, TX, USDA/PMC | 27.55 | days shortening | 11 | 4 | 18 | 134 | 1 |
| Austin, TX, UT PRC | 30.38 | days lengthening | 19 | 167 | 206 | 11 | 115 |
| Austin, TX, UT PRC | 30.38 | days shortening | 12 | 66 | 56 | 173 | 6 |
| Temple, TX, USDA/ARS | 31.04 | days lengthening | 8 | 7 | 45 | 1 | 46 |
| Temple, TX, USDA/ARS | 31.04 | days shortening | 15 | 6 | 62 | 193 | 15 |
| Stillwater, OK, OSU | 35.99 | days lengthening | 1 | 0 | 2 | 0 | 2 |
| Stillwater, OK, OSU | 35.99 | days shortening | 18 | 8 | 86 | 83 | 52 |
| Columbia, MO, MU BRF | 38.90 | days lengthening | 0 | 3 | 0 | 0 | 6 |
| Columbia, MO, MU BRF | 38.90 | days shortening | 31 | 206 | 228 | 149 | 113 |
| Lincoln, NE, UNL ARF | 41.15 | days shortening | 18 | 13 | 123 | 66 | 67 |
| Hickory Corners, MI, KBS | 42.42 | days lengthening | 0 | 1 | 0 | 0 | 0 |
| Hickory Corners, MI, KBS | 42.42 | days shortening | 36 | 181 | 199 | 198 | 109 |
| Brookings, SD, SDSU | 44.31 | days shortening | 15 | 12 | 109 | 73 | 63 |

However, cumulative GDD was not sufficient to explain the variation in flowering time in the Gulf subpopulation. Instead, we hypothesized that the Gulf subpopulation also used a photoperiod cue, given the imperfect relationship between GDD, photothermal time, and flowering in the Gulf subpopulation, and the presence of a relationship between latitude of common A screenshot of a cell phone

Description automatically generatedA close up of a map

Description automatically generatedgarden and flowering time based on cumulative GDD.

Cumulative GDD vs photoperiodicity as a function of each plant’s location of origin.