

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Agronomy & Horticulture -- Faculty Publications

Agronomy and Horticulture Department

---

1-2-2021

## GWAS revealed effect of genotype × environment interactions for grain yield of Nebraska winter wheat

Shamseldeen Eltaher

P Stephen Baenziger

Vikas Belamkar

Hamdy A. Emara

Ahmed A. Nower

*See next page for additional authors*

Follow this and additional works at: <https://digitalcommons.unl.edu/agronomyfacpub>

 Part of the [Agricultural Science Commons](#), [Agriculture Commons](#), [Agronomy and Crop Sciences Commons](#), [Botany Commons](#), [Horticulture Commons](#), [Other Plant Sciences Commons](#), and the [Plant Biology Commons](#)

---

This Article is brought to you for free and open access by the Agronomy and Horticulture Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Agronomy & Horticulture -- Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

---

**Authors**

Shamseldeen Eltaher, P Stephen Baenziger, Vikas Belamkar, Hamdy A. Emara, Ahmed A. Nower, Khaled F.M. Salem, Ahmad M. Alqudah, and Ahmed Sallam

---

RESEARCH ARTICLE

Open Access



# GWAS revealed effect of genotype × environment interactions for grain yield of Nebraska winter wheat

Shamseldeen Eltaher<sup>1,2</sup>, P. Stephen Baenziger<sup>1</sup>, Vikas Belamkar<sup>1</sup>, Hamdy A. Emara<sup>2</sup>, Ahmed A. Nower<sup>2</sup>, Khaled F. M. Salem<sup>2,3</sup>, Ahmad M. Alqudah<sup>4</sup> and Ahmed Sallam<sup>5\*</sup> 

## Abstract

**Background:** Improving grain yield in cereals especially in wheat is a main objective for plant breeders. One of the main constraints for improving this trait is the G × E interaction (GEI) which affects the performance of wheat genotypes in different environments. Selecting high yielding genotypes that can be used for a target set of environments is needed. Phenotypic selection can be misleading due to the environmental conditions. Incorporating information from phenotypic and genomic analyses can be useful in selecting the higher yielding genotypes for a group of environments.

**Results:** A set of 270 F<sub>3,6</sub> wheat genotypes in the Nebraska winter wheat breeding program was tested for grain yield in nine environments. High genetic variation for grain yield was found among the genotypes. G × E interaction was also highly significant. The highest yielding genotype differed in each environment. The correlation for grain yield among the nine environments was low (0 to 0.43). Genome-wide association study revealed 70 marker traits association (MTAs) associated with increased grain yield. The analysis of linkage disequilibrium revealed 16 genomic regions with a highly significant linkage disequilibrium (LD). The candidate parents' genotypes for improving grain yield in a group of environments were selected based on three criteria; number of alleles associated with increased grain yield in each selected genotype, genetic distance among the selected genotypes, and number of different alleles between each two selected parents.

**Conclusion:** Although G × E interaction was present, the advances in DNA technology provided very useful tools and analyzes. Such features helped to genetically select the highest yielding genotypes that can be used to cross grain production in a group of environments.

**Keywords:** Bread wheat (*Triticum aestivum* L.), Yield, LD, Association mapping, Gene annotation, Breeding programs

\* Correspondence: [amsallam@aun.edu.eg](mailto:amsallam@aun.edu.eg)

<sup>5</sup>Department of Genetics, Faculty of Agriculture, Assiut University, Assiut 71526, Egypt

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Bread wheat (*Triticum aestivum* L.) is the third most important food crop in the world after maize (*Zea mays* L.) and rice (*Oryza sativa* L.). To meet the increasing food demand of a growing population, the breeders have focused on the development of cultivars having higher yield and yield stability, and increased resistance/tolerance to biotic and abiotic stresses. Grain yield (GY) is controlled by numerous genes that interact with each other and with the environment [2, 50]. Grain yield is a complex trait that is determined by multiple yield component traits, and each component trait is a quantitative trait controlled or affected by multiple loci [2, 72]. Thus, there needs a detailed genetic dissection of the grain yield trait and its component traits to manipulate the alleles at the relevant loci to the greatest advantage.

F<sub>3.6</sub> Nebraska winter wheat genotypes are tested in nine environments (8 environments in Nebraska and one in Kansas). Multi-environment yield trials (MEYTs) are used in the last selection cycles to identify superior genotypes in plant breeding programs and to determine where the cultivars are best adapted. This task is difficult due to the frequent presence of GEI.

The GEI reduces the association between phenotype and genotype by reducing heritability, and eventually genetic progress in plant breeding programs. Means across environments are adequate indicators of genotypic performance only in the absence of GEI. If it is present, the use of means across environments ignores the fact that genotypes differ in relative performance in different environments [33]. Analysis of Variance (ANOVA) analysis is not sufficient to provide an understanding of the genotypes or environments that give rise to the interaction [33, 53]. The purpose of MEYTs is not only to classify superior genotypes for the target area, but also to determine if the target area can be divided into mega environments (MEs). Investigation of ME is a requirement for meaningful cultivar evaluation and recommendation [66]. The international Maize and Wheat Improvement Center (CIMMYT) introduced the definition of ME, defined as a broad, not essentially attached area, occurring in more than one country and frequently transcontinental, defined by similar biotic and abiotic stresses, cropping system supplies, customer favorites, and, for convenience, by volume of production [13].

Traditional wheat breeding is mostly built on phenotypic selection which is one of the most important steps for genetic improvement. Every wheat breeder chooses to have an environment at the selection nursery site that will increase the beneficial and minimize the negative aspects of natural selection. In winter wheat breeding, for example, it is common for the selection nursery to be an environment which causes the death of winter tender lines. However, it is blindness, analytical, inefficiency

and costs a long time [8]. Fortunately, the hardworking and intelligence of the breeders would have kept signatures in the wheat genome during crop improvement, and this selection signal could be detected using different methods. GWAS should be performed to annotate the signatures in detail, taking into account the selection signal could not be correlated with phenotype [37]. Wheat breeders use single nucleotide polymorphism (SNP) high-density maps to identify genomic regions associated with quantitative traits in biparental mapping experiments or in genome-wide association studies (GWAS) [6, 10, 21, 45, 59, 63, 70]. There are many reports that dissect the effect of genotype, environment, and GEI using linkage mapping effects ([39, 42]; L. [41, 65]). But, the reports on the dissection of GEI using genome-wide association mapping methods are rare [65]. Therefore, the objectives of the present study were to i) study the genetic variation and GEI using the genotype main effects and genotype × environment interaction effects (GGE-biplot analysis) for grain yield for 270 F<sub>3.6</sub> Nebraska winter wheat genotypes grown in different environments, ii) identify the highest yielding genotypes at the different environments iii) identify marker trait association (MTAs) related to grain yield trait and to dissect GEI using GWAS.

## Methods

### Plant materials

A set of 270 F<sub>3.6</sub> wheat genotypes (Nebraska Duplicate Nursery, hereafter referred to as DUP2017) which is the preliminary yield trial was derived from 800 to 1000 crosses among Nebraska's adapted cultivars or experimental genotypes [24]. The parents of these crosses are mainly from wheat breeding programs in the Great Plains, and a few crosses to globally important wheat lines. The breeding lines used in this study were derived from 85 crosses of 800–1000 that were initially made. The pedigree of all 270 genotypes was presented in (Supplementary Table S1).

In wheat growing season 2016/2017, DUP2017 was grown in nine environments [Mead (latitude 41.2286° N, and longitude 96.4892° W), Lincoln, (latitude 40.8136° N, and longitude 96.7026° W) Clay Center (latitude 40.5217° N, longitude 98.0553° W), North Platte (latitude 41.1403° N, and longitude 100.7601° W), Grant, (latitude 40.8430° N, and longitude 101.7252° W) McCook, (latitude 40.1967° N, and longitude 100.6249° W), Sidney, (latitude 41.1448° N, and longitude 102.9774° W) and Alliance (latitude 42.0930° N, and longitude 102.8702° W) in Nebraska, and one location in Kansas (latitude 39.1836° N, and longitude 96.5717° W)]. The experimental layout was incomplete augmented block design with one replication in each location. The incomplete blocks consisted of 27 experimental genotypes and three check

cultivars (Goodstreak, Camelot, and Freeman) and there were 10 incomplete blocks per trial. The check cultivars (Goodstreak, Camelot and Freeman) are adapted to diverse ecogeographic regions of Nebraska and by parentage and morphology quite diverse. At all locations, the plots ( $N = 300$ ) were planted at a seeding rate of 54 kg/ha and the plot consisted of five rows of 3 m length with 0.23 m between rows.

### Phenotyping

Grain yield was measured using a combine harvester which harvested all five rows of each plot. At Lincoln and Mead, the harvested grain was stored until dried to room humidity before weighing. At the other locations, the grain was weighed on the combine.

### Genotyping-by-sequencing and SNPs calling

DNA was extracted from the wheat leaves of 2–3 young two-week-old seedlings using BioSprint 96 DNA Plant Kits (Qiagen Valencia, California, USA) following the manufacturer's instructions. The genotyping-by-sequencing (GBS) was performed as described by Poland et al. [48]. The SNPs were called using Tassel v5.2.40 GBS analysis pipeline with default parameters [12]. The GBS-tags were aligned to the reference genome using Burrows-Wheeler Aligner [36]. The reference genome v1.0 of the 'Chinese Spring' genome assembly from the International Wheat Genome Sequencing Consortium (IWGSC) was used in SNP calling. The raw sequence data of the 270 genotypes of the current study along with 6791 other genotypes previously genotyped in our program were combined for SNP calling in order to increase the coverage of the genome and read depth at SNP sites [9, 71]. A set of 200,064 SNPs were resulted from SNP calling. SNPs were removed from the dataset if they were either monomorphic, showed more than 20% missing values, had conflicting calls from SNP or exhibited minor allele frequencies (MAF) of less than 5% [30, 71]. Interestingly, none of our lines (270  $F_{3,6}$ ) have missing information's of more than 20%. The GBS data is available in (Supplementary Table S2).

### Statistical analysis

For the field experiments, grain yield was analyzed using methods used for augmented design with replicated check cultivars (augmented incomplete block design). The augmented design is especially useful for statistically controlling spatial variability in large trials (with minimal or no replicates) to assess genotypic effects where seed is often limiting. In the early stages of a breeding program, a plant breeder is faced with evaluating the performance of large numbers of genotypes with limited seed. A general technique for unreplicated designs is the one known as "systematically spaced checks." In this technique, a

standard check(s) genotype is systematically spaced in the trial [25].

The incomplete block consisted of 27 experimental lines and the three check cultivars which were planted in ten incomplete blocks for a total of 300 plots. The liner mixed model was done using this model.

$$Y = \text{Check} + \text{Environment} + \text{Iblock (Environment)} + \text{Genotype} + \text{GXE} + \text{Error}$$

In this model all terms except checks were fit as random effect, and the check was fit as fixed effect. The residual maximum likelihood (REML) implemented in ASREML-R version 4.1 [15], was used to estimate the variance components and the associated standard errors. The likelihood ratio test using "lrt" function in ASREML-R was used to test significance for each term [16]. The variance component was used also to estimate broad sense heritability using the following formula:

$$H^2 = \text{Var}(G)/(\text{Var}(G) + \text{Var}(GXE)/E + \text{Var}(E)/\text{ExR})$$

Pearson's correlations among all pairs of environments of GY was calculated based on genotype performance of each experimental genotype for each environment using R software package "corrplot". The GGE Biplot which describes the relationship between different environments was performed using GEA-R (Genotype x Environment Analysis with R for Windows) Version 4.1 [47]. The population structure (Q matrix) for the  $F_{3,6}$  Nebraska winter wheat was performed using the criteria described in [24]. The analysis was done by STRUCTURE 3.4.0 [49] and the kinship matrix (K) was estimated using TASSEL v5.2.40 [12].

### Climate data analysis

The monthly average temperature, average rainfall and average snowfall were collected from (<https://www.usclimatedata.com/climate/united-states/us>). Principal component analysis was done for all climate factors using ClustVis online tool. This web server is freely available at <http://biit.cs.ut.ee/clustvis/> [43]. The scatter plot was visualized using excel 2016.

### Genome-wide association studies

The GWAS analysis was conducted separately for GY at each environment using 11,991 SNPs markers after filtration to remove SNPs with minor allele frequencies (MAF < 0.05) and exclude all the heterozygous SNPs which were calculated as missing values. The GY phenotypic values, Kinship matrix, Q matrix and SNPs were subjected to association analysis using a mixed linear model (MLM) in TASSEL v5.2.40 software.

The  $-\log^{10}$   $P$ -values of the MLM were later adjusted by calculating the corresponding Bonferroni correction (BC) at a significance level of 5%. Phenotypic effects at the marker loci were calculated as differences between the means of the marker classes. The phenotypic variance explained ( $R^2$ ) by significant makers was determined using TASSEL v5.2.40.

Manhattan plots for grain yield trait were visualized using the Shiny AIM application [31]. Linkage disequilibrium ( $r^2$ ) was estimated using TASSEL 5.0 between each pair of SNPs located on the same chromosome. The LD heatmap was visualized using 'LDheatmap' R package [57].

### Candidate genes linked with grain yield

The physical position of high LD genomic regions that include the significant SNPs were used to identify the high-confidence (HC) putative candidate gene models using annotations version provided by the IWGSC. We used the recently published wheat genome sequence. WheatMine web-based platform, was used to identify the gene annotations and gene ontologies for the potential candidate genes based on IWGSC v1.0 and v1.1 (<https://urgi.versailles.inra.fr/WheatMine/begin.do>).

## Results and discussion

### Effects of environment (E), genotype (G), and G × E interaction (GEI)

The ANOVA for grain yield (Table 1) identified highly significant differences among genotypes at  $P < 0.0001$ . That high genetic variation existed among genotypes is very useful for wheat breeders to efficiently select the highest yielding genotypes, in each location or across locations, to be used in breeding programs. Genotype (G) × environment (E) interactions (GEI) were significant at  $P < 0.0001$  for grain yield. Significant GEI indicated that the genotypes performed differently in different environments and that genotypes should be selected for adaptation to specific environments [3, 4, 65]. Hence, the GEI confirmed that genotypes responded differently to the variation in environmental conditions at locations, which indicated the need to test wheat cultivars at multiple locations.

The phenotypic correlation for grain yield among the nine environments is presented in (Fig. 1). No significant or very low significant correlations at  $P < 0.05$  were observed among the cultivar yield values in all environments. A moderately positive significant correlation between Lincoln and Mead ( $r = 0.42^*$ ) and between Grant and McCook ( $r = 0.43^*$ ) was expected as these pairs of locations are in similar ecogeographic zones. The results of weak or low correlations further support the diversity of the testing environments and the significant effect of GEI on the genotypes' performances.

### The performance of the genotypes in different environments

It is common for MEYTs data to represent a combination of crossover and non-crossover types of GEI. The minimum, maximum, mean of grain yield in each environment is presented in Table 2. The maximum grain yield ranged from 3503.53 (Kansas) to 8287.50 Kg/Ha (McCook). The lowest and highest average of grain yield were also accounted to the same two environments. This huge difference in grain yield for the same set of genotypes was due to the strong effect of environment and GEI.

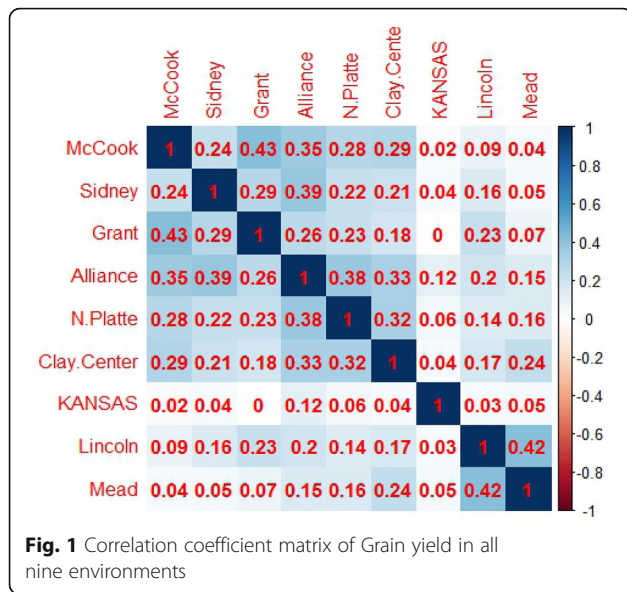
The highest yielding genotypes differed by location; NE17660 (Alliance), NE17626 (Clay Center), NE17528 (Grant), NE17588 (Kansas), NE17609 (Lincoln), NE17441 (McCook), NE17662 (Mead), NE17463 (North Platte), and NHH17447 (Sidney) (Supplementary Table S3). GEI can be caused by crossover interactions or by non-crossover interactions (e.g. changes in the magnitude of the differences among lines). As there was no common genotype that ranked as the highest yielding genotype in more than one environment, we chose the 50 highest yielding genotypes (~ 18.5% of the experimental genotypes) in each location to represent the high yielding genotypes at that environment. Then, a genotype was selected if it was among the 50 high yielding genotypes in at least two environments. As a result, 13 genotypes were marked and selected (Table 3). The same procedure was applied in selecting the high drought tolerant wheat genotypes, [52]. Those genotypes

**Table 1** Variance component and associated standard error estimated using a general linear mixed model by residual maximum likelihood (REML) for grain yield measured across eight locations in Nebraska and one location in Kansas in 2017

	Variance Estimate	Standard Error	Significance
Environment (E)	249.19	124.76	$P < 0.0001$
Iblock (Environment)	8.68	1.71	$P < 0.0001$
Genotype (G)	12.80	1.77	$P < 0.0001$
Genotype × Environment (GEI)	27.25	3.74	$P < 0.0001$
Error	38.75	3.34	$P < 0.0001$
Heritability In broad sense ( $H^2$ )	0.64		

Significance testing performed using likelihood ratio test





which were in the high yield group in multiple locations were considered as having the non-crossover GEI (Table 3). The genotype NE17625 was found among the highest 50 yielding genotypes in all the environments except Kansas. The genotype NE17626 was found among the highest 50 yielding genotypes at all the environments except Mead. Moreover, the genotype NE17443 was found to be among the highest 50 yielding in seven environments. Two genotypes NE17629 and NE17549 were found among highest 50 yielding in six environments. Remarkably the selected genotypes are heterogeneous in terms of their pedigree. For example, some of the selected genotypes shared the same parent such as NE17625, NE17626, NE17629 and NE17549 which were all reselections from NW03666 (Supplementary Table S1). Both NE17479 and NE17435 were half-sibs and shared the same parents (NE06545/NW07534). The other seven selected genotypes had different pedigrees. Pedigree information provides useful information for plant breeders to maintain diversity while making the next set of crosses using the selected genotypes.

As mentioned previously, the significant GEI often is interpreted as a specific breeding program for improving grain yield may be required optimal improvement for each environment [44, 52]. However, crossing these selected genotypes may be useful for a breeding program, with a full consideration to the pedigree information, that extends across more than one environment, especially when the environment at a location will change from year to year (e.g. may not be predictable).

Alliance and Grant had the highest number of selected top 50 genotypes in common with 11 in each. Kansas and Sidney, on the other hand, had the fewest common selected genotypes (six genotypes). This result may be

**Table 2** The Maximum, minimum and mean of grain yield trait (Kg/Ha) measured across eight locations in Nebraska and one location in Kansas in 2017

Environments	Max	Min	Mean
Alliance	4737.57	2683.52	3688.10
Clay Center	5472.09	3169.01	4377.17
Grant-D	3796.24	2509.37	3213.25
Kansas	3503.53	1663.30	2543.39
Lincoln	5676.86	2710.02	4289.65
McCook	8287.50	4952.45	6044.39
Mead	4135.17	1628.13	3046.68
North Platte	3915.26	2644.65	3281.68
Sidney	3850.33	3342.43	3608.08

due to the Kansas trial was considerably further south and in a different state while the other eight environments are in Nebraska and where the breeding program targets its new cultivars.

Analytical approaches to GEI analysis are important for enhancing the value of MEYTs and gaining an understanding of causes of GE interactions [61, 65, 68]. The methods used to understand GEI include the characterization of trial sites according to environmental factors, using either direct measurements, calculated indices, or variables derived from crop growth models [18]. The highly significant GEI were explained by the differences in the precipitation, snow cover, and temperature from one location to another location during the growing session (Supplementary Table S4). Although eight environments are geographically within Nebraska, the climate data differed by the environment.

**GGE bi-plot analysis**

The GGE-biplot approach, which was based on environment focused scaling, was used to estimate the relationships between the environments (Fig. 2). The lines that join the biplot origin and the markers of the environments are called environment vectors. The angle between the vectors of 2 environments is related to the correlation coefficient between them. The angles among most of our environments were only a little smaller than 90°; therefore, the correlation between them should be close to 0 (See Fig. 1). This GGE biplot approach (Fig. 2) suggested that Alliance and North Platte were the most closely correlated environments with Grant and McCook closely behind. However, the largest correlation coefficients were between McCook and Grant and between Lincoln and Mead (Fig. 1). Some contradictions between the figures and actual correlations were predictable because the biplot did not estimate 100% of the GGE variation [33, 67].

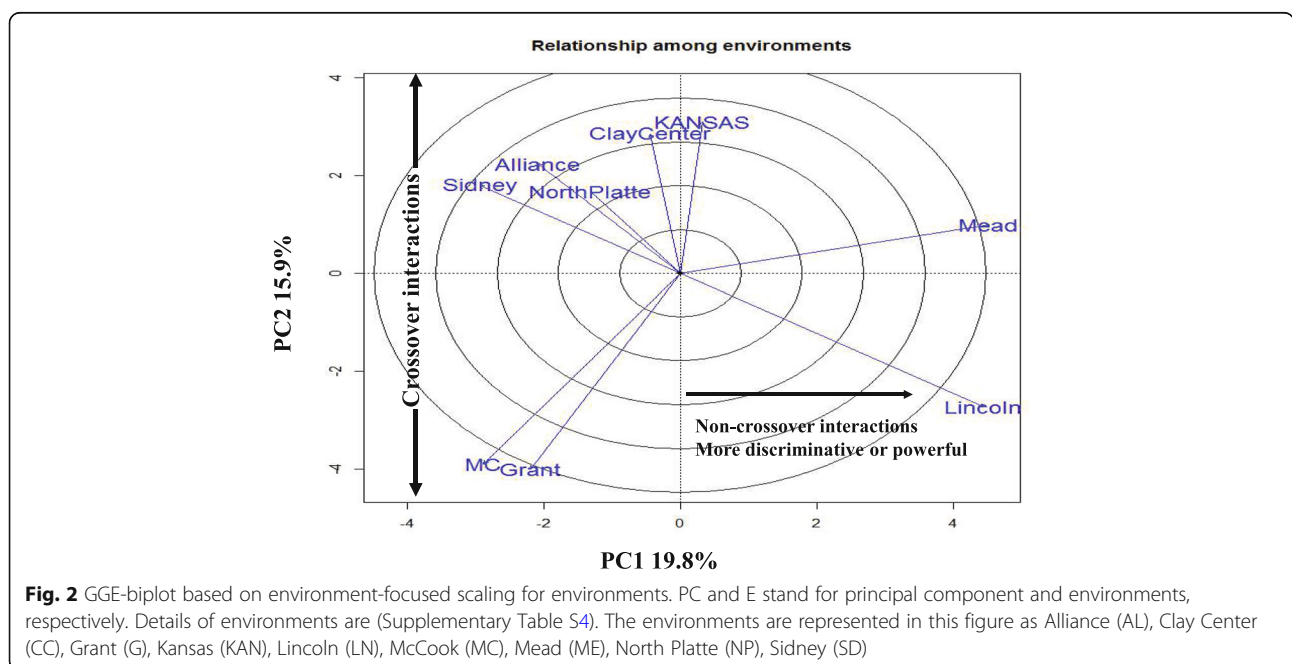
**Table 3** The best high yielding genotypes across all environments

Genotypes/Env.	Alliance	Clay Center	Grant	Kansas	Lincoln	McCook	Mead	North Platte	Sidney	Total
NE17625	*	*	*		*	*	*	*	*	8
NE17626	*	*	*	*	*	*		*	*	8
NE17443	*	*	*	*	*			*	*	7
NE17629		*	*	*		*	*	*		6
NE17549	*	*	*			*	*	*		6
NE17435	*				*	*	*			5
NE17479	*		*	*	*			*		5
NE17524		*	*		*	*	*			5
NE17533	*			*		*	*	*		5
NE17545	*		*	*		*			*	5
NE17661	*		*			*		*	*	5
NE17550	*	*	*		*				*	5
NE17624	*		*		*	*	*			5
Total	11	8	11	6	8	10	7	8	6	

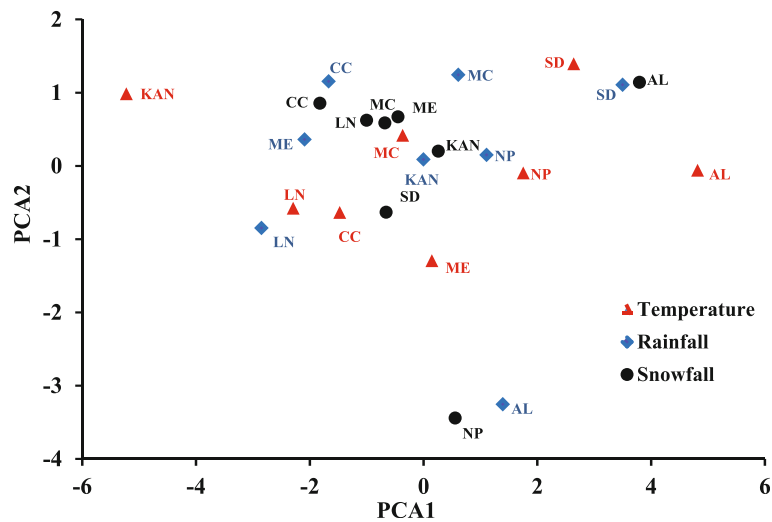
Most of our environments in this study were considered as PC2 environments except Kansas, Lincoln and Mead were included in PC1, which had positive and negative scores. PC1 represents proportional genotype yield differences across environments, which leads to a non-crossover GEI. Genotypes with superior PC1 scores can be easily identified in environments with larger PC1 scores. In contrast to environmental PC1, PC2 had both positive and negative scores (Fig. 2). Positive and negative scores are due to crossover GEI, leading to inconsistent genotype yield differences across environments [66]. A genotype may have large positive interactions

with some environments; but have large negative interactions with other environments.

In order to create a detailed climate factor (Supplementary Table S4 and Fig. 3), PCA evaluated the standardized values of the growing season mean temperature, average rainfall and average snowfall. Looking at (Fig. 3) we find that all the three climatic factors (average temperature, average rainfall, and average snowfall) were widely distributed throughout the PCA1 and PCA2. But there were several closely observed snowfall points among Lincoln, Mead, McCook and Kansas. The average temperature of Alliance, Sidney and Kansas were widely distributed across







**Fig. 3** Principal component analysis for temperature, rainfall, and snowfall in the nine environments. The environments are represented in this figure as Alliance (AL), Clay Center (CC), Kansas (KAN), Lincoln (LN), McCook (MC), Mead (ME), North Platte (NP), Sidney (SD). Weather data in Grant location is not available

the PCA. All these informative data indicated that different climate factors caused strong GE interactions.

#### Genome-wide association study for grain yield

The GWAS analysis was performed using MLM model which takes population structure into consideration [6, 69]. Due to the highly significant interaction among the genotypes and environments, the GWAS was performed for each environment, separately. The GWAS found a total of 70 MTAs associated with GY in the nine environments (Table 4 and Figure 4; Supplementary Table S5). The lowest number of significant SNPs (three SNPs) for grain yield were detected in the Grant environment, while the highest number of significant SNPs (11) was observed in three environments: Lincoln, McCook, and Sidney. The phenotypic variation ( $R^2$ ) ranged from 7.36% to 12.91%. All QTLs detected using GWAS can be considered as having minor effects on increasing grain yield. Grain yield is a complex trait controlled by many genes and affected by environment, and thus the identification of large number of associations is expected. At the genomic level, the highest number of significantly associated SNPs was observed in the D genome (30 SNPs) followed by A genome (21 SNPs) then B genome (19 SNPs) (Fig. 5). At the chromosomal level, the 71 significant SNPs associated with increased GY were distributed on all wheat chromosomes except 1A, 4B, 4D, 6A and 6B. The highest number of significant SNPs were located on the same chromosome (2D) and associated with high grain yield across environments (13 SNPs). The 13 SNPs were found in 3 environments (2 SNPs in Grant, 6 SNPs in Mead and 5 SNPs in Sidney). These valuable results reflected the importance of D genome in the GY traits. A broad comparison of marker-trait association results from the current study with

two previous studies were made using a chromosome basis because of differences in marker type and marker positions on different genetic maps. Edae [21] detected a stable QTL for grain yield on chromosome 2DS both under irrigated and rain-fed conditions using DArT markers. Also, the DArT marker wpt6531 on the short arm of chromosome 2DS, which was associated with yield is about 8 cM away from the wpt4144 marker, which was associated with grain yield in a previous study by Burguen et al. [14]. Previous studies have emphasized the importance of the D genome for grain yield using different types of markers [14, 20, 21, 23, 34].

No common markers were found among environments due to the lack of or very low significant correlations among environments for grain yield. Marker-assisted selection (MAS) can be useful for specific environments. The MTAs found in this study should be validated in additional environments and germplasm before using them in MAS. Previous studies identified SNP markers associated with GY on various chromosomes (1D, 1B, 2A, 3B, 4A, 5A, 5B, 5D, 7A, and 7B) [2, 10, 21, 32, 38, 63]. Chromosomes 3B, 5A, 5B and 7A were identified as having important yield QTL using 567 loci including RFLP, SSR, and AFLP markers [58]. El-basyoni [22] identified QTLs associated with GY in different environments on chromosomes 1B, 2A, 3A, 4A, 5A, 5B, 6A, 6D and 7B using DArT markers with a previous duplicate nursery lines of Nebraska winter wheat. Moreover, significant markers associated with GY were found on chromosomes 1B, 2A, 3A, 4A, 5A, 5B, 6A, 6D, and 7B in European winter wheat [11, 17, 19, 26, 54, 55, 58, 73]. Neumann [46], detected significant markers for GY in winter wheat on the

**Table 4** Summary of GWAS analysis for grain yield including number of SNPs, range of  $P$  value, range of  $R^2$ , and range of allele effects in the nine environments

Environments	Number of SNPs	Log <sup>10</sup> P- Value		Range of R <sup>2</sup>		Range of Allele effect	
		Min	Max	Min R <sup>2</sup>	Max R <sup>2</sup>	Min allele effect	Max allele effect
Alliance	5	9.43E-08	5.44E-06	11.50	12.74	1.65 (A and C)	4.29 (T)
Clay Center	5	9.93E-07	1.43E-06	10.6	11.68	2.60 (G and T)	4.48 (G)
Grant	3	1.31E-06	6.19E-06	8.66	9.93	1.87 (C)	2.26 (C)
Kansas	7	1.11E-07	3.07E-06	10.55	12.91	1.65 (A)	2.53(G)
Lincoln	11	1.03E-09	7.38E-06	7.36	10.44	3.04 (A)	10.58 (A)
McCook	11	1.03E-08	7.39E-06	8.72	11.58	2.36 (C)	5.19(A and C)
Mead	8	3.23E-07	8.60E-06	9.53	12.28	3.24 (T)	4.25(A and C)
North Platte	9	1.32E-07	6.05E-06	9.06	11.83	2.06 (A)	3.59 (A)
Sidney	11	1.69E-07	5.60E-06	8.89	12.73	0.35 (T)	1.00 (T)

chromosomes, 3A, 3B, 7A, 5B and 7B. Markers responsible for GY were identified on chromosomes, 4B and 7D which were reported in bi-parental QTL analyses [2, 11, 17]. Significant markers associated with GY on chromosome 2A were reported in previous studies [1, 27–29, 38, 60, 62]. A new publication of Kan et al. [32] who revealed a major quantitative trait locus (QTL) QYld.osu-1BS for grain yield in 260 F<sub>2:4</sub> winter wheat population of doubled haploid (DH) lines derived from the cross of Duster and Billings and they validated the QTLs using kompetitive allele specific PCR (KASP) markers for the unique sequences for QYld.osu-1BS allele.

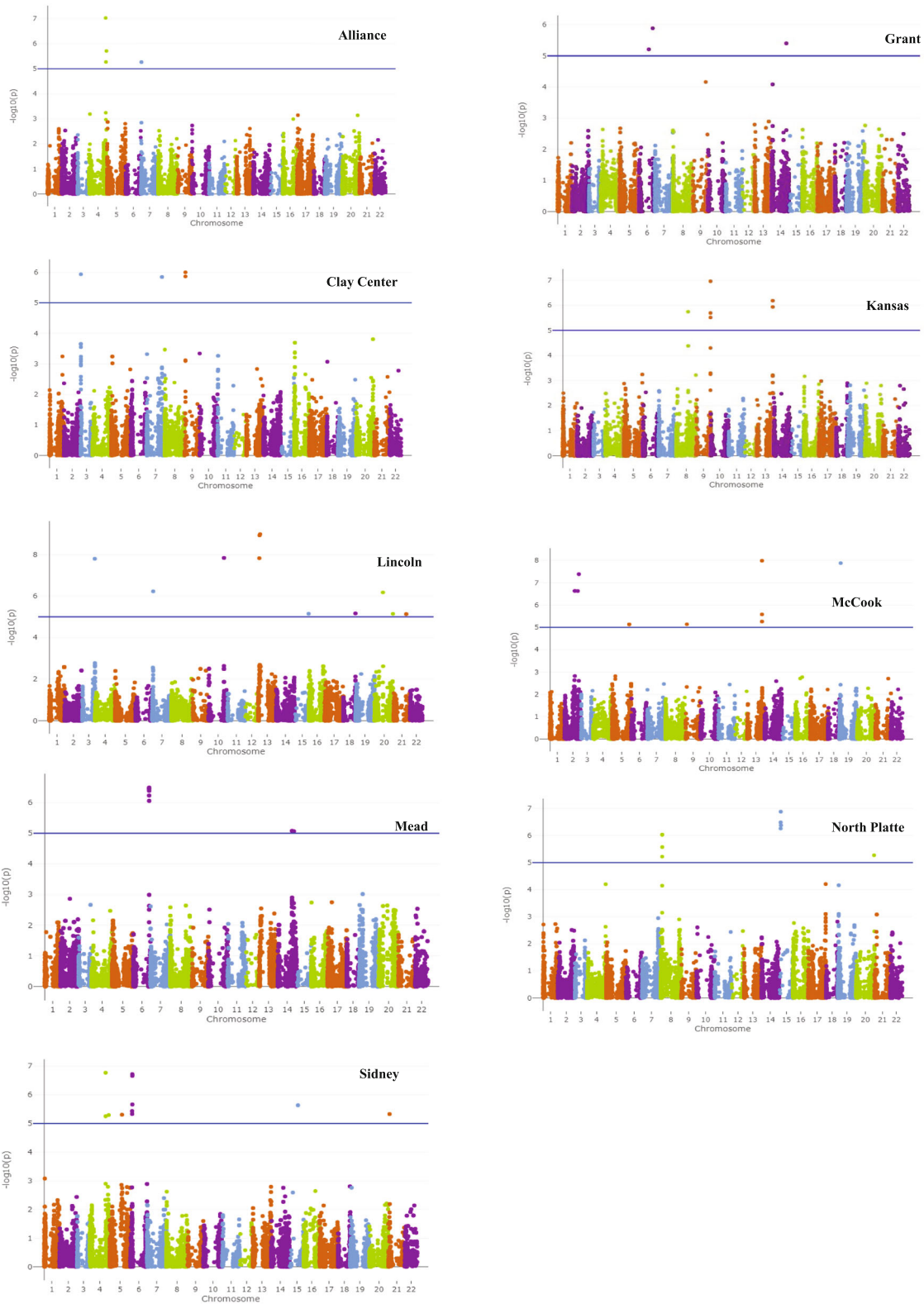
#### Linkage disequilibrium (LD) and gene annotation

The significance of linkage equilibrium ( $r^2$ ) was estimated between each pair of SNPs located on the same chromosome in each environment (Supplementary Table 5). All SNP pairs that had a high LD were considered a genomic region (GR). As a result, a set of 16 genomic regions associated with increased grain yield were identified across the nine environments. All SNPs located on the same chromosome had a high significant LD in Clay Center (two GRs), Kansas (three GRs), and North Platte (two GRs). In Alliance, there was no significant LD among SNPs located on 2A chr., while the SNPs located on 3A chr. were in significant LD. The five SNPs that are located on 2D chromosome in Mead were in a high significant LD. Among the SNPs located on the same chromosome, there were some cases which some SNP pairs had a significant LD, while the other did not have. For example, in McCook, there was a non-significant LD among SNPs on 1B chr., however, only S1B\_427530781 and S1B\_427530781 were in a complete LD. Likewise, in Sidney, the five SNPs located on 2D chromosomes had two groups in which SNPs were in high significant LD. The first group consisted of two SNPs (S2D\_67556531 and S2D\_69503850), while the

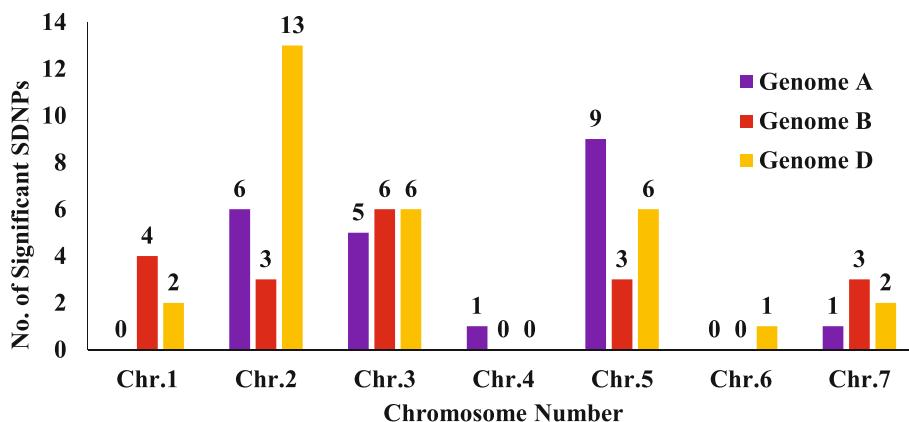
second one consisted of three SNPs (S2D\_76749441, S2D\_77122291, and S2D\_77122292). There was no significant LD between the two groups. Such information is very important to know which SNPs, located on the same chromosome, could be inherited together or individually. The significant SNPs located on the same chromosome, if the LD value among a group of target SNPs is high, then these SNPs could represent the same QTL and inherited together. If the LD is low, on the other hand, then the two significant SNPs represent two different QTLs [6, 51]. The gene annotation was identified for all genomic regions with significant LD. The candidate genes within the 16 GRs are listed in (Supplementary Table S6).

As aforementioned, 2D chromosome had the highest number of SNPs associated with high grain yield across environment. Therefore, gene annotation was described in detail on this chromosome. By studying LD pattern among these SNPs, we can identify how many genes within the physical position of these SNPs could represent. In Mead, the five SNPs represent one QTL, while, in Sidney, S2D\_67556531 and S2D\_69503850 co-inherited together and the other three SNPs on the same chromosome co-inherited together. When estimating the LD between each two pairs of the 13 SNPs, the results indicated no significant LD between any pair of SNPs from different environments (Fig. 6). This is also a further indication of the strong  $G \times E$  interaction which affects the expression of genes in response to the environment.

By looking to the gene annotation in the LD genomic regions on 2D chromosome (Supplementary Table S6). We found that *TraesCS2D01G506200* gene (genome region 8) is annotated as Zinc transporter and metal ion transport was associated with grain yield in Mead environment. There was many reports demonstrated the crucial role of Zn in improving grain yield in wheat [35, 40]. Recently, Alqudah et al. [5] found that deficiency of metal



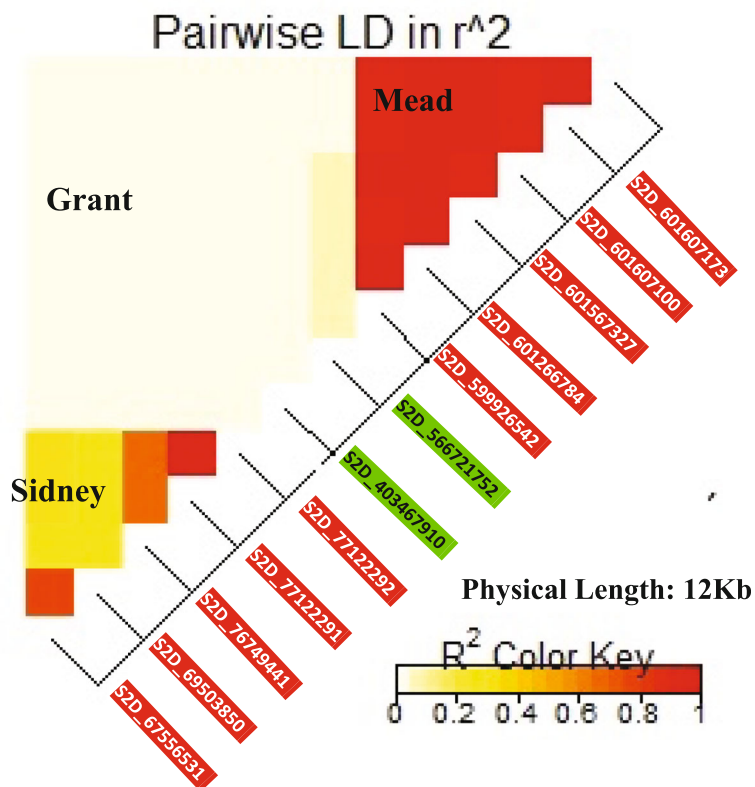
**Fig. 4** Manhattan plot displaying SNP markers-trait association identified for GY traits in GWAS using  $F_{36}$  Nebraska winter wheat. The blue line is significant threshold of 5% bonferroni correction (BC 5%)



**Fig. 5** The distribution of significant associated SNPs with GY across all environments

ion transport specially Cu and Zn may result in reducing grain yield through increasing spikelet and floret sterility/abortion. Interestingly, in Sidney environment, *TraesCS2D01G118400* and *TraesCS2D01G119200* (genomic region 15) genes are also annotated as potassium transporter and calcium exchanger respectively, demonstrating the role of the metals in grain yield. Potassium and calcium deficiency can significantly reduce crop yield but due to the complex relationship in absorption or

transport among them, their mechanisms in grain improvement is still not well-understood [56]. Other interesting candidate genes in Mead environment are involved in phytohormone e.g. auxin (*TraesCS2D01G506900*) and Jasmonate (*TraesCS2D01G507200*). In wheat, phytohormones control the spikelet's and grain development and grain filling. Slow grain development and filling rates are highly linked with low contents of the cytokinin and auxin [64] that in turn reduce grain yield. Understanding the



**Fig. 6** Heatmap of LD among SNPs located on 2D chr. Across the three environments. The red line indicates that high LD among all SNPs in Mead

mechanisms of metal transport and phytohormones which are found to be highly associated with grain yield in the current study are needed to improving wheat grain yield.

#### The promising high yielding genotypes for future breeding program

For the 13 selected genotypes (Table 3), three criteria may be considered to determine the genotypes as candidate parents for a future cross to improve grain yield. These criteria are based on:

- (1) the presence of favorable alleles associated with increased grain yield was recorded in each genotype (Supplementary Table S7). NE17435 had the highest number of favorable alleles associated with increased grain yield at 46 sites, while, NE17550 had 39 favorable alleles. Four genotypes; NE17545, NE17626, NE17524, and NE17629 had the same number of 43 alleles associated with increased grain yield. Although NE17435 had the highest number of target alleles for grain yield, but it was among the 50 highest yielding genotypes in only six environments. NE17625 and NE17626 which were among the 50 highest in eight environments had 40 and 43 target alleles for grain yield, respectively. It was useful to count the number of favorable alleles that each selected genotype carries as it helps in determining the target genotypes as parents for future crosses in the breeding program to improve grain yield in single or specific environments.
- (2) the genetic diversity among the 13 genotypes. The GD among all genotypes in this study was extensively described in [24]. This population was divided into three possible subpopulations [24]. The genetic distance among the 13 selected genotypes is presented in (Supplementary Table S8). Five genotypes were found to be assigned to subpopulation 1 (G1), five were assigned to subpopulation 3 (G3), and the remaining three genotypes were assigned to subpopulation (G2). The highest genetic distance was found between NE17624 and NE17661 (GD = 0.333), while the lowest GD was between NE17549 and NE17625 (GD = 0.007). Generally, all selected genotypes had a low level of genetic distance and they tend to be genetically similar. This is due to the fact that all genotypes represent Nebraska winter wheat breeding program [24] Moreover, many accessions were reselections from the same line (Supplementary Table S1).
- (3) the number of unique alleles associated with increased grain yield between each pair of the 13 selected genotypes (Supplementary Table S9). The highest number of different alleles (33) was found

between NE17549 and both NE17479 and NE17435. Only three different alleles, on the other hand, were found between NE17549 and NE17625. The different alleles between each two genotypes depends on the genetic distance among them. There was a positive significant correlation between different alleles and genetic distance ( $r = 0.65^{**}$ ).

Each criterion identified different candidate parents. By considering all criteria together with the priority for the number of different alleles and genetic distance, a cross between NE17625 and NE17479 may be useful for developing a cultivar that may be high yielding in five environments (Alliance, Grant, Kansas, Lincoln, and North Platte). Both parents had 40 alleles associated with increased grain yield, 32 different alleles associated with increased grain yield, and genetic distance of 0.292. Moreover, NE17625 and NE17479 were among the highest 50 genotypes in eight and five environments, respectively. Although NE17625 and NE17626 were among the highest 50 yielding genotypes in eight environments, crossing between them is not useful because it will lead to very low genetic variation in grain yield as both genotypes tend to be genetically similar with only GD of 0.007 and nine different alleles. Therefore, using genomic tools (e.g. identifying number of target alleles, number of different alleles, genetic diversity analyses, etc.) in parallel with phenotypic selection is very fruitful to improve target traits. Moving forward such information on important MTAs can be considered along with genomic selection which is now routinely performed in the breeding program for shortlisting promising parental lines for new sets of crosses for improving grain yield [9].

#### Putative SNP markers for validation and future use in MAS

For the candidate SNPs detected by GWAS, a set of 10 SNP markers were found to be present in the 13 selected genotypes, while three SNPs were not shown in any of the selected genotypes. As it is recommended in each environment to validate the SNPs that were found in the that environment, it may be useful to convert these 13 SNP to kompetitive allele specific PCR (KASP) markers and validate them for high grain yield in a different genetic background. For example, S2A\_718916923 marker was found in all selected genotypes and should be validated. Although the marker allele C was found to be significantly associated with increased grain yield in Alliance, this allele was also present in all selected genotypes in other environments. This marker allele has probably minor effects in increasing grain yield in the other environments but the GEI hinders this marker in the other environments to be significantly associated with grain yield.



Interestingly, two SNPs associated with GY in this study were also found to be associated with grain yield in a previous study [10]. This earlier study evaluated grain yield for synthetic winter wheat genotypes in Turkey for two seasons. The two SNPs (S3A\_24993796 and S3A\_24993797) are associated with GY in Alliance environment (Supplementary Table S5). Moreover, in their study, the allele A for S3A\_24993796 and C for S3A\_24993797 were associated with increased grain yield. The same two alleles were also associated with increased grain yield of Nebraska winter wheat. The amount of phenotypic variation ( $R^2$ ) explained by these markers is similar ( $\sim 11.5\%$ ) in both studies. High LD between the two SNPs has also been highlighted in both of our studies. Overall, the two SNPs are associated with grain yield in two independent genetic backgrounds (Nebraska wheat and synthetic winter wheat from Turkey) and two testing environments (Nebraska and Turkey) further validating the associations identified in this study. It is known that states in the Great Plain and Turkey share some of the same climatic features [7]. Hence, introgression of favorable alleles across germplasm which have not have been previously incorporated into their germplasm may be a possibility [7].

## Conclusion

$G \times E$  interaction needs to be investigated for identifying genotypes that can be used for improving grain yield in all tested environments. However, utilizing the advances in DNA sequencing and genetic analysis by GWAS has helped in identifying possible candidate genotypes to be used in improving grain yield in a group of environments. Three criteria were suggested to identify candidate genotypes to be used as parents for crossing in a group of tested environments. These three criteria depend on combining the information at phenotypic and genomic levels. This research provided a framework for considering how to select parents for future crosses.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-020-07308-0>.

**Additional file 1: Table S1.** Pedigree information for F<sub>3:6</sub> Nebraska Duplicate Nursery winter wheat. **Table S2.** List of SNPs generated from GBS data. **Table S3.** The genotypes performance for GY under nine environments. **Table S4.** Average high and low monthly temperatures, total monthly precipitation, and total snowfall in all environments except Grant and Kansas for growing season (2016 /2017). **Table S5.** Details of GWAS analysis for grain yield using mixed linear model (MLM) at the significance level of 5% bonferroni correction using 11,991 SNPs. **Table S6.** gene annotation and candidate genes for the high LD genomic regions. **Table S7.** The presence (1) and absence (0) of all alleles associated with high yielding detected by GWAS in 13 selected genotypes in all environments. **Table S8.** The distance matrix based on genotypic data between the best high yielding genotypes. **Table S9.** The different allele effects matrix among the 13 selected genotypes.

## Abbreviations

AFLPs: Amplified fragment length polymorphisms; ANOVA: Analysis of Variance; CIMMYT: The international Maize and Wheat Improvement Center; DUP17: Nebraska Duplicate Nursery 2017; GBS: The genotyping-by-sequencing; GEI: Genotype  $\times$  environment interaction; GRs: Genomic regions; GY: Grain yield; GWAS: Genome-wide association study; KASP: Kompetitive allele specific PCR; LD: Linkage disequilibrium; MEYTs: Multi-environment yield trials; MTAs: Marker Trait Association; PCA: Principal component analysis; QTLs: Quantitative trait loci; RFLPs: Restriction fragment length polymorphisms; RMEL: The residual maximum likelihood; SNPs: Single nucleotide polymorphisms; SSRs: Simple sequence repeats

## Acknowledgments

The authors thank Dr. Jesse Poland for performing the GBS data and Sarah Blecha for extracting the DNA.

## Authors' contributions

AS, SHE designed the study, performed all analyses, and wrote the manuscript; PSB designed the study, discussed the results, and drafted the manuscript; AMA, performed gene annotation, G $\times$ E interaction analysis, and drafted the manuscript; VB, performed the SNP calling of the GBS data; HAE, AAN, and KFMS helped in discussing the results. All authors have read and approved the manuscript.

## Funding

Not applicable.

## Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the NCBI repository, <http://www.ncbi.nlm.nih.gov/bioproject/680548>

## Ethics approval and consent to participate

Not application.

## Consent for publication

Not application.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Department of Agronomy & Horticulture, University of Nebraska-Lincoln, Lincoln, USA. <sup>2</sup>Department of Plant Biotechnology, Genetic Engineering and Biotechnology Research Institute (GEBRI), University of Sadat City (USC), Sadat City, Egypt. <sup>3</sup>Department of Biology, College of Science and Humanitarian Studies, Shaqra University, Qwaieah, Saudi Arabia. <sup>4</sup>Institute of Agricultural and Nutritional Sciences, Martin Luther University Halle-Wittenberg, Betty-Heimann-Str. 3, 06120 Halle (Saale), Germany. <sup>5</sup>Department of Genetics, Faculty of Agriculture, Assiut University, Assuit 71526, Egypt.

Received: 5 May 2020 Accepted: 7 December 2020

Published online: 02 January 2021

## References

- Ahmad MQ. Association Mapping For Drought Tolerance In Wheat (*Triticum aestivum* L.). Pakistan: University of Agriculture Faisalabad; 2013.
- Ain Q-u, Rasheed A, Anwar A, Mahmood T, Imtiaz M, He Z, Xia X, Quraishi UM. Genome-Wide Association for Grain Yield under Rainfed Conditions in Historical Wheat Cultivars from Pakistan. *Front Plant Sci.* 2015;6:743.
- Aina OO, Dixon AGO, Paul I, Akinrinde EA. G $\times$  E interaction effects on yield and yield components of cassava (landraces and improved) genotypes in the savanna regions of Nigeria. *Afr J Biotechnol.* 2009;8(19) Academic Journals (Kenya).
- Akter A, Jamil Hasan M, Umma K, Rahman MH, Khatun M, Islam MR. GGE Biplot analysis for yield stability in multi-environment trials of promising hybrid Rice (*Oryza Sativa* L.). *Bangladesh Rice J.* 2015;19(1):1–8.
- Alqudah AM, Haile JK, Alomari DZ, Pozniak CJ, Kobiljski B, Börner A. Genome-Wide and Snp Network Analyses Reveal Genetic Control of Spikelet Sterility and Yield-Related Traits in Wheat. *Sci Rep.* 2020;10(1):1–12.



6. Alqudah AM, Sallam A, Baenziger PS, Börner A. GWAS: fast-forwarding gene identification in temperate cereals: barley as a case study-A review. *J Adv Res.* 2019;(22):119–35 Elsevier.
7. Baenziger PS, Salah I, Little RS, Santra DK, Regasa T, Wang MY. Structuring an Efficient Organic Wheat Breeding Program. *Sustainability.* 2011;3(8):1190–205.
8. Baenziger, P.S. Wheat breeding and genetics. Reference Module in Food Science, no December. 2016 <https://doi.org/10.1016/b978-0-08-100596-5.03001-8>.
9. Belamkar V, Guttieri MJ, Hussain W, Diego Jarqu\in, El-basyoni I, Poland J, Lorenz AJ, Baenziger PS. Genomic Selection in Preliminary Yield Trials in a Winter Wheat Breeding Program. *G3.* 2018;8(8):2735–47.
10. Bhatta M, Morgounov A, Belamkar V, Baenziger PS. Genome-Wide Association Study Reveals Novel Genomic Regions for Grain Yield and Yield-Related Traits in Drought-Stressed Synthetic Hexaploid Wheat. *Int J Mol Sci.* 2018;19(10):3011.
11. Bossolini E, Wicker T, Knobel PA, Keller B. Comparison of Orthologous Loci from Small Grass Genomes Brachypodium and Rice: Implications for Wheat Genomics and Grass Genome Annotation. *Plant J.* 2007;49(4):704–17 Wiley Online Library.
12. 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;23(19):2633–5.
13. Braun H-J, Rajaram S, Ginkel M. "CIMMYT" Approach to Breeding for Wide Adaptation. In: *Adaptation in Plant Breeding*; Springer; 1997. p. 197–205.
14. Burguen J, Dreisigacker S, Vargas M, Herrera-foessel SA, Lillemo M, Singh RP, Trethowan R, et al. Association Analysis of Historical Bread Wheat Germplasm Using Additive Genetic Covariance of Relatives and Population Structure. 2007;1913(November):1889–913 <https://doi.org/10.1534/genetics.107.078659>.
15. Butler AD. Package 'Asreml'; 2018.
16. Butler DG, Cullis BR, Gilmour AR, Gogel BJ, Thompson R. ASReml-R reference manual version 4. ASReml-R Ref Man. 2017;176 <http://www.homepages.ed.ac.uk/iwhite/asreml/uop>.
17. Charmet G, Robert N, Perretant MR, Gay G, Sourdille P, Groos C, Bernard S, Bernard M. Marker Assisted Recurrent Selection for Cumulating QTLs for Bread-Making Related Traits. *Euphytica.* 2001;119(1–2):89–93.
18. Chinipardaz A, Karimizadeh R, Asghari A, Chinipardaz R, Sofalian O, Ghaffari A. Application of GGE Biplot Analysis to Evaluate Grain Yield Stability of Rainfed Spring Durum Wheat Genotypes and Test Locations by Climatic Factors in Iran. *Crop Breeding J.* 2016;6(2):41–9.
19. Dhaliwal AS, Mares DJ, Marshall DR. Effect of 1B/1R chromosome translocation on milling and quality characteristics of bread Wheats. *Cereal Chem.* 1987;64(2):72–6.
20. Dodig D, Zoric M, Kobijlski B, Savic J, Kandic V, Quarrie S, Barnes J. Genetic and Association Mapping Study of Wheat Agronomic Traits under Contrasting Water Regimes. *Int J Mol Sci.* 2012;13(5):6167–88.
21. 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;127(4):791–807.
22. El-basyoni IS. Association Mapping for Important Biotic & Abiotic Related Traits in Structured Wheat Population; 2012.
23. Ellis MH, Bonnett DG, Rebetzke GJ. A 192bp Allele at the Xgwm261 Locus is Not Always Associated with the Rht8 Dwarfing Gene in Wheat (*Triticum aestivum* L.). *Euphytica.* 2007;157(1–2):209–14.
24. Eltaher S, Sallam A, Belamkar V, Emara HA, Nower AA, Salem KFM, Poland J, Baenziger PS. Genetic Diversity and Population Structure of F<sub>3,6</sub> Nebraska Winter Wheat Genotypes Using Genotyping-By-Sequencing. *Front Genet.* 2018;9:76.
25. Federer WT, Crossa J. I4 Screening Experimental Designs for Quantitative Trait Loci, Association Mapping, Genotype-by Environment Interaction, and Other Investigations. *Front Physiol.* 2012;1–8 <https://doi.org/10.3389/fphys.2012.00156>.
26. Fowler DB, N'Diaye A, Laudencia-Chinguanco D, Pozniak CJ. Quantitative trait loci associated with phenological development, lowtemperature tolerance, grain quality, and agronomic characters in wheat (*Triticum aestivum* L.). *PLoS One.* 2016;11(3):e0152185.
27. Friebe BR, Tuleen NA, Gill BS. Development and Identification of a Complete Set of *Triticum aestivum-Aegilops geniculata* Chromosome Addition Lines. *Genome.* 1999;42(3):374–80.
28. Gupta PK, Rustgi S, Kumar N. Genetic and Molecular Basis of Grain Size and Grain Number and Its Relevance to Grain Productivity in Higher Plants. *Genome.* 2006;49(6):565–71.
29. Huang XQ, Kempf H, Ganai MW, Röder MS. Advanced Backcross QTL Analysis in Progenies Derived from a Cross between a German Elite Winter Wheat Variety and a Synthetic Wheat (*Triticum aestivum* L.). *Theor Appl Genet.* 2004;109(5):933–43.
30. Hussain W, Baenziger PS, Belamkar V, Guttieri MJ, Venegas JP, Easterly A, Sallam A, Poland J. Genotyping-by-Sequencing Derived High-Density Linkage Map and Its Application to QTL Mapping of Flag Leaf Traits in Bread Wheat. *Sci Rep.* 2017;7(1):16394.
31. Hussain W, Campbell M, Walia H, Morota G. ShinyAIM: Shiny-Based Application of Interactive Manhattan Plots for Longitudinal Genome-Wide Association Studies. *BioRxiv.* 2018:1–4 <https://doi.org/10.1101/383026>.
32. Kan CC, Jia H, Powers C, Carver BF, Yan L. Genetic characterization and deployment of a major gene for grain yield on chromosome arm 1BS in winter wheat. *Mol Breed.* 2020;40(3) <https://doi.org/10.1007/s11032-020-1106-0>.
33. Kaya Y, Akçura M, Taner S. GGE-Biplot Analysis of Multi-Environment Yield Trials in Bread Wheat. *Turk J Agric Forestry.* 2006;30(5):325–37.
34. 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;19(2):163–77.
35. Li F, Wen W, Liu J, Zhang Y, Cao S, He Z, Rasheed A, et al. Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biol.* 2019;19(1):168.
36. Li H, Durbin R. Fast and Accurate Short Read Alignment with Burrows–Wheeler Transform. *Bioinformatics.* 2009;25(14):1754–60.
37. Liu J, Xu Z, Fan X, Zhou Q, Cao J, Wang F, Ji G, Yang L, Feng B, Wang T. A Genome-Wide Association Study of Wheat Spike Related Traits in China. *Front Plant Sci.* 2018;9:1584.
38. Liu Y, Bowman BC, Hu Y-G, Liang X, Zhao W, Wheeler J, Klassen N, Bockelman H, Bonman JM, Chen J. Evaluation of Agronomic Traits and Drought Tolerance of Winter Wheat Accessions from the USDA-ARS National Small Grains Collection. *Agronomy.* 2017;7(3):51.
39. Lu C-f, Shen L-s, Tan Z, Xu Y, He P, Chen Y, Zhu L. Comparative Mapping of QTLs for Agronomic Traits of Rice across Environments Using a Doubled Haploid Population. *Theor Appl Genet.* 1996;93(8):1211–7.
40. Ma D, Sun D, Wang C, Ding H, Qin H, Hou J, Huang X, Xie Y, Guo T. Physiological Responses and Yield of Wheat Plants in Zinc-Mediated Alleviation of Drought Stress. *Front Plant Sci.* 2017;8:860.
41. Ma L, Yang C, Zeng D, Cai J, Li X, Ji Z, Xia Y, Qian Q, Bao J. Mapping QTLs for Heading Synchrony in a Doubled Haploid Population of Rice in Two Environments. *J Genetics Genomics.* 2009;36(5):297–304.
42. Manneh B, Stam P, Struik PC, Bruce-Oliver S, Van Eeuwijk FA. QTL-Based Analysis of Genotype-by-Environment Interaction for Grain Yield of Rice in Stress and Non-Stress Environments. *Euphytica.* 2007;156(1–2):213–26.
43. Metsalu T, Vilo J. ClustVis: A web tool for visualizing clustering of multivariate data using principal component analysis and Heatmap. *Nucleic Acids Res.* 2015;43(W1):W566–70 <https://doi.org/10.1093/nar/gkv468>.
44. Mohammadi R, Amri A. Genotypex Environment Interaction and Genetic Improvement for Yield and Yield Stability of Rainfed Durum Wheat in Iran. *Euphytica.* 2013;192(2):227–49.
45. Mohammadi R, Farshadfar E, Amri A. Interpreting Genotypex Environment Interactions for Grain Yield of Rainfed Durum Wheat in Iran. *Crop J.* 2015; 3(6):526–35.
46. Neumann K, Kobijlski B, Denčić S, Varshney RK, Börner A. Genome-Wide Association Mapping: A Case Study in Bread Wheat (*Triticum aestivum* L.). *Mol Breed.* 2011;27(1):37–58.
47. Pacheco A, Vargas M, Alvarado G, F Rodr\iguez, Crossa J, Burgueño J. GEA-R (Genotype x Environment Analysis with R for Windows) Version 4.0. Mexico: CIMMYT; 2015.
48. Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, Dreisigacker S, et al. Genomic Selection in Wheat Breeding Using Genotyping-by-Sequencing. *Plant Genome.* 2012;5(3):103–13.
49. Pritchard JK, Stephens M, Donnelly P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics.* 2000;155(2):945–59.
50. Quarrie SA, Quarrie SP, Radosevic R, Rancic D, Kaminska A, Barnes JD, Leverington M, et al. Dissecting a Wheat QTL for Yield Present in a Range of Environments : From the QTL to Candidate Genes. 2006;57(11):2627–37 <https://doi.org/10.1093/jxb/erl026>.
51. Sallam A, Martsch R. Association Mapping for Frost Tolerance Using Multi-Parent Advanced Generation Inter-Cross (MAGIC) Population in Faba Bean (*Vicia faba* L.). *Genetica.* 2015;143(4):501–14 <https://doi.org/10.1007/s10709-015-9848-z>.

52. Sallam A, Mourad AMI, Hussain W, Baenziger PS. Genetic Variation in Drought Tolerance at Seedling Stage and Grain Yield in Low Rainfall Environments in Wheat (*Triticum aestivum* L.). *Euphytica*. 2018;214(9):169.
53. Samonte SOPB, Wilson LT, McClung AM, Medley JC. Targeting Cultivars onto Rice Growing Environments Using AMMI and SREG GGE Biplot Analyses. *Crop Sci*. 2005;45(6):2414–24.
54. Shah MM, Baenziger PS, Yen Y, Gill KS, Moreno-Sevilla B, Haliloglu K. Genetic Analyses of Agronomic Traits Controlled by Wheat Chromosome 3A. *Crop Sci*. 1999;39(4):1016–21.
55. Shehabeldin S, Eltaher M. Manuscript.Docx; 2018.
56. Shen X, Yuan Y, Zhang H, Guo Y, Zhao Y, Li S, Kong F. The Hot QTL Locations for Potassium, Calcium, and Magnesium Nutrition and Agronomic Traits at Seedling and Maturity Stages of Wheat under Different Potassium Treatments. *Genes*. 2019;10(8):607.
57. Shin J-H, Blay S, McNeney B, Graham J, et al. LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J Stat Softw*. 2006;16(3):1–10.
58. Snape JW, Foulkes MJ, Simmonds J, Leverington M, Fish LJ, Wang Y, Ciavarella M. Dissecting Genex Environmental Effects on Wheat Yields via QTL and Physiological Analysis. *Euphytica*. 2007;154(3):401–8.
59. Sukumaran S, Dreisigacker S, Lopes M, Chavez P, Reynolds MP. Genome-Wide Association Study for Grain Yield and Related Traits in an Elite Spring Wheat Population Grown in Temperate Irrigated Environments. *Theor Appl Genet*. 2015;128(2):353–63.
60. Tsilo TJ, Hareland GA, Simsek S, Chao S, Anderson JA. Genome Mapping of Kernel Characteristics in Hard Red Spring Wheat Breeding Lines. *Theor Appl Genet*. 2010;121(4):717–30.
61. Voltas J, Lopez-Corcoles H, Borras G. Use of Biplot Analysis and Factorial Regression for the Investigation of Superior Genotypes in Multi-Environment Trials. *Eur J Agron*. 2005;22(3):309–24.
62. Wang L, Ge H, Hao C, Dong Y, Zhang X. Identifying Loci Influencing 1,000-Kernel Weight in Wheat by Microsatellite Screening for Evidence of Selection during Breeding. *PLoS One*. 2012;7(2):e29432.
63. Wang S-X, Zhu Y-L, Zhang D-X, Shao H, Liu P, Hu J-B, Zhang H, et al. Genome-Wide Association Study for Grain Yield and Related Traits in Elite Wheat Varieties and Advanced Lines Using SNP Markers. *PLoS One*. 2017; 12(11):e0188662.
64. Wilkinson S, Kudoyarova GR, Veselov DS, Arkhipova TN, Davies WJ. Plant Hormone Interactions: Innovative Targets for Crop Breeding and Management. *J Exp Bot*. 2012;63(9):3499–509.
65. Xu F-f, Fu-fu TANG, Ya-fang SHAO, Ya-ling CHEN, Chuan T, Jin-song BAO. Genotypex Environment Interactions for Agronomic Traits of Rice Revealed by Association Mapping. *Rice Sci*. 2014;21(3):133–41.
66. Yan W, Hunt LA, Sheng Q, Szlavnicz Z. Cultivar Evaluation and Mega-Environment Investigation Based on the GGE Biplot. *Crop Sci*. 2000;40(3):597–605.
67. Yan W, Rajcan I. Biplot Analysis of Test Sites and Trait Relations of Soybean in Ontario. *Crop Sci*. 2002;42(1):11–20.
68. Yan W, Hunt LA. Interpretation of Genotypex Environment Interaction for Winter Wheat Yield in Ontario. *Crop Sci*. 2001;41(1):19–25.
69. Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, et al. A Unified Mixed-Model Method for Association Mapping That Accounts for Multiple Levels of Relatedness. *Nat Genet*. 2006;38(2):203.
70. Zanke CD, Ling J, Plieske J, Kollers S, Ebmeyer E, Korzun V, Argillier O, et al. Whole Genome Association Mapping of Plant Height in Winter Wheat (*Triticum aestivum* L.). *PLoS One*. 2014;9(11):e113287.
71. Zhang J, Xu Y, Chen W, Dell B, Vergauwen R, Biddulph B, Khan N, Luo H, Appels R, den Ende W. A Wheat 1-FEH W3 Variant Underlies Enzyme Activity for Stem WSC Remobilization to Grain under Drought. *New Phytol*. 2015;205(1):293–305.
72. Zhang L-y, Liu D-c, Guo X-l, Yang W-l, Sun J-z, Wang D-w, Zhang A. Genomic Distribution of Quantitative Trait Loci for Yield and Yield-Related Traits in Common Wheat. 2010;52(11):996–1007 <https://doi.org/10.1111/j.1744-7909.2010.00967.x>.
73. Zhang X, P Pérez-Rodríguez, Semagn K, Beyene Y, Babu R, López-Cruz MA, Vicente FS, et al. Genomic Prediction in Biparental Tropical Maize Populations in Water-Stressed and Well-Watered Environments Using Low-Density and GBS SNPs. *Heredity*. 2015;114(3):291.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

