



# Dissecting the Genetic Architecture of Carbon Partitioning in Sorghum Using Multiscale Phenotypes

J. Lucas Boatwright<sup>1,2\*</sup>, Sirjan Sapkota<sup>1</sup>, Matthew Myers<sup>1</sup>, Neeraj Kumar<sup>1,2</sup>, Alex Cox<sup>1</sup>, Kathleen E. Jordan<sup>1</sup> and Stephen Kresovich<sup>1,3</sup>

<sup>1</sup> Advanced Plant Technology, Clemson University, Clemson, SC, United States, <sup>2</sup> Department of Plant and Environmental Sciences, Clemson University, Clemson, SC, United States, <sup>3</sup> Feed the Future Innovation Lab for Crop Improvement, Cornell University, Ithaca, NY, United States

## OPEN ACCESS

### Edited by:

Lewis Lukens,  
University of Guelph, Canada

### Reviewed by:

Cheng-Ruei Lee,  
National Taiwan University, Taiwan  
Robert Klein,  
United States Department of  
Agriculture (USDA), United States

### \*Correspondence:

J. Lucas Boatwright  
jboatw2@clemson.edu

### Specialty section:

This article was submitted to  
Plant Breeding,  
a section of the journal  
Frontiers in Plant Science

Received: 05 October 2021

Accepted: 19 April 2022

Published: 18 May 2022

### Citation:

Boatwright JL, Sapkota S, Myers M,  
Kumar N, Cox A, Jordan KE and  
Kresovich S (2022) Dissecting the  
Genetic Architecture of Carbon  
Partitioning in Sorghum Using  
Multiscale Phenotypes.  
*Front. Plant Sci.* 13:790005.  
doi: 10.3389/fpls.2022.790005

Carbon partitioning in plants may be viewed as a dynamic process composed of the many interactions between sources and sinks. The accumulation and distribution of fixed carbon is not dictated simply by the sink strength and number but is dependent upon the source, pathways, and interactions of the system. As such, the study of carbon partitioning through perturbations to the system or through focus on individual traits may fail to produce actionable developments or a comprehensive understanding of the mechanisms underlying this complex process. Using the recently published sorghum carbon-partitioning panel, we collected both macroscale phenotypic characteristics such as plant height, above-ground biomass, and dry weight along with microscale compositional traits to deconvolute the carbon-partitioning pathways in this multipurpose crop. Multivariate analyses of traits resulted in the identification of numerous loci associated with several distinct carbon-partitioning traits, which putatively regulate sugar content, manganese homeostasis, and nitrate transportation. Using a multivariate adaptive shrinkage approach, we identified several loci associated with multiple traits suggesting that pleiotropic and/or interactive effects may positively influence multiple carbon-partitioning traits, or these overlaps may represent molecular switches mediating basal carbon allocating or partitioning networks. Conversely, we also identify a carbon tradeoff where reduced lignin content is associated with increased sugar content. The results presented here support previous studies demonstrating the convoluted nature of carbon partitioning in sorghum and emphasize the importance of taking a holistic approach to the study of carbon partitioning by utilizing multiscale phenotypes.

**Keywords:** carbon partitioning, linear mixed models, Nested Association Mapping, pleiotropy, source, sink

## 1. INTRODUCTION

The integration of multi-scale phenotypes and appropriate mathematical models can assist in the identification of cross-scale interactions leading to emergent properties of dynamic biological systems (Fischer, 2008; Benes et al., 2020). Indeed, a holistic understanding of complex systems such as plant above-ground biomass and carbon partitioning requires multiscale phenotypes to

address changes in anatomical and physiological processes dictated by underlying genetic networks (Eberius and Lima-Guerra, 2009). The responsiveness of plant carbon-partitioning regimes to environmental conditions such as those induced by a changing climate can significantly affect crop yields and food security thus requiring attention both regionally (Chipanshi et al., 2003; Knox et al., 2012; Meki et al., 2013; Druille et al., 2020) and systemically—particularly under conditions of elevated CO<sub>2</sub>, heat, drought, and other severe-weather events (Michener et al., 1997; Ottman et al., 2001; Pennisi, 2009; Yan et al., 2011). Crops in the Andropogoneae tribe such as maize [*Zea mays* (L.)], miscanthus [*Miscanthus x Giganteus* (Greef et Deuter)], sorghum [*Sorghum bicolor* (L.) Moench], and sugar cane [*Saccharum officinarum* (L.)] have been the focus of continued development to serve as staple and/or energy crops under extreme weather conditions (Lobell and Field, 2007; Kakani et al., 2011; Zegada-Lizarazu et al., 2012; van der Weijde et al., 2013; Fischer et al., 2016) and limit ongoing carbon emissions from fossil fuel use (Heaton et al., 2008; David and Ragauskas, 2010; Brosse et al., 2012; Monti, 2012; Olson et al., 2012; Mullet et al., 2014). These grasses exhibit highly efficient C<sub>4</sub> photosynthetic pathways (Carpita and McCann, 2008; Prasad et al., 2009), leaf-level nitrogen-use efficiency (Gardner et al., 1994; Byrt et al., 2011), water-use efficiency (Kakani et al., 2011; Zegada-Lizarazu et al., 2012; Bhattarai et al., 2019), and high yields (Rooney et al., 2007; Byrt et al., 2011).

Sorghum, in particular, is capable of rapidly accumulating significant quantities of carbon and has been designated as an advanced biofuel feedstock by the U.S. Department of Energy. The Code of Federal Regulations (7 C. F. R. §4288.102) states that advanced biofuels may be derived from biomass in the form of cellulose, hemicellulose, or lignin as well as from sugar or starch (Boatwright et al., 2021). Sorghum meets these conditions as it exhibits great diversity in these carbon-partitioning regimes (Morris et al., 2013), and the sorghum types are further classified based on these traits as cellulosic, forage, grain, or sweet (Boatwright et al., 2021). Sorghum not only meets the requirements as an advanced biofuel feedstock but is capable of rapidly accumulating significant quantities of non-structural (Calviño and Messing, 2012) and structural carbohydrates (Zhao et al., 2009; Mullet et al., 2014; Brenton et al., 2016) necessary for biofuel (Rooney et al., 2007), forage (McCormick et al., 2018), and grain production (Peng et al., 1991). As such, sorghum represents an excellent system for the study of carbon accumulation and partitioning as well as the development of climate-resilient sources of biofuel and calories (Boatwright et al., 2021).

Sucrose is the primary source of energy and carbon in plant sink tissues (Qazi et al., 2012) as well as the primary target for ethanol-based, renewable biofuel production (Rooney et al., 2007; Brenton et al., 2020). Synthesis of sucrose occurs in the leaf cytosol after which it is transported to various sinks including both storage sinks (i.e., stems) in addition to structural vegetative and reproductive organs, which function as growth sinks (Milne et al., 2013; Cooper et al., 2019; Brenton et al., 2020). However, changes in the quantities of structural and non-structural carbohydrates do not occur in a one-to-one manner nor are

they independent (Vietor and Miller, 1990; Billings, 2015). Reduced shoot biomass associated with *dw3* has been shown to decrease grain yield *via* reduced grain size (George-Jaeggli et al., 2011), and differences in carbon partitioning in the stem contribute to tradeoffs between structural and non-structural carbohydrate content (Calviño and Messing, 2012). Carbon partitioning is also subject to environmental conditions such as those that transition plants between growth and reproductive phases as seen under drought conditions (Kakani et al., 2011). A comprehensive examination of the carbon partitioning sinks is necessary to understand the correlations and tradeoffs between these traits in the form of macroscale phenotyping of traits, such as above-ground biomass and plant height, to the microscale assessment of compositional traits using tools such as near infrared spectroscopy (NIR) (Murray et al., 2008; Brenton et al., 2016, 2020).

The sorghum Carbon-Partitioning Nested Association Mapping (CP-NAM) panel (Boatwright et al., 2021) contains 11 subpopulations generated using diverse parental accessions from the sorghum Bioenergy Association Panel (BAP) (Brenton et al., 2016) and the recurrent parent, Grassl—an accession capable of accumulating significant biomass and fermentable carbohydrates per unit time and area (Kresovich et al., 1988). NAM populations contain sets of RIL families generated from the diverse founders, and as such, benefit from recombination of the founder genotypes, high allele richness, higher statistical power, and are less sensitive to genetic heterogeneity (Yu and Buckler, 2006; Boatwright et al., 2021). As the CP-NAM covers the diversity of sorghum types and carbon-partitioning regimes, it represents an excellent source of genotypic and phenotypic diversity to elucidate the genetic architecture underlying carbon fixation, translocation, and utilization so that source/sink dynamics and compositional traits may be understood holistically while simultaneously meeting the demands dictated by a changing environment (Boatwright et al., 2021). Here, we employ quantitative trait locus (QTL) mapping, univariate linear-mixed models (LMMs), and multivariate-response linear-mixed models (MV-LMMs) to identify loci associated with the primary carbon sinks represented by structural and non-structural carbohydrate content in sorghum. Associated loci are then further examined for pleiotropic and interactive effects across these sink-dependent traits as a means of addressing the genetic interplay across the different carbon-partitioning accessions. Using publicly available genomic resources from the sorghum CP-NAM (Boatwright et al., 2021), we identify numerous putative loci involved in carbon partitioning, both known and novel, as well as the extent of pleiotropic effects across these traits. Identification of such hub genes responsible for increased carbon assimilation and partitioning lay the foundation for future network-based approaches to build optimized carbon pathways.

## 2. MATERIALS AND METHODS

### 2.1. Plant Materials and Phenotyping

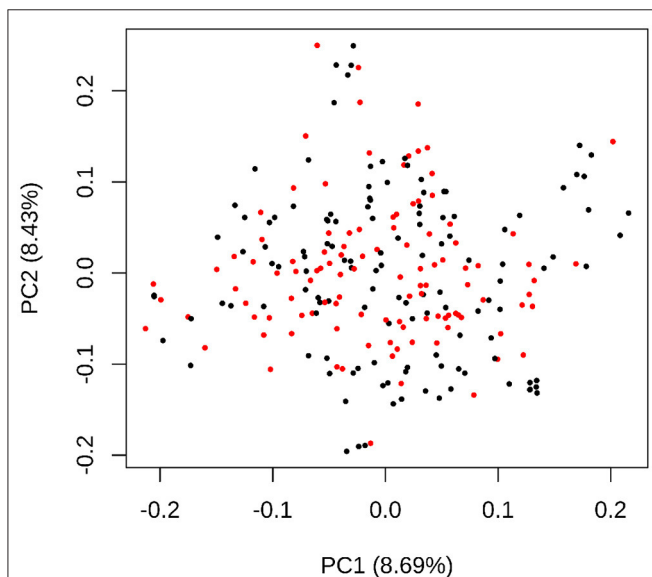
CP-NAM seeds were accessed from the Clemson University sorghum germplasm collection and planted at the Simpson Research Farm (34.64737384683981, -82.74780269784793),

South Carolina in May of 2020. While the original CP-NAM panel contained 2,489 accessions, a subset of 110 individuals were selected from each RIL family using the Partitioning Around Medoids (PAM) function in the R package *cluster* (Schubert and Rousseeuw, 2019) to reduce the field size and manual labor necessary to manage the CP-NAM field while representing most of the genetic diversity within the RIL populations (Guo et al., 2010). A total of 110 sample clusters were identified based on the genomic data from each RIL family with a medoid sample centrally located in each cluster, visualized here using multidimensional scaling (Figure 1). The medoid is the individual within each cluster that best represents the genetic diversity of that sample cluster. The 110 accessions representing the medoids for each population were selected as representatives of each cluster, resulting in 1,210 accessions across the CP-NAM, and planted for phenotyping in 2020.

The NAM RILs were planted in single-row plots that were 3.04 m in length with 0.762 m between-row spacing in a randomized complete block design with two replicates per line. Randomization was done within blocks containing RILs from a given family, and families were planted together to avoid large field effects within families. Plants were irrigated on an as-needed basis but did not occur past 90 days after planting due to plants exceeding the irrigation pivot height. Harvesting started September 14th and continued through October 19th. Due to the range in harvesting days, the maturity stage and days

to harvest were recorded for each plot to use as covariates for phenotypic and genomic analysis to avoid the confounding effect of varying maturity groups across the RIL families. Phenotypes collected included above-ground biomass, stand count, maturity at harvest, days to harvest (DTH), plant height, and dry and wet weights. Stand count was measured as the total number of emerged seedlings between 15 and 30 days after planting. Plant height was measured at harvest from the base of the stalk to the apex of the panicle, or, if no panicle was present, to the apex of the shoot apical meristem.

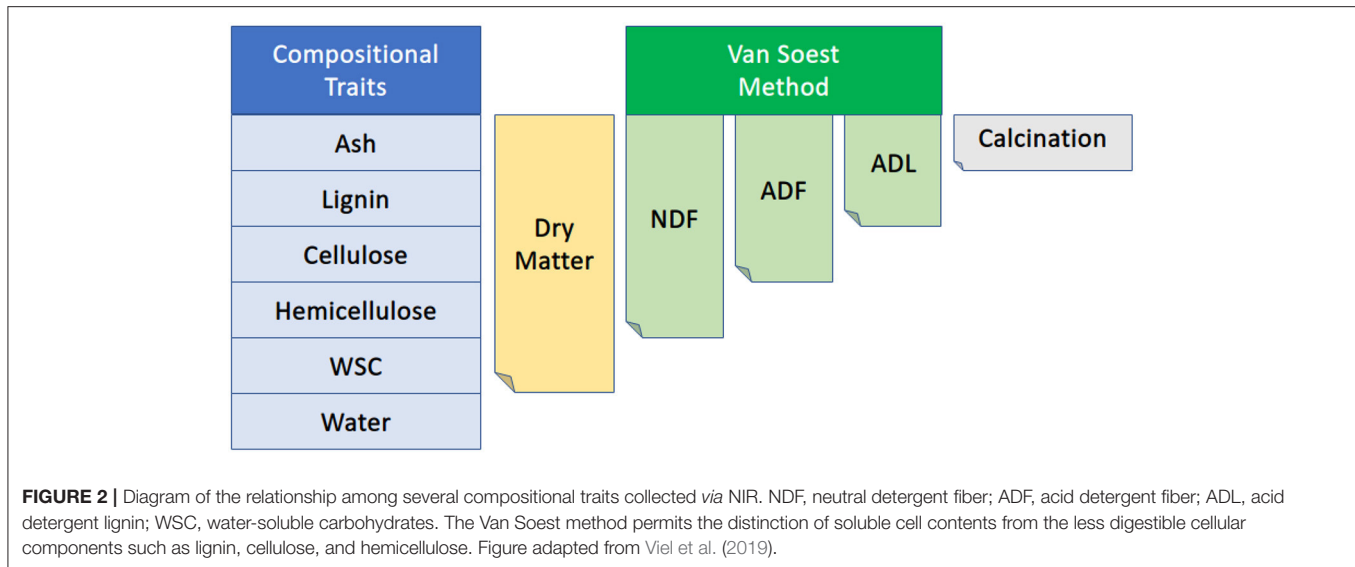
A representative meter was selected for each plot, and all plants were cut at the base within that meter and weighed (in kilograms). To remove the confounding effects of tillering on a per area basis, three representative plants were selected for subsequent phenotyping including above-ground biomass (including panicles and leaves), wet weight, and dry weight, where biomass represents a per area measure of above-ground plant weight, i.e., scaled meter weight. Based on planting density, this represents approximately 0.5 m of row length. Any panicles or partially formed panicles were removed along with all leaf matter before collecting wet weight. The stalks were then cut into billets for collection into mesh drying bags and placed into drying bins at 40°C until stalks were dried to a constant moisture content. Dried stalks were removed from the drying bins, and dry weights were taken. Stalks were then ground with a Retsch SM 300 cutting mill so that compositional traits could be measured using a PerkinElmer DA7250™ NIR instrument (<https://www.perkinelmer.com>), which uses calibration curves for spectral measurements built using wet chemistry values generated by Dairyland Laboratories, Inc. (Arcadia, WI, USA) as described in Brenton et al. (2016). These wet chemistry values were generated on accessions from the BAP grown in South Carolina from which the founders of this population were selected. Wet chemistry estimates were strongly associated with NIR estimates for the compositional traits examined here (Supplementary Table 1). Compositional traits include acid detergent fiber (ADF), adjusted crude protein (Adj. CP), neutral detergent fiber (NDF), ash-free NDF (aNDFom), ash, calcium, chloride, dietary cation-anion difference (DCAD), dry matter, potassium, lignin, magnesium, moisture, net energy growth (NEG), net energy lactation (NEL), relative feed value (RFV), non-fiber carbohydrates (NFC), and water-soluble carbohydrates (WSC) where all traits are expressed as a percent of dry matter. NEG, NEL, NEM, and TDN were also estimated using an Ohio Agricultural Research & Development Center (OARDC) summative energy equation and may appear conjugated with the OARDC abbreviation (Figure 2).



**FIGURE 1** | A multidimensional scaling plot using the genotypic data of PI229841 RIL accessions where the 110 medoid samples are colored red and the remaining samples are black. The x-axis represents the first principal component of the genotypic data, and the y-axis represents the second principal component, with variance explained by each component in parentheses. The proximity of accessions to each other indicates the approximate degree of similarity—with shorter distances indicating higher similarity. The MDS distribution represents a two-dimensional projection of the genetic diversity used for clustering.

## 2.2. Descriptive Statistics and QTL Mapping

The repeatability was estimated for all traits using the package *heritability* in the R programming language (R Core Team, 2019; Kruijer et al., 2015). The best linear unbiased predictors (BLUPs) were calculated for each trait with the R package *lme4* (Bates et al., 2015) using the *lmer* function with genotype as random effects. The resulting BLUPs were used as adjusted phenotypic values for all mapping and association analyses. Heatmaps and correlation



metrics were measured using the *seaborn* (Waskom, 2021) and *pandas* (McKinney, 2010) packages in CPython (Van Rossum and Drake, 2009), respectively.

Genotypic data from Boatwright et al. (2021) were used to perform QTL mapping and genome-wide association studies (GWAS). In summary, Genotype-By-Sequencing (GBS) data were generated using a double-digest approach (*PstI* and *MspI*), processed using the Tassel GBS version 2 pipeline (Glaubitz et al., 2014), and imputed using Beagle 5.1 (Browning et al., 2018) as described in Boatwright et al. (2021). The complete CP-NAM contains 144,087 SNPs across 11 RIL families. Given sorghum's 730 Mb genome, this corresponds to a density of approximately one variant per 5 kb. The linkage disequilibrium was shown to decay to 0.2 around 100 kb, indicating that these GBS data should capture the majority of independent haploblocks in this population (Boatwright et al., 2021). Genetic maps were constructed in Boatwright et al. (2021) for each RIL family with Haldane's mapping function (Kosambi, 2016) and a genotyping error rate of 0.0001, where the conditional probabilities of the true genotypes were estimated using a hidden Markov model. Here, we performed QTL mapping for each RIL family using the *qtl2* package (Broman et al., 2019) in R, and genomic scans were performed using Haley-Knott regression (Haley and Knott, 1992) and linear mixed model approaches (Broman et al., 2019) including both full and leave-one-chromosome-out models. All models included maturity and DTH as covariates except for DTH. QTL effects were estimated as  $100 \times (1 - 10^{-(2 \times LOD)/n})$ , where  $n$  is the number of individuals in the corresponding mapping population (Broman and Sen, 2009). However, we recognize that estimates for PVE with these population sizes may exhibit inflated values due to the Beavis effect (King and Long, 2017).

### 2.3. Genome-Wide Association

GWAS were done using both the Genome-wide Efficient Mixed Model Association (GEMMA) program version 0.98.3

(Zhou and Stephens, 2014) and the Genome Association and Prediction Integrated Tool (GAPIT) version 3 (Lipka et al., 2012; Wang and Zhang, 2021). We specifically used GEMMA for both MV-LMMs and Bayesian Sparse Linear Mixed Models (BSLMM) (Zhou et al., 2013) while GAPIT was used for BLINK (Huang et al., 2019) and LMMs. The phenotypic and genotypic data were converted to Plink format using Plink [v1.90b6.10] (Purcell et al., 2007). The genomic relatedness matrix was calculated using the VanRaden algorithm (VanRaden, 2008) and all models were run using a MAF filter of 0.01 for all 1,210 CP-NAM accessions. Univariate LMMs are of the following the form,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon};$$

$$\mathbf{u} \sim N(\mathbf{0}, \mathbf{G}) \text{ and } \boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{R}),$$

where  $\mathbf{y}$  is a vector of phenotypic values for a single trait,  $\mathbf{X}$  is the numeric genotype matrix generated from the SNPs,  $\boldsymbol{\beta}$  represents the unknown vector of fixed effects representing the effect size for each SNPs,  $\mathbf{Z}$  is the design matrix for random effects,  $\mathbf{u}$  is the unknown vector of random effects, and  $\boldsymbol{\epsilon}$  is the unknown vector of residuals. These univariate models test the alternative hypothesis  $H_1: \boldsymbol{\beta} \neq \mathbf{0}$  against the null hypothesis  $H_0: \boldsymbol{\beta} = \mathbf{0}$  for each SNP.

In addition to the frequentist univariate model described above which estimates fixed effect coefficients by selecting the optimal value minimizing the least-squared error—the equivalent of a flat prior, we ran a BSLMM which assumes fixed effects are distributed according to the sparse prior,  $\boldsymbol{\beta} \sim \pi N(\mathbf{0}, \sigma^2 \boldsymbol{\tau}^{-1}) + (1 - \pi) \delta_0$  (Zhou et al., 2013). Runs were executed using 20e6 sampling steps with a burn-in of 5e6, and the posterior inclusion probability (PIP) threshold established as 0.036 based upon a 99.95% quantile from simulated data sets across quantiles to determine the empirical significance cutoff (Sapkota et al., 2020). While more computationally intensive due

to the Markov Chain Monte Carlo sampling approach involved, this model provides shrinkage of  $\beta$  estimates to control for type I errors and provides a posterior distribution of plausible values rather than simple point estimates. Additionally, we also conducted univariate analyses using BLINK (Huang et al., 2019). BLINK approximates the maximum likelihood approach used by LMMs, instead using Bayesian Information Criteria in a fixed effect model where each SNP is iteratively associated with the phenotype of interest. Markers in linkage disequilibrium (LD) with the most significant marker are then excluded—as estimated using a Pearson correlation coefficient  $\geq 0.7$ . For subsequent markers, the next most significant SNP is selected, and the exclusion process is conducted in the same way until no markers can be excluded. Unlike SUPER and FarmCPU methods, BLINK does not assume that causal genes are evenly distributed across the genome and is faster with higher statistical power—due to its multi-locus approach—and lower type I error rates (Huang et al., 2019).

For MV-LMMs, we used GEMMA models of the form:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{U} + \mathbf{E};$$

$$\mathbf{U} \sim MVN_{n \times d}(0, \mathbf{K}, \mathbf{V}_g) \text{ and } \mathbf{E} \sim MVN_{n \times d}(0, \mathbf{I}_{n \times n}, \mathbf{V}_e),$$

where  $\mathbf{Y}$  is an  $n$  by  $d$  matrix of  $d$  phenotypes for  $n$  individuals,  $\mathbf{X}$  is the numeric genotype matrix generated from the SNPs,  $\boldsymbol{\beta}$  is a  $d$  vector of fixed effects representing the effect size for the  $d$  phenotypes,  $\mathbf{Z}$  is the design matrix for random effects,  $\mathbf{U}$  is the  $n$  by  $d$  matrix of random effects,  $\mathbf{E}$  is the  $n$  by  $d$  matrix of residuals,  $\mathbf{K}$  is a known  $n$  by  $n$  relatedness matrix,  $\mathbf{V}_g$  is a  $d$  by  $d$  symmetric matrix of genetic variance components,  $\mathbf{I}_{n \times n}$  is an  $n$  by  $n$  identity matrix, and  $\mathbf{V}_e$  is a  $d$  by  $d$  symmetric matrix of environmental variance components. As the maturity of accessions significantly affects all phenotypes, maturity and DTH were used as covariates in all QTL mapping and GWAS models except for DTH.

## 2.4. Pleiotropic and Epistatic Tests

The estimated effect sizes and standard errors for every marker in the LMMs for ADF, ash, dry matter, dry weight, height, NDF, P, wet weight, and WSC were filtered using a local false sign rate  $< 0.1$  based on a condition-by-condition analysis using ashR in R (Stephens, 2017). A control set of estimated effects and standard errors were also randomly selected for 40,000 markers to estimate the covariance between SNPs for each phenotype. A correlation matrix of the random control set was estimated and used to control for the confounding effects of correlated variation among the traits using mashr in R (Urbut et al., 2019). We utilized both canonical and data-driven covariance matrices following mashr best practices to test for pleiotropy across traits (Urbut et al., 2019). The posterior probabilities were calculated for each SNP by fitting a mash model with all tests. Bayes factors were extracted and plotted from mash results using the CDBNgenomics R package (MacQueen et al., 2020). Variants exhibiting Bayes Factors greater than 10 were considered as demonstrating significant pleiotropic effects.

Using PLINK v1.90b6.10, we performed tests for epistasis using loci exhibiting pleiotropic effects as described above. PLINK uses a linear regression in the form of:

$$\mathbf{Y} \sim \boldsymbol{\beta}_0 + \mathbf{A}\boldsymbol{\beta}_1 + \mathbf{B}\boldsymbol{\beta}_2 + \mathbf{AB}\boldsymbol{\beta}_3 + \boldsymbol{\epsilon},$$

where  $\mathbf{Y}$  is the matrix of phenotypes across the 1,210 accessions, and  $\mathbf{A}$  and  $\mathbf{B}$  represent the allele dosages of SNPs A and B, respectively. The coefficients  $\boldsymbol{\beta}_{0-3}$  represent the mean, estimated effects of SNP A, SNP B, or their interaction, respectively, while  $\boldsymbol{\epsilon}$  represents the residual deviations. Pairwise tests for interaction are based on the coefficients estimated for  $\boldsymbol{\beta}_3$ . Significant loci were filtering using a Chi-square statistic  $> 80$ , which is slightly more stringent filtering than Bonferroni correction for these data. Both the Bayes Factors from the pleiotropic analysis and the epistatic results were plotted using Circos v0.69-8 (Schoelz et al., 2021).

## 3. RESULTS

### 3.1. Trait Heritabilities and Correlations

Heritability is the proportion of phenotypic variance attributable to genetic variance, and when differences between genotypes is assumed to derive entirely from genetic effects, the measurement of consistent individual differences is called repeatability (Kruijer et al., 2015). As such, repeatability includes genetic and environmental sources of variation, thereby providing an upper bound for broad-sense heritability. We calculated the repeatability for all traits and identified many traits with repeatability greater than 0.2 (Table 1 and Supplementary Tables 2, 3) with maturity and DTH exhibiting the highest repeatabilities ( $> 0.75$ ). Agronomic phenotypes exhibited higher repeatability compared to compositional traits with all agronomic traits exceeding 0.5 repeatability. Agronomic traits also demonstrated relatively low correlation among the traits except for biomass, wet weight, and dry weight which were all highly correlated ( $\geq 0.67$ ) (Figure 3). These measures for repeatability and correlation among traits are consistent with previous estimates in sorghum (Brenton et al., 2016).

Many compositional traits exhibited strong correlation ( $> |0.5|$ ) (Figure 4), which is expected due to the aggregate nature of some traits and their dependency on maturity (Figure 2). Importantly, while many of the fiber-based compositional traits exhibited strong repeatability, only six compositional traits of 34 had Pearson's correlation coefficients  $> |0.3|$  with dry weight and none had values exceeding  $|0.5|$ . The lack of strong correlation between compositional traits and dry weight suggests that sorghum composition could be improved without significantly affecting total vegetative yield (Murray et al., 2008; Brenton et al., 2016).

### 3.2. Mapping and Associations

#### 3.2.1. Agronomic Traits

As the CP-NAM is composed of 11 RIL families, it provides the opportunity to resolve genotype-to-phenotype associations through both QTL mapping and GWAS. To this end, QTL mapping was performed using 110 accessions for each RIL

**TABLE 1** | Repeatability for agronomic (top portion of table) and compositional (bottom portion) traits.

Trait	Repeatability
Maturity	0.81
Plant height	0.53
Wet weight	0.52
Dry weight	0.55
Biomass	0.52
DTH	0.76
ADF	0.65
Adj. CP	0.38
aNDFom	0.61
Crude protein	0.35
Dry Matter	0.36
Lignin	0.45
NDF	0.55
NEG (OARDC)	0.59
WSC	0.47
TDN (ADF)	0.62

Agronomic traits include maturity at harvest, plant height, wet weight, dry weight, above-ground biomass, and days to harvest. Compositional traits include acid detergent fiber (ADF), adjusted crude protein (Adj. CP), ash-free neutral detergent fiber (aNDFom), crude protein, dry matter, lignin, neutral detergent fiber (NDF), net energy growth (NEG), total digestible nutrients (TDN) as a percentage of ADF, and water-soluble carbohydrates (WSC). NEG was estimated using an Ohio Agricultural Research & Development Center (OARDC) summative energy equation.

family in the CP-NAM for every trait using maturity and DTH as covariates except when DTH is the response variable. We identified 59 QTL across the 11 RIL families for the agronomic traits (**Supplementary Table 4**). Several known QTL were identified for plant height on chromosomes (Chr) 6 [*qHT7.1/Dw2*], Chr7 [*Dw3*], and Chr9 [*Dw1*] aggregated by RIL families derived from grain, cellulosic, sweet, and forage parents (**Figure 5** and **Supplementary Tables 4, 5**) along with several potentially novel associations on Chr1 and Chr8. The newly identified QTL on Chr1 spanned from 10 to 12 Mb (13.2 PVE) and from 22 to 56.7 Mb (14.5 PVE) (**Supplementary Table 4**). The QTL from the latter position also had a significant genome-wide association for height using the BLINK model with a peak at Chr1:50,888,855 (**Supplementary Figure 1** and **Supplementary Table 6**). Another novel QTL was identified for height on Chr8 from approximately 0.37 to 2.7 Mb (**Supplementary Table 4**) using a leave-one-chromosome-out method (**Figure 5C** and **Supplementary Figure 2**). A significant genome-wide association for height was found for the SNP at Chr8:2,033,695 using the BSLMM model (**Supplementary Figure 3**), and the associated region is within previously identified QTL for transpiration rate and efficiency of energy of PSII (Ortiz et al., 2017).

Sorghum has over 40 identified flowering time and maturity QTL (Mace et al., 2013). QTL mapping results for DTH in PI506069 RILs identified a locus on Chr4 (11.8 PVE) from 70 to 113 cM that peaked at 79 cM (**Figure 5B**). This QTL colocalizes with the flowering time gene *CN2* (Marla et al., 2019),

which is a centroradialis-like gene homologous to Terminal Flower1 (*TFL1*). An additional 11 loci were identified using BLINK including *Ma3/Ma5* [Chr1], *SbCN12* [Chr3], and *Ma1* [Chr6] (Yang et al., 2014) along with several other unidentified loci (**Supplementary Figure 4**). The identified loci include phytochromes and other flowering time modulators that mediate photoperiod sensitivity in these non-temperately adapted accessions.

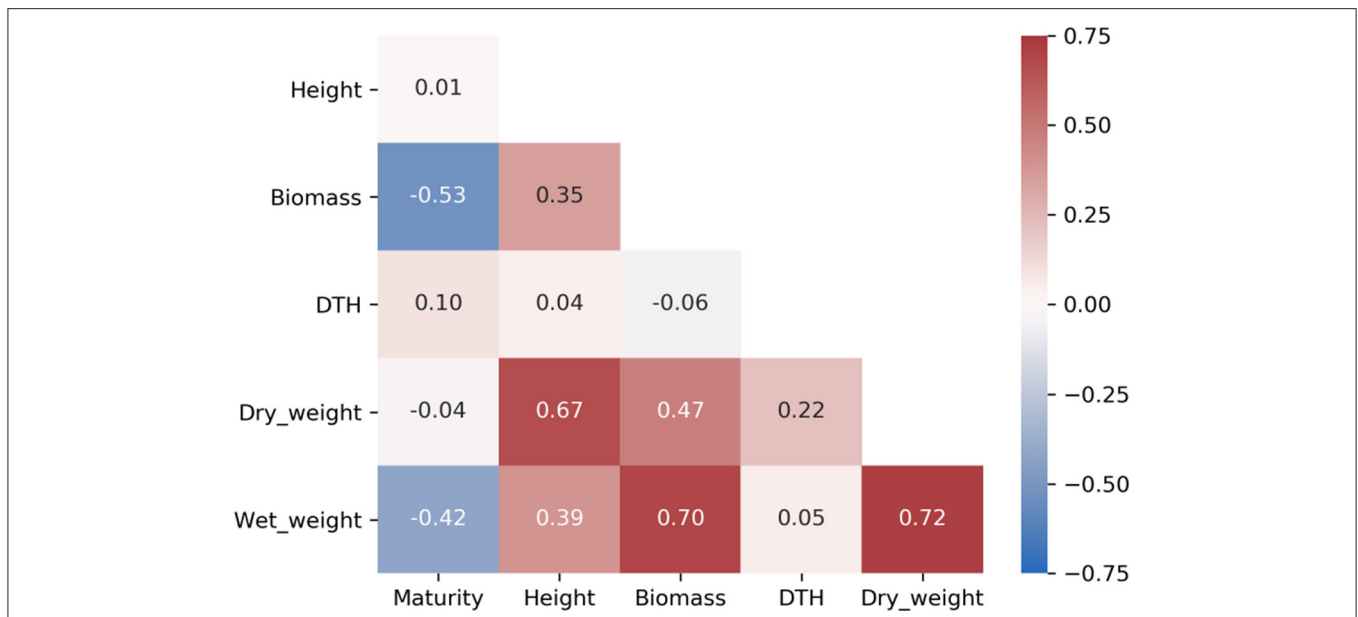
### 3.2.2. Biomass Traits

In addition to height and DTH phenotypes, various measures of biomass yield were taken including wet weight, dry weight, and above-ground biomass (abbreviated as biomass). These biomass traits were often associated with the same QTL—particularly the QTL on Chr6 and Chr7 (**Figure 5**), but significant associations from GWAS were more variable (**Figure 6** and **Supplementary Table 5**). The QTL on Chr3 (12.7 PVE) identified using wet weight spans from approximately 1–6 Mb in the sweet x cellulosic RILs of PI586454 (**Figure 5C**) and coincides with QTL associated with stem circumference and transpiration rate (Zhao et al., 2016; Ortiz et al., 2017). GWAS of wet weight also identified an association on Chr3 at approximately 62 Mb (**Figure 6D**), which colocalizes with numerous trait associations including plant height (Bouchet et al., 2017), stem circumference (Zhao et al., 2016), and days to flowering (Kong et al., 2018). However, the gene(s) mediating these phenotypes is unclear. Dry weight and wet weight were associated with several QTL on Chr6 (**Figures 5A,D**) that were also captured through GWAS (**Figure 6**) and ranged from 1 to 5 Mb and 49 to 51 Mb, respectively. The QTL spanning 1 to 5 Mb corresponds to the known maturity locus, *Ma6* (Murphy et al., 2014). These phenotypes also captured the height loci *Dw2*, which encodes a protein kinase that regulates stem internode length (Hilley et al., 2017), and *Dw3*, which encode a P-glycoprotein auxin transporter and only affects plant height below the flag leaf (Li et al., 2015). As auxin stimulates the production of hemicellulose and consequently stem elongation, the association is consistent with the known identity of the locus (Li et al., 2015).

The forage RIL family, CP-NAM PI655972, was uniquely suited for identifying a QTL (14.3 PVE) controlling biomass content on chromosome 5 (**Figure 5D**). The biomass QTL also overlapped a QTL for adjusted crude protein content (**Supplementary Table 4**). Ritter et al. (2008) identified a sucrose content QTL that falls completely within the biomass QTL and partially overlaps the adjusted crude protein QTL seen here (Ritter et al., 2008). Given the large range of the QTL or even the overlapping region, it is difficult to pin down what gene(s) may be responsible for these associations. In addition to this unique locus, biomass was associated with the same loci on Chr6 and Chr7 as height, wet weight, and dry weight (**Figure 6**).

### 3.2.3. Compositional Traits

QTL mapping was also performed for all compositional traits, and select traits were plotted for all RIL families (**Figure 7**). As described in the methods, NIR and wet chemistry estimates of these compositional traits demonstrated strong correlations



**FIGURE 3** | Heatmap of Pearson's correlation coefficients among agronomic phenotypes. Biomass represents a scaled meter weight, and DTH is days to harvest.

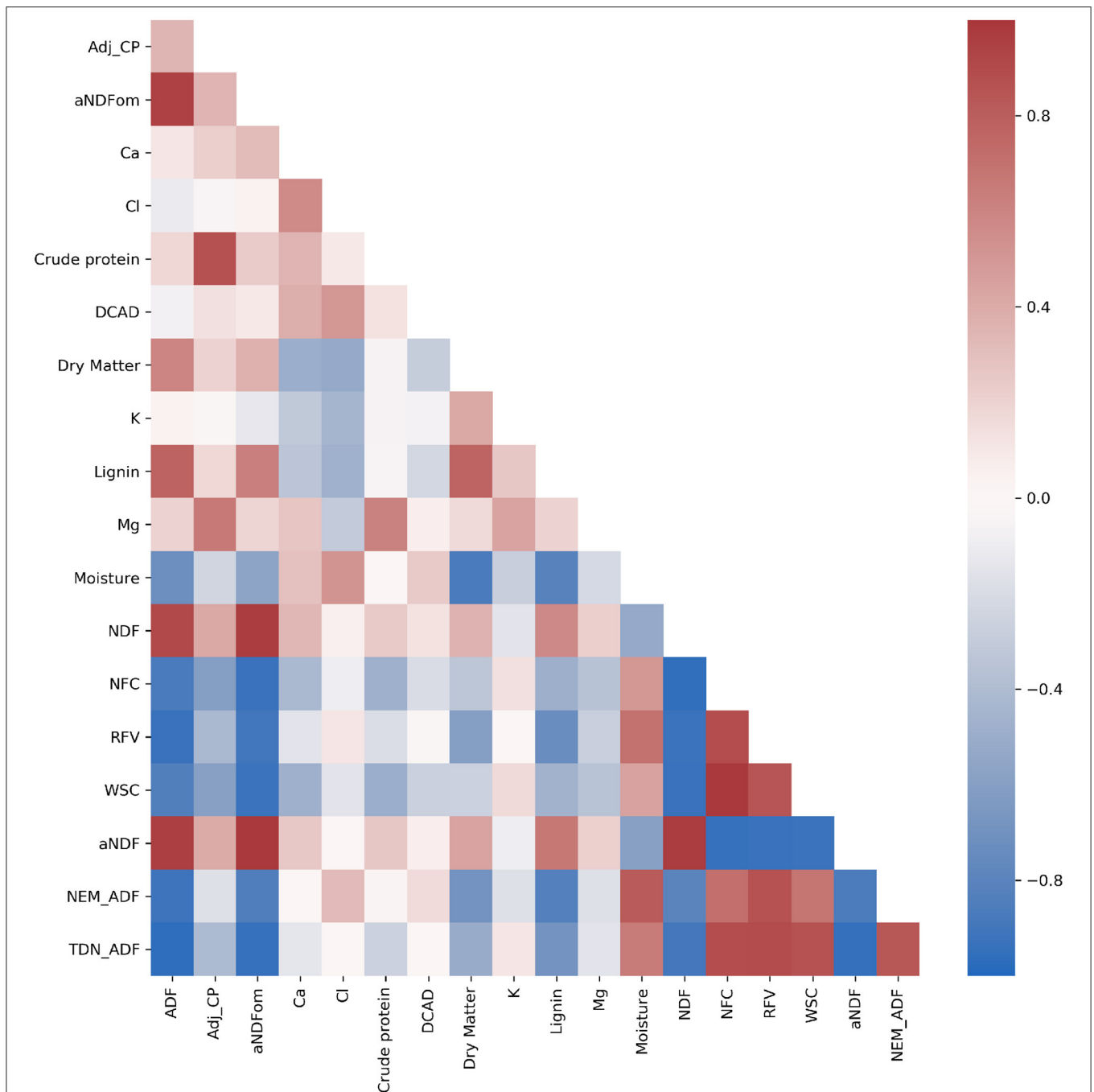
across a highly diverse panel from which the founders were derived. We identified 522 QTL across 34 compositional traits and the 11 RIL families (**Supplementary Table 7**). Several overlapping QTL were identified across traits within RIL families, and the various sorghum RIL families/types captured different QTL for the same traits. The most significant QTL on Chr6 associated with ADF was consistently identified in all RIL families. The narrowest range of this locus was obtained in PI229841 and PI508366 RILs and spanned from approximately 50.3 to 51.8 Mb (**Supplementary Table 7**).

WSC content provides an estimate of the carbon partitioned and accumulated by accessions in the stem in the form of water-soluble carbohydrates (Brenton et al., 2020). For WSC, we identified a QTL on Chr6 in the PI22913 RILs around 50 Mb, which results from a cross between the sweet sorghum accession PI22913 and the cellulosic Grassl (**Figure 7A**). The QTL occurs within the Dry Midrib (*D*) locus, which has also shown strong association with midrib color, grain yield, sugar yield, juice volume, and biomass, indicative of a pleiotropic effect of the *D* locus across these phenotypes (Burks et al., 2015; Xia et al., 2018). This QTL also overlaps with QTL identified using ADF, wet weight and dry weight phenotypes described above. Burks et al. (2015) previously demonstrated that green midrib color was more strongly associated with sugar content traits than the *D* locus genotypic data with sugar content, and therefore suggested that selecting for green midribs is a simple alternative to genetic selection for sweet sorghum breeding programs. Consistent with this observation, the *D* locus accounted for 64.2% of the variance explained for WSC in PI229841 RILs.

Following the design pattern indicated in **Figure 2**, we used a combination of LMMs with various compositional traits as covariates to deconvolute the contribution of individual

traits to phenotypic variance, and as a converse approach, we also ran multivariate-response models on constituent parts to compare with composite traits. While population genomic studies have historically utilized univariate LMMs, more recent works are finding that multivariate-response linear mixed models (MV-LMM) have higher true-positive rates particularly when correlated traits with low, medium, and high heritabilities are analyzed together in one MV-LMM (Rice et al., 2020). The use of MV-LMMs may also provide additional power to detect causal loci exhibiting pleiotropic effects across multiple traits (Mural et al., 2021). By using MV-LMMs on combinations of carbon-partitioning traits, we can better understand the interplay among these traits and predict the systemic effects of trait selection on the respective carbon sinks.

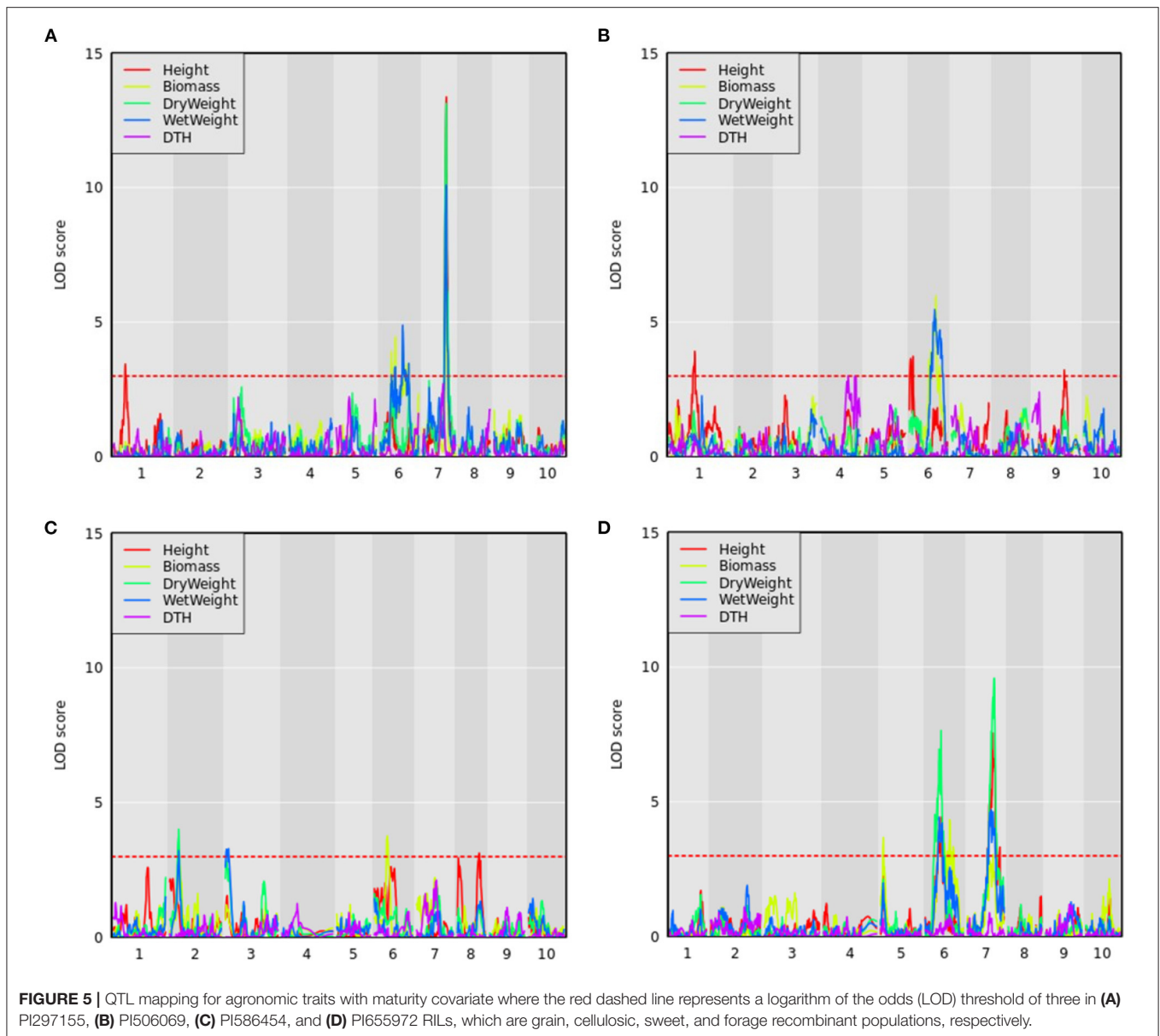
Running WSC and NDF in a multivariate-response model with maturity and DTH as covariates approximates the dry matter phenotypic variance (**Figures 2, 8**). While only one locus has significant associations, the highly significant association on Chr6 occurs broadly from approximately 49.5–52Mb with the three most significant SNPs (50,556,927; 50,558,124; and 50,574,062). As noted from the QTL mapping results for WSC, this associated corresponds to the *D* locus. The most significant SNP (Chr6:50,558,124) exhibited considerable phenotypic variation across traits but contrasting effects for NDF and WSC for each allele (**Supplementary Figures 5, 6**). The identity of the gene underlying this locus is believed to be a NAC transcription factor where recessive parents possess a premature stop codon in the NAC domain and were shown to exhibit lower lignin content but higher sugar and grain yields (Xia et al., 2018) similar to the relationship between NDF and WSC seen here. Running the same model but adding the top SNP as a covariate brings the peak on Chr4 above the significance



threshold (**Supplementary Figures 7, 8**). This SNP overlaps a QTL previously identified with dry matter growth rate, leaf appearance rate (Fiedler et al., 2014), and stem circumference (Zhao et al., 2016). Potential candidate genes in the region include two high-affinity nitrate transporter (NRT) genes

(Sobic.004G009400/Sobic.004G009500). In sorghum, increased expression of NRTs has been suggested to improve the efficiency at which inorganic and organic nitrogen is assimilated (Gelli et al., 2014) and affect both the biomass and grain yield (Gelli et al., 2017).



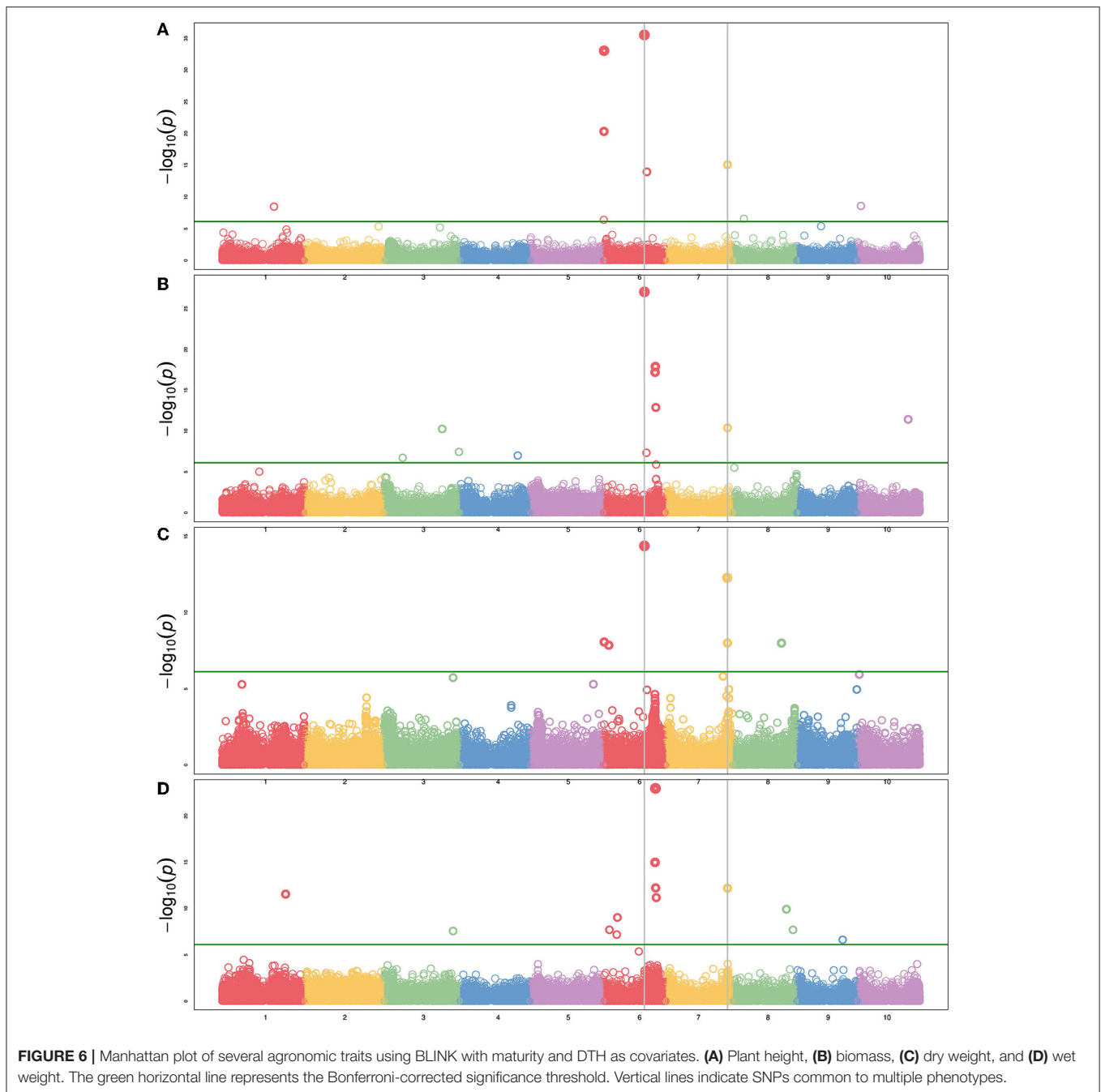


Using MV-LMM, we ran ash and lignin as response variables with maturity and DTH covariates (**Figure 9**), which may be roughly viewed as examining ADL (**Figure 2**). Though we do not have a direct measure of ADL for comparison, the significant loci are a subset of those found using ADF and NDF (**Supplementary Figures 9, 10**) indicating the utility of MV-LMM and covariate models for compositional analyses.

By examining accumulation of soluble sugars in addition to lignocellulosic content and aggregate traits such as NDF, we may obtain a broader perspective of varying sink strength across the accessions. NDF represents a measure of the total lignocellulosic content. Lower lignin content and high lignocellulose production is preferred for efficient biofuel production (Jung et al., 2015). Using dry matter as the response variable, NDF was included in the LMM as a covariate to examine the effects of keeping NDF

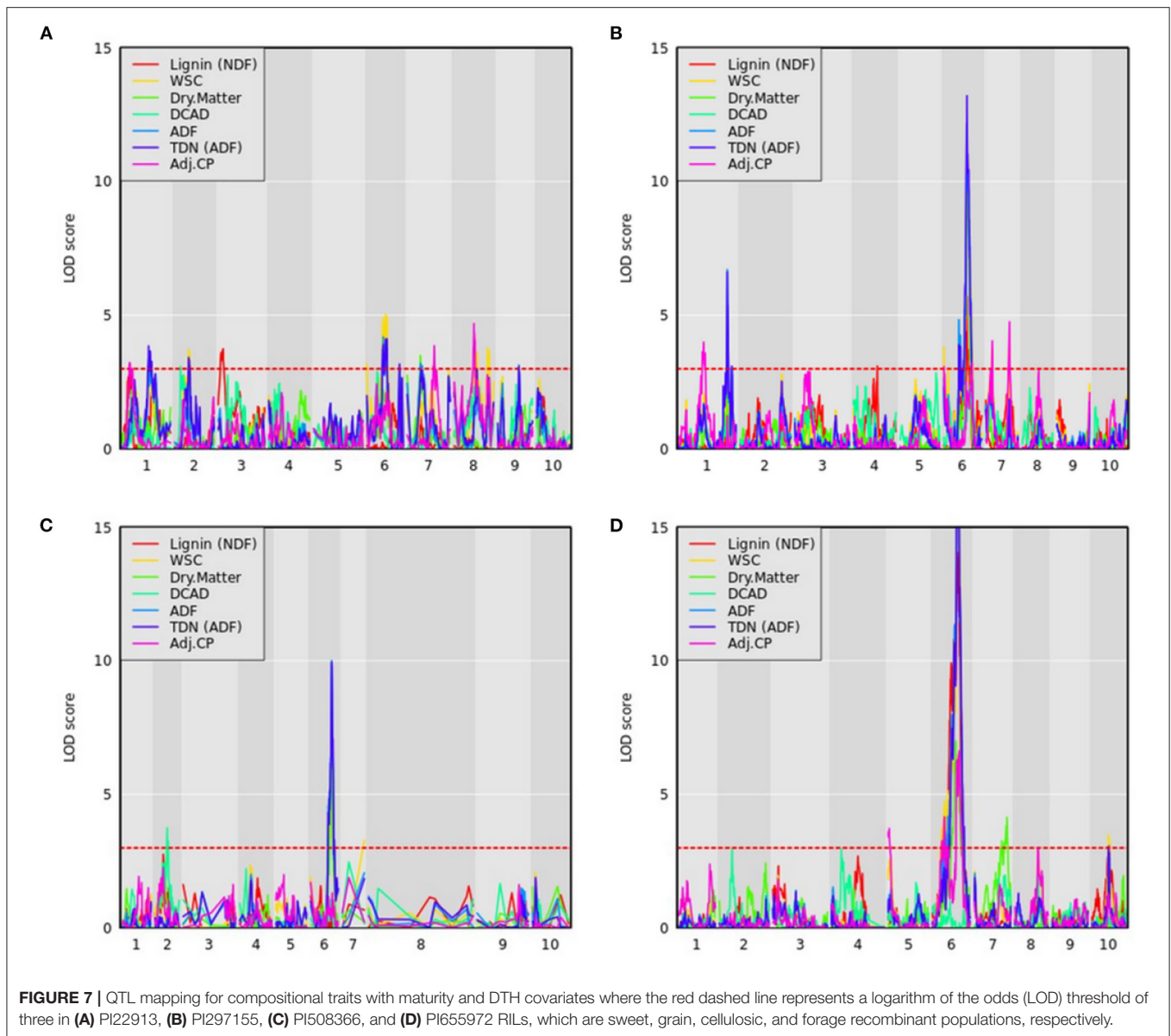
constant on the SNP significance (**Figure 10B**). Biologically, this could be seen as running a model on the phenotypic variance of WSC. Similarly, using dry matter as the response variable with ADF as a covariate examines the phenotypic variance of hemicellulose and WSC (**Figure 10C**). By using covariates or multivariate-response models, the individual (i.e., WSC) or cumulative (i.e., NDF) phenotypic variance of compositional traits may be disentangled, and the relationships among traits may be more clearly distinguished.

Interestingly, the significantly associated SNPs are not identical between models such as a WSC LMM vs. a dry matter response with NDF covariate LMM (**Figures 10B,D**) with composite traits such as dry matter often exhibiting more associated loci. As previously indicated, the integration of multi-scale phenotypes and multivariate models may be



identifying emergent properties of these biological systems as some trait associations are not merely the sum of their parts (Fischer, 2008; Benes et al., 2020), which further indicates the importance of running several different models that attempt to examine characteristics of a trait from multiple perspectives. The various dry matter models (**Figures 10A–C** and **Supplementary Figure 11**) were also associated with a previously identified WSC locus containing a putative vacuolar iron transporter (VIT) on Chr4 (Brenton et al., 2020). It has been suggested that the candidate gene underlying this locus (Sobic.004G301500) may affect sugar accumulation either

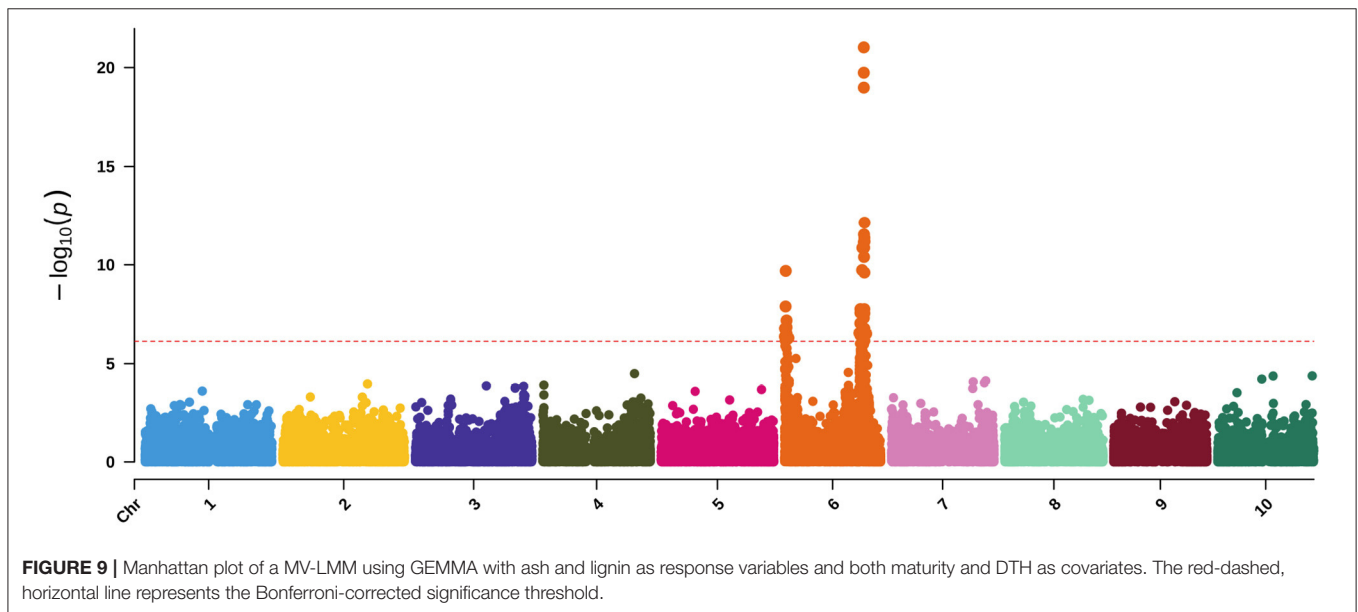
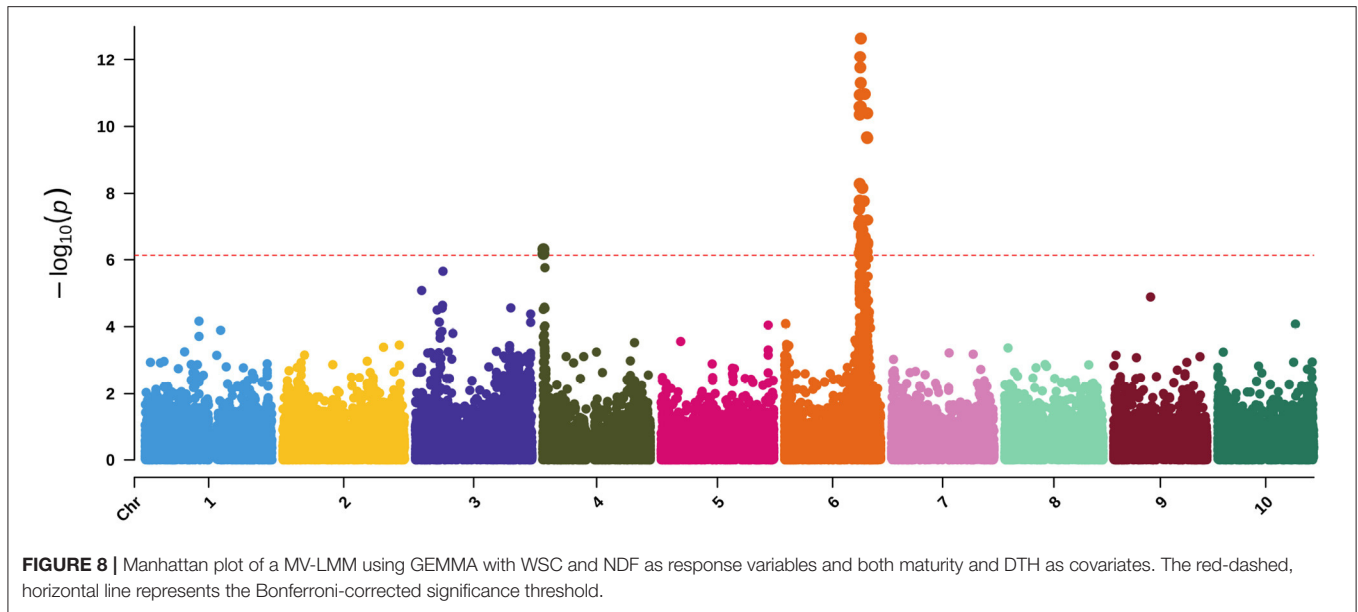
through neofunctionalization or *via* an iron-deficiency response (Brenton et al., 2020). Though interestingly, the same LD block is also hit with adjusted crude protein as well as NEG (**Supplementary Figures 12, 13**). Previous identification of a putative *Dw4* locus identified from plant height GWAS also corresponds to this locus (Li et al., 2015), and the locus has also been associated with increased total biomass and root biomass (Moghimi et al., 2019). Together, these results suggest a mechanism foundational to carbon accumulation underlying this locus, or the locus exhibits a pleiotropic effect on carbon accumulation or partitioning.



While the sorghum gene is classified as an iron transporter, further comparison with the Arabidopsis ortholog (AT3G43660) indicates a potential role in cellular manganese ion homeostasis (GO:0030026) (Gollhofer et al., 2011; Berardini et al., 2015). The CCC2-like domain of Sobic.004G301500 or one of the duplicate loci (Sobic.004G301600/Sobic.004G301650) therefore likely acts to transport manganese to vacuoles and maintain manganese homeostasis. As manganese serves to increase nitrogen assimilation (Przemeck and Schrader, 1981), is a fundamental catalyst during the water-splitting reaction of photosystem II (Fischer et al., 2015), and is necessary for respiration (Alejandro et al., 2020), a pivotal role in manganese homeostasis might better explain these associations. Further, since a role in manganese homeostasis has been described, these duplicated loci may instead demonstrate subfunctionalization

followed by tissue-specific expression of one copy or the duplication may alter gene dosage and consequently modify some rate-limiting process. The association on Chr7 at approximately 59.5 Mb found using dry matter with an ADF covariate, which is equivalent to looking at phenotypic variation due to WSC and hemicellulose, is *Dw3* (Figure 10C).

To estimate the pleiotropic effects of variants across traits, we also performed a meta-analysis of SNP effects estimated using LMMs with an empirical Bayesian multivariate adaptive shrinkage approach that included results from nine traits (Urbut et al., 2019). Multivariate shrinkage serves to regularize parameter estimates across all models, effectively shrinking potential outliers toward zero. This shrinkage effect also reduces spurious associations (i.e., false positive rates), acting as a form of multiple testing correction. Variants with parameter

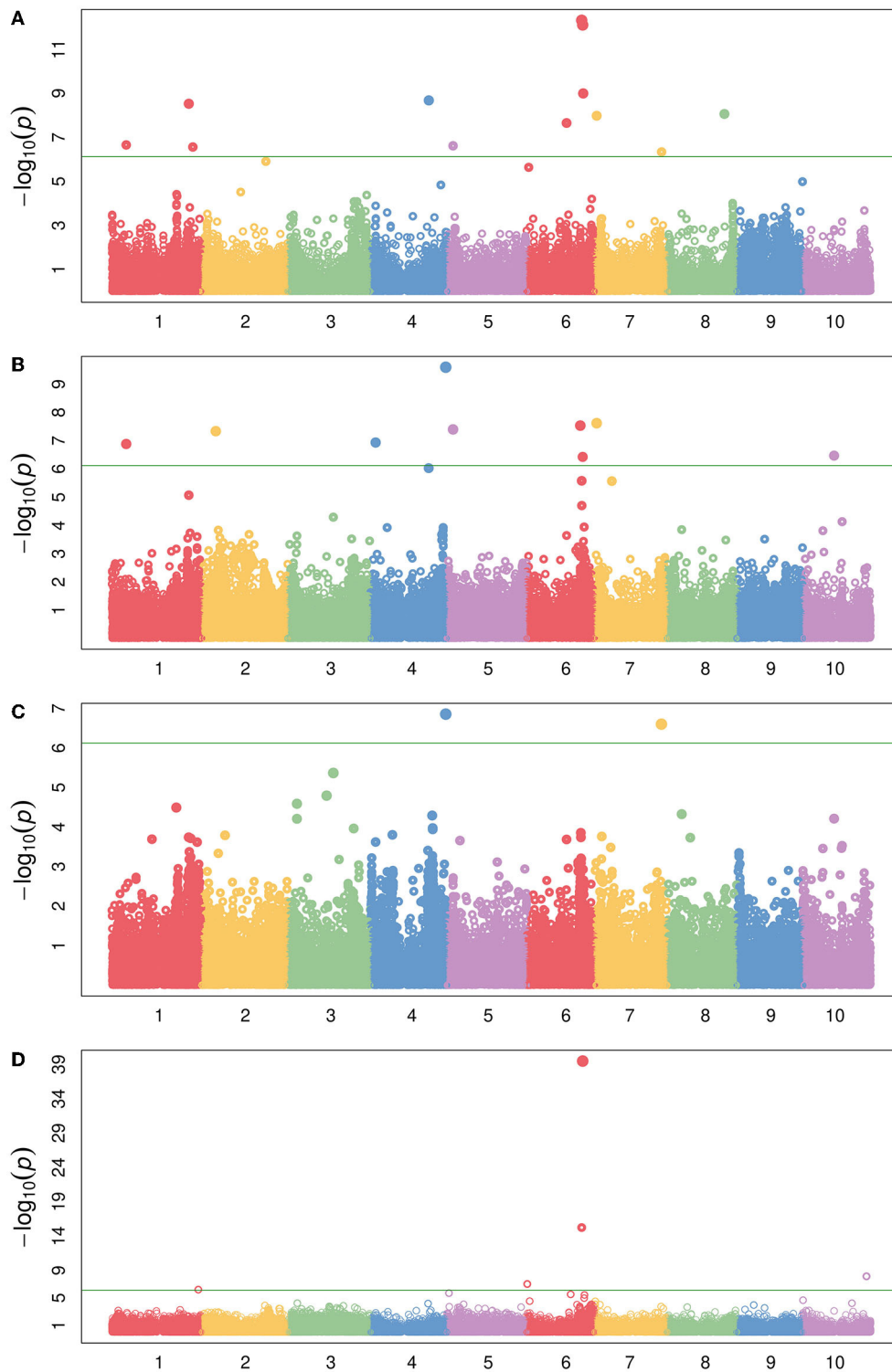


estimates still significant after shrinkage represent putatively true associations. Over 190 variants exhibited strong associations across the nine traits with most associations occurring in chromosomes six and seven (**Supplementary Figure 14**). While most associations occur within known loci including *Dw2*, *Dw3*, *Ma3*, *Ma6*, and the *D* locus, a SNP on Chr10 colocalizes with previously identified locus for fresh stem weight and juice yield (Lv et al., 2013) as well as sucrose content (Ritter et al., 2008). These loci exhibiting pleiotropic effects also demonstrated strong epistatic effects across the genome, with Chr6 exhibiting significantly more effects than other chromosomes (**Supplementary Figure 15**). Taken together, these multivariate approaches highlight the pleiotropic and epistatic effects of loci across the sorghum genome and support the

importance of collecting peripherally related phenotypes to maximize carbon accumulation and partitioning.

## 4. DISCUSSION

The diverse carbon-partitioning regimes of sorghum have the potential to provide valuable insights into the genetic control of carbon partitioning in grasses (Braun and Slewinski, 2009) from transport (Milne et al., 2013) to compartmentalization (Furbank and Kelly, 2021). Genes sensitive to carbohydrate concentration compose part of a highly conserved network necessary for cellular adjustment to nutrient availability and the partitioning of carbon among tissues and organs (Koch, 1996). A holistic understanding of these processes requires



**FIGURE 10** | Manhattan plot of several compositional traits with different covariates following the design represented by **Figure 1**. **(A)** dry matter with maturity and DTH covariates, **(B)** dry matter with maturity, DTH, and NDF covariates, **(C)** dry matter with maturity, DTH, and ADF covariates, and **(D)** WSC with maturity and DTH covariates. The green horizontal line represents the Bonferroni-corrected significance threshold.

multiscale phenotypes from molecule-specific quantification to anatomically aggregated measures of carbon. These multiscale

metrics are necessary to accurately assess traits such as biomass where optical measures are typically poorly correlated with

manually collected, macroscale phenotypes (Eberius and Lima-Guerra, 2009). Orthogonal and partially correlated measures of diverse morphology assist in resolving functional questions of plant growth and development while simultaneously improving significant associations with functional genomic data. We pair these multi-aspect traits with statistical models that mimic the biological design to better distinguish which loci correspond to particular components. In conjunction with broad-scale phenotyping, multiparameter statistical approaches improve inferences through joint consideration of genomic and phenotypic measures (Eberius and Lima-Guerra, 2009). To better resolve the broad-scale effects of carbon-partitioning in this population, we performed individual analyses of traits before performing a meta-analysis across traits. This provides a consistent framework for weighing and resolving the effects of individual variants across traits.

Using QTL mapping, we identified 59 QTL for the agronomic traits and 522 QTL for 34 compositional traits across the 11 RIL families resulting in numerous putative loci that are associated with a variety of different phenotypes across different scales. We similarly identified a median of 30 associated variants per trait for 42 traits using MLM alone, which totalled 1,163 significant hits before identifying 194 variants with significant pleiotropic effects and interactions among those loci (**Supplementary Figures 16, 17** and **Supplementary Tables 4–8**). In particular, Chr6 displays strong pleiotropic and epistatic effects within a highly localized region around the latter half of the chromosome. Several well-studied loci are present at that location including, *qHT7.1*, *Dw2*, *Mal*, and the *D* locus. The different alleles at these loci have been shown to result in broad phenotypic variation for several different carbon partitioning traits. The interactions between loci and locus pleiotropy hint at the underlying genetic architecture of these dynamic carbon-partitioning traits and may represent whole-network hub genes or intramodular hubs mediating carbon accumulation or partitioning networks. By performing these additional tests for pleiotropy and interactive effects across carbon partitioning traits, we seek to highlight genomic positions that are likely to have broad phenotypic effect in carbon networks. However, source and sink interactions can complicate the dissection of individual traits (Brenton et al., 2016). High-capacity, non-photosynthetic sinks can increase yield through sugar-responsive genes that mediate feed-forward loops that ultimately bolster systemic carbon accumulation (Bihmidine et al., 2013; Brenton et al., 2020). Similar carbon relationships have been previously identified such as the positive correlation between plant height and yield, which has been observed in sorghum for well over half a century (Graham and Lessman, 1966). Conversely, as seen here with the *D* locus NAC transcription factor exhibiting reduced lignin content but increased sugar and grain yields (Xia et al., 2018), selection for some loci can result in a tradeoff between carbon regimes. The identification of these feedback mechanisms

suggests that ongoing optimization of carbon allocation should simultaneously focus on improved source and sink strengths as a system of interconnected processes from nitrogen assimilation to photosynthetic efficiency (Brenton et al., 2016, 2020; Boatwright et al., 2021). Additionally, these findings indicate that sorghum yields (i.e., sugar, grain, forage, and biomass) may be further optimized to incorporate beneficial alleles from other sorghum types.

Using these models, we identified numerous candidate loci—both known and novel—associated with carbon-partitioning traits using the CP-NAM. Several traits, such as WSC and biomass traits, shared associated loci supporting previous observations that selection for non-target, sink-related traits may collectively increase yields across carbon-partitioning regimes. Further, we identified strong pleiotropic and interactive effects across the sorghum genome with particularly strong effects on Chr6. The identification of these broad-effect loci will set the stage for future studies to examine the individual and interactive effects of alleles on carbon-partitioning traits using methods such as allele-specific expression in hybrid systems as well as generate multi-trait, multi-environment data to further extricate the environmental and genotype-by-environment effects on carbon-partitioning traits by leveraging the power of multiscale traits and MV-LMMs (Covarrubias-Pazarán, 2016). Additionally, breeders may consider collecting peripherally related traits as a means of understanding and maximizing carbon flow in their system as selection for carbon sinks is not a zero-sum relationship as indicated by the relationships and correlations identified here.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

JB wrote the manuscript and performed all computational analyses. AC, MM, NK, and SS managed the field experiments. AC, KEJ, MM, NK, and SS collected the phenotypes. KEJ has read and approved the manuscript. JB, SS, and SK conceptualized, developed, and implemented the study design. All authors contributed to the article and approved the submitted version.

## FUNDING

This project was funded in part by the U.S. Department of Energy's Advanced Research Project Agency award number DE-AR0001134. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.

## ACKNOWLEDGMENTS

Computational analyses were performed on Clemson University's Palmetto Cluster, and we thank the staff who assisted with cluster and software maintenance.

## REFERENCES

- Alejandro, S., Höller, S., Meier, B., and Peiter, E. (2020). Manganese in plants: from acquisition to subcellular allocation. *Front. Plant Sci.*, 11, 300. doi: 10.3389/fpls.2020.00300
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Benes, B., Guan, K., Lang, M., Long, S. P., Lynch, J. P., Marshall-Colón, A., et al. (2020). Multiscale computational models can guide experimentation and targeted measurements for crop improvement. *Plant J.* 103, 21–31. doi: 10.1111/tpj.14722
- Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., et al. (2015). The arabidopsis information resource: making and mining the “gold standard” annotated reference plant genome. *Genesis* 53, 474–485. doi: 10.1002/dvg.22877
- Bhattarai, B., Singh, S., West, C. P., and Saini, R. (2019). Forage potential of pearl millet and forage sorghum alternatives to corn under the Water-Limiting conditions of the texas high plains: a review. *Crop Forage Turfgrass Manage.* 5, 1–12. doi: 10.2134/cftm2019.08.0058
- Bihmidine, S., Hunter, 3rd, C. T., Johns, C. E., Koch, K. E., and Braun, D. M. (2013). Regulation of assimilate import into sink organs: update on molecular drivers of sink strength. *Front. Plant Sci.* 4, 177. doi: 10.3389/fpls.2013.00177
- Billings, M. (2015). *Biomass Sorghum and Sweet Sorghum Data Gathering Report*. W&A Crop Insurance. USDA-RMA, CTOR: Jaime.
- Boatwright, J. L., Brenton, Z. W., Boyles, R. E., Sapkota, S., Myers, M. T., Jordan, K. E., et al. (2021). Genetic characterization of a sorghum bicolor multiparent mapping population emphasizing carbon-partitioning dynamics. *G3* 11, 1–12. doi: 10.1093/g3journal/jkab060
- Bouchet, S., Olatoye, M. O., Marla, S. R., Perumal, R., Tesso, T., Yu, J., et al. (2017). Increased power to dissect adaptive traits in global sorghum diversity using a nested association mapping population. *Genetics* 206, 573–585. doi: 10.1534/genetics.116.198499
- Braun, D. M., and Slewinski, T. L. (2009). Genetic control of carbon partitioning in grasses: roles of sucrose transporters and tie-dyed loci in phloem loading. *Plant Physiol.* 149, 71–81. doi: 10.1104/pp.108.129049
- Brenton, Z. W., Cooper, E. A., Myers, M. T., Boyles, R. E., Shakoob, N., Zielinski, K. J., et al. (2016). A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. *Genetics* 204, 21–33. doi: 10.1534/genetics.115.183947
- Brenton, Z. W., Juengst, B. T., Cooper, E. A., Myers, M. T., Jordan, K. E., Dale, S. M., et al. (2020). Species-Specific duplication event associated with elevated levels of nonstructural carbohydrates in sorghum bicolor. *G3* 10, 1511–1520. doi: 10.1534/g3.119.400921
- Broman, K. W., Gatti, D. M., Simecek, P., Furlotte, N. A., Prins, P., Sen, S., et al. (2019). R/qtl2: Software for mapping quantitative trait loci with High-Dimensional data and multiparent populations. *Genetics* 211, 495–502. doi: 10.1534/genetics.118.301595
- Broman, K. W., and Sen, S. (2009). *A Guide to QTL Mapping with R/QTL*, Vol. 46. New York, NY: Springer. doi: 10.1007/978-0-387-92125-9
- Brosse, N., Dufour, A., Meng, X., Sun, Q., and Ragauskas, A. (2012). Miscanthus: a fast-growing crop for biofuels and chemicals production. *Biofuels Bioprod. Biorefin.* 6, 580–598. doi: 10.1002/bbb.1353
- Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* 103, 338–348. doi: 10.1016/j.ajhg.2018.07.015
- Burks, P. S., Kaiser, C. M., Hawkins, E. M., and Brown, P. J. (2015). Genomewide association for sugar yield in sweet sorghum. *Crop Sci.* 55, 2138–2148. doi: 10.2135/cropsci2015.01.0057
- Byrt, C. S., Grof, C. P. L., and Furbank, R. T. (2011). C4 plants as biofuel feedstocks: Optimising biomass production and feedstock quality from a lignocellulosic perspective. *Integr. Plant Biol.* 53, 120–135. doi: 10.1111/j.1744-7909.2010.01023.x
- Calviño, M., and Messing, J. (2012). Sweet sorghum as a model system for bioenergy crops. *Curr. Opin. Biotechnol.* 23, 323–329. doi: 10.1016/j.copbio.2011.12.002
- Carpita, N. C., and McCann, M. C. (2008). Maize and sorghum: genetic resources for bioenergy grasses. *Trends Plant Sci.* 13, 415–420. doi: 10.1016/j.tplants.2008.06.002
- Chipanshi, A. C., Chanda, R., and Totolo, O. (2003). Vulnerability assessment of the maize and sorghum crops to climate change in botswana. *Clim. Change* 61, 339–360. doi: 10.1023/B:CLIM.000004551.55871.eb
- Cooper, E. A., Brenton, Z. W., Flinn, B. S., Jenkins, J., Shu, S., Flowers, D., et al. (2019). A new reference genome for sorghum bicolor reveals high levels of sequence similarity between sweet and grain genotypes: implications for the genetics of sugar metabolism. *BMC Genomics* 20, 420. doi: 10.1186/s12864-019-5734-x
- Covarrubias-Pazarán, G. (2016). Genome-assisted prediction of quantitative traits using the R package sommer. *PLoS ONE* 11, e0156744. doi: 10.1371/journal.pone.0156744
- David, K., and Ragauskas, A. J. (2010). Switchgrass as an energy crop for biofuel production: a review of its ligno-cellulosic chemical properties. *Energy Environ. Sci.* 3, 1182–1190. doi: 10.1039/b926617h
- Druille, M., Williams, A. S., Torrecillas, M., Kim, S., Meki, N., and Kiniry, J. R. (2020). Modeling climate warming impacts on grain and forage sorghum yields in argentina. *Agronomy* 10, 964. doi: 10.3390/agronomy10070964
- Eberius, M., and Lima-Guerra, J. (2009). “High-throughput plant phenotyping-data acquisition, transformation, and analysis,” in *Bioinformatics*, eds D. Edwards, J. Stajich, and D. Hansen (New York, NY: Springer), 259–278. doi: 10.1007/978-0-387-92738-1\_13
- Fiedler, K., Bekele, W. A., Duensing, R., Gründig, S., Snowdon, R., Stützel, H., et al. (2014). Genetic dissection of temperature-dependent sorghum growth during juvenile development. *Züchter Genet. Breed. Res.* 127, 1935–1948. doi: 10.1007/s00122-014-2350-7
- Fischer, H. P. (2008). Mathematical modeling of complex biological systems: from parts lists to understanding systems behavior. *Alcohol Res. Health* 31, 49–59.
- Fischer, H. W., Reddy, N. L. N., and Rao, M. L. S. (2016). Can more drought resistant crops promote more climate secure agriculture? Prospects and challenges of millet cultivation in Ananthapur, Andhra Pradesh. *World Dev. Perspect.* 2, 5–10. doi: 10.1016/j.wdp.2016.06.005
- Fischer, W. W., Hemp, J., and Johnson, J. E. (2015). Manganese and the evolution of photosynthesis. *Orig. Life Evol. Biosph.* 45, 351–357. doi: 10.1007/s11084-015-9442-5
- Furbank, R. T., and Kelly, S. (2021). Finding the C4 sweet spot: cellular compartmentation of carbohydrate metabolism in C4 photosynthesis. *J. Exp. Bot.* 72, 6018–6026. doi: 10.1093/jxb/erab290
- Gardner, J. C., Maranville, J. W., and Paparozzi, E. T. (1994). Nitrogen use efficiency among diverse sorghum cultivars. *Crop Sci.* 34, 728–733. doi: 10.2135/cropsci1994.0011183X003400030023x
- Gelli, M., Duo, Y., Konda, A. R., Zhang, C., Holding, D., and Dweikat, I. (2014). Identification of differentially expressed genes between sorghum genotypes with contrasting nitrogen stress tolerance by genome-wide transcriptional profiling. *BMC Genomics* 15, 179. doi: 10.1186/1471-2164-15-179
- Gelli, M., Konda, A. R., Liu, K., Zhang, C., Clemente, T. E., Holding, D. R., et al. (2017). Validation of QTL mapping and transcriptome profiling for identification of candidate genes associated with nitrogen stress tolerance in sorghum. *BMC Plant Biol.* 17, 123. doi: 10.1186/s12870-017-1064-9

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.790005/full#supplementary-material>

- George-Jaeggli, B., Jordan, D. R., van Oosterom, E. J., and Hammer, G. L. (2011). Decrease in sorghum grain yield due to the DW3 dwarfing gene is caused by reduction in shoot biomass. *Field Crops Res.* 124, 231–239. doi: 10.1016/j.fcr.2011.07.005
- Glaubitx, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014). TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS ONE* 9, e90346. doi: 10.1371/journal.pone.0090346
- Gollhofer, J., Schläwicke, C., Jungnick, N., Schmidt, W., and Buckhout, T. J. (2011). Members of a small family of nodulin-like genes are regulated under iron deficiency in roots of *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 49, 557–564. doi: 10.1016/j.plaphy.2011.02.011
- Graham, D., and Lessman, K. (1966). Effect of height on yield and yield components of two isogenic lines of sorghum vulgare pers. 1. *Crop Sci.* 6, 372–374. doi: 10.2135/cropsci1966.0011183X000600040024x
- Guo, B., Sleper, D. A., and Beavis, W. D. (2010). Nested association mapping for identification of functional markers. *Genetics* 186, 373–383. doi: 10.1534/genetics.110.115782
- Haley, C. S., and Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69, 315–324. doi: 10.1038/hdy.1992.131
- Heaton, E. A., Dohleman, F. G., and Long, S. P. (2008). Meeting US biofuel goals with less land: the potential of miscanthus. *Glob. Chang. Biol.* 14, 2000–2014. doi: 10.1111/j.1365-2486.2008.01662.x
- Hilley, J. L., Weers, B. D., Truong, S. K., McCormick, R. F., Mattison, A. J., McKinley, B. A., et al. (2017). Sorghum DW2 encodes a protein kinase regulator of stem internode length. *Sci. Rep.* 7, 4616. doi: 10.1038/s41598-017-04609-5
- Huang, M., Liu, X., Zhou, Y., Summers, R. M., and Zhang, Z. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience* 8, 1–12. doi: 10.1093/gigascience/giy154
- Jung, S.-J., Kim, S.-H., and Chung, I.-M. (2015). Comparison of lignin, cellulose, and hemicellulose contents for biofuels utilization among 4 types of lignocellulosic crops. *Biomass Bioenergy* 83, 322–327. doi: 10.1016/j.biombioe.2015.10.007
- Kakani, V. G., Vu, J. C. V., Allen, L. H., and Boote, K. J. (2011). Leaf photosynthesis and carbohydrates of CO<sub>2</sub>-enriched maize and grain sorghum exposed to a short period of soil water deficit during vegetative development. *J. Plant Physiol.* 168, 2169–2176. doi: 10.1016/j.jplph.2011.07.003
- King, E. G., and Long, A. D. (2017). The beavis effect in next-generation mapping panels in *Drosophila melanogaster*. *G3* 7, 1643–1652. doi: 10.1534/g3.117.041426
- Knox, J., Hess, T., Daccache, A., and Wheeler, T. (2012). Climate change impacts on crop productivity in Africa and South Asia. *Environ. Res. Lett.* 7, 034032. doi: 10.1088/1748-9326/7/3/034032
- Koch, K. E. (1996). Carbohydrate-modulated gene expression in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 509–540. doi: 10.1146/annurev.arplant.47.1.509
- Kong, W., Kim, C., Zhang, D., Guo, H., Tan, X., Jin, H., et al. (2018). Genotyping by sequencing of 393 sorghum bicolor BTx623 × IS3620C recombinant inbred lines improves sensitivity and resolution of QTL detection. *G3* 8, 2563–2572. doi: 10.1534/g3.118.200173
- Kosambi, D. D. (2016). “The estimation of map distances from recombination values,” in *D.D. Kosambi: Selected Works in Mathematics and Statistics*, ed R. Ramaswamy (New Delhi: Springer India), 125–130. doi: 10.1007/978-81-322-3676-4\_16
- Kresovich, S., Miller, F. R., Monk, R. L., Dominy, R. E., and Broadhead, D. M. (1988). Registration of ‘grass’ sweet sorghum. *Crop Sci.* 28, 194–195. doi: 10.2135/cropsci1988.0011183X002800010060x
- Kruijjer, W., Boer, M. P., Malosetti, M., Flood, P. J., Engel, B., Kooke, R., et al. (2015). Marker-based estimation of heritability in immortal populations. *Genetics* 199, 379–398. doi: 10.1534/genetics.114.167916
- Li, X., Li, X., Fridman, E., Tesso, T. T., and Yu, J. (2015). Dissecting repulsion linkage in the dwarfing gene *dw3* region for sorghum plant height provides insights into heterosis. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11823–11828. doi: 10.1073/pnas.1509229112
- Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444
- Lobell, D. B., and Field, C. B. (2007). Global scale climate-crop yield relationships and the impacts of recent warming. *Environ. Res. Lett.* 2, 014002. doi: 10.1088/1748-9326/2/1/014002
- Lv, P., Ji, G., Han, Y., Hou, S., Li, S., Ma, X., et al. (2013). Association analysis of sugar yield-related traits in sorghum [*Sorghum bicolor* (L.)]. *Euphytica* 193, 419–431. doi: 10.1007/s10681-013-0962-7
- Mace, E. S., Tai, S., Gilding, E. K., Li, Y., Prentis, P. J., Bian, L., et al. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa’s indigenous cereal crop sorghum. *Nat. Commun.* 4, 2320. doi: 10.1038/ncomms3320
- MacQueen, A. H., White, J. W., Lee, R., Osorno, J. M., Schmutz, J., Miklas, P. N., et al. (2020). Genetic associations in four decades of multi-environment trials reveal agronomic trait evolution in common bean. *Genetics* 215, 267–284. doi: 10.1534/genetics.120.303038
- Marla, S. R., Burrow, G., Chopra, R., Hayes, C., Olatoye, M. O., Felderhoff, T., et al. (2019). Genetic architecture of chilling tolerance in sorghum dissected with a nested association mapping population. *G3* 9, 4045–4057. doi: 10.1534/g3.119.400353
- McCormick, R. F., Truong, S. K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., et al. (2018). The sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J.* 93, 338–354. doi: 10.1111/tpj.13781
- McKinney, W. (2010). “Data structures for statistical computing in Python,” in *Proceedings of the 9th Python in Science Conference, Proceedings of the Python in Science Conference* (Austin, TX: SciPy), 56–61. doi: 10.25080/Majora-92bf1922-00a
- Meki, M. N., Kemanian, A. R., Potter, S. R., Blumenthal, J. M., Williams, J. R., and Gerik, T. J. (2013). Cropping system effects on sorghum grain yield, soil organic carbon, and global warming potential in central and south Texas. *Agric. Syst.* 117, 19–29. doi: 10.1016/j.agry.2013.01.004
- Michener, W. K., Blood, E. R., Bildstein, K. L., Brinson, M. M., and Gardner, L. R. (1997). Climate change, hurricanes and tropical storms, and rising sea level in coastal wetlands. *Ecol. Appl.* 7, 770–801. doi: 10.1890/1051-0761(1997)007(0770:CCHATS)2.0.CO;2
- Milne, R. J., Byrt, C. S., Patrick, J. W., and Grof, C. P. L. (2013). Are sucrose transporter expression profiles linked with patterns of biomass partitioning in sorghum phenotypes? *Front. Plant Sci.* 4, 223. doi: 10.3389/fpls.2013.00223
- Moghimi, N., Desai, J. S., Bheemanahalli, R., Impa, S. M., Vennapusa, A. R., Sebel, D., et al. (2019). New candidate loci and marker genes on chromosome 7 for improved chilling tolerance in sorghum. *J. Exp. Bot.* 70, 3357–3371. doi: 10.1093/jxb/erz143
- Monti, A. (2012). *Switchgrass: A Valuable Biomass Crop for Energy*. London: Springer Science & Business Media.
- Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., et al. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc. Natl. Acad. Sci. U.S.A.* 110, 453–458. doi: 10.1073/pnas.1215985110
- Mullet, J., Morishige, D., McCormick, R., Truong, S., Hilley, J., McKinley, B., et al. (2014). Energy sorghum—a genetic model for the design of C4 grass bioenergy crops. *J. Exp. Bot.* 65, 3479–3489. doi: 10.1093/jxb/eru229
- Mural, R. V., Grzybowski, M., Miao, C., Damke, A., Sapkota, S., Boyles, R. E., et al. (2021). Meta-analysis identifies pleiotropic loci controlling phenotypic trade-offs in sorghum. *Genetics* 218, 1–15. doi: 10.1093/genetics/iyab087
- Murphy, R. L., Morishige, D. T., Brady, J. A., Rooney, W. L., Yang, S., Klein, P. E., et al. (2014). GHD7 (MA 6) represses sorghum flowering in long days: GHD7 alleles enhance biomass accumulation and grain production. *Plant Genome* 7, 1–10. doi: 10.3835/plantgenome2013.11.0040
- Murray, S. C., Sharma, A., Rooney, W. L., Klein, P. E., Mullet, J. E., Mitchell, S. E., et al. (2008). Genetic improvement of sorghum as a biofuel feedstock: I. QTL for stem sugar and grain nonstructural carbohydrates. *Crop Sci.* 48, 2165–2179. doi: 10.2135/cropsci2008.01.0016
- Olson, S. N., Ritter, K., Rooney, W., Kemanian, A., McCarl, B. A., Zhang, Y., et al. (2012). High biomass yield energy sorghum: developing a genetic model for C4 grass bioenergy crops. *Biofuels Bioprod. Biorefin.* 6, 640–655. doi: 10.1002/bbb.1357
- Ortiz, D., Hu, J., and Salas Fernandez, M. G. (2017). Genetic architecture of photosynthesis in sorghum bicolor under non-stress and cold stress conditions. *J. Exp. Bot.* 68, 4545–4557. doi: 10.1093/jxb/erx276



- Ottman, M. J., Kimball, B. A., Pinter, P. J., Wall, G. W., Vanderlip, R. L., Leavitt, S. W., et al. (2001). Elevated CO<sub>2</sub> increases sorghum biomass under drought conditions. *New Phytol.* 150, 261–273. doi: 10.1046/j.1469-8137.2001.00110.x
- Peng, S., Krieg, D. R., and Girma, F. S. (1991). Leaf photosynthetic rate is correlated with biomass and grain production in grain sorghum lines. *Photosynth. Res.* 28, 1–7. doi: 10.1007/BF00027171
- Pennisi, E. (2009). Plant genetics. How sorghum withstands heat and drought. *Science* 323, 573. doi: 10.1126/science.323.5914.573
- Prasad, P. V. V., Vu, J. C. V., Boote, K. J., and Allen, L. H. (2009). Enhancement in leaf photosynthesis and upregulation of rubisco in the C4 sorghum plant at elevated growth carbon dioxide and temperature occur at early stages of leaf ontogeny. *Funct. Plant Biol.* 36, 761–769. doi: 10.1071/FP09043
- Przemek, E., and Schrader, B. (1981). “The effect of manganese nutrition on nitrogen assimilation in roots,” in *Structure and Function of Plant Roots*, eds R. Brouwer, O. Gašparíková, J. Kolek, and B. C. Loughman (Dordrecht: Springer Netherlands), 123–127. doi: 10.1007/978-94-009-8314-4\_23
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Qazi, H. A., Paranjpe, S., and Bhargava, S. (2012). Stem sugar accumulation in sweet sorghum - activity and expression of sucrose metabolizing enzymes and sucrose transporters. *J. Plant Physiol.* 169, 605–613. doi: 10.1016/j.jplph.2012.01.005
- R Core Team. (2019). R: a language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria. Available online at: <https://www.R-project.org/>.
- Rice, B. R., Fernandes, S. B., and Lipka, A. E. (2020). Multi-Trait genome-wide association studies reveal loci associated with maize inflorescence and leaf architecture. *Plant Cell Physiol.* 61, 1427–1437. doi: 10.1093/pcp/pcaa039
- Ritter, K. B., Jordan, D. R., Chapman, S. C., Godwin, I. D., Mace, E. S., and Lynne McIntyre, C. (2008). Identification of QTL for sugar-related traits in a sweet × grain sorghum (*Sorghum bicolor* L. *moench*) recombinant inbred population. *Mol. Breed.* 22, 367–384. doi: 10.1007/s11032-008-9182-6
- Rooney, W. L., Blumenthal, J., Bean, B., and Mullet, J. E. (2007). Designing sorghum as a dedicated bioenergy feedstock. *Biofuels Bioprod. Biorefin.* 1, 147–157. doi: 10.1002/bbb.15
- Sapkota, S., Boyles, R., Cooper, E., Brenton, Z., Myers, M., and Kresovich, S. (2020). Impact of sorghum racial structure and diversity on genomic prediction of grain yield components. *Crop Sci.* 60, 132–148. doi: 10.1002/csc2.20060
- Schoelz, J. M., Feng, J. X., and Riddle, N. C. (2021). The Drosophila HP1 family is associated with active gene expression across chromatin contexts. *Genetics.* 219, 1–14. doi: 10.1093/genetics/iyab108
- Schubert, E., and Rousseeuw, P. J. (2019). “Faster k-medoids clustering: improving the PAM, CLARA, and CLARANS algorithms,” in *Similarity Search and Applications*, eds G. Amato, C. Gennaro, V. Oria, M. Radovanović (Cham: Springer International Publishing), 171–187. doi: 10.1007/978-3-030-32047-8\_16
- Stephens, M. (2017). False discovery rates: a new deal. *Biostatistics* 18, 275–294. doi: 10.1101/038216
- Urbat, S. M., Wang, G., Carbonetto, P., and Stephens, M. (2019). Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. *Nat. Genet.* 51, 187–195. doi: 10.1038/s41588-018-0268-8
- van der Weijde, T., Alvim Kamei, C. L., Torres, A. F., Vermerris, W., Dolstra, O., Visser, R. G. F., et al. (2013). The potential of C4 grasses for cellulosic biofuel production. *Front. Plant Sci.* 4, 107. doi: 10.3389/fpls.2013.00107
- Van Rossum, G., and Drake, F. L. (2009). *Introduction to Python 3: (Python Documentation Manual Part 1)*. Godalming, England: CreateSpace Independent Publishing Platform.
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980
- Viel, M., Collet, F., Prétot, S., and Lanos, C. (2019). Hemp-straw composites: gluing study and multi-physical characterizations. *Materials* 12, 1199. doi: 10.3390/ma12081199
- Vietor, D. M., and Miller, F. R. (1990). Assimilation, partitioning, and nonstructural carbohydrates in sweet compared with grain sorghum. *Crop Sci.* 30, 1109. doi: 10.2135/cropsci1990.0011183X003000050030x
- Wang, J., and Zhang, Z. (2021). GAPIT version 3: boosting power and accuracy for genomic association and prediction. *Genomics, Proteomics & Bioinformatics.* 19, 629–640. doi: 10.1016/j.gpb.2021.08.005
- Waskom, M. (2021). seaborn: statistical data visualization. *J. Open Source Softw.* 6, 3021. doi: 10.21105/joss.03021
- Xia, J., Zhao, Y., Burks, P., Pauly, M., and Brown, P. J. (2018). A sorghum NAC gene is associated with variation in biomass properties and yield potential. *Plant Direct* 2, e00070. doi: 10.1002/pld3.70
- Yan, K., Chen, P., Shao, H., Zhang, L., and Xu, G. (2011). Effects of short-term high temperature on photosynthesis and photosystem II performance in sorghum. *J. Agron. Crop Sci.* 197, 400–408. doi: 10.1111/j.1439-037X.2011.00469.x
- Yang, S., Murphy, R. L., Morishige, D. T., Klein, P. E., Rooney, W. L., and Mullet, J. E. (2014). Sorghum phytochrome B inhibits flowering in long days by activating expression of SbPRR37 and SbGHD7, repressors of SbEHD1, SbCN8 and SbCN12. *PLoS ONE* 9, e105352. doi: 10.1371/journal.pone.0105352
- Yu, J., and Buckler, E. S. (2006). Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* 17, 155–160. doi: 10.1016/j.copbio.2006.02.003
- Zegada-Lizarazu, W., Zatta, A., and Monti, A. (2012). Water uptake efficiency and above- and belowground biomass development of sweet sorghum and maize under different water regimes. *Plant Soil* 351, 47–60. doi: 10.1007/s11104-011-0928-2
- Zhao, J., Mantilla Perez, M. B., Hu, J., and Salas Fernandez, M. G. (2016). Genome-wide association study for nine plant architecture traits in sorghum. *Plant Genome* 9, 1–14. doi: 10.3835/plantgenome2015.06.0044
- Zhao, Y. L., Dolat, A., Steinberger, Y., Wang, X., Osman, A., and Xie, G. H. (2009). Biomass yield and changes in chemical composition of sweet sorghum cultivars grown for biofuel. *Field Crops Res.* 111, 55–64. doi: 10.1016/j.fcr.2008.10.006
- Zhou, X., Carbonetto, P., and Stephens, M. (2013). Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genet.* 9, e1003264. doi: 10.1371/journal.pgen.1003264
- Zhou, X., and Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* 11, 407. doi: 10.1038/nmeth.2848

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Boatwright, Sapkota, Myers, Kumar, Cox, Jordan and Kresovich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.