**The genetic basis for panicle trait variation in switchgrass (*Panicum virgatum*)**

**Running head: Panicle genetics in switchgrass**

Li Zhang1\*, Xiaoyu Weng1\*, Alice MacQueen1\*, Kathrine D. Behrman1, Jason Bonnette1, John L. Reilley2, Francis M. Rouquette Jr3, Philip A. Fay4, Yanqi Wu5, Felix B. Fritschi6, Robert B. Mitchell7, David B. Lowry8, Arvid R. Boe9, Thomas E. Juenger1\*

1Department of Integrative Biology, University of Texas at Austin, Austin TX 78712; 2Kika de la Garza Plant Materials Center, National Resources Conservation Service, US Department of Agriculture, Kingsville, TX 78363; 3Texas A&M AgriLife Research and Extension Center, Texas A&M University, Overton, TX 75684; 4Grassland, Soil and Water Research Laboratory, Agricultural Research Service, US Department of Agriculture, Temple, TX 76502; 5Department of Plant and Soil Sciences, Oklahoma State University, Stillwater, OK 74078; 6Division of Plant Sciences, University of Missouri, Columbia, MO 65211; 7Wheat, Sorghum, and Forage Research Unit, Agricultural Research Service, US Department of Agriculture, University of Nebraska–Lincoln, Lincoln, NE 68583; 8Department of Plant Biology and DOE Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI 48824; 9Departmentof Agronomy, Horticulture & Plant Science, South Dakota State University, Brookings, SD 57007

\* Corresponding authors. Email: lz5943@utexas.edu or [tjuenger@austin.utexas.edu](mailto:tjuenger@austin.utexas.edu)

\* These authors contribute equally.

Orcid-ID (Zhang, L): <https://orcid.org/0000-0001-8625-8042>

Orcid-ID (Juenger, T): https://orcid.org/0000-0001-9550-9288

**Abstract**

Grass species exhibit large diversity in panicle architecture influenced by genes, the environment, and their interaction. The genetic study of panicle architecture in perennial grasses is limited. In this study, we evaluate the genetic basis of panicle architecture including panicle length, primary branching number, and secondary branching number in an outcrossing switchgrass population grown across ten field sites in the central United States through multi-environment mixed QTL analysis. We also evaluate genetic effects in a diversity panel of switchgrass grown at three of the ten field sites using genome-wide association (GWAS) and multivariate adaptive shrinkage. Furthermore, we search for candidate genes underlying panicle traits in both of these independent mapping populations. Overall, 18 QTL were detected for the three panicle traits, and 187 unlinked genomic regions in the diversity panel affected one or more panicle trait. Twelve of the QTL exhibited consistent effects (i.e., no QTL by environment interactions or no QTL x E), and most (four of six) of the effects with QTL x E exhibited site-specific effects. Most (74.5%) significant diversity panel SNPs had significant effects in all panicle traits and all field sites and showed pervasive pleiotropy and limited environment interactions. Panicle QTL co-localized with significant SNPs found using GWAS, providing additional power to distinguish between true and false associations in the diversity panel.

Key words: panicle, QTL x E interaction, phenotypic plasticity, site-specific effect, *Panicum virgatum*, switchgrass

**Declarations**

**Funding**

This research was supported and funded by the National Science Foundation Plant Genome Research Program (IOS-1444533) and by the US Department of Energy, Office of Science, Office of Biological and Environmental Research Award DESC0014156 to T.E.J.

**Conflicts of interest**

We declare that there is no conflicts of interest.

**Data Availability**

Whenever possible, plant material will be shared upon request. Source data and code to replicate these analyses are available at: <https://github.com/lzhangUT/PanicleData.git>. Large genetic data files to replicate these analyses are available from the UT dataverse at: *make\_stable\_doi\_link*.

**Author contributions**

D.B.L., J.B., F.B.F., and T.E.J. designed research; D.B.L., J.B., P.A.F., R.B.M., J.L.-R., A.R.B., Y.W., F.M.R., R.L.W., X.W., K.D.B., A.L., D.B., A.S., F.B.F., and T.E.J. performed research; L.Z., X.W., A.H.M. and K.D.B. analyzed data; and L. Z., X.W. and A.H.M. wrote the paper with comments and editing by all authors.

**Key Message**

We investigate the genetic basis of panicle architecture in switchgrass in two mapping populations across a latitudinal gradient, and find many stable, repeatable genetic effects and limited genetic interactions with the environment.

**Introduction**

As the bearers of grain, grass panicles (or inflorescences) have been targets of selection for thousands of years (Doust 2007). There is enormous diversity in panicle architecture within and among grass species (Coen and Nugent 1994). Panicle architecture is a critical determinant of interspecies differences in plant morphology and life history, which is often measured as variation in panicle length, branching structure (number, length, and pattern), and flower number and size borne on each branch type. Simple panicles may have only primary branches, while complex panicles can possess many secondary and tertiary branches (Bommert and Whipple 2018; Glemin and Bataillon 2009). In wild grasses, branching pattern plays an important role in wind pollination and affects the number and size of seeds, which ultimately influences seed yield and plant fitness (Brown et al. 2006; Friedman and Harder 2004). In domesticated species, there is a direct association between panicle architecture and seed productivity (Brown et al. 2006; Crowell et al. 2016; Wang and Li 2005). Analysis of the phylogenetic distribution of panicle variation in the grasses suggests that different panicle architecture have arisen independently many times, and homoplasy across the grass phylogeny has obscured the mechanisms of panicle diversity (Doust and Kellogg 2002; Kellogg 2000). These traits likely evolve in response to natural selection mediated by aspects of wind pollination (Friedman and Harder 2004) and environmental variation such as light (Vogler et al. 1999), drought (Mal and Lovett-Doust 2005; Caruso 2006), nutrient availability (Dorken and Barrett 2004), and intraspecific competition (Wolfe and Mazer 2005). Given the importance of inflorescence architecture to the fitness of wild species and the productivity of domesticated species, it is of great interest to understand genetic variation in panicle architecture.

Plant reproductive components (inflorescences, flowers, seeds) often exhibit phenotypic plasticity in response to the environment, perhaps related to evolved reproductive allocation tradeoffs (Bazzaz and Grace 1997). In domesticated crops, selection has likely favored stability in grain yield possibly through reducing genetic variation in plasticity or favoring certain allocation tradeoffs (Gage et al. 2017). Genetic variation in phenotypic plasticity in response to the environment is better known as genotype-by-environment interactions (GxE) (Des Marais et al. 2013). As such, insights in the genetics of plasticity or stability may be helpful for understanding the evolution of panicle traits or the impact of artificial selection in domestication. Quantitative genetic studies of GxE in many plant species (e.g., maize, rice) have identified important quantitative trait loci (QTL) impacting many panicle traits (Adriani et al. 2016; Doust et al. 2005; Leng et al. 2017; Liu et al. 2008; Miura et al. 2010). For example, Doust et al. (2005) detected 14 QTL for four inflorescence traits under two trials with high and low density of plants that represent genes underlying divergencebetween foxtail and green millet. They further found significant GxE for primary branch number and bristle number per primary branch with joint QTL analysis. GxE is common in QTL studies and identifying GxE and the pattern of interactions has implications for the role of genetic architecture underlying phenotypic traits and their response to selection. -Genome wide association studies (GWAS) and studies of GxE of panicle architecture have been common in various crops (Zhao et al. 2016; Bai et al. 2016; Liu et al. 2018; Ta et al. 2018; Thapa et al. 2021; Zhong et al. 2021; Wang et al. 2021). For example, Wang et al. (2021) performed association mapping of panicle morphology-related traits in the sorghum mini core panel measured in multiple environments. They identified several loci that were related to panicle traits and suggested a number of candidate genes that resided in the loci. Their study also suggested that GWAS study of GxE may facilitate the molecular identification of panicle morphology-related genes and the enhancement of yield and adaptation in sorghum. Identifying the genetic basis of panicle traits in additional species, and evaluating the evidence for GxE using both quantitative studies and GWAS, will increase our understanding of the genetic regions responsible for panicle architecture.

Switchgrass (*Panicum virgatum* L.) has been championed as a potential biofuel crop since its selection by the US Department of Energy (US DOE) as a model grass species for bioenergy in the early 1990s (Hohenstein and Wright 1994; McLaughlin 1993). The potential for high biomass production on marginal land, adaptation to a wide range of environments, and ecosystem service such as carbon sequestration, water flow management, and erosion control, make switchgrass an excellent candidate for meeting bioenergy needs (Mitchell et al. 2012; Robertson et al. 2017). Switchgrass is a warm-season C4 perennial grass native to the North America, with a range that extends from the eastern seaboard west to the Rocky Mountains and from southern Canada south to the Texas Coastal Plain and Northern Mexico (Casler 2007; Hopkins 1995). Two major distinctive ecotypes, northern upland and southern lowland ecotypes, have been classified in the past based on morphology and habit preference (Porter Jr 1966). A recent study based on a resequenced switchgrass diversity panel defined a third coastal ecotype, which is broadly sympatric with the lowland ecotype but possesses upland leaf characters and lowland plant morphotype (Lovell et al. 2021).

Information on panicle morphology is limited in switchgrass, although panicle length differences have been reported between switchgrass ecotypes and cultivars (Porter Jr 1966; Price 2014; Van Esbroeck 2003). We hypothesize that panicle evolution in switchgrass may be related to selection on aspects of mating system and degree of investment in vegetative versus sexual reproduction, especially in the context of seedling establishment in differing habitats. For example, lowland switchgrass has a restricted bunch grass growth form and occurs primarily in patchy distributions along riparian areas. In contrast, upland switchgrass has a rhizomatous spreading growth form that occurs in many prairie habitats. Pattern of pollen dispersal across patches, or aspects of seed establishment (e.g. seed size/number tradeoffs or disturbance regimes) likely differ in these habitats and may have driven divergence in panicle form. Panicle morphology and its relationship to seed quality may be important targets of selection and breeding, as consistent seed production will be critical to meet the demands for large-scale biofuel production (Das and Taliaferro 2009; Vogel 2000).

In this study, we evaluated the genetic architecture of switchgrass panicle traits in two mapping populations: a pseudo-F2 mapping population (hereafter, “four-way”) grown across ten field sites (or common gardens) in the central US and a natural population of switchgrass (hereafter, “diversity panel”) grown at three of the ten sites. We assessed three panicle traits for each population at the end of their respective growing seasons: panicle length (PL), primary branching number (PBN) per panicle, and secondary branching number (SBN). For these phenotypes, we assessed 1) the genetic architecture underlying the trait, 2) the sensitivity of QTL and their effects across different environments from the four-way, 3) the single nucleotide polymorphism (SNP) effects on the traits in the three common gardens from the diversity panel, and 4) the candidate genes that were found for both two populations potentially involved in the regulation of panicle architecture in switchgrass.

**Materials and Methods**

**Field experiment and phenotyping of the four-way**

The details of the creation of the four-way population are described in Milano et al. (2016). Briefly, the grandparents of the mapping population were derived from highly divergent southern lowland and northern upland ecotypes. The population was developed by initial crosses between AP13 (A) x DAC6 (B) and WBC3 (C) x VS16 (D). AP13 and WBC3 are genotypes clonally derived from an individual selected from the lowland cultivar ‘Alamo’ (southern Texas accession) and an individual from naturally occurring population ‘West Bee Cave’ (central Texas accession), respectively. DAC6 and VS16 are genotypes clonally derived from individuals selected from the upland cultivars ‘Dacotah’ and ‘Summer’ (both northern upland accessions), respectively. The F1 hybrids of each of those crosses were then intercrossed reciprocally to produce the four-way outbred mapping population.

The grandparents, F1 hybrid parents, and the F2 progeny were propagated by dividing plants manually to produce 10 clones, each of which was maintained in a 3.8-L pot at the Brackenridge Field Laboratory, Austin, TX in 2013-2015. One replicate of each of the mapping progeny genotypes (i.e., 380 core genotypes), along with multiple replicates of grandparents and F1 parents, were transplanted from May to July of 2015 at 10 field sites. The 10 field sites cover 17 degrees of latitude from South Texas to South Dakota (Figure 1A). Detailed information of the 10 field sites, including latitude, longitude and soil type is provided in Table 1. The annual mean temperature at the 10 sites in 2016 ranged from 10.4 °C in the north to 20.7 °C in the south, and the total rainfall varied from 574 mm to 1440 mm (Figure 1B, data are from local weather station or from NOAA if local weather data are not available; the weather station or NOAA link is included in Table 1). To control weeds, each field site was covered with one layer of weed barrier cloth (Dewitt, Sikeston, MO). Holes were cut into the weed cloth in a honeycomb fashion. Plants were randomized into the holes, with each plant having four nearest neighbors each located 1.56m away from each other. A row of border plants was planted at every edge position of the field to minimize edge effects. The border plants were derived from rhizome plugs obtained from an approximately 10-year-old stand of Alamo switchgrass. Plants were well watered in the field during the summer of 2015 to facilitate establishment and all phenotypes were collected in 2016.

Three panicles were cut from each plant at full maturity. Panicle length (PL in mm), primary branching number (PBN), and secondary branching number (SBN) were assessed at the end of the 2016 growing season. A diagram depicting these phenotypes is presented in Figure 2a, with representative images of panicles from the four grandparents. PL was measured on the primary panicle from the base of the first primary branch to the top of the panicle. PBN was counted as the total number of branches along the primary rachis. Due to the numerous secondary branches in switchgrass, SBN in our study referred to the total number of secondary branches on the lowest primary branch of the panicle (Figure 2). In total, over 10,000 separate panicle morphology measurements were collected from the four-way population. The phenotypic data (i.e., average values) for each genotype at each field site are provided in Supplemental Table S1.

**Genotyping and multi-environment QTL modelling of the four-way**

Illumina fragment paired end libraries, representing each of the four grandparents (A: AP13, B: DAC; C: WBC; D: VS16) were sequenced and genotyped in the context of the *P.virgatum* reference genome v5 as detailed in (Bragg et al. 2020). The genetic map spans 750 recombinant 4-way progeny genotyped at 4700 markers. Details on the genetic map construction, map polishing and fine-scale reordering can be accessed on https://datadryad.org/stash/dataset/doi:10.5061/dryad.ghx3ffbjv (Lovell et al., 2020). For computational efficiency in GxE analysis, the genetic map was reduced into 738 markers, with an average intermarker distance of to 2cM.

We estimated quantitative genetic variation for the panicle traits within our full sib family using marker-based realized relationship matrices and linear mixed models implemented in the Sommer package ([Covarrubias-Pazaran 2016](javascript:;)) in [R Core Team (2020)](javascript:;). Since additive genetic variance and dominance are not orthogonal in a full-sib family like the four-way, it was not feasible to cleanly partition additive from nonadditive components of variance ([Hill 2013](javascript:;)). As such, we report our estimates from the kinship matrix as genetic variance (*Vg*), and our heritabilities as broad-sense heritability (*H2*), which was calculated as *Vg/Vp*, where *Vp* is the total phenotypic variance.

Details of the mapping scheme and application in the outbred four-way population are described in Malosetti et al. (2013) and Lowry et al. (2019). In brief, ‘single trait under multiple environments’ QTL mapping for each panicle trait in the cross-pollinated (CP) family was implemented in Genstat (2019). The QTL approach with CP family resulted in four possible QTL alleles designated *A* and *B* corresponding to marker alleles of the first pair of grandparents (AP13 x DAC) and QTL alleles *C* and *D* corresponding to marker alleles of the second pair of grandparents (WBC x VS16). A multienvironment mixed model was fit for each trait as shown in Eq. 1:

Eq. (1),

where *μ* is the population mean; *E* represents the environment effect; , denoting the total effect from the additive effect from the first grandparent (i.e., the difference between *A* and *B* alleles, , the second grandparent (i.e., the difference between *C* and *D* alleles, , and the dominance effect (i.e., the intralocus interaction, ; represents the QTL × environment interactions; and *e* represents the error term that was modeled by an unstructured variance–covariance matrix. The unstructured model was used to specify the data structure in the genome-wide QTL scan of simple interval mapping (SIM) and composite interval mapping (CIM). A backward selection procedure was used to retain significant fixed terms (*p*< 0.05) after three consecutive runs of CIM to confirm stability of QTL. The QTL with highest LOD peaks were considered as the most significant QTL, and the flanking markers associated with 1.5 LOD drop around the most significant QTL were considered as confidence interval for the QTL peaks.

**Genome-wide association and multivariate adaptive shrinkage in the switchgrass diversity panel**

The formation and resequencing of the switchgrass diversity panel has been described previously (Lovell et al. 2021). Briefly, hundreds of tetraploid switchgrass plants were resequenced, and these genotypes were clonally replicated and planted at multiple common gardens spanning a latitudinal gradient across the continental United States. We phenotyped panicle length (PL in mm), primary branch number (PBN), and secondary branch number (SBN) as above for three panicles cut from each plant at full maturity at the end of the 2019 growing season, in a subset of genotyped individuals and common gardens. A diagram depicting the different panicle architectures from several representative genotypes is presented in Figure 2b. We phenotyped 382 genotyped individuals that had clones present at each of three common garden locations (Austin, TX or PKLE; Columbia, MO or CLMB; and Hickory Corners, MI or KBSM), which overlapped with the three sites in the four-way and cross a large latitude of Central US. We took three measurements per individual at each field site, then found phenotypic BLUPs (Best Linear Unbiased Predictions) for each genotype in ASReml-R (Butler et al. 2017) using the model: trait ~ genotype + error, where genotype was a random factor and both genotype and error were fitted as identity variance models (~idv()). The phenotypic data (raw values and phenotypic BLUPs) for each genotype at each field site for the diversity panel are provided in Table S1.

We analyzed SNP (single nucleotide polymorphism) effects on three panicle traits in three common gardens using multivariate adaptive shrinkage (mash), using effect estimates from univariate genome-wide association studies (GWAS). GWAS were conducted using the bigsnpr R package (Privé et al. 2018), which performs fast statistical analysis of large SNP arrays encoded as matrices, and which implements the current best practices in human genetics for principal component analysis of population genetic data (Privé et al. 2020). Only SNPs with <20% missing data and minor allele frequencies >0.05 at all three gardens were used in univariate GWAS, resulting in 18.7M SNPs retained for the analysis. We used singular value decomposition (SVD) on all 18.7M SNPs for all 382 genotyped individuals to create 15 genetic principal components (PCs) for population structure correction using the snp\_autoSVD() function in bigsnpr. To choose the number of PCs that best controlled for population structure and reduced genomic inflation, we ran univariate linear regressions for each combination of phenotype and common garden including the range of 0 to 15 PCs as covariates, then selected either the smallest number of PCs that made λGC, the genomic inflation factor, less than 1.05, or else selected the number of PCs that minimized λGC when the first criterion could not be met (Supplemental Table S2). In practice, because the PCs are orthogonal by definition, GWAS results are not sensitive to the number of PCs used, as long as a sufficient number of PCs are included to capture true population structure effects (Price et al. 2006). Our first two PCs accounted for population structure, and our third and fourth the effects of ecotype within genetic subpopulations; further PCs captured additional structure between the genetic subpopulations not tied to known phenotypic differences, suggesting at least four PCs should be used (Figure S1). Our univariate GWAS all used at least four PCs to account for population structure, and all had λGC < 1.043; our nine univariate GWAS had strong associations and appeared free of obvious population structure confounding issues (Figure S2).

We then ran mash on the effect estimates and standard errors generated from univariate GWAS, following mash documentation (Urbut et al. 2019): first, 100K SNPs unlinked at an r2 of 0.2 were used as a ‘random’ set to learn the background correlation structure; second, 5K SNPs with the maximum -log10p-values in any of the univariate GWAS were used to construct data-driven covariance matrices; third, the random set was used to fit the mashr model; fourth, posterior summaries using the model fit on the random set were computed on all 18.7M SNPs. We determined which SNPs had evidence of significant phenotypic effects using local false sign rates (lfsr), which are analogous to false discovery rates but more conservative (in that they also reflect the uncertainty in the estimation of the sign of the effect) (Urbut et al. 2019). These lfsr were condition-specific; for an overall measure of significance for each SNP, we used the log10 (Bayes Factor) computed by mash, which measures the overall significance of a SNP on the trait effects included in mash.

**Enrichment tests to find candidate genes in both mapping populations**

To determine if SNPs with significant trait effects on panicles in our diversity panel (assessed using mash) were enriched in panicle QTL intervals in our four-way mapping population, we compared SNP enrichment in the QTL intervals to SNP enrichment of 1000 permutations of the QTL regions. First, to reduce enrichments due only to partially linked SNPs within a QTL region, the 18.7M SNPs used in mash were clumped to keep only the most significant SNP in each LD block, using a linkage threshold of r2 < 0.2. Significance was assessed using the log10 (Bayes Factor). SNP clumping resulted in 2.7M SNPs unlinked at an r2 of 0.2. Second, 1000 permutations of the QTL regions were created of the same size (in bp) of the 18 QTL found using the four-way mapping cross. For both the QTL intervals and these 1000 permutations, we assessed the number of QTL that had significant enrichments of mash SNPs in the top 1% percentile of the 2.7M partially linked SNPs using hypergeometric tests. We also explored SNP effects in a completely unlinked set of SNPs, where the 18.7M SNPs used in mash were clumped to keep only the most significant SNP in each LD block, using a linkage threshold of r2 < 0. For SNPs with a lfsr < 0.05 for a trait and site combination, we calculated the variance explained by each SNP using the equation Var(SNPi) = 2\*MAFSNP\_i\*(1-MAFSNP\_i)\*EffectSNP\_i^2, where MAFSNP\_i is the minor allele frequency of SNPi, and EffectSNP\_i is the point estimate for the phenotypic effect estimated using mash. We summed these estimates for the unlinked SNPs with significant effects on each phenotype to determine the total amount of phenotypic variation explained by significant, unlinked mash SNPs.

Third, we identified genes that were located both in the confidence intervals of the discovered QTL from the four-way and within 20kb of the 6149 mash SNPs with log10 (Bayes Factor) > 1.3 from the diversity panel. Because these genes were identified in two independent mapping panels, we have increased confidence that these genes are involved in panicle architecture in switchgrass. We used the pvdiv\_table\_topsnps() function of the switchgrassGWAS R package (<https://github.com/Alice-MacQueen/switchgrassGWAS>) to find genes within 20kb of mash SNPs, a distance consistent with a 50% linkage disequilibrium decay in this species (Grabowski et al. 2017; Lovell et al. 2021). These genes were compared with the rice (v7, accessed from phytozome https://phytozome-next.jgi.doe.gov/info/Osativa\_v7\_0) and Arabidopsis annotation databases (TAIR 10, accessed from phytozome https://phytozome-next.jgi.doe.gov/info/Athaliana\_TAIR10) to further identify candidate genes with functional validation in panicle architecture, or bolt architecture after the transition to flowering, in other species (Bouché et al. 2016; Yao et al. 2018). The annotation file for switchgrass was accessed on JGI (Joint Genome Institute) Phytozome 13 website: https://njp-spin.jgi.doe.gov/.

**Results**

**Phenotypic variation and heritability of the four-way**

Values for the three measured panicle traits increased in F2 individuals from the four-way as latitude of the common garden increased (Figure 3). Each trait showed a continuous, unimodal distribution within sites, and transgressive behavior in the F2 generation. The lowland genotype F0 individuals, AP13 and WBC, always had larger values of panicle length (PL in mm), primary branching number (PBN), and secondary branching number (SBN) than the upland genotype F0 individuals, DAC and VS16 (Figure 3).

The heritability (*H2*) for PL, PBN and SBN varied significantly by site and were typically moderate (0.2 to 0.5) or high (> 0.5) (Table 2). The *H2* for PL ranged from 0.20 to 0.71, with an average of 0.46 and values greater than 0.50 at four of five northern sites. The *H2* for PBN ranged between 0.45 and 0.66 for nine out of the ten sites, with Stillwater, OK (STIL) having low heritability (*H2*=0.20). The *H2* for SBN ranged from 0.02 to 0.62, where Stillwater, OK (STIL) had *H2* close to zero (*H2*=0.02), Columbia, MO (CLMB) had low heritability (*H2*=0.15), and four sites had heritability point estimates of approximately 0.50. The phenotypic correlations between traits were generally positive but varied by site, ranging from 0.21 to 0.63 for phenotypic correlation (Table 3).

**Multi-environment mixed QTL model**

A total of 18 QTL were identified for panicle morphology traits using multi-environment mixed model analyses (Figure 4, Table 4). Seven QTL were identified for PL, distributed across seven chromosomes. Among these, five QTL (2K@77.89, 4K@26.26, 5K@76.02, 5N@36.27 and 9N@38.02) had consistent effects across field sites (Figure 5a), while two QTL (3N@62.06 and 6N@54.19) showed interaction with the environment (QTL x E). The additive effects for QTL 3N@62.06 changed in magnitude across geographic regions. QTL 6N@54.19 (C x D cross) had the largest effects at the most northern and southern site and smaller effect at mid-latitude sites. There is also consistent directionality for these effects, that is, lowland alleles always make longer panicles (Figure 5a, A x B and C x D crosses of 3N@62.06, and C x D cross of 6N@54.19). The A x B cross of QTL 6N@54.19 had a trade-off pattern, with a sign change (aka. antagonistic pleiotropy) in allelic effects between three northern sites and the southernmost site.

Seven QTL were identified for PBN distributed across seven chromosomes. Four QTL (2K@74.02, 2N@66.12, 5N@84.04 and 9N@26.03) had consistent effects across locations, while three QTL (3K@38, 5K@14.06, and 7N@54.06) had QTL x E interactions, including both changes of magnitude (3K@38 and 7N@54.06) and direction (5K@14.06) of the allelic effect across geographic regions (Figure 5b). Four QTL were identified for SBN. Three QTL (2N@72.03, 5K@95.5 and 9N@36.02) had consistent effects across locations, while there was a magnitude changing interactions for QTL 9K@51.96 (Figure 5c). Similarly, there is consistent directionality for the QTL effects with magnitude change-lowland alleles often make more branches, including primary and secondary branches. We also observed that two QTL for PBN (2K@74.02 and 9N@26.03) co-localized with PL QTL on chromosome 2K and 9N, based on overlapping confidence intervals (Figure 4). QTL 9N@38.02 for SBN co-localized with the QTL of PL and PBN on chromosome 9N (Figure 4). The majority of QTL (12 of 18) did not show significant QTL x E interactions.

**Genome-wide association and multivariate adaptive shrinkage in the switchgrass diversity panel**

Three panicle traits were measured in 382 clonal propagates of a diversity panel grown in three common gardens. Trait values showed continuous distributions within sites which were bimodal for PL and PBN at all sites and for SBN at PKLE (Figure 6a). These bimodal distributions were caused by unimodal trait distributions within ecotypes that had trait means and trait distributions which differed significantly between the upland ecotype and the coastal and lowland ecotypes at all sites (Figure S3); coastal and lowland ecotype trait means differed only for PBN, driven mostly by PBN differences between the Atlantic and Gulf genetic subpopulations (Figure S3). Narrow-sense heritabilities were high (>0.5) for all traits at all sites (Figure 6b), marginally higher in Texas than at the northern sites, and higher for PL (75.2-83.9%) and PBN (78.1-84.1%) than SBN (58.8-68.1%). Phenotypic correlations were high (>0.5) and positive for all traits within sites (Figure 6c); between sites, correlations were highest for PBN (77.6-85.0%) (Figure S4).

We next explored how genetic effects for the three panicle traits at the three sites varied across the genome using mash. We first explored the full mash model, which was fitted using 100K sampled SNPs, then had posterior summaries computed for all 18.7M SNPs. Most of the 18.7M SNPs (78.5%) had mash model posterior weights on covariance matrices with equal effects in all nine trait-site conditions; most remaining SNPs (21.3%) had high model weights on covariance matrices with no effects in any condition (Figure 6d). The remaining SNPs (0.24%) had high model weights on one of three data-driven matrices, two of which showed patterns of negative covariance between traits or sites (Figure S5). Across all 18.7M SNPs, 48271 (0.26% of 18.7M) had a significant effect in at least one trait-site condition; most significant SNPs (35971, 74.5%) had significant effects in all nine trait-site conditions (Table S3). Thus, the majority of SNP effects in the diversity panel data showed little evidence for GxE for panicle traits at these three gardens, supporting our QTL findings that there was little QTL x E for panicle traits across the ten common gardens.

**Enrichment tests to find candidate genes in both mapping populations**

Because QTL regions were large enough that they could contain one or more effect on panicle architecture that were partially linked or even unlinked in the diversity panel, we made comparisons between QTL regions and mash SNPs using a partially linked SNP set with SNPs; however, major results were similar when the full SNP set was used (data not shown). For SNPs clumped at r2 ≤ 0.2 (hereafter ‘partially linked SNPs’) with the most significant marker selected, 6149 (0.23% of 2.7M) had a significant effect in at least one trait-site condition; for SNPs clumped at r2 ≤ 0 (hereafter ‘unlinked SNPs’) with the most significant marker selected, 187 (61.2% of 303) had a significant effect in at least one trait-site condition. All QTL regions from the four-way cross contained significant partially linked SNPs in the diversity panel that fell within 20kb of genes that had functionally validated roles in panicle, spikelet, or grain traits in rice (Table S4). Because many regions of the genome had significant SNPs in the diversity panel, we considered the possibility that QTL regions could be enriched with significant, partially linked SNPs from the diversity panel by chance alone. We first determined that 10 of the 18 QTL regions also had a significant enrichment of partially linked SNPs (p hypergeometric test < 0.05, Table S5). 0.2% of permuted genomic intervals had as many or more permuted QTL regions enriched for partially linked SNPs (p = 0.002, Figure 6e), while no permuted genomic intervals had more than 10 regions significantly enriched for partially linked SNPs. Thus, even with the very different population makeup of the GWAS panel, we could confirm a higher than expected overlap between SNP effects and QTL effects on panicle traits.

Finally, we examined specific SNP effects across traits and sites for unlinked SNPs with at least one significant effect (187 SNPs). The variance explained by these SNPs, summed just for effects that were significant for that trait-site condition, was between 17.4 and 84.6% of the total trait variation, highest for PBN (75-84.6%) and modest for PL and SBN (17.4-23.4%) except at PKLE, where the variance explained for PL was 36.3% (Table S6). For PL and SBN, variance explained by just these unlinked SNPs was between 23-43% of the total heritability, while these SNPs explained 94-104% of the total heritability for PBN. The alternate allele of these SNPs typically had a positive effect on panicle traits at all three sites for all three panicle traits, with the largest effects on PBN and the smallest effects for SBN, and a MAF between 0.025 and 0.075 (Figure 6f). Rarely, antagonistic pleiotropy between CLMB and the other two sites was observed, as for the SNP at 47.94Mb on Chr02K; effects on PL were commonly larger at PKLE than at the other two sites, which contributed to the higher phenotypic variance explained by these SNPs at this site.

We found 497 candidate genes by filtering for genes in the confidence intervals of both the QTL from the four-way population and within 20kb of significant, partially linked SNPs from the diversity panel (Table S4). Among these overlapping candidate genes, we identified key hormone-related genes associated with panicle development. For example, a homolog of the rice DELLA protein SLR1 (Pavir.9NG141800) was found in the overlapping interval on Chr9N for both PL and PBN. SLR1 is a component involved in GA signaling pathway and regulates panicle length and branch number via the DELLA–KNOX signaling pathway (Su et al. 2021). Another candidate gene, Pavir.2KG521100, is the homolog of OsGH3.8 and was found in the overlapping interval on Chr2K for both PL and PBN. In rice, OsGH3.8 mediates crosstalk between miR156-SPL7 and auxin pathways to regulate panicle architecture (Dai et al. 2018). These candidates suggest an important role for auxins and gibberellins in panicle development in switchgrass. Two flowering time genes, a homolog of rice Hd16 (Pavir.5NG232181) and a DOF transcription factor (Pavir.5NG191200), were found in the interval affecting PBN on Chr05N, which also overlaps a QTL interval for flowering in this population (Lowry et al. 2019). These genes are known to be involved in the photoperiodic flowering pathway and control panicle morphology in rice (Hori et al. 2013; Wu et al. 2015). Interestingly, Hd16 encodes a casein kinase I and phosphorylates the DELLA protein SLR1, suggesting a potential interaction between candidate genes in the overlapping intervals on Chr05N and Chr9N (Dai and Xue, 2010).

**Discussion**

There has been considerable interest in the molecular mechanisms of GxE across a diversity of phenotypes, species, and environments. GxE is common and is often driven by differential sensitivity of alleles and may play an important role in adaptive plasticity and local adaptation (Des Marais et al. 2013). With its large scale, our study evaluated the genetic basis and examined the GxE of panicle morphological traits in switchgrass from a four-way mapping population which were grown at ten field sites in the central United States and from a diversity panel which were grown at three of the ten sites. Overall, we detected moderate heritability (except for the field site Stillwater, OK from the four-way) for panicle traits and positive phenotypic correlations between traits at each site forboth populations. These data suggest considerable standing genetic variation in inflorescence characteristics available for natural or artificial selection to act upon. Our study identified genomic regions (QTL) that contribute to panicle trait variation across a broad latitudinal gradient. These QTL exhibited constant effects (i.e., no QTL x E: 12 QTL), antagonistic pleiotropy (i.e., sign change: 2 QTL), or site-specific effects (i.e., magnitude change: 4 QTL) across the studied environmental gradients. Most QTL with QTL x E are conditionally neutral. This is consistent with a recent meta-analyses which found that asymmetry of QTL effects are more often caused by conditional neutrality than by trade-offs (Wadgymar et al. 2017). We also did GWAS analyses and enrichment tests to find overlapping candidate genes using an independent switchgrass diversity panel, increasing our confidence in the genomic regions and candidate genes influencing panicle traits in switchgrass.

We were only able to measure panicle traits for three field sites using the diversity panel, compared to ten sites for the four-way. However, the diversity panel contains hundreds of representatives from three distinct genetic subpopulations of switchgrass, and thus captures substantially more natural variation than the four parents of the four-way, which came from two genetic subpopulations of switchgrass. In addition, we were able to obtain a balanced sample of 382 switchgrass genotypes grown at all three field sites for the diversity panel. We consider these panels complementary, and using both increases our power to distinguish true from false positives in GWAS mapping, while amplifying signals of causal QTL in the four-way that may be rare in the GWAS population (Brachi et al. 2010). In the four-way, most of the identified QTL showed no GxE effect. In the diversity panel, we found that most (99.8%) SNP effect patterns had high posterior weights on correlation matrices of equal effects or no effects across all panicle traits and field sites, corresponding to patterns of consistent, stable effects with no GxE, or patterns of non-significant effects. In addition, most (78.5%) of significant SNPs had similar effects across all panicle traits and field sites. Thus, we found stronger evidence for pleiotropic effects on panicle traits in the diversity panel than in the four-way, and weaker evidence for effects with GxE.

GWAS analyses on panicle morphology-related traits have been conducted in other crops like rice and sorghum, either in single environment or multiple environments (i.e., different locations, different growing seasons, and/or different managements) (Zhao et al. 2016; Thapa et al. 2021; Zhong et al. 2021; Wang et al. 2021). QTL identification in GWAS with multiple environments often considers QTL detected in at least two environments as significant QTL, while rarely focusing on the effects of QTL in different environments (Ta et al. 2018; Wang et al. 2021). For example, Wang et al., (2021) considered the association to be strong when it reached the Bonferroni correction P-value in at least two environments, while their experiment studying sorghum was actually grown in 11 environments. In our study, we used a more formal approach to quantify the effects across environments in the GWAS panel and found strikingly consistent effects across sites for most loci. We also completed a search for the overlapping SNPs with genomic regions identified from the four-way. After univariate GWAS, we re-estimated effects of SNPs on panicle traits while sharing information across all panicle traits and field sites using mash. Then, we conducted a permutation analysis to ask if significant SNP effects estimated using mash were enriched in QTL regions in the four-way, and if so, if the QTL regions were enriched more than random sets of genomic intervals. Ten QTL had significant enrichments of significant SNPs from the diversity panel (Table S5), more than 99.8% of random genomic intervals. In addition, we identified 6149 LD blocks within 20kb of 497 candidate genes in regions identified by both four-way and diversity panel mapping (Table S4).

Many candidate genes for panicle traits have been reported in various crop plants and model systems (Doust 2007; Doust et al. 2005; McSteen 2006; Miura et al. 2010; Vollbrecht et al. 2005; Wang et al., 2021). In our study, candidate gene *Pavir.9NG141800,* a homolog of the rice DELLA protein SLR1,was found on Chr9N for a QTL affecting both panicle length and primary branching number. SLR1 was shown to physically interact with the meristem identity class I KNOTTED1-LIKE HOMEOBOX (KNOX) protein OSH1 to repress OSH1-mediated activation of downstream genes that are related to panicle development, providing a mechanistic link between gibberellin and panicle architecture morphogenesis (Su et al., 2021). The candidate gene Pavir.5NG191200, a homolog of rice Dof (DNA binding with one finger) transcription factor, was found on Chr05N. Wu et al. (2015) found that overexpressing *OsDof12* led to smaller panicles by decreasing primary and secondary branch numbers. They further performed the Brassinosteroid (BR)-responsive tests and found that overexpression of *OsDof12* could also result in BR hyposensitivity, suggesting that *OsDof12* is involved in rice plant architecture formation by suppressing BR signaling.

In addition to the candidates compared with rice and Arabidopsis, we also compared the candidate genes with Setaria from the study of Doust et al. (2005), in which they identified the genomic regions controlling panicle traits between foxtail millet and green millet. Seraria is a grass in the subfamily Panicoidean, the same subfamily as maize and sorghum, and closer to switchgrass than rice. We found seven corresponding candidates from the four-way and one candidate from the diversity panel, mainly for panicle length and primary branching number. These candidates are the homologs of *ba1* (*barren stalk1*), *tb1* (*teosinte branched1*), and *bif2* (*barren inflorescence2*) genes found in other crops. *ba1* represents one of the earliest genes involved in the patterning of maize inflorescences. The mutant phenotype of *ba1* causes a reduction or elimination of branches and spikelets (Gallavotti et al. 2004). *tb1* encodes a non-canonical basic helix-loop-helix protein required for the initiation of all aerial lateral meristems in maize (Studer et al. 2017), and is a key regulator of apical dominance and inflorescence architecture in bread wheat (Dixon et al. 2018). Together with the *tb1* gene, it regulates vegetative lateral meristem development (Studer et al. 2017). *bif2* affects axillary meristems in the maize inflorescence. Mutants of *bif2* make fewer branches owing to a defect in branch meristem initiation, exhibiting reduced formation of all axillary structures including tassel branches, spikelets, and ear shoots (McSteen and Hake 2001). These comparisons provide more candidates for consideration. Together with other candidate genes, they might be targets for future switchgrass molecular research and breeding for panicle architecture.

In summary, our results suggest that variation of panicle traits in switchgrass is predominantly due to stable, consistent QTL that do not display GxE, with a minority of QTL displaying different effects across geographic regions (i.e., GxE). Future work focusing on the few QTL with GxE could identify rarer drivers of QTL by environment interactions in panicle traits, to help facilitate the selection of suitable genotypes of switchgrass for specific environments. Molecular research on candidate genes could provide insights to the pathways and mechanisms in panicle development in switchgrass.

**Supplemental Files**

The phenotyping data (panicle length, PL; primary branching number, PBN; and secondary branching number, SBN) for genotypes in the four-way at each of the ten field sites and in the diversity panel at each of the three sites (Table S1), the GWAS parameters (Table S2), the number of significant mash SNPs for each trait-site combination (Table S3), the candidate gene lists (Table S4), the enrichment test results for diversity panel SNPs within the four-way QTL regions (Table S5), and the variance explained by the unlinked diversity panel SNPs for each phenotype (Table S6) are included in the supplemental excel files. A visualization of the population structure correction in the diversity panel (Figure S1), the univariate Manhattans and QQ-plots from genome-wide association in the diversity panel (Figure S2), ecotype-specific phenotypic distributions in the diversity panel (Figure S3), phenotypic correlations between all traits and sites in the diversity panel (Figure S4), and the data-driven covariance matrices specified by mash (Figure S5) are also provided in the Supplementary Information file.

**Acknowledgements**

We thank the numerous field technicians, students, and postdocs who worked in the field helping out with data collection. Special thanks go to the Juenger Lab members for helping improve the manuscript with their comments. This research was supported and funded by the National Science Foundation Plant Genome Research Program (IOS-1444533) and by the US Department of Energy, Office of Science, Office of Biological and Environmental Research Award DESC0014156 to T.E.J. This material is based upon work supported in part by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494. Support for this research was provided by the National Science Foundation Long-term Ecological Research Program (DEB 1832042) at the Kellogg Biological Station and by Michigan State University AgBioResearch. We thank the Joint Genome Institute and collaborators for prepublication access to the *Panicum virgatum* v5 AP13 genome reference.

**References**

Adriani DE, Dingkuhn M, Dardou A, Adam H, Luquet D, Lafarge T (2016) Rice panicle plasticity in Near Isogenic Lines carrying a QTL for larger panicle is genotype and environment dependent. Rice 9:28

Bai X, Zhao H, Huang Y, Xie W, Han Z, Zhang B, Guo Z, Yang L, Dong H, Xue W, Li G, Hu G, Hu Y, Xing Y (2016) Genome-Wide Association Analysis Reveals Different Genetic Control in Panicle Architecture Between Indica and Japonica Rice. The Plant Genome 9:plantgenome2015.11.0115

Bommert P, Whipple C (2018) Grass inflorescence architecture and meristem determinacy. Seminars in Cell & Developmental Biology 79:37-47

Bouché F, Lobet G, Tocquin P, Perilleux C (2016) FLOR-ID: an interactive database of flowering-time gene networks in Arabidopsis thaliana, Nucleic Acids Research 44 (D1): D1167-D1171

Butler DG, Cullis BR, Gilmour AR, Gogel BG and Thompson R (2017) ASReml-R Reference Manual Version 4. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, et al. (2010) Linkage and Association Mapping of *Arabidopsis thaliana* Flowering Time in Nature. PLOS Genetics 6(5): e1000940

Bragg J, Tomasi P, Zhang L, Williams T, Wood D, Lovell JT, Healey A, Schmutz J, Bonnette JE, Cheng P, Chanbusarakum L, Juenger T, Tobias CM (2020) Environmentally responsive QTL controlling surface wax load in switchgrass. Theoretical and Applied Genetics 133:3119-3137

Brown PJ, Klein PE, Bortiri E, Acharya CB, Rooney WL, Kresovich S (2006) Inheritance of inflorescence architecture in sorghum. Theoretical and Applied Genetics 113:931-942

Caruso CM (2006) Plasticity of Inflorescence Traits in Lobelia siphilitica (Lobeliaceae) in Response to Soil Water Availability. American journal of botany 93:531-538

Casler MD (2007) Genetic Diversity, Plant Adaptation Regions, and Gene Pools for Switchgrass. Crop science 47:2261-2273

Coen ES, Nugent JM (1994) Evolution of flowers and inflorescences. Development 1994:107

Covarrubias-Pazaran G (2016) Genome-assisted prediction of quantitative traits using the R package sommer. PLOS ONE 11:e0156744

Crowell S, Korniliev P, Falcão A, Ismail A, Gregorio G, Mezey J, McCouch S (2016) Genome-wide association and high-resolution phenotyping link Oryza sativa panicle traits to numerous trait-specific QTL clusters. Nature Communications 7:10527

Dai C, Xue HW (2010) Rice early flowering1, a CKI, phosphorylates DELLA protein SLR1 to negatively regulate gibberellin signalling. EMBO J 29(11):1916-1927

Dai Z, Wang J, Yang X, Lu H, Miao X, Shi Z (2018) Modulation of plant architecture by the miR156f–OsSPL7–OsGH3.8 pathway in rice, Journal of Experimental Botany 69 (21): 5117–5130

Das MK, Taliaferro CM (2009) Genetic variability and interrelationships of seed yield and yield components in switchgrass. Euphytica 167:95-105

Des Marais DL, Hernandez KM, Juenger TE (2013) Genotype-by-environment interaction and plasticity: exploring genomic responses of plants to the abiotic environment. Annual Review of Ecology, Evolution, and Systematics 44:5-29

Dixon LE, Greenwood JR, Bencivenga S, Zhang P, Cockram P, Mellers G, Ramm K, Cavanagh C, Swain SM, Boden SA (2018) *TEOSINTE BRANCHED1* Regulates Inflorescence Architecture and Development in Bread Wheat (*Triticum aestivum*). The Plant Cell 30(3):563–581

Dorken ME, Barrett SCH (2004) Phenotypic plasticity of vegetative and reproductive traits in monoecious and dioecious populations of Sagittaria latifolia (Alismataceae): a clonal aquatic plant. Journal of Ecology 92:32-44

Doust A (2007) Architectural evolution and its implications for domestication in grasses. Ann Bot 100:941-950

Doust AN, Devos KM, Gadberry MD, Gale MD, Kellogg EA (2005) The Genetic Basis for Inflorescence Variation Between Foxtail and Green Millet (Poaceae). Genetics 169:1659

Doust AN, Kellogg EA (2002) Inflorescence diversification in the panicoid "bristle grass" clade (Paniceae, Poaceae): evidence from molecular phylogenies and developmental morphology. American journal of botany 89:1203-1222

El-Soda M, Malosetti M, Zwaan BJ, Koornneef M, Aarts MGM (2014) Genotype × environment interaction QTL mapping in plants: lessons from Arabidopsis. Trends in Plant Science 19:390-398

Friedman J, Harder LD (2004) Inflorescence architecture and wind pollination in six grass species. Functional Ecology 18:851-860

Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK, Doebley JF, Pè ME, Schmidt RJ (2004) The role of barren stalk1 in the architecture of maize. Nature 432(7017):630-5

Glemin S, Bataillon T (2009) A comparative view of the evolution of grasses under domestication. New Phytol 183:273-290

Grabowski PP, Evans J, Daum C, Deshpande S, Barry KW, Kennedy M, Ramstein G, Kaeppler SM, Buell CR, Jiang Y and Casler MD (2017) Genome-wide associations with flowering time in switchgrass using exome-capture sequencing data. New Phytol 213: 154-169

Hohenstein WG, Wright LL (1994) Biomass energy production in the United States: an overview. Biomass and Bioenergy 6:161-173

Hopkins AA (1995) Genotypic Variability and Genotype × Environment Interactions among Switchgrass Accessions from the Midwestern USA. Crop science v. 35:pp. 565-560-1995 v.1935 no.1992

Hori K, Ogiso-Tanaka E, Matsubara K, Yamanouchi U, Ebana K, Yano M (2013) Hd16, a gene for casein kinase I, is involved in the control of rice flowering time by modulating the day-length response. Plant J 76(1):36-46

Kellogg EA (2000) Molecular and morphological evolution in the Andropogoneae, CSIRO, Melbourne, Australia

Leng Y, Xue D, Huang L, Chen L, Ren D, Yang Y, Zhang G, Hu J, Zhu L, Guo L, Lin Y, Qian Q, Zeng D (2017) Mapping QTL with main effect, digenic epistatic and QTL × environment interactions of panicle related traits in rice (Oryza sativa). International Journal of Agriculture and Biology 19:1608-1614

Liu G, Zhang Z, Zhu H, Zhao F, Ding X, Zeng R, Li W, Zhang G (2008) Detection of QTLs with additive effects and additive-by-environment interaction effects on panicle number in rice (Oryza sativa L.) with single-segment substitution lines. Theor Appl Genet 116:923-931

Liu, K, Sun, X, Ning, T et al. (2018) Genetic dissection of wheat panicle traits using linkage analysis and a genome-wide association study. Theor Appl Genet 131: 1073–1090

Lovell JT, Healey A, Schmutz J, Juenger T (2020) Switchgrass v5 4-way (AP13 x DAC, WBC x VS16) genetic map, Dryad, Dataset, <https://doi.org/10.5061/dryad.ghx3ffbjv>

Lovell JT, MacQueen AH, Mamidi S et al. (2021) Genomic mechanisms of climate adaptation in polyploid bioenergy switchgrass. Nature 590: 438–444

Lowry DB, Lovell JT, Zhang L, Bonnette J, Fay PA, Mitchell RB, Lloyd-Reilley J, Boe AR, Wu Y, Rouquette FM, Wynia RL, Weng X, Behrman KD, Healey A, Barry K, Lipzen A, Bauer D, Sharma A, Jenkins J, Schmutz J, Fritschi FB, Juenger TE (2019) QTL × environment interactions underlie adaptive divergence in switchgrass across a large latitudinal gradient. Proceedings of the National Academy of Sciences 116:12933

Mal TK, Lovett-Doust J (2005) Phenotypic plasticity in vegetative and reproductive traits in an invasive weed, Lythrum salicaria (Lythraceae), in response to soil moisture. American journal of botany 92:819-825

Malosetti M, Ribaut J-M, van Eeuwijk FA (2013) The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. Frontiers in Physiology 4

McLaughlin S (1993) New switchgrass biofuels research program for the southeast. In: Proceedings of the annual automative technology development contractors coordinating meeting Nov. 2–5, 1992, Dearborn:111–115

McSteen P (2006) Branching Out: [The ramosa](https://academic.oup.com/plcell/article-abstract/18/3/518/6114806) pathway and the evolution of grass inflorescence morphology. Plant Cell 18:518

McSteen P, Hake S (2001) barren inflorescence2 regulates axillary meristem development in the maize inflorescence. Development 128(15):2881-91

Milano ER, Lowry DB, Juenger TE (2016) The genetic basis of upland/lowland ecotype divergence in switchgrass (*Panicum virgatum*). G3 (Bethesda) 6:3561-3570

Mitchell R, Vogel KP, Uden DR (2012) The feasibility of switchgrass for biofuel production. Biofuels 3:47-59

Miura K, Ikeda M, Matsubara A, Song XJ, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M (2010) OsSPL14 promotes panicle branching and higher grain productivity in rice. Nat Genet 42:545-549

Porter Jr CL (1966) An analysis of variation between upland and lowland switchgrass, *Panicum virgatum* L., in central Oklahoma. Ecology 47:980-992

Price A, Patterson N, Plenge R et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38: 904–909

Price DL (2014) Predictive Relationships between Plant Morphological Traits and Biomass Yield in Switchgrass. Crop science v. 54:pp. 637-630-2014 v.2054 no.2012

Privé F, Luu K, Blum M, McGrath J, Vilhjálmsson B (2020) Efficient toolkit implementing best practices for principal component analysis of population genetic data, Bioinformatics 36(16): 4449–4457

Privé F, Aschard H, Ziyatdinov A, Blum G (2018) Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr, Bioinformatics 34(16): 2781–2787

Robertson GP, Hamilton SK, Barham BL, Dale BE, Izaurralde RC, Jackson RD, Landis DA, Swinton SM, Thelen KD, Tiedje JM (2017) Cellulosic biofuel contributions to a sustainable energy future: Choices and outcomes. Science 356

Studer AJ, Wang H, Doebley JF (2017) Selection During Maize Domestication Targeted a Gene Network Controlling Plant and Inflorescence Architecture. Genetics 207(2):755-765

Su S, Hong J, Chen X, Zhang C, Chen M, Luo Z, Chang S, Bai S, Liang W, Liu Q, Zhang D (2021) Gibberellins orchestrate panicle architecture mediated by DELLA-KNOX signalling in rice. Plant Biotechnol J. doi: 10.1111/pbi.13661

Ta KN, Khong N, Ha TL et al.(2018) A genome-wide association study using a Vietnamese landrace panel of rice (*Oryza sativa*) reveals new QTLs controlling panicle morphological traits. BMC Plant Biol 18:282

Thapa, R, Tabien, RE & Septiningsih, EM (2021) Genome-wide association study to identify chromosomal regions related to panicle architecture in rice (*Oryza sativa*). Genet Resour Crop Evol 68:2849–2865

Urbut SM, Wang G, Carbonetto P et al. (2019) Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. Nat Genet 51:187–195

Van Esbroeck GA (2003) Variation between Alamo and Cave-in-Rock Switchgrass in Response to Photoperiod Extension. Crop science 43

Vogel K (2000) Improving warm-season forage grasses using selection, breeding, and biotechnology. In: Moore KJ, Anderson BE (eds) Native warm-season grasses: research trends and issues 30:83-106

Vogler DW, Peretz S, Stephenson AG (1999) Floral plasticity in an iteroparous plant: the interactive effects of genotype, environment, and ontogeny in Campanula rapunculoides (Campanulaceae). American journal of botany 86:482-494

Vollbrecht E, Springer PS, Goh L, Buckler Iv ES, Martienssen R (2005) Architecture of floral branch systems in maize and related grasses. Nature 436:1119-1126

VSN International (2019) Genstat for Windows 19th Edition. VSN International, Hemel Hempstead, UK

Wadgymar SM, Lowry DB, Gould BA, Byron CN, Mactavish RM, Anderson JT (2017) Identifying targets and agents of selection: innovative methods to evaluate the processes that contribute to local adaptation. Methods in Ecology and Evolution 8:738-749

Wang L, Upadhyaya HD, Zheng J, Liu Y, Singh SK, Gowda CLL, Kumar R, Zhu Y, Wang Y-H and Li J (2021) Genome-Wide Association Mapping Identifies Novel Panicle Morphology Loci and Candidate Genes in Sorghum. Front Plant Sci 12:743838

Wang Y, Li J (2005) The plant architecture of rice (Oryza sativa). Plant molecular biology 59:75-84

Wolfe LM, Mazer SJ (2005) Patterns of Phenotypic Plasticity and Their Fitness Consequences in Wild Radish (Raphanus sativus: Brassicaceae). International Journal of Plant Sciences 166:631-640

Wu Q, Li D, Li D, et al (2015) Overexpression of OsDof12 affects plant architecture in rice (Oryza sativa L.). Front Plant Sci 6:833

Yao W, Li G, Yu Y, Ouyang Y (2018) funRiceGenes dataset for comprehensive understanding and application of rice functional genes. GigaScience 7 (1): 1-9

Zhao J, Mantilla Perez MB, Hu J, Salas Fernandez MG (2016) Genome-Wide Association Study for Nine Plant Architecture Traits in Sorghum. Plant Genome 9(2). doi: 10.3835/plantgenome2015.06.0044.

Zhong, H, Liu, S, Meng, X et al. (2021) Uncovering the genetic mechanisms regulating panicle architecture in rice with GPWAS and GWAS. BMC Genomics 22: 86

Table 1. The latitude, longitude, site code, soil texture, and source of weather data for the ten experimental fields of the four-way and the three overlapping sites (marked as \*) of the diversity panel in the study.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Field Site | Site Code | Latitude | Longitude | Soil Texture | Weather Data Source |
| Brookings, SD | BRKG | 44.307 | -96.67 | Clay loam | https://www.ncdc.noaa.gov/cdo-web/results |
| Hickory Corners, MI | KBSM\* | 42.42 | -85.37 | Loam | https://lter.kbs.msu.edu/datatables/7 |
| Lincoln, NE | LINC | 41.154 | -96.42 | Loam | https://www.ncdc.noaa.gov/cdo-web/results |
| Manhattan, KS | MNHT | 39.141 | -96.64 | Sandy loam | mesonet.k-state.edu/weather/historical |
| Columbia, MO | CLMB\* | 38.897 | -92.22 | Loam | http://agebb.missouri.edu/weather/history/index.asp?station\_prefix=bfd |
| Stillwater, OK | STIL | 35.991 | -97.05 | Sandy loam | https://www.mesonet.org/index.php/weather/local/perk |
| Overton, TX | OVTN | 32.303 | -94.98 | Sandy loam | https://www.ncdc.noaa.gov/cdo-web/results |
| Temple, TX | TMPL | 31.043 | -97.35 | Clay | https://www.ars.usda.gov/plains-area/temple-tx/grassland-soil-and-water-research-laboratory/docs/temple-climatic-data/ |
| Austin, TX | PKLE\* | 30.384 | -97.73 | Clay | https://www.ncdc.noaa.gov/cdo-web/results |
| Kingsville, TX | KING | 27.55 | -97.88 | Sandy clay loam | https://www.ncdc.noaa.gov/cdo-web/results |

Table 2. Broad-sense heritability (*H2*), and its one standard error (±1SE), for panicle length (PL), primary branching number (PBN), and secondary branching number (SBN) at each of the ten field sites (ordered from north to south) for the four-way.

|  |  |  |  |
| --- | --- | --- | --- |
| Sites/Traits | PL | PBN | SBN |
| BRKG | 0.55±0.07 | 0.64±0.06 | 0.40±0.08 |
| KBSM | 0.71±0.06 | 0.62±0.07 | 0.56±0.07 |
| LINC | 0.58±0.07 | 0.65±0.07 | 0.47±0.08 |
| MNHT | 0.38±0.08 | 0.66±0.06 | 0.38±0.08 |
| CLMB | 0.57±0.07 | 0.64±0.07 | 0.15±0.08 |
| STIL | 0.20±0.08 | 0.13±0.08 | 0.02±0.07 |
| OVTN | 0.44±0.08 | 0.63±0.06 | 0.62±0.07 |
| TMPL | 0.49±0.08 | 0.54±0.07 | 0.30±0.08 |
| PKLE | 0.24±0.09 | 0.45±0.09 | 0.29±0.09 |
| KING | 0.47±0.08 | 0.58±0.07 | 0.48±0.08 |

Table 3. The phenotypic correlation between panicle traits within sites. PL, panicle length, PBN, primary branching number, and SBN, secondary branching number for the four-way.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sites |  | Phenotypic correlation | | |  | Sites | | Phenotypic correlation | | |
|  |  | PL | PBN | SBN |  |  | | PL | PBN | SBN |
|  | PL | 1 | - | - |  | | PL | 1 | - | - |
| BRKG | PBN | 0.41 | 1 | - | STIL | | PBN | 0.42 | 1 | - |
|  | SBN | 0.42 | 0.39 | 1 |  | | SBN | 0.42 | 0.57 | 1 |
|  | PL | 1 | - | - |  | | PL | 1 | - | - |
| KBSM | PBN | 0.30 | 1 | - | OVTN | | PBN | 0.52 | 1 | - |
|  | SBN | 0.42 | 0.54 | 1 |  | | SBN | 0.59 | 0.58 | 1 |
|  | PL | 1 | - | - |  | | PL | 1 | - | - |
| LINC | PBN | 0.43 | 1 | - | TMPL | | PBN | 0.35 | 1 | - |
|  | SBN | 0.52 | 0.60 | 1 |  | | SBN | 0.53 | 0.46 | 1 |
|  | PL | 1 | - | - |  | | PL | 1 | - | - |
| MNHT | PBN | 0.56 | 1 | - | PKLE | | PBN | 0.27 | 1 | - |
|  | SBN | 0.68 | 0.67 | 1 |  | | SBN | 0.21 | 0.51 | 1 |
|  | PL | 1 | - | - |  | | PL | 1 | - | - |
| CLMB | PBN | 0.36 | 1 | - | KING | | PBN | 0.52 | 1 | - |
|  | SBN | 0.40 | 0.50 | 1 |  | | SBN | 0.63 | 0.61 | 1 |

Table 4. The identified QTL, along with their marker name (chromosome with physical distance in mega base pair), maximum LOD values, and flanking markers with a LOD drop of 1.5 for panicle morphology traits (PL: panicle length; PBN: number of primary branches; SBN: number of secondary branches) for the four-way. The presence of genotype by environmental interaction is marked as ‘Yes’ or ‘No’ in column QxE.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| trait | QTL | MARKER | LOD | Left flanking marker | Right flanking\_marker | QxE |
| PL | 2K@77.89 | Chr02K\_62.598826 | 6.18 | Chr02K\_60.739957 | Chr02K\_64.045891 | No |
| PL | 3N@62.06 | Chr03N\_26.099983 | 4.29 | Chr03N\_24.521536 | Chr03N\_30.30364 | Yes |
| PL | 4K@26.26 | Chr04K\_13.041487 | 4.66 | Chr04K\_9.916183 | Chr04K\_29.613196 | No |
| PL | 5K@76.02 | Chr05K\_56.620419 | 4.82 | Chr05K\_44.678143 | Chr05K\_58.488157 | No |
| PL | 5N@36.27 | Chr05N\_16.511689 | 3.63 | Chr05N\_11.735767 | Chr05N\_47.154718 | No |
| PL | 6N@54.19 | Chr06N\_48.768076 | 3.56 | Chr06N\_43.871788 | Chr06N\_51.935176 | Yes |
| PL | 9N@38.02 | Chr09N\_18.617122 | 5.82 | Chr09N\_10.880731 | Chr09N\_20.831824 | No |
| PBN | 2K@74.02 | Chr02K\_59.503978 | 4.09 | Chr02K\_56.436103 | Chr02K\_63.664705 | No |
| PBN | 2N@66.12 | Chr02N\_55.500715 | 5.52 | Chr02N\_50.387752 | Chr02N\_56.445418 | No |
| PBN | 3K@38 | Chr03K\_17.77051 | 8.83 | Chr03K\_13.323286 | Chr03K\_20.786505 | Yes |
| PBN | 5K@14.06 | Chr05K\_7.188103 | 4.90 | Chr05K\_4.388419 | Chr05K\_8.204815 | Yes |
| PBN | 5N@84.04 | Chr05N\_64.047349 | 4.73 | Chr05N\_60.974614 | Chr05N\_65.990782 | No |
| PBN | 7N@54.06 | Chr07N\_49.904749 | 4.17 | Chr07N\_49.035214 | Chr07N\_49.904749 | Yes |
| PBN | 9N@26.03 | Chr09N\_12.531268 | 4.89 | Chr09N\_7.913256 | Chr09N\_21.588445 | No |
| SBN | 2N@72.03 | Chr02N\_58.696003 | 9.46 | Chr02N\_54.556579 | Chr02N\_60.798034 | No |
| SBN | 5K@95.5 | Chr05K\_60.232411 | 6.22 | Chr05K\_58.583292 | Chr05K\_60.232411 | No |
| SBN | 9K@51.96 | Chr09K\_24.465322 | 10.37 | Chr09K\_19.959778 | Chr09K\_28.697896 | Yes |
| SBN | 9N@38.02 | Chr09N\_18.617122 | 9.29 | Chr09N\_17.684245 | Chr09N\_19.333648 | No |

**List of figures**

Figure 1. The geographic location and climate at the ten field sites. (a) The ten sites across the latitudinal gradients from southern Texas to North Dakota. The experimental sites in this study span much of the natural range of switchgrass (the green layer with buffered points). (B) The mean temperature and annual rainfall of the ten sites for the study year in 2016 for the four-way (ordered from south to north).

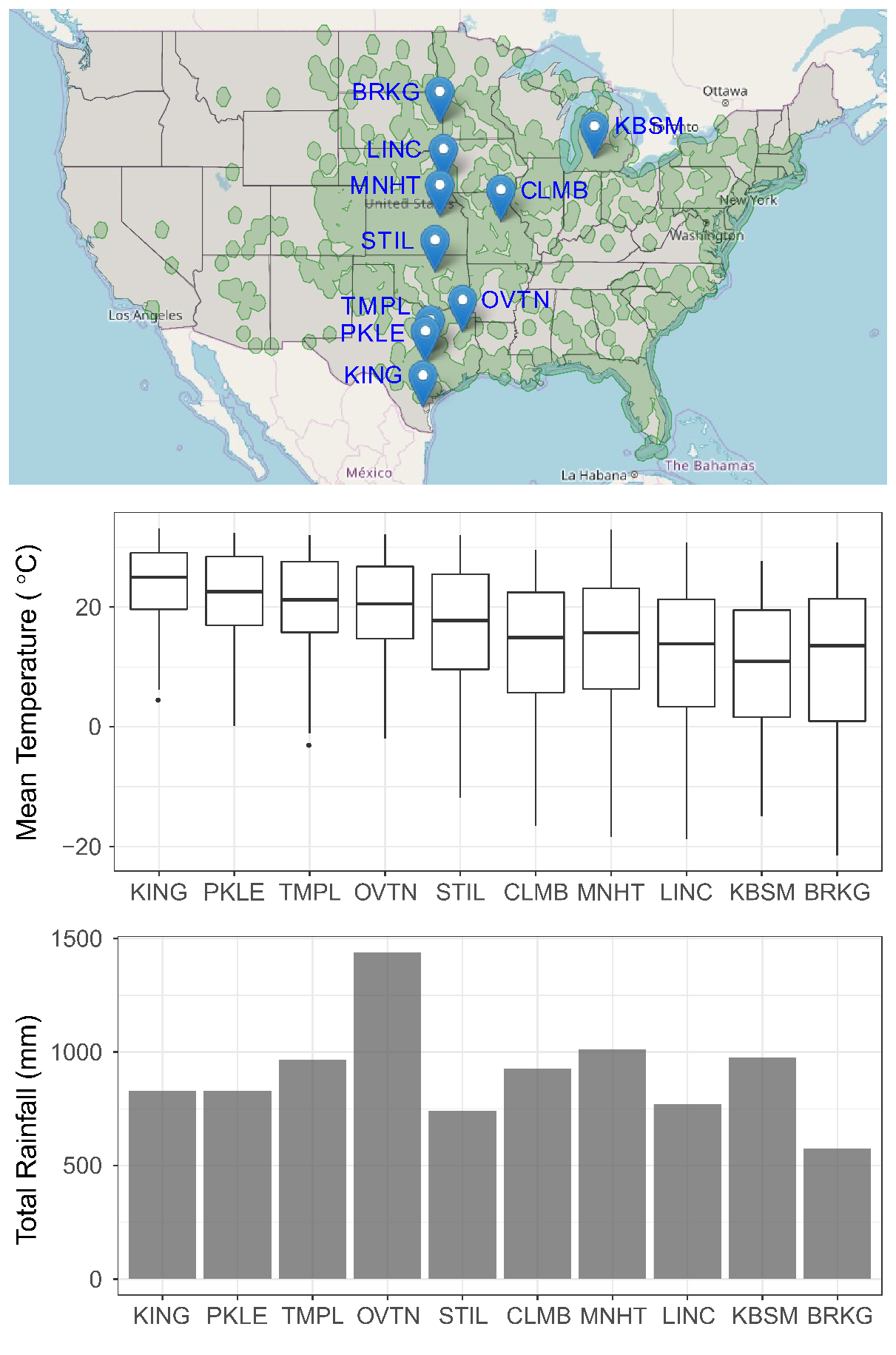
Figure 2. The diagrams exemplifies the paniclesof switchgrass from (a) the four grandparents (WBC, AP13, VS16 and DAC) of the four-way and (b) ten representative genotypes from the diversity panel, and numbers are the IDs for the genotypes. PL is the panicle length (in mm), PBN is the primary branching number, and SBN is the secondary branching number.

Figure 3. The phenotypic distribution of the F2 population, and the phenotypic means of the four grandparents (lowland AP13, WBC and upland DAC and VS16) of the four-way for panicle length (PL in mm), primary branching number (PBN), and secondary branching number (SBN) across the ten field sites (ordered from south to north).

Figure 4. The summary of QTL and significant QTL-by-environment interaction (marked as \* in red) identified from the four-way population for panicle length (PL in mm), primary branching number (PBN), and secondary branching number (SBN).

Figure 5. The additive effects of each QTL identified from the four-way for (a) panicle length (PL in mm), (b) primary branching number (PBN), and (c) secondary branching number (SBN) across geographic regions ordered by from south to north. Genstat reports all of the means as equal when there is not a significant effect of the environment in the model.

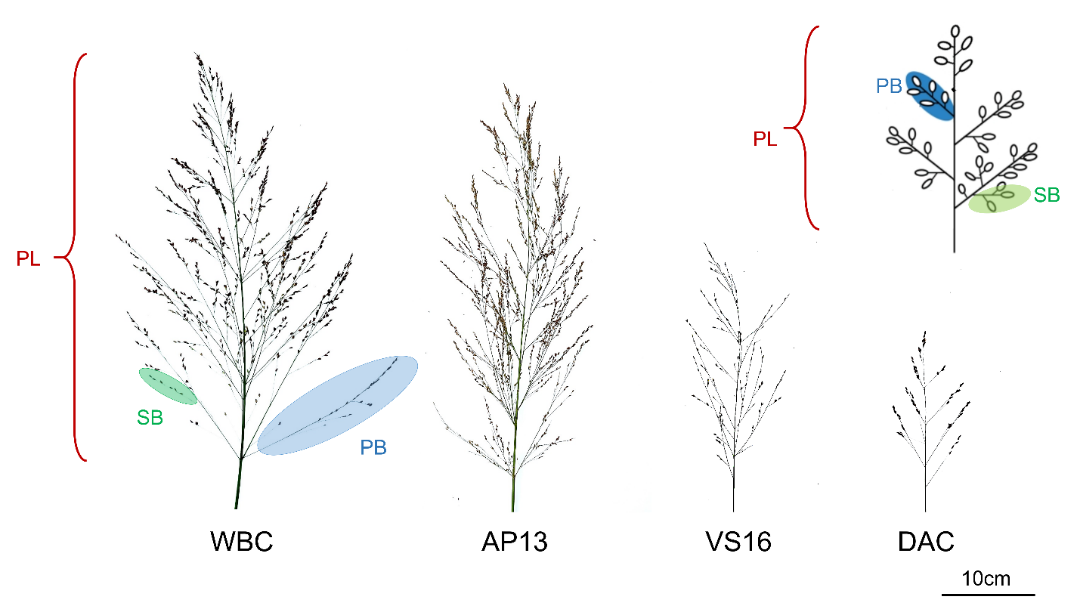
Figure 6. Phenotypic data and genetic effects on panicle traits in a 382 individual diversity panel, analyzed at the three field sites (PKLE, CLMB, and KBSM in Figure 1) using multivariate adaptive shrinkage, or mash. (a) The phenotypic distribution of the diversity panel (violin plots), and the phenotypic means for the three ecotypes (upland, lowland, and costal) present in the panel (colored points and error bars) for panicle length (PL in mm), primary branching number (PBN), and secondary branching number (SBN) across the three field sites (ordered from south to north). (b) Narrow-sense heritability (*h*2) estimates for each panicle trait estimated separately at each site. Error bars represent two times the standard error. (c) Phenotypic correlations for each panicle trait within sites. (d) Posterior weights (mass) on covariance matrices showing different patterns of effects, for the fitted phenotypic effects of 18.7M SNPs estimated using mash. Equal effects matrix has positive covariances of 1 between all trait-site pairs, and the no effects matrix has covariances of 0 between all trait-site pairs. Data-driven matrices are indicated by “ED” and can be seen in Figure S5. (e) Number of QTL with a significant enrichment of partially linked mash SNPs is indicated by the dashed red line; number of permuted genomic regions with significant enrichment of the same SNPs is indicated by the histogram. (f) SNP effects estimated using mash for unlinked SNPs significant above a log10 (Bayes Factor) of 5. Points and error bars represent means and standard deviations estimated by mash, respectively; colors indicate the minor allele frequency for the SNP in the diversity panel.



b

a

Figure 1



a



b

Figure 2

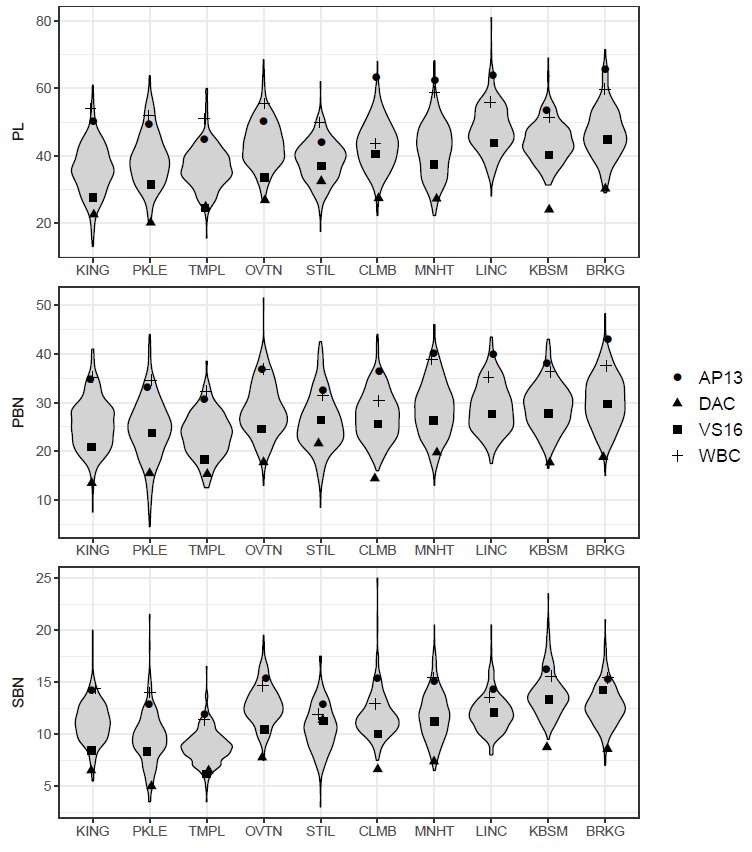


Figure 3

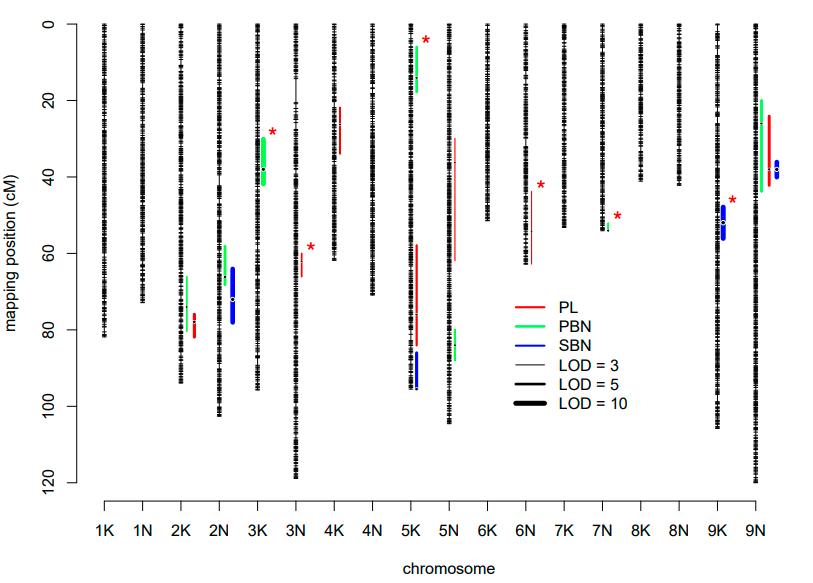


Figure 4

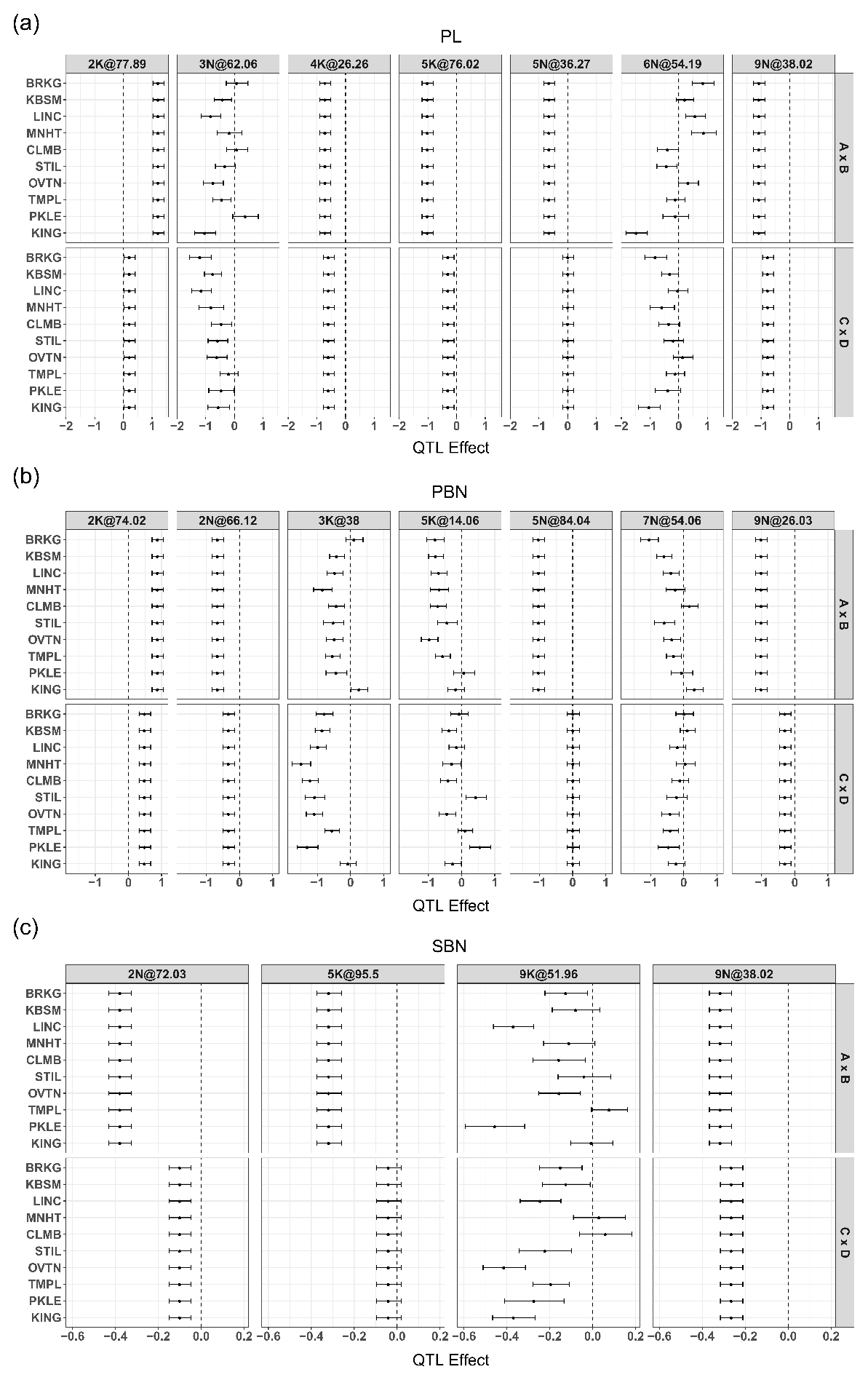


Figure 5

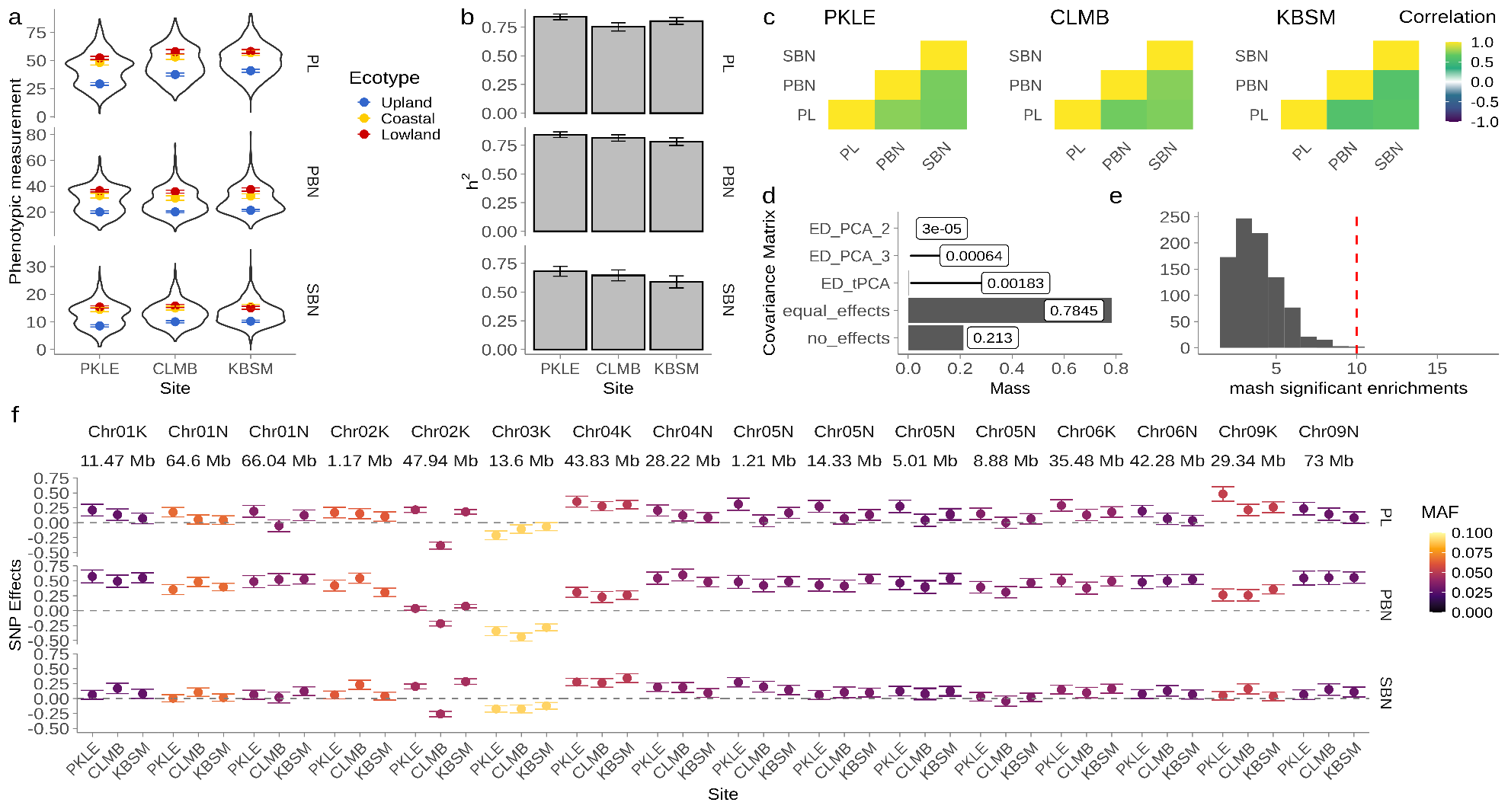


Figure 6