

Article

Transferability of Marker Trait Associations in Wheat Is Disturbed Mainly by Genotype $\times$ Year Interaction

Blake Russell ¹, Gina Brown-Guedira ², Clay H. Sneller ³,
Mohsen Mohammadi ^{1,*}

¹ Department of Agronomy, Purdue University, 915 West State Street, West Lafayette, IN 47907, USA

² US Department of Agriculture, Agricultural Research Services, Small Grains Genotyping Laboratory, Raleigh, NC 27695, USA

³ Ohio Agricultural Research and Development Center, The Ohio State University, Wooster, Ohio 446931, USA

* Correspondence: Mohsen Mohammadi, Email: mohamm20@purdue.edu; Tel.: +1 (765) 496-6885.

ABSTRACT

Grain yield is the primary target for increasing soft red winter wheat production in the United States. Genome wide association studies were performed for yield-related traits using field-based evaluation of 270 elite breeding lines. Yield ranged from 3.9 to 7.5 tons·ha⁻¹. The effect of genotype by year interaction was significant for most traits. Heritability estimates ranged from 0.21 for biomass to 0.84 for plant height. Biomass, kernel weight, and grain weight per spike were positively correlated with grain yield while days to heading and maturity were negatively correlated with grain yield. We used FarmCPU method and 45K markers to identify marker-trait associations. Fifty-nine MTAs were identified based on two-year phenotypic estimates. An association for grain yield on chromosome 7D showed a $-\log P$ value of 16.35 and explained 18% of the phenotypic variation. Associations were identified at $-\log P$ values of 6.30 and 4.08 on chromosomes 5A for thousand kernel weight and grain per spike that explained 10% and 6% of phenotypic variations, respectively. Grain yield and days to heading data from previous trials in other states were used to validate and evaluate the transferability of associations. We observed that for grain yield, only 7 out of 28 associations and for days to heading, only 8 out of 47 associations were validated across seemingly homogenous environments, indicating the majority of associations are not stable QTL across environments. This study suggests that achieving genetic gains appears to require utilizing genotype by environment interaction and local field based germplasm testing.

KEYWORDS: wheat; genome wide association studies; marker trait associations; yield components; grain yield; days to heading

Open Access

Received: 12 May 2020

Accepted: 19 June 2020

Published: 28 June 2020

Copyright © 2020 by the author(s). Licensee Hapres, London, United Kingdom. This is an open access article distributed under the terms and conditions of [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

INTRODUCTION

Over 35% of the world's population relies on wheat as a main source of food [1]. Wheat is widely cultivated around the world for its adaptability to diverse growing regions and environmental conditions. The United States produced 51.3 million metric tons of wheat during the 2018-19 season [2], where most wheat harvested occurs west of the Mississippi River and in the Great Plains [3]. However, soft red winter (SRW) wheat is grown mainly in the Midwestern and eastern United States, accounting for 15–20% of US wheat production [2]. Specifically, the growing region extends from 30° N to 45° N in latitude and about 73° E to 96° W in longitude [4]. In the Midwest and eastern wheat region, grain yield and resistance to *Fusarium* head blight disease are the underlying traits for profitability. The soft wheat products require minimal gluten protein, and lower protein levels than hard wheats [5]. Therefore, producing more grain is the first focus of most SRW wheat breeding programs.

Wheat provides approximately 20% of the protein and calories for human consumption worldwide [6]. In order to meet the needs of the growing population, food supplies from major cereals such as maize, rice and wheat will need to increase by 2–3% annually, and wheat has shown the lowest rate of increases [7]. Ray et al. [8] estimated wheat yields are increasing at 0.9% per year, much less than the 2.4% required to double global production by 2050. With future food security and climate challenges ahead, wheat breeding efficiency and genetic gains must improve significantly to develop stable, adapted, and high-yielding wheat varieties.

Global demand for wheat is growing faster than genetic gains in yield potential [9]. In the Great Plains region, the annual rate of genetic gain was estimated at 0.44%, mainly due to traits contributing to an increase in grain number [10]. The USDA winter wheat regional performance nurseries for the Great Plains region displayed similar results over a 50-year period, with estimated genetic gain for grain yield at 0.79% per year. From 1919 to 2008, the genetic gains in SRW wheat in multiple environments ranged from 0.56% to 1.41% [11].

Much of hereto forth yield increases were due to increases in the number of spikes per area, the number of seeds per spike and spikelet, and harvest index—producing more grain from increasing yield components but maintaining the same biomass [11]. With harvest index approaching its theoretical maximum biologic limits [12], increasing biomass can provide an opportunity to increase the photosynthetic tissues for fixing carbon and a productive canopy to capture radiation energy and convert it into dry matter. Reynolds et al. [13] reported that an increase in radiation use efficiency, grain number, and grain yield were positively associated with an increase in above ground plant biomass.

Breeding methodologies and techniques have changed drastically over the years. Further advances in statistical methodology and molecular markers led to the construction of genetic maps, evaluating complex traits,

and associating the phenotypic variation with molecular markers [14,15]. The genetic maps facilitated the identification of quantitative trait loci (QTL)—the genomic region responsible for trait variation [16]. In wheat, QTL mapping has been performed for traits including yield components [17], plant height [18], heat tolerance [19], grain quality [20], and disease tolerance [21–23], among others. The identification of QTL has led to the use of molecular markers in screening germplasm for trait improvement [24,25]. Bi-parental mapping is a powerful tool. However, the limited number of recombination events in bi-parental populations restricts the phenotypic diversity [16,26] and leads to a low mapping resolution [27].

The need to dissect complex traits within a large, diverse population led to the development of statistical methods that gave rise to genome-wide association studies (GWAS). Unlike bi-parental mapping, GWAS consists of genetically diverse germplasm that harbor many historical and ancestral recombination events. GWAS is based on the strength of linkage disequilibrium (LD) between the markers and the observable phenotypes in a population [27,28]. The statistical power to detect causal polymorphisms is based on the extension of LD in the population [29]. Wheat, being a self-pollinating species, experiences relatively slow LD decay. Selection on wheat, as is practiced in breeding programs, leads to relatively slower rates of LD decay, as Liu et al. [30] displayed that the extent of high LD islands is much greater in cultivars (1053 kb) than landraces (785 kb) due to the effect of artificial selection.

GWAS has been used previously to study wheat kernel size and milling quality [31–33], spike traits [30], root traits [34], and grain yield and yield components traits [35,36]. These studies implemented the various GWAS mapping approaches such as mixed linear model (MLM) [37] and compressed mixed linear model (CMLM) [38] to appropriately account for the underlying population structure and kinship. Recent studies have shown that single locus models, such as MLM and CMLM, generate more false negatives due to overfitting [39,40]. The multi-locus Fixed and Random Model Circulating Probability Unification (FarmCPU) model was shown to better control false positives and false negatives [39,41], improving statistical power to identify true marker trait associations (MTAs).

In this study, our goal was to identify MTAs for yield and yield component traits in an elite SRW wheat population developed by eastern and midwestern public breeding programs. Previous work by Gaire et al. [33] in this population identified MTAs concerning SRW wheat end use quality traits in this population, but no work to date has explored yield related traits in the context of GWAS. We achieve this goal by field-based phenotyping and high-throughput genotyping.

MATERIALS AND METHODS

Experimental design. The Triticeae Coordinated Agricultural Project (TCAP) population consists of lines developed from breeding programs in

Illinois, Kentucky, Maryland, Missouri, New York, Ohio, Indiana, and Virginia. The pedigree of lines are detailed in Huang et al. [42]. The germplasm were grown in two growing seasons 2016-17 (WL17) and 2017-18 (WL18) at the Purdue Agronomy Center for Research and Education (ACRE) in West Lafayette, IN (40.43° N, 86.99° W) after a previous soybean crop. Similar field layouts and germplasm were planted in both years. Trials were planted in late September and harvest in late June of the following year. The experiments were planted using a Hege (Wintersteiger, Australia) drill planter and harvested with a Wintersteiger plot harvester at physiological maturity. In each year, two replications were planted. Each replicate was a 13-row $\times$ 24-column layout, consisting of eight incomplete blocks, each accommodating 39 plots. Each plot measured 1.22 m $\times$ 1.22 m and we planted 20 grams seed per plot, which amounts to approximately plant density of 370–420 seeds per square meter. Before planting, 336 kg·ha⁻¹ of mono-ammonium phosphate (11-52-0) was applied. A spring nitrogen top-dress application of 112 kg·ha⁻¹ in the form of liquid urea ammonium nitrate (28-0-0) was applied as recommended by crop management practices in the region. Trials were rainfed and did not rely on any form of irrigation. Monthly precipitation and temperature obtained from iClimate (2019) are detailed in Supplementary Table S1.

Trait measurements. We measured grain yield (YLD), days to heading (HD), days to maturity (MD), thousand kernel weight (TKW), biomass (BIO), number of spikes per area (NS), number of grains per spike (GPS), grain weight per spike (GWS), and plant height (PH). YLD was measured at harvest, adjusted for 13% seed moisture, and was expressed as kg·ha⁻¹. HD was determined by complete emergence of heads (Feekes 10.5, Zadoks 58) in more than 50% of individual plants in a plot and expressed as the number of days after January 1st. Similarly, MD was determined when more than 50% of plot reached physiological maturity (Feekes 11.3, Zadoks 91) and expressed as the number of days after January 1st. At maturity, PH was recorded by four random measurements per plot, from the ground to the top spikelet, excluding the awns, and expressed in centimeter (cm). Yield components were evaluated by measuring traits from an area of 0.25 m $\times$ 0.3048 m that was cut from the ground level after physiological maturity. First aboveground BIO was dried to constant weight, measured and expressed in grams (g). Next effective tiller numbers per unit area were counted from the cut sample and represented as number of spikes (NS). Then, five random spikes were randomly sampled from the total cut area to measure the number of grains per spike (GPS), and grain weight per spikes (GWS)—also expressed in grams. TKW was measured by counting and weighing 1000 kernels, which was expressed in grams.

Description of genotypic data. This population was initially genotyped by using the 90K SNP chip array [43], and the marker density was later increased by completing genotyping-by-sequencing method, as explained in Poland et al. [44]. Briefly, reduced genomic libraries were

created using *Pst1-Msp1* restriction enzyme combination consistent with Poland et al. [44]. The samples were pooled together at 96-plex to create libraries and each library was sequenced on a single lane of Illumina Hi-Seq 2500 (Illumina, San Diego, USA). Variant calling was performed using the TASSEL 5 GBSv2 pipeline [45] with 64 base k-mer length and minimum k-mer count of five. Reads were aligned to the wheat genome sequence assembly IWGSCv1.0 [46], using aln method of Burrows-Wheeler aligner (BWA) version 0.7.10 [47]. For filtering of both 90K SNP chip array and GBS markers, we excluded any markers missing $\geq 10\%$ data and those with minor allele frequency less than 0.05. We then used Linkage Disequilibrium K-number neighbor imputation (LD-kNNi) algorithm [48] implemented in TASSEL 5 [45] to impute the missing markers. Markers that were not mapped to any specific chromosome were excluded from further analysis. The final genotypic dataset that was used in this study consisted of 45K variants of which 13K were produced from the 90K SNP chip array pipeline and 32K were produced from GBS pipeline.

Statistical analysis of phenotypic data. In order to test the significance of genotypes, year, and genotype $\times$ year interaction, analysis of variance (ANOVA) was performed in R environment [49]. For each trait, the following ANOVA model was fitted:

$$Y_{ijkl} = \mu + G_i + Yr_j + R_k(Yr_j) + GYr_{ij} + B_l(RYr_{kj}) + \varepsilon_{ijkl} \quad (1)$$

Where the response variable Y_{ijkl} is the observed phenotypic value of the i th genotype, in the j th year, in the k th replicate, and the l th incomplete block; μ is the overall mean, G_i is the effect of the i th genotype, Yr_j is the effect of the j th year, $R_k(Yr_j)$ is the effect of the k th replicate within the j th year, GYr_{ij} is the interaction effect of the i th genotype by the j th year, and $B_l(RYr_{kj})$ is the effect of the l th incomplete block within the k th replicate and the j th year. The ε_{ijkl} represents the residual error.

To produce phenotypic values of each line for GWAS analysis, the best linear unbiased estimate (BLUE) values were derived by implementing a mixed model [37] using the “*lme4*” package [50] in R environment [49] in equation [1], where genotype was considered as fixed effect and other terms were considered as random effects. The Pearson correlation coefficient was calculated by *cor* function in R by using BLUE values. Path analysis was performed on BLUE values by using the latent variable analysis “*lavaan*” package [51] in R environment [49].

Estimating heritability estimates. Estimation of heritability based on experimental design requires a balanced design where all experimental entries are included in each replicate. Therefore, for producing variance components for estimating the broad sense heritability (H^2), we used a reduced model as follows:

$$Y_{ijkl} = \mu + G_i + Yr_j + R_k(Yr_j) + GYr_{ij} + \varepsilon_{ijkl} \quad (2)$$

Where the response variable Y_{ijkl} is the observed phenotypic value of the i th genotype, in the j th year, in the k th replicate, and the l th incomplete

block; μ is the overall mean, G_i is the effect of the i th genotype, Yr_j is the effect of the j th year, $R_k(Yr_j)$ is the effect of the k th replicate within the j th year, and GYr_{ij} is the interaction effect of the i th genotype by the j th year. The ε_{ijkl} represents the residual error. In this model all terms were considered as random effect. The broad sense heritability (H^2) on an entry-mean basis was estimated following the equation [52,53]:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{gy/y}^2 + \sigma_{\varepsilon/yr}^2} \quad (3)$$

where σ_g^2 is the genetic variance, σ_ε^2 is the error variance, and y is the number of years ($y = 2$), and r is the number of replications per year ($r = 2$).

Genome-wide association studies (GWAS). Principal component analysis (PCA) of marker data was used to visualize the underlying population structure. We used the first three principal components (PCs) to produce a 3D scatter plot. Pair-wise LD estimates between adjacent markers were calculated, as the squared coefficient of correlation (r^2), using TASSEL 5 [45] with a sliding window of 1000 markers. The pairwise LD estimates were plotted against the physical distance to determine the decay of LD against physical range on each chromosome, and in particular around the regions, where marker-trait associations were identified in GWAS. LD decay plots generated in R using the Hill and Weir [54] method and loess regression with assessment at r^2 value of 0.2 [55,56].

GWAS was performed using the GAPIT software [57] in R for each trait using the Fixed and Random Model Circulating Probability Unification (FarmCPU) model [41] and first 3 PCs were used to control the population structure [58]. We reported MTAs that were identified at $-\log P \geq 4.0$ (p -value ≤ 0.0001). If a genomic region was identified with multiple MTAs close to each other, we only report a representative MTA. We also identified MTAs that passed a 5% false discovery rate (FDR) for controlling multiple testing [59]. The coefficient of determination (R^2) for each identified MTA was determined by fitting a linear model in R environment with the contrasting alleles of the marker and the 3 PCs as the covariates using an ordinary least squares regression.

Transferability and validation of GWAS results. YLD and HD data obtained from trials conducted in previous years and other states during the Triticeae CAP project were used to validate the transferability in other environments of the MTAs we identified in Indiana. This data comes from diverse environments i.e., five different locations and two growing seasons 2011-12 and 2012-13, as described by Huang et al. [42]. These environments are: moderate nitrogen in Kentucky 2011-12 (KYM12), moderate nitrogen in Maryland 2011-12 (MDM12), moderate nitrogen in Missouri 2011-12 (MOM12) and 2012-13 (MOM13), low nitrogen in Ohio 2011-12 (OWL12) and 2012-13 (OWL13), moderate nitrogen in Ohio 2011-12 (OWM12) and 2012-13 (OWM13), low nitrogen in Virginia 2011-12 (VAL12) and in 2012-13 (VAL13), and moderate nitrogen in Virginia 2011-

12 (VAM12) and 2012-13 (VAM13). We abbreviated grain yield and heading date we obtained from our 2016-17 and 2017-18 seasons as WL17 and WL18. In total, we assembled data from 14 environments for validation and transferability examination. For WL17 and WL18 environments, we first accounted for incomplete block design and then included the data in the multi-environment data analysis. Multi-dimensional scaling and linear discriminant analysis were used to cluster environments into seemingly homogeneous groups based on YLD or HD data. Then the accuracy of grouping was examined by cross validation. The *cmdscale* function in R was used to perform multidimensional scaling with Euclidean distances extracted using the *dist* function. Eigenvalues from three dimensions were extracted and incorporated into the *lda* function in the *MASS* package [60] for linear discriminant analysis and cross validation by setting *CV* = TRUE. Upon confirmation of groupings, BLUEs were obtained for each homogeneous group following the same model [2] and GWAS analysis was completed for each homogeneous group. We considered a MTA as validated or transferable if identified with a $-\log P > 1.3$ (p -value < 0.05) in another homogenous group of environments. We chose this threshold because for validation, we are only interested in one specific marker and there is no need to control for testing of multiple hypotheses.

RESULTS

Phenotypic analysis and relationship among traits. We evaluated grain yield and yield components of a soft red winter wheat population in West Lafayette, Indiana for two years. For all traits, the effect of genotype was significant at 0.001, indicating the presence of noticeable genetic variation in the germplasm. In addition, the effect of year, and replicate within years were significant at 0.001 for all traits except for GPS, where the effect of year, and rep within years were significant at 0.01. More importantly, the genotype $\times$ year interaction effect was significant at 0.001 for YLD, BIO, PH, HD, and MD, at 0.01 for NS, and at 0.05 for GWS, but not significant for TKW and GPS (Table 1). The significant effect of genotype $\times$ year interaction will be further discussed in the GWAS section.

Table 1. Analysis of Variance. Abbreviations as follows: GY = grain yield, TKW = thousand kernel weight, BIO = biomass, NS = number of spikes, GPS = grain per spike, GWS = grain weight per spike, PH = plant height, HD = days to heading, and MD = days to maturity.

Source of Variance	df	Trait								
		YLD	TKW	BIO	NS	GPS	GWS	PH	HD	MD
Genotype	269	***	***	***	***	***	***	***	***	***
Year	1	***	***	***	***	***	***	***	***	***
Genotype $\times$ Year	269	***	ns	***	**	ns	*	***	***	***
Rep(Year)	2	***	***	***	***	**	***	***	***	***
Block(Rep $\times$ Year)	28	**	ns	***	***	ns	ns	***	***	***

Significant values: *** < 0.001 , ** < 0.01 , * < 0.05 , ns > 0.05 .

Grain yield ranged from 3900 to 7500 kg·ha⁻¹ with a mean of 5830 kg·ha⁻¹ and heritability of 0.50 (Table 2). The top 10% highest yielding lines in the population averaged at 6940 kg·ha⁻¹, while the 10% lowest yielding lines averaged 4650 kg·ha⁻¹—a 1.5-fold difference. Not all of the 10% highest yielding entries were developed by one breeding program, indicating a potential for achieving genetic gains via germplasm exchange. Among these, high yielding lines developed from public breeding programs at Purdue (10 lines), Illinois (7 lines), Missouri (3 lines), Ohio State (2 lines), Kentucky (2 lines), and Maryland (2 lines) were identified.

Table 2. Summary Statistics and Heritabilities (H^2) based on WL1718.

Trait	Unit	Mean	SD	Minimum	Maximum	H^2
Grain yield	kg·ha ⁻¹	5827	678	3905	7500	0.50
Thousand kernel weight	grams	32	1.79	27	38	0.49
Biomass	grams	201	20	151	255	0.21
Number of spikes	count	108	13	74	152	0.48
Grain per spike	count	37	4	25	50	0.44
Grain weight per spike	grams	0.98	0.14	0.68	1.38	0.41
Plant height	centimeters	90	6	76	111	0.84
Days to heading	Julian Days (from Jan 1)	133	2.01	128	139	0.69
Days to maturity	Julian Days (from Jan 1)	171	1.28	168	175	0.62

The traits with the greatest and significant positive phenotypic correlation to YLD were BIO ($r = 0.29$ ***), TKW ($r = 0.29$ ***), and GWS ($r = 0.29$ ***). BIO had an average of 201 grams per cut area and heritability of 0.21. TKW ranged from 27.8 to 38.8 grams with a mean of 32.3 grams and heritability of 0.49 (Table 2). The three lines with the greatest kernel weight were MD04W249-11-7, 04702A1-18, and MD03W64-10-3 and the three lines with the smallest kernel weight were OH08-178-52, VA09W-188WS, and MO080584. However, looking at the top 10% high yielding entries, the range of thousand kernel weight was narrower (30-35 grams), and around the average value for kernel weight. Total grain number in wheat is the cumulative effect of spike number per unit area and the number of grains per spike. The NS per cut area ranged from 74 to 152 spikes. The lines with greatest number of spikes were OH08-172-42, TRIBUTE, and IL08-12174 and the lines with lowest number of spikes per measured area were INW1021, 0566A1-3-1-67, and 05251A1-1-136-9-5. GPS and GWS had a mean of ~37 grains per spike and 0.98 grams, respectively and similar heritability estimates (Table 2). A significant negative correlation ($r = -0.22$ ***) was observed between NS and GPS (Table 3). This negative correlation has been observed previously in multiple experiments [61,62]. PH had the highest heritability of 0.84, averaged at 90 cm, and showed a standard deviation of 6.2 cm. The tallest lines were CAYUGA, MO101329, and MO100647 while the shortest lines were 03207A1-7-3-1, 9346A1-2-5-5-2-1, and MD03W665-10-5. The height of the

10% shortest lines averaged 80 cm. Lastly, the HD and MD had a mean of 133 and 171 days, respectively, and were highly correlated with one another ($r = 0.68$ ***) (Table 3). Lines that headed later (>138 days) and matured later (>173 days) included NY103-208-7263, NY99066-3444, CAYUGA, NY96009-3037, and MEDINA, all varieties developed in New York and adapted to the eastern climate region. Both traits were significant and negatively correlated with YLD ($r = -0.19$ **, $r = -0.18$ **) (Table 3) as this relationship has been documented previously [63].

Table 3. Phenotypic correlations of BLUEs of nine measured traits. Abbreviations as follows: GY = grain yield, TKW = thousand kernel weight, BIO = biomass, NS = number of spikes, GPS = grain per spike, GWS = grain weight per spike, PH = plant height, HD = days to heading, and MD = days to maturity.

Trait	TKW	BIO	NS	GPS	GWS	PH	HD	MD
GY	0.29 ***	0.29 ***	0.04	0.07	0.29 ***	0.05	-0.19 **	-0.18 **
TKW		0.04	-0.18 *	-0.16 **	0.28 ***	0.01	-0.08	-0.16 **
BIO			0.47 ***	0.24 ***	0.22 ***	0.31 ***	0.09	0.13 *
NS				-0.22 ***	-0.31 ***	-0.14 *	-0.16 **	-0.10
GPS					0.72 ***	0.28 ***	0.22 ***	0.24 ***
GWS						0.26 ***	0.12 *	0.10
PH							0.38 ***	0.28 ***
HD								0.68 ***
MD								

Significance: *** <0.001, ** <0.01, * <0.05.

Path coefficient analysis. The correlation magnitudes were further broken down by using path analysis, following Dewey and Lu [64]. Path analysis parses out the correlation magnitude to direct and indirect components of influence by measuring the direct influence of one variable upon another [64]. In Supplementary Figure S1, the single arrow lines indicate direct influence as measured by path coefficients (P_{xx}) and the indirect effects are the association between variables measured by correlation coefficients (r_{xx}). The indirect effects are the product of the path coefficients and correlation coefficients. The sum of the path coefficients and indirect effects of correlation coefficients equal the phenotypic correlations, thus breaking down the reasoning for positive and negative correlations observed. In the a priori model, grain yield is directly affected by traits with significant phenotypic correlation (Table 3). These traits are thousand kernel weight, grain weight per spike, biomass, heading date, and maturity date. Biomass had the largest direct path coefficient of 0.27, followed by grain weight per spike and thousand kernel weight coefficients of 0.21, and 0.19, respectively. (Supplementary Table

S2). The indirect effect of thousand kernel weight on grain weight per spike represents almost one-sixth of the phenotypic correlation (Table 3) and direct path coefficient between grain weight per spike and yield (Supplementary Table S2). Biomass and grain weight per spike are correlated ($r = 0.22$ ***; Table 3) and positively contribute to correlations with grain yield. Days to heading showed a negative direct effect on grain yield with path coefficient of -0.15 (Supplementary Figure S1), consistent with its negative correlation with grain yield ($r = -0.19$ **; Table 3). Similar patterns were observed for days to maturity.

Genome-wide association studies. The objectives of this study were to identify MTAs that control grain yield and other agronomic traits in this population in the Indiana environment and examine the transferability of MTA results across other environments. Of the 45K variants used in this study, approximately 17K, 22K, and 5.7K were located on sub-genome A, B, and D, respectively. The first three principal components (PCs) of all marker data explained 6.5%, 5.2%, and 3.8% of the total variation (Supplementary Figure S2). Consistent with the reports of Gaire et al. [33] and Huang et al. [42], PCs separated two distinct groups, which were previously attributed to whether germplasm is progeny, close relative, or descendants of the soft red winter wheat variety “Truman” or not [42]. Linkage disequilibrium persisted variably across different chromosomes and the half decay distance (in base pairs) are presented in Supplementary Table S3 for each chromosome. For example, LD persisted the longest physical range on chromosomes 2B (~125 mega base pairs:Mbp) and 7D (109 Mbp). In contrast, chromosomes 5D (0.74 Mbp) and 6D (0.71 Mbp) displayed the fastest LD decay.

We used the first 3 PCs to account for the underlying population structure in GWAS analysis for all traits evaluated in West Lafayette, IN, USA. For GWAS we used estimates of phenotypic data based on two years of study i.e., WL17 and WL18, termed WL1718 throughout, and 45K genome-wide variants. In this study, we reported and discussed MTAs that were identified at $-\log P \geq 4.0$ (p -value ≤ 0.0001) threshold. A total of 62 MTAs were identified for eight traits in WL1718 except for NS on 20 chromosomes (all excluding 3D). Based on their physical distances and the LD decay, the 62 MTAs were resolved in 59 independent loci (Figure 1). Of the 59 loci, 11 passed the 5% FDR threshold for grain yield, days to heading, days to maturity, and plant height. Chromosome 3B showed the highest number of loci. Regions on chromosome 5A were found to be associated with four phenotypic traits including grain per spike, days to maturity, plant height, and thousand kernel weight (Figure 1; Table 4). Plant height showed maximum number of MTAs among traits. None of the MTAs were associated with multiple traits.

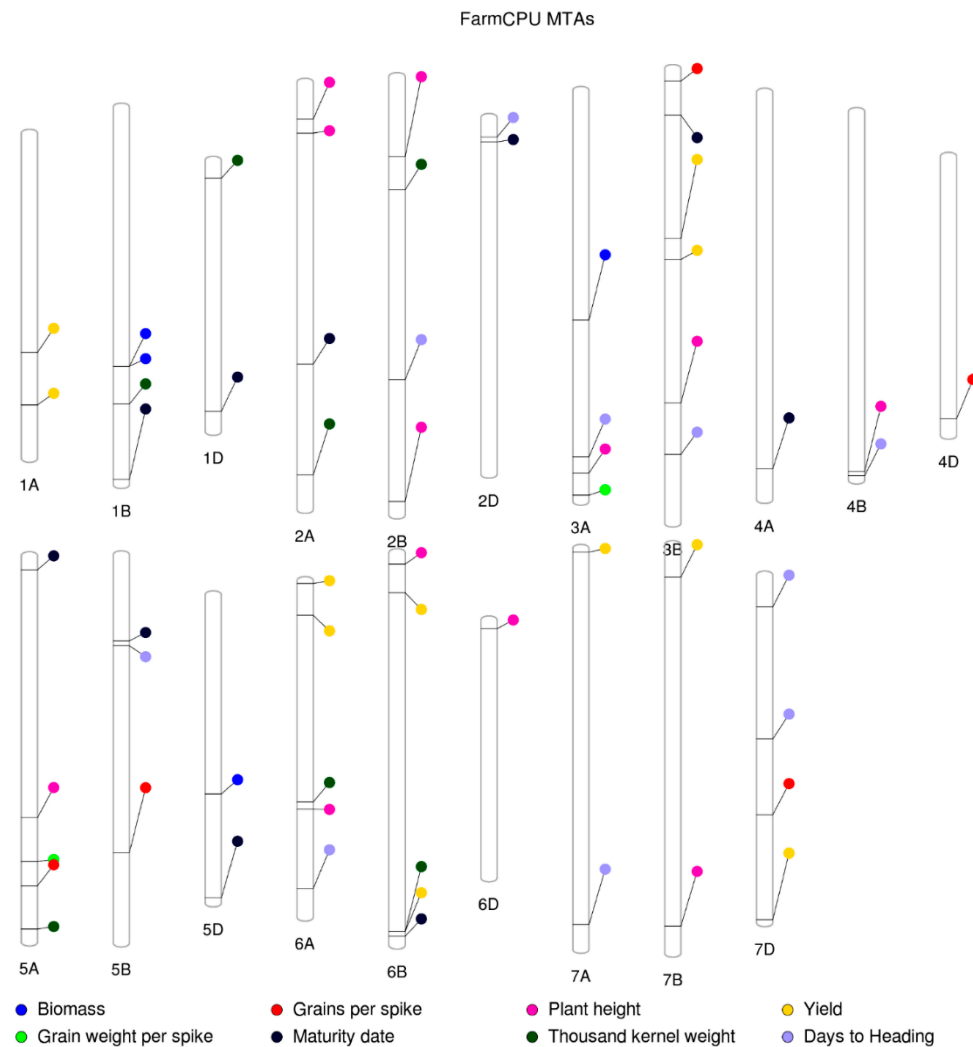


Figure 1. Genetic map of all significant MTAs identified by FarmCPU method for yield component traits. Chromosome size extracted from Ensembl. Figure created by phenogram software (<http://visualization.ritchielab.org/phenograms/plot#>).

For YLD, eleven MTAs were reported on chromosomes 1A, 3B, 6A, 6B, 7A, 7B, and 7D (Figure 2). The MTA with the largest $-\log P$ value of 16.35 on chromosome 7D located at 633,027,374 base pairs (bp) explained 18% of phenotypic variation for grain yield. The next largest signal on chromosome 1A of $-\log P = 8.27$ had allele effect of $174 \text{ kg}\cdot\text{ha}^{-1}$ (Table 4).

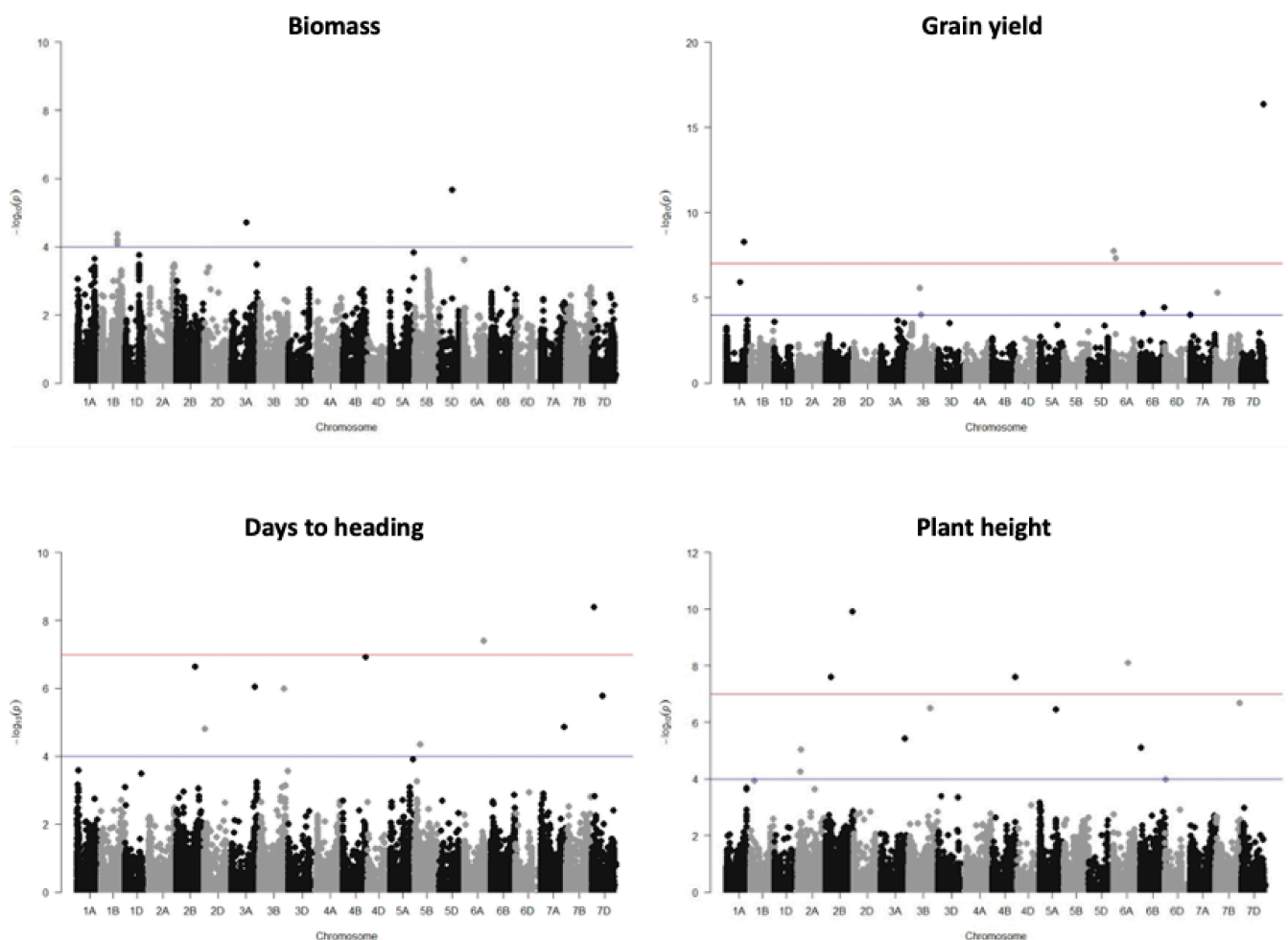


Figure 2. Manhattan plots of traits based on FarmCPU method. Blue horizontal line indicates $-\log P = 4.0$, and red horizontal line indicates 5% FDR threshold.

Five MTA were identified for GPS on 3B, 4D, 5A, 5B, and 7D (Table 4). These marker effects accounted for approximately 2 grain per spike and explained 4–7% of the phenotypic variation. One MTAs for GWS was found on chromosome 3B. Marker *gbs_3A_739555657* explained 8% of the phenotypic variation and accounted for an allele effect of 63 milligrams of grain weight per spike (Table 4). Lastly, TKW had 7 MTAs on chromosomes 1D, 2A, 3B, 5A, 6A, and 6B (Table 4). The strongest signal for TKW was identified on 5A at position 685,795,509 bp. This region exerted an effect of 540 mg and covered 10% of total phenotypic variation. The next largest signal was observed at position 206,962,855 bp on chromosome 2B with an effect of 690 mg and phenotypic explanation of 8%.

Table 4. Marker trait associations for yield and yield component traits in WL1718 environment.

Trait	Chr	Marker	Position	Alleles ^a	MAF	<i>-logP</i>	Effect	<i>R</i> ²	Units
Biomass	1B	IWA6758	474118005	<u>A</u> /G (226/20)	0.12	4.07	12.30	0.04	grams
	1B	IWB72708	473529137	<u>T</u> /C (227/19)	0.11	4.36	12.80	0.05	grams
	3A	3A_419257151	419257151	<u>A</u> /G (241/9)	0.07	4.69	13.41	0.07	grams
	5D	5D_365732020	365732020	A/ <u>C</u> (212/51)	0.21	5.65	8.00	0.09	grams
Grain per spike	3B	3B_22698880	22698880	A/ <u>G</u> (246/20)	0.08	4.07	1.99	0.06	count
	4D	4D_479593371	479593371	G/ <u>A</u> (229/39)	0.15	4.33	1.75	0.07	count
	5A	5A_606524326	606524326	C/ <u>G</u> (221/44)	0.17	4.08	1.67	0.06	count
	5B	5B_546826603	546826603	G/ <u>A</u> (242/24)	0.1	4.61	2.33	0.04	count
Grain weight per spike	7D	7D_440881288	440881288	G/ <u>A</u> (233/34)	0.13	4.04	1.87	0.05	count
	3A	3A_739555657	739555657	T/ <u>C</u> (238/31)	0.12	4.60	0.063	0.08	grams
Days to maturity	1B	1B_680465515	680465515	G/ <u>A</u> (241/6)	0.06	4.92	0.64	0.05	Julian days
	1D	1D_458723021	458723021	G/ <u>A</u> (247/21)	0.08	4.30	0.48	0.03	Julian days
	2A	2A_515253009	515253009	<u>T</u> /C (230/3)	0.08	6.24	0.59	0.03	Julian days
	2D	2D_44485665	44485665	<u>A</u> /G (237/31)	0.12	9.41	0.74	0.16	Julian days
	3B	3B_85344544	85344544	<u>C</u> /T (238/7)	0.07	6.46	0.71	0.04	Julian days
	4A	4A_688222191	688222191	<u>A</u> /G (244/7)	0.06	4.72	0.49	0.03	Julian days
	5A	5A_26153196	26153196	G/ <u>A</u> (135/126)	0.48	4.28	0.21	0.02	Julian days
	5B	5B_158399441	158399441	G/ <u>A</u> (239/28)	0.11	4.91	0.32	0.02	Julian days
	5D	IWB54292	556553226	T/ <u>G</u> (173/96)	0.36	4.31	0.19	0.05	Julian days
	6B	IWA3268	705159045	<u>T</u> /C (232/35)	0.14	4.09	0.33	0.07	Julian days
Plant height	2A	2A_66985350	66985350	C/ <u>A</u> (248/19)	0.08	4.24	2.39	0.06	centimeters
	2A	IWB51951	92797308	<u>T</u> /G (227/43)	0.16	5.01	1.46	0.08	centimeters
	2B	2B_146441175	146441175	A/ <u>G</u> (230/34)	0.14	7.60	1.77	0.04	centimeters
	2B	2B_776795892	776795892	G/ <u>A</u> (215/5)	0.11	9.90	3.00	0.04	centimeters
	3A	3A_699195908	699195908	T/ <u>C</u> (233/6)	0.08	5.41	2.28	0.07	centimeters
	3B	IWB9589	611497265	<u>T</u> /C (211/52)	0.21	6.49	1.30	0.10	centimeters

Table 4. *Cont.*

Trait	Chr	Marker	Position	Alleles ^a	MAF	<i>-logP</i>	Effect	<i>R</i> ²	Units
Plant height	4B	IWB43355	657825660	A/ <u>G</u> (150/114)	0.43	7.59	1.45	0.07	Centimeters
	5A	5A_480705790	480705790	<u>C</u> /A (255/13)	0.05	6.45	2.64	0.03	centimeters
	6A	6A_419959989	419959989	T/ <u>G</u> (149/113)	0.43	8.09	1.29	0.07	centimeters
	6B	6B_21208064	21208064	G/ <u>A</u> (239/29)	0.11	5.08	1.51	< 0.01	centimeters
	7B	IWA4750	701186266	<u>A</u> /G (172/93)	0.35	6.67	1.20	0.01	centimeters
Thousand kernel weight	1B	1B_542725487	542725487	A/ <u>C</u> (212/53)	0.21	4.77	0.37	0.03	grams
	1D	1D_32441418	32441418	<u>T</u> /C (232/27)	0.12	5.08	0.49	0.02	grams
	2A	2A_718459754	718459754	T/ <u>G</u> (253/16)	0.06	4.37	0.57	0.04	grams
	2B	2B_206962855	206962855	A/ <u>G</u> (236/14)	0.09	5.89	0.69	0.08	grams
	5A	5A_685795509	685795509	C/ <u>A</u> (226/42)	0.16	6.30	0.54	0.10	grams
Grain yield	6A	6A_406733069	406733069	<u>A</u> /G (247/8)	0.06	4.68	0.82	0.07	grams
	6B	6B_695913077	695913077	G/ <u>A</u> (225/44)	0.16	5.67	0.45	0.02	grams
	1A	IWA5011	400311021	<u>T</u> /C (235/33)	0.13	5.90	191	0.04	kg·ha ⁻¹
	1A	1A_496309488	496309488	<u>G</u> /A(199/59)	0.24	8.27	174	0.06	kg·ha ⁻¹
	3B	IWB32652	349636369	<u>A</u> /G (172/96))	0.36	4.00	97	0.03	kg·ha ⁻¹
	3B	3B_310333182	310333182	<u>G</u> /A (221/19)	0.13	5.56	213	0.03	kg·ha ⁻¹
	6A	IWB26414	5326425	<u>A</u> /G (236/30)	0.12	7.73	224	0.03	kg·ha ⁻¹
	6A	IWB63176	63563014	<u>A</u> /G (193/74)	0.28	7.30	163	0.15	kg·ha ⁻¹
	6B	IWB38887	696150409	A/ <u>G</u> (143/126)	0.47	4.41	92	0.05	kg·ha ⁻¹
	6B	6B_73187805	73187805	<u>G</u> /A (251/15)	0.06	4.07	250	0.14	kg·ha ⁻¹
	7A	IWB59141	6499010	A/ <u>C</u> (194/69)	0.27	4.00	117	0.11	kg·ha ⁻¹
	7B	IWB6720	59632081	A/ <u>C</u> (227/39)	0.15	5.28	154	0.05	kg·ha ⁻¹
	7D	7D_633027374	633027374	C/ <u>T</u> (236/17)	0.09	16.35	492	0.18	kg·ha ⁻¹
Days to heading	2B	IWB34502	553613770	<u>T</u> /C (144/126)	0.44	6.62	0.44	0.01	Julian days
	2D	2D_35683268	35683268	<u>G</u> /A (172/90)	0.35	4.81	0.37	0.02	Julian days
	3A	IWB6009	669524837	T/ <u>G</u> (144/126)	0.47	6.03	0.39	0.05	Julian days

Table 4. *Cont.*

Trait	Chr	Marker	Position	Alleles ^a	MAF	<i>-logP</i>	Effect	<i>R</i> ²	Units
Days to heading	3B	3B_705185712	705185712	<u>C</u> /T (142/121)	0.46	5.98	0.39	0.02	Julian days
	4B	4B_665871684	665871684	<u>C</u> /T (241/10)	0.07	6.92	1.11	0.02	Julian days
	5B	5B_167440402	167440402	A/ <u>G</u> (241/26)	0.10	4.33	0.42	0.02	Julian days
	6A	6A_565344991	565344991	C/ <u>A</u> (151/110)	0.42	7.39	0.46	0.02	Julian days
	7A	7A_690860911	690860911	G/ <u>A</u> (185/66)	0.28	4.86	0.38	0.05	Julian days
	7D	7D_301325415	301325415	<u>G</u> /T (243/21)	0.09	5.77	0.74	0.10	Julian days
	7D	7D_58927880	58927880	C/ <u>T</u> (204/61)	0.24	8.38	0.58	0.08	Julian days

^a The underlined nucleotide represents the favorable allele. For days to heading, days to maturity, and plant height, the favorable allele was reducing whereas all other traits the favorable allele was considered as increasing;

Chr: chromosome;

MAF: minor allele frequency;

*R*²: coefficient of determination.

For BIO, we identified 4 MTAs on chromosomes 1B, 3A, and 5D (Figure 2, Table 4). The largest signal for biomass was identified on chromosome 5D at position 365,732,020 bp with $-\log P$ of 5.65 that explained 9% of variation observed in biomass. The next large signal for biomass was $-\log P$ of 4.69 on chromosome 3A. Independent MTAs for BIO represented 4–9% of the phenotypic variation with positive allele effects between 8.00 and 13.41 grams.

Ten MTAs were identified for days to heading for WL1718 across nine chromosomes (Figure 2). Two MTAs on 7D had $-\log P$ values of 5.77 and 8.38 with allele effects of 0.74 and 0.58 earlier heading date, respectively. Ten MTAs were identified with $-\log P$ up to 9.41 for days to maturity (Table 4). The most significant signal was identified at 44,485,665 bp position of chromosome 2D, which explained 16% of variation. Eleven MTAs were identified with $-\log P$ up to 9.90 for PH (Figure 2). One marker on 6D explained 16% of the phenotypic variation for plant height and had a minor allele frequency of 0.07 (Table 4).

Transferability of GWAS results. We used existing YLD and HD data that were generated from the same germplasm in other states and seasons. Altogether, we assembled 14-environment datasets, of which WL17 and WL18 are from our field testing in Indiana. Linear discriminant analysis (LDA) on grain yield resulted in three homogeneous groups (Figure 3A) and on heading date resulted in four homogenous groups (Figure 3B). Strikingly, we observed that year-to-year variations resulted in different groupings in some cases (Supplementary Table S4). For example, for grain yield, LDA group 1 included WL17, KYM12, MDM12, MOM12, and MOM13, group 2 included WL18, OWL12, OWM12, VAL12, OWL13, and OWM13, and group 3 consisted of VAL13 and VAM13. We observed that for example, VAL12 and VAL13 are categorized in different groups (Figure 3A). Similar observation was true for WL17 and WL18. In addition, we noticed that groupings were different for grain yield and heading date. LDA for grain yield and heading date had a percent separation above 87% for each discriminant function and cross-validation confirmed successful separation of environments.

We performed GWAS for YLD and HD based on phenotypic observations from four environments: WL1718, Group 1, Group 2, and Group 3. In Group 1, twelve MTAs were identified in chromosomes 1B, 2B, 5A, 5B, 6A, 6B, 7A, 7B, and 7D. Three MTAs were present on chromosome 6B and two MTAs on 7A. For Group 3, eight MTAs were identified on chromosomes 3B, 5B, and 7B, however, applying the same standard for markers in LD as above, resolved to five independent MTAs. No MTAs were identified in Group 2. When we compared YLD signals among the three homogenous groups, there was not any MTA identified in more than one group, indicating that QTL are specific to each group.

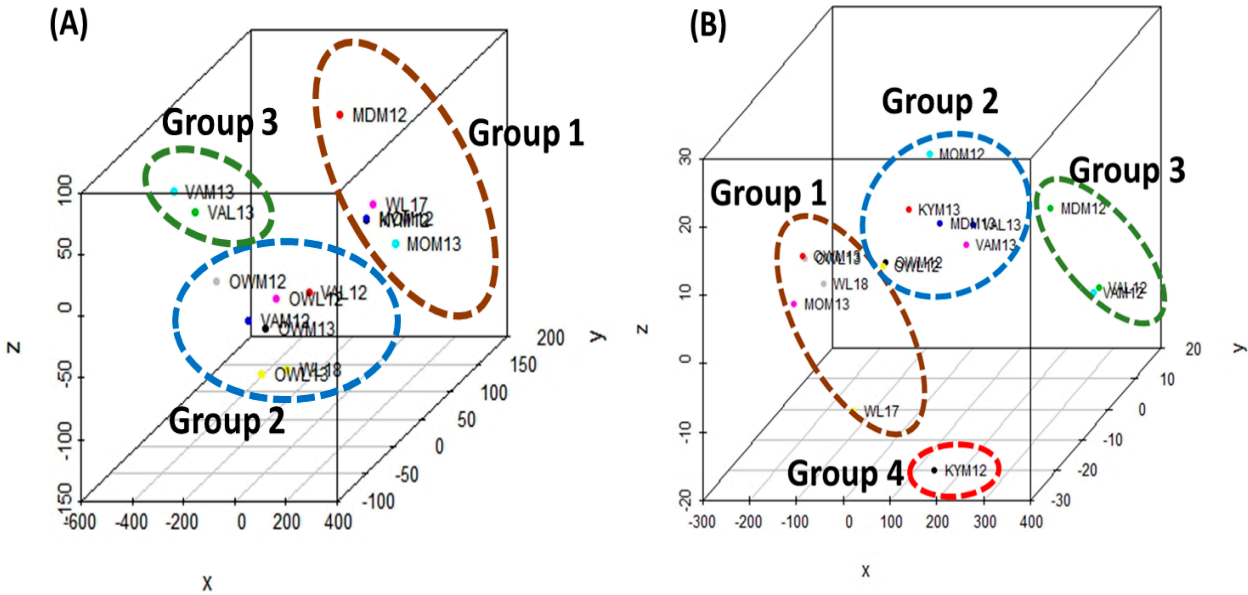


Figure 3. Grouping of environment and year based on linear discriminant analysis. 3D plot of multi-dimensional scaling to visually observe groupings based on (A) grain yield and (B) days to heading.

Table 5. $-\log P$ value and marker effect for significant multi-environment marker trait associations.

Trait	Chr	SNP	Environment							
			WL1718		Group 1		Group 2		Group 3	
			$-\log P$	Effect	$-\log P$	Effect	$-\log P$	Effect	$-\log P$	Effect
Grain yield	1A	IWA5011	5.90	191	1.55	64	1.68	96	-	-
	1A	1A_496309488	8.27	174	2.73	71	-	-	-	-
	3B	3B_310333182	5.56	213	-	-	-	-	1.69	170
	6A	IWB63176	7.30	163	-	-	-	-	1.57	94
	6B	6B_73187805	4.07	250	7.75	238	-	-	2.38	235
	7A	IWB59141	4.00	117	1.72	56	-	-	3.04	150
	7D	7D_633027374	16.35	492	20.87	393	-	-	1.64	184
Days to heading	2B	IWB34502	6.62	0.44	5.53	0.29	1.94	0.30	-	-
	3A	IWB6009	6.03	0.39	1.85	0.16	-	-	-	-
	3B	3B_705185712	5.98	0.39	-	-	1.40	0.24	-	-
	4B	4B_665871684	6.92	1.11	6.74	0.84	2.08	0.81	-	-
	5B	5B_167440402	4.33	0.42	1.46	0.22	-	-	-	-
	7A	7A_690860911	4.85	0.38	9.10	0.40	2.23	0.33	-	-
	7D	7D_301325415	5.77	0.74	7.48	0.68	5.50	0.98	-	-
	7D	7D_58927880	8.38	0.57	4.44	0.30	1.48	0.29	-	-

Represented MTAs based on the accepted threshold ($-\log P$ value > 1.3).

A total of 28 independent MTAs were identified across environmental groupings for YLD but we only noticed seven MTAs that were identified in at least two environments which are indicative of transferability across environments. Two of these MTAs are located on chromosomes 6B and 7D. The MTA on chromosome 6B for YLD at position 73,187,805 bp was identified in WL1718, Group 1, and Group 3 environments with $-\log P$ of 4.07, 7.75, and 2.38, respectively (Table 5). The marker effect for this

validated MTA showed an effect size of 238–250 kg·ha⁻¹ across environments. The MTA on chromosome 7D for YLD is at 633,027,374 bp, and was identified in WL1718, Group 1, and Group 3 environments with $-\log P$ of 16.35, 20.87, and 1.64, respectively (Table 5). The marker effect of this MTA was approximately 492 kg·ha⁻¹ in WL1718, 393 kg·ha⁻¹ in Group 1, and 184 kg·ha⁻¹ in Group 3 (Table 5).

For HD, LDA grouping clustered environments into four groups. LDA group 1 included WL17, WL18, MOM13, OWL13, and OWM13, group 2 included MOM12, OWL12, OWM12, VAL13, and VAM13, and group 3 consisted of MDM12, VAL12, and VAM12 (Supplementary Table S4). KYM12 was a singleton Group 4, with no other group member (Figure 3B), and was left out of the analysis. GWAS was performed for these three groups. Group 1 had 35 MTAs that were grouped into 26 independent loci. Eleven of these loci were located on chromosome 7A and five were located on chromosome 7D. For Group 2, eleven MTAs were identified on chromosomes 1A, 1B, 3A, 4B, 5B, 6A, 6B, 7A, 7B, and 7D. Lastly, Group 3 did not show any significant MTAs for HD. When we compared HD signals among the three homogenous groups, only one MTA, marker 7D_301325415 on chromosome 7D, was present in more than one group (Table 5).

A total of 47 MTAs were detected for heading date across environments but we only noticed eight MTAs that were identified in at least two environments which are indicative of transferability across environments. These MTAs were identified on chromosomes 2B, 3A, 3B, 4B, 5B, 7A, and 7D (Table 5). For HD, one marker from the SNP chip array, IWB34502 located at 553,613,770 bp on chromosome 2B was associated with days to heading (Table 5), in WL1718, Group 1, and Group 2 environments with allele effect of 0.44, 0.29, and 0.30 days, respectively. A marker with similar effects in the same environments was identified at 690,860,911 bp on chromosome 7A with a $-\log P$ of 4.86, 9.10, and 2.23 in WL1718, Group 1, and Group 2, respectively (Table 5). Chromosome 7D contained two markers significant for days to heading. The positive allele associated with this marker (301,325,415 bp) on 7D showed effect sizes of 0.74, 0.68, and 0.98 days for WL1718, Group 1, and Group 2, respectively (Table 5). The marker at position 58,927,880 bp on chromosome 7D was found to be associated with heading date in environment WL1718, Group 1, and Group 2 with $-\log P$ of 8.38, 4.44, and 1.48 (Table 5).

For YLD, seven out of 28 MTAs and for HD, eight out of 47 MTAs were found to be transferable across seemingly homogenous environments. Therefore, we concluded that not all marker-trait associations are transferable and MTAs are often environment specific.

DISCUSSION

In this study, we analyzed associations between genotypes and phenotypes in a US SRW wheat elite population, consisting of breeding lines that were developed by breeding programs in the Midwest and east.

Marker-trait associations for this population have been previously identified for *Fusarium* head blight [65], days to heading [66], and grain quality [33] from plants grown in Ohio and Virginia. We dissected the genetic architecture of grain yield and related traits based on phenotypes observed in Indiana. In addition, we examined the transferability of SNPs across environments for the traits of YLD and HD.

Phenotypic correlations among traits and deciphering their relationship can give insight into identifying selection criteria for improving traits of interest. Our study showed that grain components including TKW, BIO, and GWS were significantly and positively correlated with YLD. Previous studies have documented positive relationship between TKW and YLD as well [67,68]. In wheat breeding research, biomass is often referred to as the whole aboveground plant parts. The pre-Green Revolution wheat germplasm were tall, and their height was the driver of plant aboveground weight. Therefore, during the Green Revolution the main force that led to increases in harvest index and productivity was only reducing plant height. In this population, although variation in biomass was observed, we think that in this era a “useful biomass” is one that can lead to non-competing multiple well-grown culms (tillers) with the potential to lead to a fertile spike. Increasing tiller numbers or protecting tillers in soft red winter wheat is one approach that can produce useful biomass. Our data showed that NS and BIO were significantly correlated ($r = 0.47$) and that NS is distributed in a wide range from 74 to 152. For example, the varieties OH08-172-42, IL08-12174, MD05W1292-11-1, 05264A1-1-3-2, and IL07-20728 showed averages above 240 grams for BIO and 134 NS. Other traits that can lead to useful biomass are smaller leaves with enhanced photosynthetic capacity and the levels of spike fertility, among others.

While TKW, BIO, and GWS showed positive correlation with YLD, the duration of vegetative growth period, indicated by days to heading (and similarly days to maturity) negatively correlated with YLD. Similar negative correlation was reported by Addison et al. [63]. Addison et al. [63] noticed this trend in a SRW wheat recombinant inbred line (RIL) population across nine environments in the southern US, with the population segregating for photoperiod and vernalization loci. Grain number is the main driver of grain yield but no correlation was observed. This is a population of elite lines therefore; loci influencing traits relating to grain number could be potentially fixed in the population of elite germplasm.

There are reports in the literature that shows positive correlations between days to heading and grain yield, [69], especially under cooler temperatures for hard red spring wheat [70]. The primary reason for the observed negative correlation between days to heading and yield in this population could be that most of the late heading germplasm were developed by and adapted to the state of New York. Therefore, a hidden G $\times$ E interaction works contrary to the yield formation. Path analysis affirms

the consequence of heading later is indirectly decreasing grain development as observed in thousand kernel weight and grain yield. Therefore, a practical consideration for future characterization of populations that are mixture of germplasm from multiple crop breeding programs is that the experimenters can use days to heading as biomarkers because a the shift in phenology could mask yield traits. When a drastic change between native germplasm and others is observed, yield differences are likely expected. Tessmann et al. [71] used QTL markers for plant height, vernalization, and photoperiod genes along with the actual heading date trait as covariates in the GWAS model to account for the latitude differences. This method is also routinely performed for maize but including flowering time (days to anthesis) as covariates [72,73].

One major concern in GWAS discoveries is marker density. Wheat is a self-pollinated crop and the germplasm has been under selection. Therefore, in the beginning of the experiment, 45K markers seemed unnecessarily dense. We found evidence to the contrary. Significant MTAs with SNPs from GBS were 32 for yield components measured in WL1718 and 10 multi-environment MTAs for GY and HD. In contrast, markers from the SNP chip array contributed 8 MTA for yield components and 5 multi-environment MTA for GY and HD. This data indicates that MTAs were identified from both sets of SNP markers. In addition, we examined the inter-marker spaces for SNP chip markers located between two GBS markers. For example, the SNP IWB72708 (identified for biomass at $-\log P = 4.36$) is located 897,950 bp downstream of gbs_1B_472631187 while 296,024 bp upstream of gbs_1B_473,825,161. The SNP IWB51951 (identified for plant height at $-\log P = 5.01$) is located 458,542 bp downstream of gbs_2A_92,338,766 while 1,095,722 bp upstream of gbs_2A_93,893,030. While these distances must be judged based on the local LD decay rates in each region, our conclusion is that the 45K marker set, combined from chip array and GBS methods, is not in excess for this germplasm and the combination of both marker sets can be complementary in GWAS applications.

FarmCPU is a multiple loci linear mixed model that eliminates confounding effects between markers and kinship by iterating between both fixed and random effect models. In the fixed effect model, individual SNPs are tested while using pseudo-QTNs as covariates to control false positives. The FarmCPU model controls false positives, false negatives, and provides greater statistical power than alternative models used for association mapping [39,41]. Based on quantile-quantile (Q-Q) plots, FarmCPU effectively controlled false positives and false negatives based on the population structure and significant associations (Supplementary Figure S3). The Q-Q plot line holds close to the 1:1 line of expected versus observed association probabilities, with a slight upward tail indicating deviation from expected distribution. A deviation in the tail area indicates properly controlling false positives and false negatives, where any inflation line upward would indicate false positives or downward indicate

false negatives [39]. Other researchers reported similar claims. Xu et al. [74] and Vanous et al. [75] concluded the multi-locus model of FarmCPU provided more statistical power than single locus models with less over or under fitting. The FarmCPU model identifies the most significant single SNP at a specific genomic location instead of a large peak of SNPs with other MLM models [39].

In the target environment of Indiana, several loci affected yield and yield components traits were identified. Germplasm were also identified that harbor those favorable alleles. The 17 lines that harbored the favorable yield QTL for the region on 7D were all developed by Purdue's small grains breeding program. It is possible that all 17 of these lines contain a 7E translocation for resistance to barley yellow dwarf and cereal yellow dwarf virus, and are descendants of the Purdue line "P107" [76]. However, we could identify the 7E translocation harboring line in the pedigrees of only 11 out of the 17 lines. This translocation could explain the slow LD decay rate in over 100 Mbp on chromosome 7D.

QTL expressed in one environment may not equally or ever be expressed in other environments. To a large degree, this can be associated with the key environmental clues that are critical regulatory event for the mode of action and expression of traits and QTL. For example, if the mode of action of a growth QTL is via tiller development before winter that are later on sensitive to freezing temperature, then two environments differing in winter temperature would results in different number of tillers that are counted in the spring. Therefore, QTL could go unnoticed in the colder environment. Similar examples can be given for kernel weight QTL expression under two hot and mild grain-fill period temperatures. Such QTL by environment interaction effects can vary depending on the location and specific year. To identify stable QTLs, GWAS on the basis of combined analysis of years and locations is suggested, which is often known as multi-environment GWAS [36,77] for future QTL implementation in marker assisted selection [78,79]. However, our results showed that majority of MTAs are environment specific. Even when we contained GWAS analysis within homogenously environments, the majority of MTAs we identified in WL site for YLD and HD and were not observed in other environments. Even when markers were significant across environments, there was differences for phenotypic variations that was explained by each marker and the size of marker effect. For some traits such as grain yield the magnitude of variance component due to $G \times Y$ was 20% greater than the magnitude of variance component due to G . Since winter wheat is grown over nine months, variation in climate and weather can directly impact the year to year variability and effect of the environment. For example, the WL site in 2017 showed significantly higher monthly temperature than WL18 site from February through April (Supplementary Table S1), which is a critical time in winter wheat development. With the increase in temperature, the vernalization period for 2017 was shorter than 2018, resulting in a decrease in yield. This could

be one potential reason for the difference in classifying the WL17 and WL18 site into different groups for YLD. Previous work is a mix of success and failure in the transferability of QTL across environments. Guan et al. [80] identified 226 QTL controlling yield component traits and heat susceptibility in a Chinese elite double haploid winter wheat population. Across the 12 environments in northern China, only 39 of these QTL were deemed “stable” based on detection in at least three individual environments. Further explanation could be the significant source of variance based on effect of environment and effect of genotype by environment on all measured traits. In the United Kingdom, a double haploid population was developed from favorable bread making hexaploid winter wheat cultivars to detect QTL controlling yield variation. The population was evaluated and phenotyped at five field trials across multiple years in England, Scotland, Germany, and France. Two QTL were mapped on chromosome 6B for grain size and yield, *Qtgw-jic.6A* and *Qyld-jic.6A*, that were stable across nine of the twelve environments [81]. These favorable QTL validated with near isogenic lines displayed improvements of 5.5% and 5.1% for grain yield and grain weight.

Seeking stable QTLs for yield determining traits may not be the most thoughtful approach to improve stability and genetic gains for wheat breeding. QTL transferability is challenging, and we suggest proceeding with caution to identify QTLs across multiple environments. In our case, detecting MTAs in homogenous environments showed minimal opportunities for making progress across regions or even for developing biomarkers for marker assisted selection. We suggest performing GWAS and evaluating MTAs in the targeted breeding environment. The ability to utilize past data is powerful for predictability and examining transferability, however, the effect of the environment could be the leading issue in non-transferable QTLs controlling significant MTAs.

SUPPLEMENTARY MATERIALS

The following supplementary materials are available online at <https://doi.org/10.20900/cbgg20200013>:

Supplementary Table S1. Monthly precipitation and temperature in West Lafayette, Indiana, for the two cropping seasons of the study;

Supplementary Table S2. Path coefficient analysis;

Supplementary Table S3. LD decay half distance per chromosome and genome;

Supplementary Table S4. Grouping of environments from linear discriminant analysis for GWAS based on grain yield and days to heading;

Supplementary Figure S1. Path coefficient analysis diagram;

Supplementary Figure S2. Principal component analysis based on all SNPs for 270 lines. 3D scatterplot of first three principal components (PCs);

Supplementary Figure S3. Q-Q plots from Manhattan plots of phenotypic traits from Figure 2.

AUTHOR CONTRIBUTIONS

BR and MM planned and designed the research. BR collected field measurements and analyzed the data. CHS provided seed and additional datasets. GBG performed genotyping. BR and MM wrote, edited, and revised the manuscript. All authors read and approved the manuscript.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest. This work is a chapter of Dr. Blake Russell's PhD dissertation at Purdue University.

FUNDING

Financial support from USDA Hatch grant 1013073 via Purdue College of Agriculture to MM is greatly appreciated.

ACKNOWLEDGEMENTS

We thank Purdue's Agronomy Farm and James Beaty for land preparation, weed control, and Jason Adams for planting assistance. For past data implemented, the authors are thankful to all researchers under the Triticeae CAP project, funded by NIFA AFRI Grant no. 2011-68002-30029 from the USDA National Institute of Food and Agriculture, who created the grain yield and days to heading validation datasets in multiple states. We also thank the Eastern Regional Small Grains Genotyping Lab for genotyping of germplasm.

REFERENCES

1. Bushuk W. Wheat breeding for end-product use. In: Braun H-J, Altay F, Kronstad WE, Beniwal SPS, McNab A, editors. *Wheat: Prospects for global improvement*. Dordrecht (The Netherlands): Springer; 1997. p. 203-11. https://doi.org/10.1007/978-94-011-4896-2_27
2. USDA. Economics, Statistics, and Market Information System. 2019. Available from: <https://usda.library.cornell.edu/concern/publications/jo98zbo9z>. Accessed 2019 Jun 28.
3. Leff B, Ramankutty N, Foley JA. Geographic distribution of major crops across the world. *Global Biogeochem Cycles*. 2004 Mar;18(1):GB1009. <https://doi.org/10.1029/2003GB002108>
4. Bacon RK. US Soft Winter Wheat Pool. In: Bonjean AP, Angus WJ, editors. *The World Wheat Book: A History of Wheat Breeding*. New York (US): Intercept; 2001. p. 469-78.
5. Kiszonas AM, Fuerst EP, Morris CF. A comprehensive survey of soft wheat grain quality in US germplasm. *Cereal Chem*. 2013 Jan;90(1):47-57. <https://doi.org/10.1094/CCHEM-06-12-0073-R>
6. Reynolds M, Foulkes J, Furbank R, Griffiths S, King J, Murchie E, et al.

- Achieving yield gains in wheat. *Plant Cell Environ.* 2012 Oct;35(10):1799-823. <https://doi.org/10.1111/j.1365-3040.2012.02588.x>
7. Hawkesford MJ, Araus JL, Park R, Calderini D, Miralles D, Shen T, et al. Prospects of doubling global wheat yields. *Food Energy Secur.* 2013 May;2(1):34-48. <https://doi.org/10.1002/fes3.15>
 8. Ray DK, Mueller ND, West PC, Foley JA. Yield trends are insufficient to double global crop production by 2050. *PLoS One.* 2013 Jun 19;8(6):e66428. <https://doi.org/10.1371/journal.pone.0066428>
 9. Reynolds MP, Rajaram S, Sayre KD. Physiological and genetic changes of irrigated wheat in the post-green revolution period and approaches for meeting projected global demand. *Crop Sci.* 1999 Nov;39(6):1611-21.
 10. Donmez E, Sears RG, Shroyer JP, Paulsen GM. Genetic gain in yield attributes of winter wheat in the Great Plains. *Crop Sci.* 2001 Sep;41(5):1412-9. <https://doi.org/10.2135/cropsci2001.4151412x>
 11. Green AJ, Berger G, Griffey CA, Pitman R, Thomason W, Balota M, et al. Genetic yield improvement in soft red winter wheat in the Eastern United States from 1919 to 2009. *Crop Sci.* 2012 Sep;52(5):2097-108. <https://doi.org/10.2135/cropsci2012.01.0026>
 12. Austin RB, Bingham J, Blackwell RD, Evans LT, Ford MA, Morgan CL, et al. Genetic improvements in winter wheat yields since 1900 and associated physiological changes. *Journal Agric Sci.* 1980 Jun;94(3):675-89.
 13. Reynolds MP, Pellegrineschi A, Skovmand B. Sink-limitation to yield and biomass: a summary of some investigations in spring wheat. *Ann Appl Biol.* 2005 Jan;146(1):39-49.
 14. Devos KM, Atkinson MD, Chinoy CN, Liu CJ, Gale MD. RFLP-based genetic map of the homoeologous group 3 chromosomes of wheat and rye. *Theor Appl Genet.* 1992 May 1;83(8):931-9. <https://doi.org/10.1007/BF00232953>
 15. Helentjaris T, Slocum M, Wright S, Schaefer A, Nienhuis J. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theor Appl Genet.* 1986 Sep 1;72(6):761-9. <https://doi.org/10.1007/BF00266542>
 16. Doerge RW. Mapping and analysis of quantitative trait loci in experimental populations. *Nat Rev Genet.* 2002 Jan;3(1):43-52. <https://doi.org/10.1038/nrg703>
 17. Kumar N, Kulwal PL, Balyan HS, Gupta PK. QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. *Mol Breed.* 2007 Feb 1;19(2):163-77. <https://doi.org/10.1007/s11032-006-9056-8>
 18. Cui F, Li J, Ding A, Zhao C, Wang L, Wang X, et al. Conditional QTL mapping for plant height with respect to the length of the spike and internode in two mapping populations of wheat. *Theor Appl Genet.* 2011 May 1;122(8):1517-36. <https://doi.org/10.1007/s00122-011-1551-6>
 19. Paliwal R, Röder MS, Kumar U, Srivastava JP, Joshi AK. QTL mapping of terminal heat tolerance in hexaploid wheat (*T. aestivum* L.). *Theor Appl Genet.* 2012 Aug 1;125(3):561-75. <https://doi.org/10.1007/s00122-012-1853-3>
 20. Olmos S, Distelfeld A, Chicaiza O, Schlatter AR, Fahima T, Echenique V, et al. Precise mapping of a locus affecting grain protein content in durum wheat.

- Theor Appl Genet. 2003 Nov 1;107(7):1243-51. <https://doi.org/10.1007/s00122-003-1377-y>
21. Löffler M, Schön CC, Miedaner T. Revealing the genetic architecture of FHB resistance in hexaploid wheat (*Triticum aestivum* L.) by QTL meta-analysis. Mol Breed. 2009 Apr 1;23(3):473-88. <https://doi.org/10.1007/s11032-008-9250-y>
 22. Shen X, Zhou M, Lu W, Ohm H. Detection of *Fusarium* head blight resistance QTL in a wheat population using bulked segregant analysis. Theor Appl Genet. 2003 Apr 1;106(6):1041-7. <https://doi.org/10.1007/s00122-002-1133-8>
 23. Zwart RS, Thompson JP, Milgate AW, Bansal UK, Williamson PM, Raman H, et al. QTL mapping of multiple foliar disease and root-lesion nematode resistances in wheat. Mol Breed. 2010 Jun 1;26(1):107-24. <https://doi.org/10.1007/s11032-009-9381-9>
 24. Anderson JA, Stack RW, Liu S, Waldron BL, Fjeld AD, Coyne C, et al. DNA markers for *Fusarium* head blight resistance QTLs in two wheat populations. Theor Appl Genet. 2001 Jun 1;102(8):1164-8. <https://doi.org/10.1007/s001220000509>
 25. Kirigwi FM, Van Ginkel M, Brown-Guedira G, Gill BS, Paulsen GM, Fritz AK. Markers associated with a QTL for grain yield in wheat under drought. Mol Breed. 2007 Nov 1;20(4):401-13. <https://doi.org/10.1007/s11032-007-9100-3>
 26. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, et al. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell. 2009 Aug 1;21(8):2194-202. <https://doi.org/10.1105/tpc.109.068437>
 27. Zhu C, Gore M, Buckler ES, Yu J. Status and prospects of association mapping in plants. Plant Genom. 2008 Jul;1(1):5-20. <https://doi.org/10.3835/plantgenome2008.02.0089>
 28. Yu J, Buckler ES. Genetic association mapping and genome organization of maize. Curr Opin Biotechnol. 2006 Apr 1;17(2):155-60. <https://doi.org/10.1016/j.copbio.2006.02.003>
 29. Ersoz ES, Yu J, Buckler ES. Applications of linkage disequilibrium and association mapping in crop plants. In: Rajeev K. Varshney RK, Tuberosa R, editors. Genomics-Assisted Crop Improvement. Cham (Switzerland): Springer; 2007. Volume 1, p. 97-119. https://doi.org/10.1007/978-1-4020-6295-7_5
 30. Liu J, Xu Z, Fan X, Zhou Q, Cao J, Wang F, et al. A genome-wide association study of wheat spike related traits in China. Front Plant Sci. 2018 Oct 31;9:1584. <https://doi.org/10.3389/fpls.2018.01584>
 31. Breseghello F, Sorrells ME. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics. 2006 Feb 1;172(2):1165-77. <https://doi.org/10.1534/genetics.105.044586>
 32. Daba SD, Tyagi P, Brown-Guedira G, Mohammadi M. Genome-wide association studies to identify loci and candidate genes controlling kernel weight and length in a historical United States wheat population. Front Plant Sci. 2018 Aug 3;9:1045. <https://doi.org/10.3389/fpls.2018.01045>
 33. Gaire R, Huang M, Sneller C, Griffey C, Brown-Guedira G, Mohammadi M. Association analysis of baking and milling quality traits in an elite soft red

- winter wheat population. *Crop Sci.* 2019 May;59(3):1085-94. <https://doi.org/10.2135/cropsci2018.12.0751>
34. Beyer S, Daba S, Tyagi P, Bockelman H, Brown-Guedira G, Mohammadi M. Loci and candidate genes controlling root traits in wheat seedlings—a wheat root GWAS. *Funct Integr Genomics.* 2019 Jan 1;19(1):91-107. <https://doi.org/10.1007/s10142-018-0630-z>
35. Lozada DN, Mason RE, Sukumaran S, Dreisigacker S. Validation of grain yield QTLs from soft winter wheat using a CIMMYT spring wheat panel. *Crop Sci.* 2018 Sep;58(5):1964-71. <https://doi.org/10.2135/cropsci2018.04.0232>
36. Sukumaran S, Lopes M, Dreisigacker S, Reynolds M. Genetic analysis of multi-environmental spring wheat trials identifies genomic regions for locus-specific trade-offs for grain weight and grain number. *Theor Appl Genet.* 2018 Apr 1;131(4):985-98. <https://doi.org/10.1007/s00122-017-3037-7>
37. Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet.* 2006 Feb;38(2):203-8. <https://doi.org/10.1038/ng1702>
38. Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, et al. Mixed linear model approach adapted for genome-wide association studies. *Nat Genet.* 2010 Apr;42(4):355-60. <https://doi.org/10.1038/ng.546>
39. Kaler AS, Gillman JD, Beissinger T, Purcell LC. Comparing different statistical models and multiple testing corrections for association mapping in soybean and maize. *Front Plant Sci.* 2020 Feb 25;10:1794. <https://doi.org/10.3389/fpls.2019.01794>
40. Wen YJ, Zhang H, Ni YL, Huang B, Zhang J, Feng JY, et al. Methodological implementation of mixed linear models in multi-locus genome-wide association studies. *Briefings Bioinf.* 2018 Jul;19(4):700-12. <https://doi.org/10.1093/bib/bbw145>
41. Liu X, Huang M, Fan B, Buckler ES, Zhang Z. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 2016 Feb 1;12(2):e1005767. <https://doi.org/10.1371/journal.pgen.1005767>
42. Huang M, Cabrera A, Hoffstetter A, Griffey C, Van Sanford D, Costa J, et al. Genomic selection for wheat traits and trait stability. *Theor Appl Genet.* 2016 Sep 1;129(9):1697-710. <https://doi.org/10.1007/s00122-016-2733-z>
43. Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, et al. Characterization of polyploid wheat genomic diversity using a high-density 90000 single nucleotide polymorphism array. *Plant Biotechnol J.* 2014 Aug;12(6):787-96. <https://doi.org/10.1111/pbi.12183>
44. Poland JA, Brown PJ, Sorrells ME, Jannink JL. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One.* 2012 Feb 28;7(2):e32253. <https://doi.org/10.1371/journal.pone.0032253>
45. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics.* 2007 Oct 1;23(19):2633-5.

- <https://doi.org/10.1093/bioinformatics/btm308>
46. Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*. 2018 Aug 17;361(6403):eaar7191. <https://doi.org/10.1126/science.aar7191>
 47. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009 Jul 15;25(14):1754-60. <https://doi.org/10.1093/bioinformatics/btp324>
 48. Money D, Gardner K, Migicovsky Z, Schwaninger H, Zhong GY, Myles S. LinkImpute: fast and accurate genotype imputation for nonmodel organisms. *G3 (Bethesda)*. 2015 Nov 1;5(11):2383-90. <https://doi.org/10.1534/g3.115.021667>
 49. R Core Team. R: A language and environment for statistical computing. Auckland (New Zealand): R Core Team; 2019.
 50. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Software*. 2015;67(1). <https://doi.org/10.18637/jss.v067.i01>
 51. Rosseel Y. Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *J Stat Software*. 2012 Dec 19;48(2):1-36.
 52. Nyquist WE, Baker RJ. Estimation of heritability and prediction of selection response in plant populations. *Crit Rev Plant Sci*. 1991 Jan 1;10(3):235-322. <https://doi.org/10.1080/07352689109382313>
 53. Piepho HP, Möhring J. Computing heritability and selection response from unbalanced plant breeding trials. *Genetics*. 2007 Nov 1;177(3):1881-8. <https://doi.org/10.1534/genetics.107.074229>
 54. Hill WG, Weir BS. Variances and covariances of squared linkage disequilibria in finite populations. *Theor Popul Biol*. 1988 Feb 1;33(1):54-78. [https://doi.org/10.1016/0040-5809\(88\)90004-4](https://doi.org/10.1016/0040-5809(88)90004-4)
 55. Edae EA, Byrne PF, Haley SD, Lopes MS, Reynolds MP. Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. *Theor Appl Genet*. 2014 Apr 1;127(4):791-807. <https://doi.org/10.1007/s00122-013-2257-8>
 56. Vos PG, Paulo MJ, Voorrips RE, Visser RG, van Eck HJ, van Eeuwijk FA. Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. *Theor Appl Genet*. 2017 Jan 1;130(1):123-35. <https://doi.org/10.1007/s00122-016-2798-8>
 57. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, et al. GAPIT: genome association and prediction integrated tool. *Bioinformatics*. 2012 Sep 15;28(18):2397-9. <https://doi.org/10.1093/bioinformatics/bts444>
 58. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006 Aug;38(8):904-9. <https://doi.org/10.1038/ng1847>
 59. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995 Jan;57(1):289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
 60. Venables WN, Ripley BD. *Modern Applied Statistics with S*, 4th ed. New York (US): Springer; 2002. ISBN 0-387-95457-0.

- <http://www.stats.ox.ac.uk/pub/MASS4>
61. Kotal BD, Arpita D, Choudhury BK. Genetic variability and association of characters in wheat (*Triticum aestivum* L.). Asian J Crop Sci. 2010;2(3):155-60. <https://doi.org/10.3923/ajcs.2010.155.160>
 62. Philipp N, Weichert H, Bohra U, Weschke W, Schulthess AW, Weber H. Grain number and grain yield distribution along the spike remain stable despite breeding for high yield in winter wheat. PLoS One. 2018 Oct 10;13(10):e0205452. <https://doi.org/10.1371/journal.pone.0205452>
 63. Addison CK, Mason RE, Brown-Guedira G, Guedira M, Hao Y, Miller RG, et al. QTL and major genes influencing grain yield potential in soft red winter wheat adapted to the southern United States. Euphytica. 2016 Jun 1;209(3):665-77. <https://doi.org/10.1007/s10681-016-1650-1>
 64. Dewey DR, Lu K. A Correlation and Path-Coefficient Analysis of Components of Crested Wheatgrass Seed Production. Agron J. 1959 Sep;51(9):515-8. <https://doi.org/10.2134/agronj1959.00021962005100090002x>
 65. Arruda MP, Brown P, Brown-Guedira G, Krill AM, Thurber C, Merrill KR, et al. Genome-wide association mapping of *Fusarium* head blight resistance in wheat using genotyping-by-sequencing. Plant Genom. 2016 Mar;9(1):1-4. <https://doi.org/10.3835/plantgenome2015.04.0028>
 66. Huang M, Mheni N, Brown-Guedira G, McKendry A, Griffey C, Van Sanford D, et al. Genetic analysis of heading date in winter and spring wheat. Euphytica. 2018 Aug 1;214(8):128. <https://doi.org/10.1007/s10681-018-2199-y>
 67. Arguello MN, Mason RE, Roberts TL, Subramanian N, Acuna A, Addison CK, et al. Performance of soft red winter wheat subjected to field soil waterlogging: grain yield and yield components. Field Crops Res. 2016 Aug 1;194:57-64. <https://doi.org/10.1016/j.fcr.2016.04.040>
 68. Sharma RC, Tiwary AK, Ortiz-Ferrara G. Reduction in kernel weight as a potential indirect selection criterion for wheat grain yield under terminal heat stress. Plant Breed. 2008 Jun;127(3):241-8. <https://doi.org/10.1111/j.1439-0523.2007.01460.x>
 69. Godoy J, Gizaw S, Chao S, Blake N, Carter A, Cuthbert R, et al. Genome-wide Association Study of Agronomic Traits in a Spring-Planted North American Elite Hard Red Spring Wheat Panel. Crop Sci. 2018 Sep;58(5):1838-52. <https://doi.org/10.2135/cropsci2017.07.0423>
 70. Lanning SP, Kephart K, Carlson GR, Eckhoff JE, Stougaard RN, Wichman DM, et al. Climatic change and agronomic performance of hard red spring wheat from 1950 to 2007. Crop Sci. 2010 May;50(3):835-41. <https://doi.org/10.2135/cropsci2009.06.0314>
 71. Tessmann EW, Dong Y, Van Sanford DA. GWAS for *Fusarium* head blight traits in a soft red winter wheat mapping panel. Crop Sci. 2019 Sep;59(5):1823-37. <https://doi.org/10.2135/cropsci2018.08.0492>
 72. Bian Y, Yang Q, Balint-Kurti PJ, Wissner RJ, Holland JB. Limits on the reproducibility of marker associations with southern leaf blight resistance in the maize nested association mapping population. BMC Genomics. 2014 Dec 1;15(1):1068. <https://doi.org/10.1186/1471-2164-15-1068>
 73. Poland JA, Bradbury PJ, Buckler ES, Nelson RJ. Genome-wide nested

- association mapping of quantitative resistance to northern leaf blight in maize. *Proc Natl Acad Sci U S A*. 2011 Apr 26;108(17):6893-8. <https://doi.org/10.1073/pnas.1010894108>
74. Xu Y, Yang T, Zhou Y, Yin S, Li P, Liu J, et al. Genome-wide association mapping of starch pasting properties in maize using single-locus and multi-locus models. *Front Plant Sci*. 2018 Sep 5;9:1311. doi: 10.3389/fpls.2018.01311
 75. Vanous A, Gardner C, Blanco M, Martin-Schwarze A, Wang J, Li X, et al. Stability Analysis of Kernel Quality Traits in Exotic-Derived Doubled Haploid Maize Lines. *Plant Genom*. 2019 Mar;12(1):1-4. doi: 10.3835/plantgenome2017.12.0114
 76. Ohm HW, Anderson JM, Sharma HC, Ayala L, Thompson N, Uphaus JJ. Registration of Yellow Dwarf Viruses Resistant Wheat Germplasm Line P961341. *Crop Sci*. 2005 Mar 1;45(2):805-6. <https://doi.org/10.2135/cropsci2005.0805>
 77. Gutiérrez L, Germán S, Pereyra S, Hayes PM, Pérez CA, Capettini F, et al. Multi-environment multi-QTL association mapping identifies disease resistance QTL in barley germplasm from Latin America. *Theor Appl Genet*. 2015 Mar 1;128(3):501-16. <https://doi.org/10.1007/s00122-014-2448-y>
 78. Collard BC, Mackill DJ. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos Trans R Soc B Biol Sci*. 2008 Feb 12;363(1491):557-72. <https://doi.org/10.1098/rstb.2007.2170>
 79. Ribaut JM, Ragot M. Marker-assisted selection to improve drought adaptation in maize: the backcross approach, perspectives, limitations, and alternatives. *J Exp Bot*. 2007 Jan 1;58(2):351-60. <https://doi.org/10.1093/jxb/erl214>
 80. Guan P, Lu L, Jia L, Kabir MR, Zhang J, Lan T, et al. Global QTL analysis identifies genomic regions on chromosomes 4A and 4B harboring stable loci for yield-related traits across different environments in wheat (*Triticum aestivum* L.). *Front Plant Sci*. 2018 Apr 25;9:529. <https://doi.org/10.3389/fpls.2018.00529>
 81. Simmonds J, Scott P, Leverington-Waite M, Turner AS, Brinton J, Korzun V, et al. Identification and independent validation of a stable yield and thousand grain weight QTL on chromosome 6A of hexaploid wheat (*Triticum aestivum* L.). *BMC Plant Biol*. 2014 Dec 1;14(1):191. <https://doi.org/10.1186/s12870-014-0191-9>

How to cite this article:

Russell B, Brown-Guedira G, Sneller CH, Mohammadi M. Transferability of Marker Trait Associations in Wheat Is Disturbed Mainly by Genotype × Year Interaction. *Crop Breed Genet Genom*. 2020;2(3):e200013. <https://doi.org/10.20900/cbgg20200013>