Check for updates

Genome-wide association study for grain yield and component traits in bread wheat (*Triticum aestivum* L.)

Hanif Khan[1], Gopalareddy Krishnappa[1,2]*, Satish Kumar[1],
Chandra Nath Mishra[1], Hari Krishna[3], Narayana Bhat Devate[3],
Nagenahalli Dharmegowda Rathan[3], Om Parkash[1],
Sonu Singh Yadav[1], Puja Srivastava[4], Suma Biradar[5],
Monu Kumar[6] and Gyanendra Pratap Singh[1]*

[1]ICAR-Indian Institute of Wheat and Barley Research, Karnal, India, [2]ICAR-Sugarcane Breeding
Institute, Coimbatore, India, [3]ICAR-Indian Agricultural Research Institute, New Delhi, India, [4]Punjab
Agricultural University (PAU), Ludhiana, India, [5]University of Agricultural Sciences, Dharwad, India,
[6]ICAR-Indian Agricultural Research Institute, Jharkhand, India

Genomic regions governing days to heading (DH), grain filling duration (GFD),
grain number per spike (GNPS), grain weight per spike (GWPS), plant height (PH),
and grain yield (GY) were investigated in a set of 280 diverse bread wheat
genotypes. The genome-wide association studies (GWAS) panel was genotyped
using a 35K Axiom Array and phenotyped in five environments. The GWAS
analysis showed a total of 27 Bonferroni-corrected marker-trait associations
(MTAs) on 15 chromosomes representing all three wheat subgenomes. The GFD
showed the highest MTAs (8), followed by GWPS (7), GY (4), GNPS (3), PH (3), and
DH (2). Furthermore, 20 MTAs were identified with more than 10% phenotypic
variation. A total of five stable MTAs (*AX-95024590, AX-94425015, AX-
95210025 AX-94539354*, and *AX-94978133*) were identified in more than
one environment and associated with the expression of DH, GFD, GNPS,
and GY. Similarly, two novel pleiotropic genomic regions with associated
MTAs i.e. *AX-94978133* (4D) and *AX-94539354* (6A) harboring co-localized
QTLs governing two or more traits were also identified. *In silico* analysis
revealed that the SNPs were located on important putative candidate genes
such as *F-box-like domain superfamily, Lateral organ boundaries, LOB,
Thioredoxin-like superfamily Glutathione S-transferase*, *RNA-binding domain
superfamily*, *UDP-glycosyltransferase family*, *Serine/threonine-protein kinase*,
Expansin, *Patatin, Exocyst complex component Exo70, DUF1618 domain,
Protein kinase domain* involved in the regulation of grain size, grain number,
growth and development, grain filling duration, and abiotic stress tolerance. The
identified novel MTAs will be validated to estimate their effects in different
genetic backgrounds for subsequent use in marker-assisted selection (MAS).

KEYWORDS

wheat, GWAS, SNPs, candidate genes, mapping, yield component traits

Introduction

Bread wheat (*Triticum aestivum* L., 2n = 6x = 42) is one of the most important staple food and the world's highest-grown and traded cereal. It provides about 21% of calories and 19% of day-to-day protein to approximately 4.5 billion global populations (Braun et al., 2010). The annual gain in wheat yield should be increased from the current level of around 1%–1.6% to meet the food demand of the estimated global population of 9 billion by the year 2050 (Wheat Initiative, 2013; FAO, 2017). The available resources likely be reduced to a great extent; the problem will be further complicated by the erratic rainfall, reduced water table, change in temperature, and reduced soil health. For sustainable crop production, we need to increase the yield potential, and multiple stress tolerance and improve input use efficiency along with climate-smart agronomic practices (Giraldo et al., 2019). Integration of modern plant breeding tools like marker-aided selection (MAS), marker-assisted recurrent selection (MARS), genomic selection, and speed breeding with conventional breeding approaches is of paramount importance to enhance yield gain in wheat (Krishnappa et al., 2021).

Grain yield is a genetically complex trait and is an outcome of the combined effect of several agro-morphological and physiological traits (Chen et al., 2012; Sukumaran et al., 2018). Agro-morphological traits include grain number per spike, thousand kernel weight, biomass, harvest index, productive tillers number, spike length, grain weight per spike, and plant height has a significant effect on wheat grain yield along with phenological traits like days to heading, maturity, and grain filling duration (Sun et al., 2017; Wang et al., 2017; Liu et al., 2018; Ma et al., 2018; Jamil et al., 2019; Li et al., 2020). Unlike grain yield, many of the yield component traits have high heritability and are easier to select particularly during the early stages of breeding cycles. The yield plateaus may be avoided by selecting the yield components, as they offer additional avenues for genetic gain enhancement. It is suggested for trait-based breeding using elite and genetically complementary genotypes to enhance wheat yield improvement (Liu et al., 2015; Reynolds et al., 2017). The grain yield and its component traits are complex and quantitative, as each of these traits is controlled by several genes with small effects. Furthermore, most of the traits have low to moderate heritability with significant genotype × environment interactions (Kaya and Akcura, 2014). Molecular breeding is a potential strategy to improve complex traits like yield and its contributing traits, but a better understanding of genetic architecture is important for the effective utilization of molecular tools. Therefore, genetic dissection of agro-morphological traits is essential for the improvement of wheat yield.

Two approaches i.e. genome-wide association studies (GWAS) and quantitative trait loci (QTL) mapping are widely used methods to dissect the genetic basis of complex quantitative traits in crop plants. In the past decade, extensive efforts have been made to identify QTLs associated with grain yield and its component traits (Gao et al., 2015; Zhang et al., 2016; Jin et al., 2020; Isham et al., 2021; Kang et al., 2021; Li et al., 2022) in wheat through bi-parental populations based QTL mapping. Conventional QTL mapping mainly depends on structured populations like recombinant inbred lines (RIL), back-crosses (BC), and doubled haploids (DH). The several shortcomings associated with QTL mapping are low resolution due to one or few cross-overs and low marker density (Korte and Farlow, 2013). Recent advances in sequencing technologies created valuable genomic resources including high-quality genome data (Brenchley et al., 2012; Chapman et al., 2015), as result, several high throughput SNP arrays have been developed and utilized in wheat. GWAS becomes a complementary strategy to QTL mapping to dissect complex traits, particularly after the availability of large-scale genomic resources (Saidou et al., 2014). Unlike bi-parental population-based QTL mapping, GWAS consists of more genetically diverse lines that harbor several historical and ancestral recombination events. Additionally, the use of diverse germplasms as study materials has the potential to capture superior alleles that have been missed by breeding practices. GWAS is based on the linkage disequilibrium (LD) that has formed in a population over the generations, the genomic regions harboring QTLs can be detected even in the absence of inclusion of causal mutations among the set of available molecular markers.

The two common limitations (i.e. limited allelic diversity and low genomic resolution) associated with the bi-parental QTL mapping methods can be overcome in the GWAS approach (Brachi et al., 2011). However, the major challenge for GWAS is the control of false positives caused by population structure and family relatedness (Kaler et al., 2020). Incorporation of these two confounding factors as covariates in the mixed linear model (MLM) addressed the issue of false positives (Price et al., 2006), however, false negatives have been significantly increased which might exclude the important loci. To overcome the false negatives, multi-locus GWAS methods like multi-loci mixed linear model (MLMM), fixed and random model circulating probability unification (FarmCPU), and Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) have been developed (Zhang et al., 2019). The statistical power of BLINK is superior and gives reduced false-positive discovery compared to many available GWAS models including SUPER and FarmCPU, as BLINK removes the assumption of equal distribution of causal genes in the whole genome (Huang et al., 2019).

GWAS has been successfully used in wheat to dissect the genetic basis of yield and its component traits. In previous studies, GWAS panels have been phenotyped in a range of production conditions including drought, irrigated, and salt stress to identify QTLs. Several drought-tolerance QTLs associated with grain yield and its component traits have been

identified (Edae et al., 2014; Ain et al., 2015; Gahlaut et al., 2019; Suliman et al., 2021; Bennani et al., 2022; Said et al., 2022). Similarly, genomic regions governing yield and its attributing traits in normal irrigated production conditions were also identified (Sukumaran et al., 2015; Sun et al., 2017; Godoy et al., 2018; Liu et al., 2018; Ma et al., 2018; Sukumaran et al., 2018; Bajgain et al., 2019; Jamil et al., 2019; Li et al., 2019; Rahimi et al., 2019; Sheoran et al., 2019; Ward et al., 2019; Ali et al., 2020; Alqudah et al., 2020; Pang et al., 2020; Alemu et al., 2021; Eltaher et al., 2021; Gao et al., 2021; Malik et al., 2021; Saini et al., 2022; Zhang et al., 2022). Also, QTLs were identified in hostile soils under salt stress conditions for yield and related traits (Hu et al., 2021). Similarly, MTAs were also identified for biotic stresses (Vikas et al., 2022) and quality traits (Sandhu et al., 2021; Rathan et al., 2022) in wheat. Although several marker-trait associations (MTAs) were identified in different GWAS studies for yield and its component traits, there might be several false positives in most of the studies due to a very low threshold ($-\log_{10}$ p-value ≥ 3.0) fixation to consider the MTA as a significant. In only a few GWAS (Gahlaut et al., 2019; Eltaher et al., 2021; Malik et al., 2021; Zhang et al., 2022), the threshold$-\log_{10} p$ values were adjusted by the calculation of the corresponding Bonferroni correction at a significance level of 5% to reduce the false positives. In wheat, many QTLs/MTAs have been identified; however, additional genetic studies are warranted using different genetic materials, as we have not reached a saturation point (Singh et al., 2021). Thus, more efforts are required to dissect the genetic mechanisms of yield and component traits in wheat and to devise marker-based breeding approaches that involve marker-assisted selection or genome-wide selection to obtain increased genetic gains. The present study aimed to identify the genomic region (s) associated with grain yield and component traits i.e. days to heading (DH), grain filling duration (GFD), grain number per spike (GNPS), grain weight per spike (GWPS), grain yield (GY), and plant height (PH) in a panel of diverse bread wheat genotypes in a range of environments through the GWAS approach and the putative candidate genes associated with the SNPs.

Materials and methods

Plant material and field experiments

A set of 280 diverse bread wheat genotypes (Supplementary Table S1) consisting of advanced breeding lines and commercial cultivars were used for GWAS analysis. The GWAS panel of 280 genotypes was selected from the All India Coordinated Research Project on Wheat and Barley. The GWAS panel was evaluated at five different environments: E1- University of Agricultural Sciences, research farm, Dharwad (15°29′20″N, 74°59′3″E, 750 m AMSL), E2- ICAR-Indian Agricultural Research Institute, New Delhi (28°38′30″N, 77°09′58″E, 228 m

AMSL), E3- ICAR-Indian Agricultural Research Institute, Jharkhand (24°16′58.4″N, 85°21′16.1″E, 651 m AMSL), E4- ICAR-Indian Institute of Wheat and Barley, Karnal (29°41′8″N, 76°59′25″E, 250 m AMSL), and E5- Punjab Agricultural University, Ludhiana (30°54' N, 75°48′E, 247 m AMSL). The crop was sown in the first fortnight of November during the 2020–21 Rabi (winter) season under irrigated conditions. The genotypes were planted in an augmented block design with only the checks (DBW187, MACS6222, WH1124, and WH1142) repeated in a 2 row of 2 m length with a row spacing of 20 cm.

Phenotyping and phenotypic data analysis

All the genotypes of a GWAS panel were phenotyped for six quantitative traits i.e. GWPS (gm), GNPS (number), GY (gm), DH (days), GFD (days); PH (cm) at Dharwad (E1), IARI-Delhi (E2) (except GNPS), IARI Jharkhand (E3), Karnal (E4) (except GNPS) locations. However, the GWAS panel was phenotyped for only three traits i.e. GWPS (gm), GY (gm), and PH (cm) at the Ludhiana location. Plant height (PH) was recorded at physiological maturity as the average of randomly selected three plants of each genotype by measuring from the soil surface to the spike tip excluding awns. Days to heading (DH) were recorded as the number of days from the planting when more than 50% of the plants in each plot showed the emergence of spikes. Physiological maturity was recorded when the majority of plants in the plot showed a complete loss of green colour from the flag leaf. The difference between the days to physiological maturity and the days to heading was used to compute the grain filling duration (GFD). Grain number per spike (GNPS) was calculated as the average of grain number in the main stem spikes of ten randomly selected plants from each genotype. Similarly, grains of all the randomly selected 10 spikes of each entry were weighed separately and the average of 10 spikes was recoded as grain weight per spike (GWS). Grain yield (GY) in grams for each genotype was recorded after harvesting the whole plot. Phenotypic data were analyzed using the R package 'augmentedRCBD' (Aravind et al., 2021).

Genotyping and quality control (QC)

Cetyl Trimethyl Ammonium Bromide (CTAB) method (Murray and Thompson, 1980) was used to extract the genomic DNA from the leaves of 21 days-old seedlings. The GWAS panel was genotyped using Axiom Wheat Breeder's Genotyping Array (Affymetrix, Santa Clara, CA, United States) having 35,143 genome-wide SNPs. The monomorphic, markers with minor allele frequency (MAF) of <5%, missing data of >20%, and heterozygote frequency >25% were removed from the analysis. The

remaining set of 14,790 high-quality SNPs was used in GWAS analysis (Supplementary Table S2).

Population statistics and GWAS

The pair-wise LD values (r^2) between the SNPs located in each chromosome were calculated with Trait Analysis by aSSociation Evolution and Linkage (TASSEL) version 5.0 (Bradbury et al., 2007). The LD block size in the whole genome and three subgenomes was estimated by keeping the r^2 threshold at half LD decay. The principal component analysis (PCA) and Kinship relationship were done through GAPIT (Lipka et al., 2012) to understand the structure of the population used in the GWAS model.

The phenotypic values of GWPS, GNPS, GY, DH, GFD, and PH of 280 diverse genotypes along with corresponding genotyping data were used in GWAS analysis. Significant MTAs were identified using the BLINK (Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway) model (Huang et al., 2019) implemented in Genome Association and Prediction Integrated Tool (GAPIT) version 3.0 (Wang and Zhang, 2021) in the R software package. Determining the correct p-value threshold for statistical significance is critical to differentiate the true positives from false positives. To determine the statistical significance threshold in GWAS, Bonferroni correction has been employed. To estimate Bonferroni correction, α was set to 0.05 which is divided by the total number of SNPs. The Bonferroni-corrected SNPs were considered for significant association and R^2 was used to describe the percentage variation explained (PVE) by significant MTAs.

In silico analysis

The sequence information of the significant SNPs was used to search for putative candidate genes with the Basic Local Alignment Search Tool (BLAST) using default parameters in the Grain Gene database (https://wheat.pw.usda.gov/GG3/) of the bread wheat genome (Wheat Chinese Spring IWGSC RefSeq v2.1 genome assembly (2021)). The genes found in the overlapping region and within the region of 0.1 Mb intervals flanking either side of the associated marker were considered putative candidate genes and their molecular functions were determined. In addition, their expression patterns were investigated using the Wheat Expression database (http://www.wheat-expression.com/) and potential links to phenotypes were determined using the Knetminer tool integrated with the Wheat Expression database. The role of the identified putative candidate genes in the regulation of GWPS, GNPS, GY, DH, GFD, and PH was also determined by the previous reports.

Results

The environment-wise heritability and variance components of the GWAS panel are presented in Table 1. All the studied traits recorded a wider distribution across the environments i.e. DH, GFD, GNPS, GWPS, PH, and GY ranging from 50.4 to 116.4 days, 19.6–55.3 days, 11.6–80.1 number, 0.2–4.6 gm, 57.6–134.8 cm, and 133.8–752.3 gm, respectively. The percent CV of all the studied environments is less than 10.0%, except GFD in E3 (11.1%) and GWPS in E2 (10.8%). The trend of heritability is more environment-specific than trait *per se*, as none of the environments recorded either only low or high performance for the studied traits. There is much variation in the trait's heritability, which ranged from 50.5% to 97.2%.The genotypic variance (σ_G^2) and environmental variance (σ_E^2) are presented in Table 1.

The trait and environment-wise mean values are illustrated graphically through boxplots and presented in Figure 1. The location means of DH were recorded as similar and highest for E2 and E4 followed by E3, and E1, whereas, E1 and E2 were recorded as similar and highest followed by E4 and E3 for GFD. The E4 recorded the highest mean for GWPS followed by E3, E1, E5, and E2. The expression of PH is also similar to DH, as the highest and lowest are recorded by E4 and E1 respectively. The highest yields were recorded by E4 and E5, which is higher than the grand pooled mean, followed by E2, E1, and E3. The general yield trend of E4 and E5 are similar and higher than the pooled mean, the trend was exactly the opposite in E1 and E3 as both of them are similar, which are lower than the grand pooled mean. Whereas, the grain yield of E2 hovers around the pooled mean (Figure 1). The frequency distribution of grain yield and component traits in the GWAS panel evaluated at E1–E5 during 2020–2021 is presented in Figure 2. The genotypes in the GWAS panel showed continuous frequency distributions for all the studied traits. Pearson's correlation coefficient (r^2) of DH, GFD, GNPS, GWPS, PH, and GY was determined and presented in Table 2. The grain-related traits i.e. GNPS and GWPS were a significant positive association with GY in all the environments and pooled mean except E2, where the association was neutral. A similar trend of significant positive association was observed between GY and PH in all the environments and pooled mean, except E4 and E5, where the association was neutral. However, the correlation between DH and GFD is consistent and significant negative in all the environments and pooled mean.

Genome-wide SNP markers distribution

The 35K SNP array was processed to obtain high-quality SNPs, as a result, a set of 14,790 cured genome-wide SNPs was selected. These high-quality set of SNPs along with phenotypic data were further used for GWAS analysis. The highest number of SNPs were mapped on the B subgenome (5649) followed by the D subgenome (4590), and the A subgenome (4551). Chromosome-wise highest SNPs were mapped on 1B (1077), followed by 2B (992), 1D (986), 2D (951), 5B (863), 6B

TABLE 1 Descriptive statistics, variance, and heritability estimates of grain yield and component traits in GWAS panel evaluated at Dharwad, IARI Delhi, IARI Jharkhand, Karnal, and Ludhiana during 2020−2021.

Trait	Env.	Mean ± SD	Range	CV (%)	LSD	h^2_{BS}	GV	EV
DH	E1	59.2 ± 3.8	50.4–72.2	2.4	4.1	85.2	11.9	2.1
	E2	95.8 ± 5.6	81.6–116.4	1.5	4.0	93.5	29.2	2.0
	E3	85.7 ± 4.8	69.8–98.8	1.0	2.3	97.2	23.1	0.7
	E4	94.7 ± 5.2	81.6–116.6	1.9	5.0	88.2	23.7	3.2
GFD	E1	45.0 ± 0.1	31.3–49.1	5.5	6.9	65.1	15.3	6.0
	E2	44.9 ± 0.1	26.9–51.9	4.9	6.2	60.5	13.5	4.8
	E3	31.6 ± 0.3	19.6–46.3	11.1	9.8	50.5	12.4	12.1
	E4	42.7 ± 0.3	28.3–55.3	5.6	6.7	70.9	13.7	5.6
GNPS	E1	52.3 ± 8.6	24.2–80.1	5.3	6.6	96.7	161.9	5.5
	E3	43.9 ± 12.9	11.6–79.4	4.1	6.0	93.9	70.6	4.6
GWPS	E1	2.1 ± 0.6	0.4–3.5	8.7	0.5	91.6	0.4	0.03
	E2	1.9 ± 0.6	0.2–3.9	10.8	0.6	88.3	0.3	0.04
	E3	2.4 ± 0.3	1.2–3.3	8.7	0.6	61.8	0.1	0.04
	E4	2.9 ± 0.5	1.5–4.6	8.9	0.8	73.4	0.2	0.07
	E5	2.1 ± 0.8	0.4–3.71	8.3	0.5	95.4	0.6	0.03
PH	E1	72.9 ± 6.9	57.6–94.1	2.5	5.2	92.9	44.9	3.4
	E2	102.2 ± 7.4	82.0–134.8	2.2	6.5	90.6	50.4	5.2
	E3	96.2 ± 7.6	77.5–115.8	2.2	5.9	91.5	46.4	4.3
	E4	110.8 ± 6.9	94.8–134.5	2.2	6.7	88.2	42.2	5.7
	E5	103.9 ± 5.9	81.3–120.1	2.5	7.2	80.6	27.0	6.5
GY	E1	312.7 ± 100.7	143.3–611.8	6.12	54.1	96.4	9835.2	366.1
	E2	466.8 ± 88.9	232.6–689.3	4.16	55.1	95.1	7348.2	379.6
	E3	256.3 ± 39.3	133.8–372.8	4.67	34.0	89.2	1199.6	144.9
	E4	556.9 ± 90.5	297.8–752.3	6.85	108.1	79.7	5746.2	1461.0
	E5	532.9 ± 82.8	243.2–734.2	5.1	77.2	88.9	5932.8	744.6

DH: days to heading (days); GFD: grain filling duration (days); GNPS: grain number per spike (number); GWPS: grain weight per spike (gm); PH: plant height (cm); GY: grain yield (gm). E1: Dharwad; E2: IARI, Delhi; E3: IARI, Jharkhand; E4: Karnal; E5: Ludhiana; SD: standard deviation; CV: coefficient of variation; h^2BS: broad sense heritability; GV: genotypic variation; EV: environmental variation.

(766), 7B (760), 2A (756), 1A (751), 7A (750), 3B (726), 5A (699), 5D (657), 3D (648), 7D (625), 3A (587), 6A (515), 4A (493), 4B (465), 6D (459), and 4D (264).

Population structure and linkage disequilibrium

The PCA analysis (Figure 3) indicated that there were no clear distinct sub-populations in the GWAS panel. The LD was estimated by calculating the squared correlation coefficient (r^2) for all the SNPs. The LD decay for the whole genome was 4.9 cM and it was found that the decay was rapid in the A genome

(3.6 cM) followed by the D genome (5.2 cM) and B genome (5.7 cM).

Genome-wide association studies

A total of 27 Bonferroni-corrected MTAs were identified for DH, GFD, GNPS, GWPS, PH, and GY. The details of the identified MTAs are presented in Table 3 and illustrated in Manhattan plots in Figures 4A,B. The Q-Q plots depicting the observed associations of SNPs of DH, GFD, GNPS, GWPS, PH, and GY compared to the expected associations after accounting for population structure are presented in Figures 4A,B.

FIGURE 1
Boxplots of grain yield and component traits in GWAS panel evaluated at Dharwad, IARI Delhi, IARI Jharkhand, Karnal, and Ludhiana during 2020−2021.
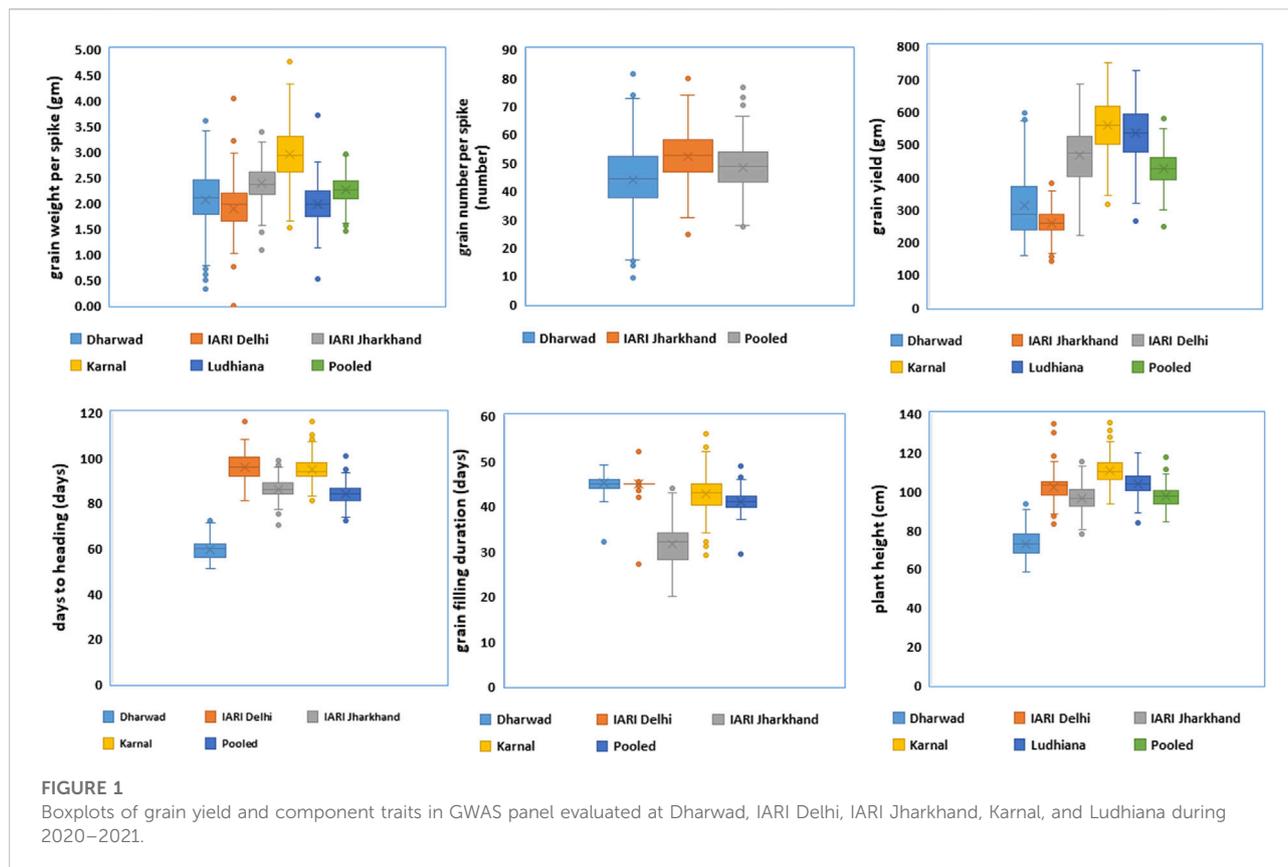
MTAs for grain-related traits

A total of seven significant MTAs were identified for GWPS in E1, E3, and E4 along with pooled mean on chromosomes 1D, 4D, 5A, 6A, 6B, and 7A, explaining the phenotypic variation ranging from 6.6% to 17.1%. The major MTA (*AX-94539354*) on 6A chromosome located at 599.2 Mb explained the highest phenotypic variation of 17.1%. Similarly, two more MTAs (*AX-94602474* and *AX-94883693*) were identified in E1, which were mapped at 15.6 Mb and 40.5 Mb and explained 11.4% and 14.0% of PVE, which were located on 7A and 1D, respectively. Two MTAs i.e. *AX-94978133* and *AX-95105308* were respectively mapped at 465.7 and 113.4 Mb on 4D and 6B chromosomes in the E4 environment. One MTA each on chromosome 6B (*AX-94387482*) and 5A (*AX-94469473*) were mapped at 337.9 and 521.0 Mb, respectively with the phenotypic variation of 11.0% and 11.4% at pooled mean and E3 environment. Genome-wise, a total of three significant MTAs were identified in the A genome while B and D genomes had two each.
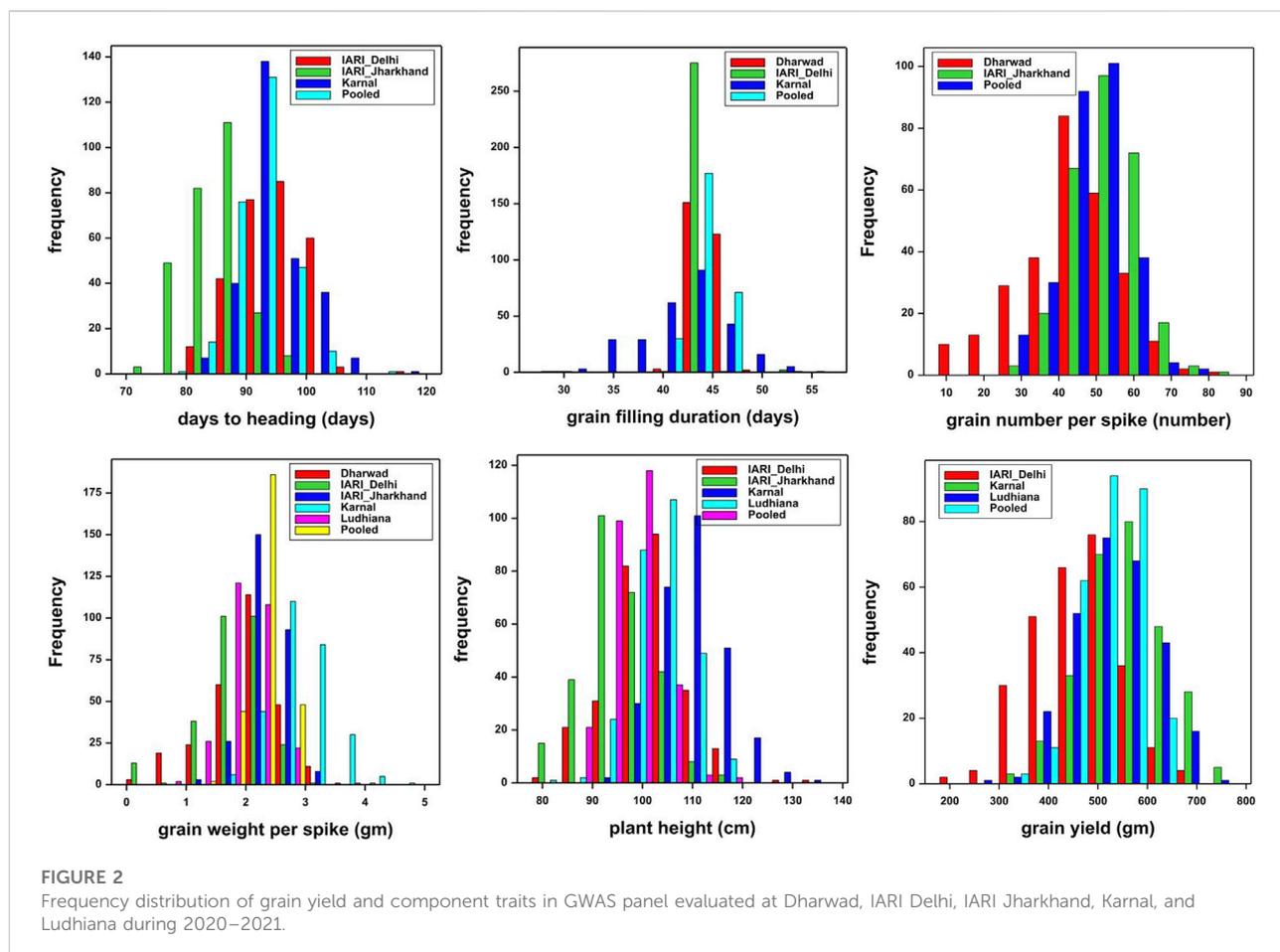
A total of three significant MTAs were identified for GNPS in E1 and pooled mean on chromosomes 4A, 4D, and 6A. One MTA (*AX-94539354*) was mapped at 599.2 Mb on 6A in both E1 and pooled mean and explained 16.3% and 15.6% PVE, respectively. The remaining two MTAs (*AX-94658573* and

AX-94978133) were identified in E1 and mapped at 715.5 and 465.7 Mb on 4A and 4D, explaining more than 10.0% PVE. Genome-wise, two significant MTAs were identified in the A subgenome and one in the D subgenome, however, there is no representation of the B subgenome.

For GY, four significant MTAs (*AX-94473624, AX-94483483, AX-94709904,* and *AX-94978133*) were identified on 1D, 4D, and 7A chromosomes in E3 and E5 along with pooled mean. Three MTAs i.e. *AX-94483483, AX-94709904,* and *AX-94978133* were identified in E3 and mapped at 206.1 Mb, 728.2 Mb, and 465.7 Mb with percent PVE of 8.1, 10.8, and 9.3, respectively. The remaining MTA (*AX-94473624*) was identified in E5 and mapped at 19.3 Mb with a percent PVE of 12.5. Also, one consistent MTA (*AX-94978133*) was identified in both E3 and pooled mean. Genome-wise, three significant MTAs were identified in the D subgenome and one in the A subgenome, however, there is no representation of the B subgenome.

MTAs for agro-morphological traits

A total of two MTAs (*AX-94724456* and *AX-95024590*) were identified for DH in E4 and E3 along with pooled mean. One MTA i.e. *AX-95024590* was identified in two environments

FIGURE 2
Frequency distribution of grain yield and component traits in GWAS panel evaluated at Dharwad, IARI Delhi, IARI Jharkhand, Karnal, and Ludhiana during 2020−2021.

(E3 and E4) along with pooled mean. The other MTA (*AX-94724456*) was identified in E4 and mapped at 484.4 Mb on the 5D chromosome, which explained a PVE of 14.6%. All the MTAs explained more than 14.0% PVE and were located on the D subgenome only. For PH, all three MTAs (*AX-94452759, AX-94498579,* and *AX-94796636*) were identified in E4 with low percent PVE, which ranged from 5.0 to 7.2. Genome-wise, two significant MTAs were identified in the D subgenome and one in the B subgenome, however, there is no representation of the A subgenome.

The highest number of MTAs (8) were identified for GFD in E2, E3, and E4 along with pooled mean on 1A, 2B, 4A, 4B, 5A, and 5B chromosomes. A maximum of four MTAs i.e. *AX-94425015, AX-94702510, AX-94794189,* and *AX-95210025* were identified in E4 and mapped at 2.0 Mb, 776.2 Mb, 356.1 Mb, and 585.4 Mb on 4B, 2B, 5B, 5A chromosomes, respectively with PVE ranging from 18.2% to 20.4%. A total of three MTAs i.e. *AX-94598412, AX-94691261*, and *AX-95181791* were identified in E3 and mapped at 5.0 Mb, 505.5, and 584.6 Mb on 4A, 1A, and 5A chromosomes, respectively with PVE ranging from 15.9% to 18.1%. Two MTAs (*AX-94425015* and *AX-95210025*), which was identified for pooled mean, which was mapped at 2.0 and

585.4 Mb on 4B and 5A chromosomes, respectively with PVE of 16.7% and 15.4%. E2 environment is represented with one MTA (*AX-95107750*) on the 1A chromosome and mapped at 112.9 Mb with a PVE of 9.1%. Genome-wise, five significant MTAs were identified in the D subgenome and three in the B subgenome, however, there is no representation of the D subgenome.

Stable and co-localized MTAs

A total of five consistent MTAs were identified on 3D, 4B, 5A, 6A, and 4D chromosomes for DH, GFD, GNPS, and GY. One MTA i.e. *AX-95024590* was identified in two environments (E3 and E4) along with pooled mean for DH, which was mapped at 152.5 Mb with PVE ranging from 14.4% to 21.1%. Similarly, two MTAs i.e. *AX-94425015* and *AX-95210025* were mapped at 2.0 and 585.4 Mb with the PVE ranging from 15.4% to 20.4%, respectively, which were consistently identified both in E4 and pooled mean for GFD. One each consistent MTA was identified for GNPS (*AX-94539354*) and GY (*AX-94978133*) on 6A and 4D chromosomes, respectively, which were mapped at

TABLE 2 Estimates of phenotypic correlation coefficients for grain yield and component traits in GWAS panel evaluated at Dharwad, IARI Delhi, IARI Jharkhand, Karnal, and Ludhiana during 2020−2021.

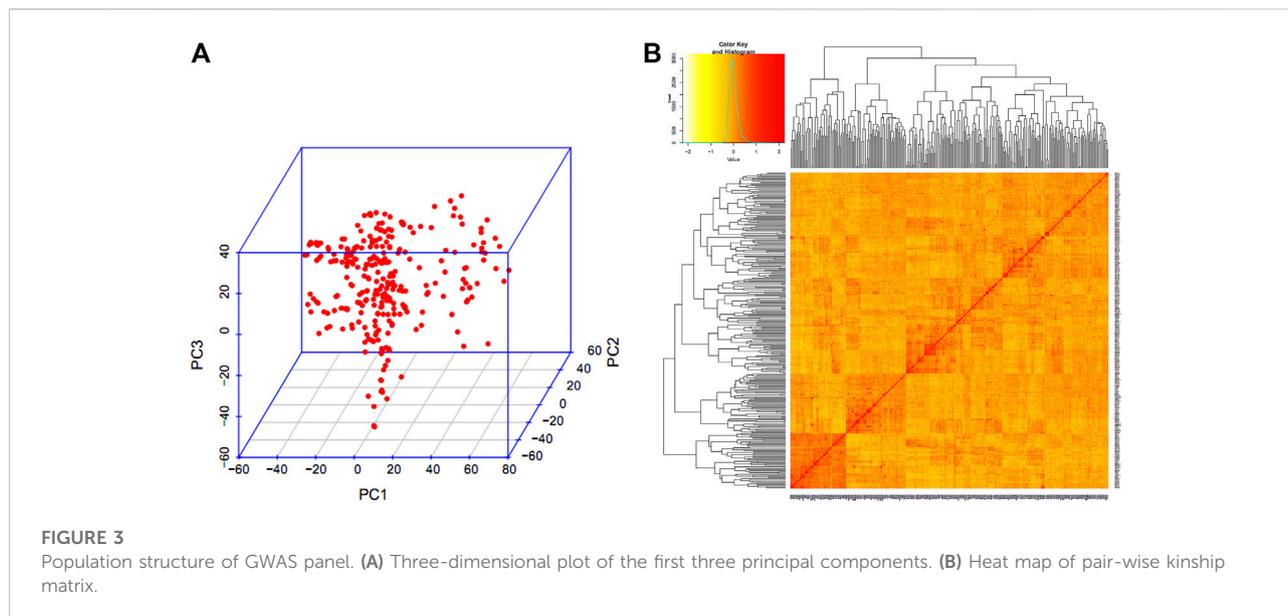| Environments	Traits	GWPS	GNPS	GY	DH	GFD	PH
Dharwad (E1)	GWPS	1.00	0.89**	0.45**	−0.25**	0.00	0.23**
	GNPS		1.00	0.30**	−0.11	0.01	0.13*
	GY			1.00	−0.11	−0.09	0.35**
	DH				1.00	−0.13**	0.00
	GFD					1.00	0.042
	PH						1.00
IARI-Delhi (E2)	GWPS	1.00	—	−0.04	0.050	−0.02	−0.02
	GY			1.00	0.04	0.00	0.13**
	DH				1.00	−0.22**	0.18**
	GFD					1.00	−0.09
	PH						1.00
IARI Jharkhand (E3)	GWPS	1.00	0.55**	0.27**	0.03	0.01	0.25**
	GNPS		1.00	0.12*	0.25**	−0.13*	0.18**
	GY			1.00	−0.01	−0.1	0.15*
	DH				1.00	−0.39**	0.10
	GFD					1.00	−0.29**
	PH						1.00
Karnal (E4)	GWPS	1.00	—	0.21**	−0.04	−0.00	0.25**
	GY			1.00	0.04	−0.08	0.06
	DH				1.00	−0.95**	0.21**
	GFD					1.00	−0.21**
	PH						1.00
Ludhiana (E5)	GWPS	1.00	—	0.89**	—	—	0.12
	GY			1.00	—	—	0.09
	PH						1.00
Across Environment	GWPS	1.00	0.61**	0.44**	−0.09	−0.07	−0.21**
	GNPS		1.00	0.31**	0.09	−0.19**	0.09
	GY			1.00	0.04	−0.03	0.18**
	DH				1.00	−0.71**	0.19**
	GFD					1.00	−0.25**
	PH						1.00

GWPS, grain weight per spike (gm); GNPS, grain number per spike (number); GY, grain yield (gm); DH, days to heading (days); GFD, grain filling duration (days); PH, plant height (cm).

599.3 and 465.7 Mb with PVE ranging from 9.3% to 16.3%. The GNPS MTA was identified in both E1 and pooled mean, whereas, GY MTA was identified in E3 and pooled mean. Furthermore, two co-localized MTAs were identified on 6A and 3D chromosomes. One co-localized MTA (*AX-94978133*) was identified on the 4D chromosome for all the three-grain related traits (GNPS, GWPS, and GY), which was mapped at 465.7 Mb with the PVE ranging from 9.3% to 15.4%. The other co-localized MTA (*AX-94539354*) was identified on the 6A

chromosome for GNPS and GWPS, which was mapped at 599.2 Mb with the PVE ranging from15.6% to 17.1%.

Putative candidate genes associated with MTAs

The significant SNPs associated with GWPS, GNPS, GY, DH, GFD, and PH were used to identify the putative candidate genes

FIGURE 3
Population structure of GWAS panel. **(A)** Three-dimensional plot of the first three principal components. **(B)** Heat map of pair-wise kinship matrix.

using the annotated wheat reference sequence (Wheat Chinese Spring IWGSC RefSeq v2.1 genome assembly (2021)) and are presented in Table 4. *AX-94724456* associated with DH found to encode *F-box-like domain superfamily* (TraesCS5D03G0939000). Similarly, SNPs i.e. *AX-95107750, AX-94598412, AX-94598412, AX-95181791, AX-95181791, AX-94691261* associated with GFD and encodes *Lateral organ boundaries, LOB* (TraesCS1A03G0271300), *Thioredoxin-like superfamily* (TraesCS4A03G0014500), *Glutathione S-transferase* (TraesCS4A03G0014500), *RNA-binding domain superfamily* (TraesCS5A03G0929100), *UDP-glycosyltransferase family* (TraesCS5A03G0929200), *Serine/threonine-protein kinase* (TraesCS1A03G0779800), respectively. Another SNP (*AX-94978133*) associated with GNPS found to encode *Expansin* (TraesCS4D03G0700700). Also, one SNP i.e. *AX-94883693* associated with GWPS found to encode *Patatin* (TraesCS1D03G0131800). Two SNPs (*AX-94452759* and *AX-94796636*) associated with PH found to encode *DUF1618 domain* (TraesCS2D03G0036000) and *Protein kinase domain* (TraesCS5D03G0455800). Similarly, *AX-94473624* associated with GY found to encode *Exocyst complex component Exo70* (TraesCS1D03G0081100).

Discussion

Although the phenotype-based selection in conventional breeding has improved wheat yield for several decades, genotype-based strategies may further complement the varietal improvement programmes. Recent efforts to sequence the wheat genome could promote the rapid improvement of varieties through molecular breeding by using genetic resources. In wheat, many QTLs/MTAs have been identified for yield and

component traits, however additional genetic studies are warranted using different genetic materials, as we have not reached a saturation point (Singh et al., 2021). Due to the genetic complexity of the wheat genome, there is always possibility to identify novel genomic regions with different genetic materials. Understanding the genetic basis of complex traits such as grain yield and component traits through GWAS with a diverse panel of genotypes can significantly improve QTL mapping resolution compared to bi-parental populations-based QTL mapping. Using the genome-wide SNPs and multi-environment data, several significant SNPs were identified in this study. Furthermore, stable and co-localized MTAs were also identified.

The significant effect of environment and genotype-environment interactions (GEI) was observed in the expression of all the studied traits. Among all traits, GFD was the most environment-sensitive trait, whereas, GNPS was relatively the most stable with minimum environmental influence. The greater magnitude of the environment and GEI have also been reported in previous studies for the expression of yield and component traits in wheat (Eltaher et al., 2021; Malik et al., 2021; Said et al., 2022). The GWAS panel has been tested in diverse production conditions, as the magnitude of GEI is a key factor in the identification of environment-specific QTL(s) as well as stable QTL(s). The highest and lowest heritability was recorded for GNPS and GFD, respectively, the trend for the percent contribution of environmental variation for the expression of GNPS and GFD was also exactly similar. Generally, the grain-related traits (GNPS and GWPS) and PH showed a significant positive association with GY in all the environments and pooled mean. However, the association between DH and GFD is consistent and significant negative in
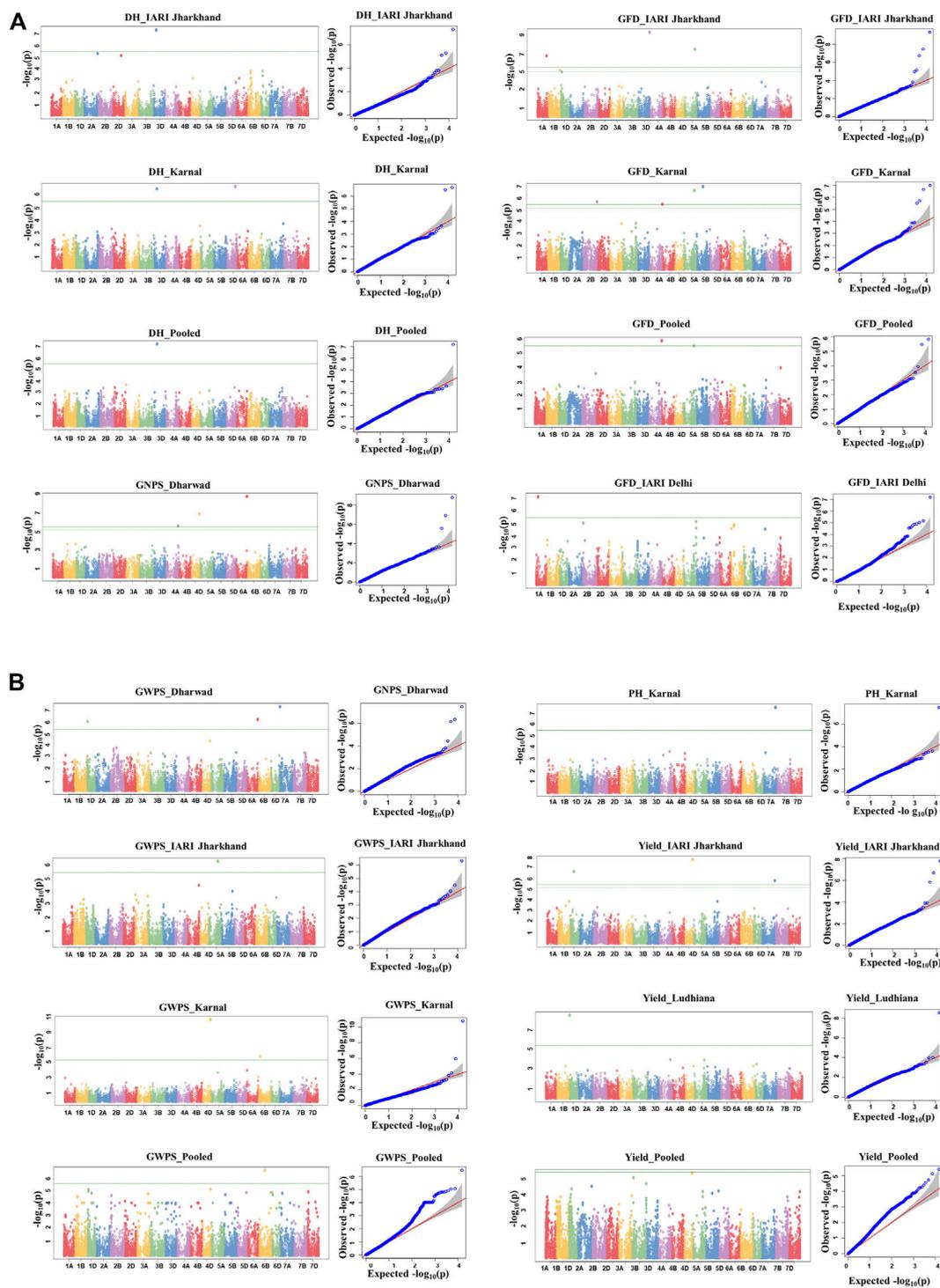
TABLE 3 Marker trait associations for grain yield and component traits in GWAS panel evaluated at Dharwad, IARI Delhi, IARI Jharkhand, Karnal, and Ludhiana during 2020−2021.

Trait	Environment	SNPs	Chr.	Position	p.value	PVE (%)
DH	E3	AX-95024590	3D	152,556,482	5.00E-08	21.1
	E4	AX-94724456	5D	484,426,977	2.04E-07	14.6
		AX-95024590	3D	152,556,482	3.17E-07	14.4
	Pooled	AX-95024590	3D	152,556,482	6.60E-08	17.3
GFD	E2	AX-95107750	1A	112,941,690	6.51E-08	9.1
	E3	AX-94598412	4A	5,068,136	4.67E-10	17.7
		AX-95181791	5A	584,672,964	3.44E-08	18.1
		AX-94691261	1A	505,524,324	1.82E-07	15.9
	E4	AX-94794189	5B	356,188,192	1.05E-07	18.2
		AX-95210025	5A	585,412,855	2.25E-07	19.0
		AX-94702510	2B	776,211,899	1.99E-06	19.9
		AX-94425015	4B	2,036,666	3.13E-06	20.4
	Pooled	AX-94425015	4B	2,036,666	1.47E-06	16.7
		AX-95210025	5A	585,412,855	3.37E-06	15.4
GNPS	E1	AX-94539354	6A	599,237,214	1.70E-09	16.3
		AX-94978133	4D	465,771,817	1.26E-07	11.1
		AX-94658573	4A	715,512,644	2.65E-06	10.1
	Pooled	AX-94539354	6A	599,237,214	2.26E-12	15.6
GWPS	E1	AX-94602474	7A	15,695,625	3.51E-08	11.4
		AX-94539354	6A	599,237,214	4.67E-07	17.1
		AX-94883693	1D	40,579,284	7.14E-07	14.0
	E3	AX-94469473	5A	521,025,733	4.76E-07	11.4
	E4	AX-94978133	4D	465,771,817	1.62E-11	12.5
		AX-95105308	6B	113,453,322	1.09E-06	6.6
	Pooled	AX-94387482	6B	337,941,809	3.39E-07	11.0
PH	E4	AX-94452759	2D	8,890,070	5.04E-08	5.4
		AX-94498579	3B	82,448,998	1.51E-07	5.0
		AX-94796636	5D	294,585,572	2.23E-06	7.2
GY	E3	AX-94978133	4D	465,771,817	1.66E-08	9.3
		AX-94483483	1D	206,101,173	2.09E-07	8.1
		AX-94709904	7A	728,244,062	1.49E-06	10.8
	E5	AX-94473624	1D	19,313,725	2.74E-09	12.5
	Pooled	AX-94978133	4D	465,771,817	3.87E-06	15.4

DH: days to heading (days); GFD: grain filling duration (days); GNPS: grain number per spike (number); GWPS: grain weight per spike (gm); PH: plant height (cm); GY: grain yield (gm); E1: Dharwad; E2: IARI, Delhi; E3: IARI, Jharkhand; E4: Karnal; E5: Ludhiana; SNPs: single nucleotide polymorphisms; PVE%: percent phenotypic variation explained.Gene and marker names were italicized.

all the environments and pooled mean. The strong positive association of GY with GNPS, GWPS, and PH was further supported by the identification of two co-localized MTAs (*AX-94978133* and *AX-94539354*), associated with the same traits i.e. GNPS, GWPS, and GY. Significant correlations found in this study have also been reported in earlier studies (Juliana et al., 2018; Baye et al., 2020; Ullah et al., 2021). In crop improvement programmes, significant associations of yield and

FIGURE 4
(A) Manhattan and respective-QQ plots for days to heading, grain number per spike, and grain filling duration in GWAS panel phenotyped at Dharwad, IARI Delhi, IARI Jharkhand, Karnal, and Ludhiana during 2020−2021. **(B)** Manhattan and respective-QQ plots for grain weight per spike, plant height, and yield in GWAS panel phenotyped at Dharwad, IARI Delhi, IARI Jharkhand, Karnal, and Ludhiana during 2020−2021.

TABLE 4 Putative candidate genes for grain yield and component traits.

Trait	SNP ID	Position	TransID	Putative candidate genes	Function
DH	AX-94724456	5D:487240908..487243412 (+strand)	TraesCS5D03G0939000	*F-box-like domain superfamily*	Heat Stress tolerance in wheat (Li et al., 2018). Regulate root growth and abiotic stress tolerance in rice (Yan et al., 2011)
GFD	AX-95107750	1A:115131649..115133537 (-strand)	TraesCS1A03G0271300	*Lateral organ boundaries, LOB*	Floral organs development in Arabidopsis (Shuai et al., 2002). Stress tolerance in potato (Liu et al., 2019a) and wheat (Wang et al., 2021)
	AX-94598412	4A:5076587..5077492 (-strand)	TraesCS4A03G0014500	*Thioredoxin-like superfamily*	Seed germination (Guo et al., 2013) and disease resistance in wheat (Shi et al., 2021)
	AX-94598412	4A:5076587..5077492 (-strand)	TraesCS4A03G0014500	*Glutathione S-transferase*	Growth and development, salt and drought stress tolerance in wheat (Wang et al., 2019)
	AX-95181791	5A:586571409..586575276 (+strand)	TraesCS5A03G0929100	*RNA-binding domain superfamily*	Extend grain filling duration and improve malt barley agronomic performance (Alptekin et al., 2021)
	AX-95181791	5A:586576183..586579881 (+strand)	TraesCS5A03G0929200	*UDP-glycosyltransferase family*	Regulation of grain size and abiotic stress tolerance in Rice (Dong et al., 2020)
	AX-94691261	1A:507053551..507056672 (+strand)	TraesCS1A03G0779800	*Serine/threonine-protein kinase, active site*	High thousand kernel weight and grains per spike in wheat (Ur Rehman et al., 2019)
GNPS	AX-94978133	4D:465931660..465932025 (-strand)	TraesCS4D03G0700700	*Expansin*	Capsule number in tobacco (Chen et al., 2016). Grain size in wheat (Lizana et al., 2010; Calderini et al., 2020)
GWPS	AX-94883693	1D:43497907..43499982 (+strand)	TraesCS1D03G0131800	*Patatin*	Seed size in Arabidopsis (Huang et al., 2001)
GY	AX-94473624	1D:20838814..20851719 (+strand)	TraesCS1D03G0081100	*Exocyst complex component Exo70*	Tissue-specific expression in wheat for biotic and abiotic stress (Zhao et al., 2018). Role in seed development in soybean (Wang et al., 2016). Pollen development in Arabidopsis (Markovic et al., 2020)
PH	AX-94452759	2D:9070785..9073543 (+strand)	TraesCS2D03G0036000	*DUF1618 domain*	Role in development and fitness in rice (Wang et al., 2014)
	AX-94796636	Chr5D: 297449170..297458745 (+strand)	TraesCS5D03G0455800	*Protein kinase domain*	OstMAPKKK5 regulates plant height and yield in rice (Liu et al., 2019)

DH: days to heading (days); GFD: grain filling duration (days); GNPS: grain number per spike (number); GWPS: grain weight per spike (gm); PH: plant height (cm); GY: grain yield (gm).Gene and marker names were italicized.

component traits in desired direction are always beneficial for simultaneous improvement of the associated traits. Furthermore, for negatively associated traits, it is advised to adopt breeding methods that could break the undesirable linkages, so that the traits can be independently improved.

PCA result of the filtered SNP data showed the allele frequencies of the genotypes were evenly distributed without any distinct sub-populations in the GWAS panel. The even distribution of allele frequencies in the GWAS panel was achieved by carefully selecting advanced breeding lines for different wheat growing zones in India, representing five agro-ecological zones, namely the Northern Hills Zone and North Western-Plains, North-Eastern Plains Zone, Central Zone, and Peninsular Zone. The various factors including size of the population, genetic drift, admixtures, selection, mutation, non-random mating, pollination behavior,

and recombination frequency may affect the LD, therefore, LD may vary in different populations (Gupta et al., 2005; Vos et al., 2017). Self-pollinated crops like wheat usually have larger LD blocks and hence decay slowly (Yu et al., 2014), whereas, LD decays rapidly in outcross crop species such as maize (Dinesh et al., 2016). The presence of high LD across the genome would reduce the QTL mapping resolution and vice versa (Dadshani et al., 2021). Under such situations, a better QTL resolution may be achieved by using genome-wide SNPs. The LD decay was found to be high and comparable in the B and D subgenomes (~5 cM) compared to the A subgenome, which had a shorter decay distance of around ~3 cM. A similar pattern of LD decay was also observed in other GWAS studies in wheat (Sukumaran et al., 2015; Rahimi et al., 2019; Sheoran et al., 2019). Therefore, marker density and population size are two

important determinants in GWAS studies and vary in self and cross-pollinated crops due to varied LD decay.

A total of 27 Bonferroni-corrected MTAs were identified for GWPS (7), GNPS (3), GY (4), DH (2), GFD (8), and PH (3). The highest number of MTAs were identified for A subgenome (11) followed by the D subgenome (10) and the B subgenome (6). A similar trend on MTAs identified in the A subgenome for yield and yield-contributing traits (Ain et al., 2015; Godoy et al., 2018; Alemu et al., 2021). A high level of stringency through Bonferroni-correction has been followed to consider MTA as significant, therefore, these MTAs could be valuable for their further validation in different genetic backgrounds to use them in MAS.

The identified MTAs (7) for GWPS on chromosomes 1D, 4D, 5A, 6A, 6B, and 7A in this study were novel as the earlier reported MTAs on the same chromosomes were identified at different positions. Although many grain-related traits like thousand kernel weight have been thoroughly studied, GWPS is comparatively less explored. Edae et al. (2014) identified an MTA on the 1D chromosome located at 88.5 cM and on the 7A chromosome located at 107.1 cM. A total of three significant MTAs were identified for another grain-related trait i.e. GNPS on 4A (599.2 Mb), 4D (715.5 Mb), and 6A (465.7 Mb) chromosomes. MTAs for GNPS in the different chromosomes were identified in different GWAS panels in previous experiments (Edae et al., 2014; Sun et al., 2017; Godoy et al., 2018; Jamil et al., 2019). However, Russell et al. (2020) identified an MTA in the same chromosome of 4D at 479.5 Mb, which was similar to that of AX-94978133 located on the 4D chromosome and mapped at 465.7 Mb, which explained 11.1% of phenotypic variation. All the identified MTAs explained more than 10% PVE for both the traits (GNPS and GWPS) except AX-95105308, which explained only 6.6% PVE.

The fundamental breeding objective of any wheat breeding program is the higher gains for GY, a highly complex and environmentally-sensitive economic trait. In the present study, a total of four significant MTAs were identified on 1D, 4D, and 7A chromosomes for GY. MTAs in the same chromosomes were also identified in different GWAS panels in previous studies on 1D (Bajgain et al., 2019; Jamil et al., 2019), 7A (Jamil et al., 2019; Ward et al., 2019; Russell et al., 2020) at different chromosomal locations. However, one MTA (S7A_720744946) on 7A chromosome was mapped at 720.7 Mb position, which is similar to the MTA identified in the present study i.e. (AX-94709904) on the 7A chromosome, which was mapped at 728.2 Mb. Furthermore, MTAs for GY were also identified in different chromosomes in different GWAS studies with diverse genetic material (Edae et al., 2014; Ain et al., 2015; Godoy et al., 2018; Li et al., 2019; Rahimi et al., 2019; Alemu et al., 2021; Suliman et al., 2021).

GWAS of yield component traits including DH, PH, and GFD led to the detection of 13 genetic loci associated with these traits. Two MTAs were identified for DH on 3D (152.5 Mb) and 5D (484.4 Mb) chromosomes. Previous reports identified MTAs mostly on different chromosomes, for instance, 3B (Edae et al., 2014; Ain et al., 2015; Gahlaut et al., 2019; Russell et al., 2020), 2D (Jamil et al., 2019), 1A, 4A, 5A, and 6A (Godoy et al., 2018) chromosomes. Three MTAs for PH were identified on 2D (88.9 Mb), 3B (824.4 Mb), and 5D (294.5 Mb) chromosomes. Previous reports also identified MTAs on the same chromosomes 2D (Ward et al., 2019; Alemu et al., 2021), 3B (Edae et al., 2014; Gahlaut et al., 2019) and also on different chromosomes (Ain et al., 2015; Sun et al., 2017; Godoy et al., 2018; Li et al., 2019; Russell et al., 2020) for PH. Jamil et al. (2019) identified an MTA on the same chromosome 3B at 824.6 Mb, which was similar to that of AX-94498579 located on the 3B chromosome and mapped at 824.4 Mb. The rate of grain filling and the length of grain filling period are two important determinants of final grain yield under different production conditions. In the present study, the maximum number of MTAs (8) was identified for GFD on 1A, 2B, 4A, 4B, 5A, and 5B chromosomes. Previous studies also identified on same chromosomes i.e. 1A (Edae et al., 2014; Jamil et al., 2019), 2B (Jamil et al., 2019), 5B (Edae et al., 2014; Jamil et al., 2019; Alemu et al., 2021) but different positions and different chromosomes (Godoy et al., 2018; Rahimi et al., 2019) for GFD.

A total of two co-localized MTA were identified, which are associated with multiple traits including GNPS, GWPS, GY, and DH. One co-localized MTA (AX-94978133) was identified on 4D associated with three traits (GNPS, GWPS, and GY). This MTA encodes expansin genes found to have a key role in wheat grain growth dynamics in wheat (Lizana et al., 2010; Calderini et al. (2020), increasing capsule number in tobacco (Chen et al., 2016). The other co-localized MTA (AX-94539354) identified on 6A was associated with two traits (GNPS and GWPS). Pleiotropic MTAs that are associated with multiple traits were also identified for grain yield and the biological yield on 1A, 4B, and 6B (Ain et al., 2015). Similarly, Co-localized QTLs associated with yield and component traits were detected (Alemu et al., 2021; Bennani et al., 2022). Such co-mapped SNPs will be much useful in marker-assisted selection for simultaneous improvement of correlated traits. Similarly, five consistent MTAs were also identified for grain yield and component traits in the present study. These co-located and stable MTAs will be suitable candidates for further validation and utilization in MAS-based varietal improvement programmes.

The various putative candidate genes underlying MTAs with high phenotypic variation for DH, GFD, GNPS, and GWPS were identified through BLAST search (Table 4). The MTAs identified in various chromosomes were located in gene coding regions related to transcription factors, a transmembrane protein, and kinase-like superfamilies. For example, AX-94978133 associated with GNPS encodes

expansin (TraesCS4D03G0700700) genes found to have a role in wheat grain growth dynamics including grain size (Lizana et al., 2010). Calderini et al. (2020) demonstrated that the targeted over-expression of an α-expansin in early developing wheat seeds leads to a significant increase in grain size without a negative effect on grain number, resulting in a yield boost under field conditions. Similarly, constitutive expression of *TaEXPA2*, an α-expansin gene in tobacco improved seed production by increasing capsule number without having any effect on plant growth patterns (Chen et al., 2016).

One SNP i.e. *AX-95181791* associated with GFD encodes an important RNA-binding domain superfamily (TraesCS5A03G0929100) that extends grain filling duration in barley. Glycine-rich RNA-binding protein (*HvGR-RBP1*) and a NAC transcription factor (*HvNAM1*) extend grain filling duration and improve agronomic performance in malt barley (Alptekin et al., 2021). Similarly, *AX-94691261* was associated with GFD encodes *Serine/threonine-protein kinase* (TraesCS1A03G0779800). The role of wheat protein kinase gene i.e. *TaSnRK2.9-5A* was studied and found to be significantly associated with high thousand kernel weight, whereas, *Hap-5A-4* was associated with high grains per spike (Ur Rehman et al., 2019). Another SNP (*AX-95181791*) for GFD encoding *UDP-glycosyltransferase family* (TraesCS5A03G0929200) has also been identified. The role of *UDP-glucosyltransferase* studied by Dong et al. (2020) suggests that *UDP-glucosyltransferase* regulates grain size and abiotic stress tolerance in rice. One MTA (*AX-95107750*) on 1A associated with GFD which encodes *Lateral organ boundaries, LOB* (TraesCS1A03G0271300) has a role in floral organs development in Arabidopsis (Shuai et al., 2002). One MTA (AX-94883693) on a 1D chromosome associated with GWPS encodes *Patatin* (TraesCS1D03G0131800). The role of *Patatin* was studied by Huang et al. (2001) and found its role in seed size in Arabidopsis. One SNP (*AX-94473624*) on the 1D chromosome associated with grain yield encodes *Exocyst complex component Exo70* (TraesCS1D03G0081100) has been found to have a role in plant growth and development including tissue-specific expression in wheat for biotic and abiotic stress (Zhao et al., 2018), seed development in soybean (Wang et al., 2016), and pollen development in Arabidopsis (Markovic et al., 2020). Similarly, two SNPs (*AX-94452759* and *AX-94796636*) on 2D and 5D chromosomes associated with plant height encodes the *DUF1618 domain* (TraesCS2D03G0036000) and the *protein kinase domain* (TraesCS5D03G0455800) are involved in various plant developmental processes. *DUF1618 domain* has been found to have a role in the development and fitness of rice (Wang et al., 2019), and *protein kinase domain* regulates plant height and yield in rice (Liu Y. et al., 2019).

Conclusion

The study with 280 diverse set of bread wheat GWAS panel has shown that DH, GFD, GNPS, GWPS, PH, and GY were quantitatively inherited traits. The strong positive correlation between GY and GNPS, GWPS, and PH suggested the possibility of improving these traits simultaneously. A total of 27 MTAs including 7 for GWPS, 3 for GNPS, 4 for GY, 2 for DH, 8 for GFD, and 3 for PH were identified through the GWAS approach. A total of five stable MTAs were identified in more than one environment and associated with the expression of DH, GFD, GNPS, and GY. Also, two novel pleiotropic genomic regions harboring co-localized QTLs governing two or more traits were also identified. The environment-specific and pooled-data MTAs identified in the present investigation represented novel genomic regions associated with trait expression. Several putative candidate genes encoding important molecular functions such as regulation of grain size, grain number, growth and development, grain filling duration, and abiotic stress tolerance were identified. Further validation and functional characterization of the candidate genes to elucidate the role of these genes in wheat is envisaged. The identified SNPs, particularly stable (*AX-95024590, AX-94425015, AX-95210025, AX-94539354, AX-94978133*) and pleiotropic SNPs (*AX-94978133* and *AX-94539354*) could be useful in marker-assisted selection programs to develop wheat varieties with increased grain yield.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://datadryad. org/stash/share/i2XTnmGIHO6F-LIlTjUhN1krZWActEJxWcZIbBSHAGo.

Author contributions

HKh, GS, and GK conceptualized the investigation and drafted the original manuscript; HKh, GK, HKr, SK, CM, OP, PS, SB, MK, and SY conducted the field experiments; HKr, HKh generation of genotypic data; ND, NR did the phenotypic and GWAS analysis. All authors contributed to the article editing and approved the submitted version.

Funding

foundation (BMGF) under the project ICAR-BMGF (Project Code:1011099).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene. 2022.982589/full#supplementary-material

SUPPLEMENTARY TABLE 1
List of 280 genotypes used in the study along with pedigree and other details are given as Supplementary Table S1.

SUPPLEMENTARY TABLE 2
Genotyping and phenotyping information of the 280 genotypes used in the GWAS are given as Supplementary Table S2 and also deposited in DRYAD (an international open-access repository of research data). The international open-access data repository link is as follows: https://datadryad. org/stash/share/i2XTnmGlHO6F-LIlTjUhN1krZWActEJxWcZlbBSHAGo.

References

Ain, Q. U., Rasheed, A., Anwar, A., Mahmood, T., Imtiaz, M., Mahmood, T., et al. (2015). Genome-wide association for grain yield under rainfed conditions in historical wheat cultivars from Pakistan. *Front. Plant Sci.* 6, 743. doi:10.3389/fpls.2015.00743

Alemu, A., Suliman, S., Hagras, A., Thabet, S., Al-Abdallat, A., Abdelmula, A. A., et al. (2021). Multi-model genome-wide association and genomic prediction analysis of 16 agronomic, physiological and quality related traits in ICARDA spring wheat. *Euphytica* 217, 205. doi:10.1007/s10681-021-02933-6

Ali, M., Zhang, Y., Rasheed, A., Wang, J., and Zhang, L. (2020). Genomic prediction for grain yield and yield-related traits in Chinese winter wheat. *Int. J. Mol. Sci.* 21 (4), 1342. doi:10.3390/ijms21041342

Alptekin, B., Mangel, D., Pauli, D., Blake, T., Lachowiec, J., Hoogland, T., et al. (2021). Combined effects of a glycine-rich RNA-binding protein and a NAC transcription factor extend grain fill duration and improve malt barley agronomic performance. *Theor. Appl. Genet.* 134 (1), 351–366. doi:10.1007/s00122-020-03701-1

Alqudah, A. M., Haile, J. K., Alomari, D. Z., Pozniak, C. J., Kobiljski, B., Borner, B., et al. (2020). Genome-wide and SNP network analyses reveal genetic control of spikelet sterility and yield-related traits in wheat. *Sci. Rep.* 10, 2098. doi:10.1038/s41598-020-59004-4

Aravind, J., MukeshSankar, S., Wankhede, D. P., and Kaur, V. (2021). *augmentedRCBD: Analysis of augmented randomised complete block designs*. R package version 0.1.5.9000, https://aravind-j.github.io/augmentedRCBD https://cran.rproject.org/package=augmentedRCBD (Accessed May 25, 2022).

Bajgain, P., Zhang, X., and Anderson, J. A. (2019). Genome-wide association study of yield component traits in intermediate wheatgrass and implications in genomic selection and breeding. *G3* 9, 2429–2439. doi:10.1534/g3.119.400073

Baye, A., Berihun, B., Bantayehu, M., and Derebe, B. (2020). Genotypic and phenotypic correlation and path coefficient analysis for yield and yield-related traits in advanced bread wheat (*Triticum aestivum* L.) lines. *Cogent Food Agric.* 6 (1), 1752603. doi:10.1080/23311932.2020.1752603

Bennani, S., Birouk, A., Jlibene, M., Sanchez-Garcia, M., Nsarellah, N., Gaboun, F., et al. (2022). Drought-tolerance QTLs associated with grain yield and related traits in spring bread wheat. *Plants* 11, 986. doi:10.3390/plants11070986

Brachi, B., Morris, G. P., and Borevitz, J. O. (2011). Genome-wide association studies in plants: The missing heritability is in the field. *Genome Biol.* 12, 232. doi:10.1186/gb-2011-12-10-232

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). Tassel: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23 (19), 2633–2635. doi:10.1093/bioinformatics/btm308

Braun, H. J., Atlin, G., and Payne, T. (2010). *Multi-location testing as a tool to identify plant response to global climate change. climate change and crop production.* (Wallingford, UK: CABI), 115–138.

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G. L. A., D'Amore, R., Allen, A. M., et al. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* 491, 705–710. doi:10.1038/nature11650

Calderini, D. F., Castillo, F. M., Arenas-M, A., Molero, G., Reynolds, M. P., Craze, M., et al. (2020). Overcoming the trade-off between grain weight and number in wheat by the ectopic expression of expansin in developing seeds leads to increased yield potential. *New Phytol.* 230, 629–640. doi:10.1111/nph.17048

Chapman, J. A., Mascher, M., Buluç, A., Barry, K., Georganas, E., Session, A., et al. (2015). A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. *Genome Biol.* 16, 26. doi:10.1186/s13059-015-0582-8

Chen, X., Min, D., Yasir, T. A., and Hu, Y. G. (2012). Evaluation of 14 morphological, yield-related, and physiological traits as indicators of drought tolerance in Chinese winter bread wheat revealed by analysis of the membership function value of drought tolerance (MFVD). *Field Crops Res.* 137, 195–201. doi:10.1016/j.fcr.2012.09.008

Chen, Y., Han, Y., Zhang, M., Zhou, S., Kong, X., and Wang, W. (2016). Overexpression of the wheat expansin gene TaEXPA2 improved seed production and drought tolerance in transgenic tobacco plants. *PLoS ONE* 11 (4), e0153494. doi:10.1371/journal.pone.0153494

Dadshani, S., Mathew, B., Ballvora, A., Mason, A. S., and Leon, J. (2021). Detection of breeding signatures in wheat using a linkage disequilibrium-corrected mapping approach. *Sci. Rep.* 11 (1), 5527. doi:10.1038/s41598-021-85226-1

Dinesh, A., Patil, A., Zaidi, P. H., Kuchanur, P. H., Vinayan, M. T., and Seetharam, K. (2016). Genetic diversity, linkage disequilibrium and population structure among CIMMYT maize inbred lines, selected for heat tolerance study. *Maydica* 61 (3), 1–7.

Dong, N. Q., Sun, Y., Guo, T., Shi, C. L., Zhang, Y. M., Kan, Y., et al. (2020). UDP-glucosyltransferase regulates grain size and abiotic stress tolerance associated with metabolic flux redirection in rice. *Nat. Commun.* 11 (1), 2629. doi:10.1038/s41467-020-16403-5

Edae, E. A., Byrne, P. F., Haley, S. D., Lopes, M. S., and Reynolds, M. P. (2014). Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. *Theor. Appl. Genet.* 127 (4), 791–807. doi:10.1007/s00122-013-2257-8

Eltaher, S., Baenziger, P. S., Belamkar, V., Emara, H. A., Nower, A. A., Salem, K. F. M., et al. (2021). GWAS revealed effect of genotype × environment interactions for grain yield of Nebraska winter wheat. *BMC Genomics* 22, 2. doi:10.1186/s12864-020-07308-0

FAO (2017). The future of food and agriculture—Trends and challenges. Rome. Available at: www.fao.org/3/a-i6583e.pdf (Accessed June 25, 2022).

Gahlaut, V., Jaiswal, V., Singh, S., Balyan, H. S., and Gupta, P. K. (2019). Multi-locus genome wide association mapping for yield and Its Contributing traits in hexaploid wheat under different water regimes. *Sci. Rep.* 9, 19486–19515. doi:10.1038/s41598-019-55520-0

Gao, F., Wen, W., Liu, J., Rasheed, A., Yin, G., Xia, X., et al. (2015). Genome-wide linkage mapping of QTL for yield components, plant height and yield-related physiological traits in the Chinese wheat cross zhou 8425B/Chinese spring. *Front. Plant Sci.* 6, 1099. doi:10.3389/fpls.2015.01099

Gao, L., Meng, C., Yi, T., Xu, K., Cao, H., Zhang, S., et al. (2021). Genome-wide association study reveals the genetic basis of yield- and quality-related traits in wheat. *BMC Plant Biol.* 21, 144. doi:10.1186/s12870-021-02925-7

Giraldo, P., Benavente, E., Manzano-Agugliaro, F., and Gimenez, E. (2019). Worldwide research trends on wheat and barley: A bibliometric comparative analysis. *Agronomy* 9, 352. doi:10.3390/agronomy9070352

Godoy, J., Gizaw, S., Chao, S., Blake, N., Carter, A., Cuthbert, R., et al. (2018). Genome-wide association study of agronomic traits in a spring-planted North American elite hard red spring wheat panel. *Crop Sci.* 58, 1838–1852. doi:10.2135/cropsci2017.07.0423

Guo, H., Wang, S., Xu, F., Li, Y., Ren, J., Wang, X., et al. (2013). The role of thioredoxin h in protein metabolism during wheat (*Triticum aestivum* L.) seed germination. *Plant physiol. biochem.* 67, 137–143. doi:10.1016/j.plaphy.2013.03.006

Gupta, P. K., Rustgi, S., and Kulwal, P. L. (2005). Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Mol. Biol.* 57 (4), 461–485. doi:10.1007/s11103-005-0257-z

Hu, P., Zheng, Q., Luo, Q., Teng, W., Li, H., Li, B., et al. (2021). Genome-wide association study of yield and related traits in common wheat under salt-stress conditions. *BMC Plant Biol.* 21, 27. doi:10.1186/s12870-020-02799-1

Huang, M., Liu, X., Zhou, Y., Summers, R. M., and Zhang, Z. (2019). Blink: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* 8 (2), giy154. doi:10.1093/gigascience/giy154

Huang, S., Cerny, R. E., Bhat, D. S., and Brown, S. M. (2001). Cloning of an Arabidopsis patatin-like gene, STURDY, by activation T-DNA tagging. *Plant Physiol.* 125 (2), 573–584. doi:10.1104/pp.125.2.573

Isham, K., Wang, R., Zhao, W., Wheeler, J., Klassen, N., Akhunov, E., et al. (2021). QTL mapping for grain yield and three yield components in a population derived from two high-yielding spring wheat cultivars. *Theor. Appl. Genet.* 134, 2079–2095. doi:10.1007/s00122-021-03806-1

Jamil, M., Ali, A., Gul, A., Ghafoor, A., Napar, A. A., Ibrahim, A. M., et al. (2019). Genome-wide association studies of seven agronomic traits under two sowing conditions in bread wheat. *BMC Plant Biol.* 19 (1), 149–218. doi:10.1186/s12870-019-1754-6

Jin, J., Liu, D., Qi, Y., Ma, J., and Zhen, W. (2020). Major QTL for seven yield-related traits in common wheat (*Triticum aestivum* L.). *Front. Genet.* 11, 1012. doi:10.3389/fgene.2020.01012

Juliana, P., Montesinos-Lopez, O. A., Crossa, J., Mondal, S., Perez, L. G., Poland, J., et al. (2018). Integrating genomic-enabled prediction and high-throughput phenotyping in breeding for climate-resilient bread wheat. *Theor. Appl. Genet.* 132, 177–194. doi:10.1007/s00122-018-3206-3

Kaler, A. S., Gillman, J. D., Beissinger, T., and Purcell, L. C. (2020). Comparing different statistical models and multiple testing corrections for association mapping in soybean and maize. *Front. Plant Sci.* 25 (10), 1794. doi:10.3389/fpls.2019.01794

Kang, C. S., Mo, Y. J., Kim, K. M., Kim, K. H., Chun, J. B., Park, C. S., et al. (2021). Mapping quantitative trait loci for yield potential traits in wheat recombinant inbred lines. *Agronomy* 11 (1), 22. doi:10.3390/agronomy11010022

Kaya, Y., and Akcura, M. (2014). Effects of genotype and environment on grain yield and quality traits in bread wheat (*T. aestivum* L.). *Food Sci. Technol.* 34, 386–393. doi:10.1590/fst.2014.0041

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* 9, 29. doi:10.1186/1746-4811-9-29

Krishnappa, G., Savadi, S., Tyagi, B. S., Singh, S. K., Mamrutha, H. M., Kumar, S., et al. (2021). Integrated genomic selection for rapid improvement of crops. *Genomics* 113 (3), 1070–1086. doi:10.1016/j.ygeno.2021.02.007

Li, F., Wen, W., Liu, J., Zhang, Y., Cao, S., He, Z., et al. (2019). Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biol.* 19, 168. doi:10.1186/s12870-019-1781-3

Li, Q., Wang, W., Wang, W., Zhang, G., Liu, Y., Wang, Y., et al. (2018). Wheat F-box protein gene TaFBA1 is involved in plant tolerance to heat stress. *Front. Plant Sci.* 9, 521. doi:10.3389/fpls.2018.00521

Li, T., Li, Q., Wang, J., Yang, Z., Tang, Y., Su, Y., et al. (2022). High-resolution detection of quantitative trait loci for seven important yield-related traits in wheat (*Triticum aestivum* L.) using a high-density SLAF-seq genetic map. *BMC Genom. Data* 23, 37. doi:10.1186/s12863-022-01050-0

Li, X., Xu, X., Liu, W., Li, X., Yang, X., Ru, Z., et al. (2020). Dissection of superior alleles for yield-related traits and their distribution in important

cultivars of wheat by association mapping. *Front. Plant Sci.* 11, 175. doi:10.3389/fpls.2020.00175

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). Gapit: Genome association and prediction integrated tool. *Bioinformatics* 28 (18), 2397–2399. doi:10.1093/bioinformatics/bts444

Liu, H., Cao, M., Chen, X., Ye, M., Zhao, P., Nan, Y., et al. (2019). Genome-wide analysis of the lateral organ boundaries domain (LBD) gene family in *Solanum tuberosum. Int. J. Mol. Sci.* 20 (21), 5360. doi:10.3390/ijms20215360

Liu, H., Searle, I. R., Mather, D. E., Able, A. J., and Able, J. A. (2015). Morphological, physiological and yield responses of durum wheat to pre-anthesis water-deficit stress are genotype-dependent. *Crop Pasture Sci.* 66, 1024–1038. doi:10.1071/CP15013

Liu, J., Wu, B., Singh, R. P., and Velu, G. (2019a). QTL mapping for micronutrients concentration and yield component traits in a hexaploid wheat mapping population. *J. Cereal Sci.* 88, 57–64. doi:10.1016/j.jcs.2019.05.008

Liu, J., Xu, Z., Fan, X., Zhou, Q., Cao, J., Wang, F., et al. (2018). A genome-wide association study of wheat spike related traits in China. *Front. Plant Sci.* 9, 1584. doi:10.3389/fpls.2018.01584

Liu, Y., Zhu, Y., Xu, X., Sun, F., Yang, J., Cao, L., et al. (2019b). OstMAPKKK5, a truncated mitogen-activated protein kinase kinase kinase 5, positively regulates plant height and yield in rice. *Crop J.* 7 (5), 707–714. doi:10.1016/j.cj.2019.03.001

Lizana, X. C., Riegel, R., Gomez, L. D., Herrera, J., Isla, A., McQueen-Mason, S. J., et al. (2010). Expansins expression is associated with grain size dynamics in wheat (*Triticum aestivum* L.). *J. Exp. Bot.* 61 (4), 1147–1157. doi:10.1093/jxb/erp380

Ma, F., Xu, Y., Ma, Z., Li, L., and An, D. (2018). Genome-wide association and validation of key loci for yield-related traits in wheat founder parent Xiaoyan 6. *Mol. Breed.* 38, 91. doi:10.1007/s11032-018-0837-7

Malik, P., Kumar, J., Singh, S., Sharma, S., Meher, P. K., Sharma, M. K., et al. (2021). Single-trait, multi-locus and multi-trait GWAS using four different models for yield traits in bread wheat. *Mol. Breed.* 41 (7), 46–21. doi:10.1007/s11032-021-01240-1

Markovic, V., Cvrckova, F., Potocky, M., Kulich, I., Pejchar, P., Kollarova, E., et al. (2020). EXO70A2 is critical for Exocyst complex function in pollen development. *Plant Physiol.* 184 (4), 1823–1839. doi:10.1104/pp.19.01340

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8 (19), 4321–4325. doi:10.1093/nar/8.19.4321

Pang, Y., Liu, C., Wang, D., St Amand, P., Bernardo, A., Li, W., et al. (2020). High-resolution genome-wide association study identifies genomic regions and candidate genes for important agronomic traits in wheat. *Mol. Plant* 13, 1311–1327. doi:10.1016/j.molp.2020.07.008

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38 (8), 904–909. doi:10.1038/ng1847

Rahimi, Y., Bihamta, M. R., Taleei, A., Alipour, H., and Ingvarsson, P. K. (2019). Genome-wide association study of agronomic traits in bread wheat reveals novel putative alleles for future breeding programs. *BMC Plant Biol.* 19, 541. doi:10.1186/s12870-019-2165-4

Rathan, N. D., Krishna, H., Ellur, R. K., Sehgal, D., Govindan, V., Ahlawat, A. K., et al. (2022). Genome-wide association study identifies loci and candidate genes for grain micronutrients and quality traits in wheat (*Triticum aestivum* L.). *Sci. Rep.* 12, 7037. doi:10.1038/s41598-022-10618-w

Reynolds, M. P., Pask, A. J. D., Hoppitt, W. J. E., Sonder, K., Sukumaran, S., Molero, G., et al. (2017). Strategic crossing of biomass and harvest index-source and sink-achieves genetic gains in wheat. *Euphytica* 213, 257. doi:10.1007/s10681-017-2040-z

Russell, B., Brown-Guedira, G., Sneller, C. H., and Mohammadi, M. (2020). Transferability of marker trait associations in wheat is disturbed mainly by genotype × year interaction. *Crop Breed. Genet. Genom* 2 (3), e200013. doi:10.20900/cbgg20200013

Said, A. A., MacQueen, A. H., Shawky, H., Reynolds, M., Juenger, T. E., and El-Soda, M. (2022). Genome-wide association mapping of genotype-environment interactions affecting yield-related traits of spring wheat grown in three watering regimes. *Environ. Exp. Bot.* 194, 104740. doi:10.1016/j.envexpbot.2021.104740

Saïdou, A. A., Thuillet, A. C., Couderc, M., Mariac, C., and Vigouroux, Y. (2014). Association studies including genotype by environment interactions: Prospects and limits. *BMC Genet.* 15 (1), 3–12. doi:10.1186/1471-2156-15-3

Saini, D. K., Chopra, Y., Singh, J., Sandhu, K. S., Kumar, A., Bazzer, S., et al. (2022). Comprehensive evaluation of mapping complex traits in wheat using

genome-wide association studies. *Mol. Breed.* 42, 1. doi:10.1007/s11032-021-01272-7

Sandhu, K. S., Mihalyov, P. D., Lewien, M. J., Pumphrey, M. O., and Carter, A. H. (2021). Genomic selection and genome-wide association studies for grain protein content stability in a nested association mapping population of wheat. *Agronomy* 11(12), 2528. doi:10.3390/agronomy11122528

Sheoran, S., Jaiswal, S., Kumar, D., Raghav, N., Sharma, R., Pawar, S., et al. (2019). Uncovering genomic regions associated with 36 agro-morphological traits in Indian spring wheat using GWAS. *Front. Plant Sci.* 10, 527. doi:10.3389/fpls.2019.00527

Shi, B., Zhao, X., Li, M., Dong, Z., Yang, Q., Wang, Y., et al. (2021). Wheat thioredoxin (TaTrxh1) associates with RD19-like cysteine protease TaCP1 to defend against stripe rust fungus through modulation of programmed cell death. *Mol. Plant. Microbe. Interact.* 34 (4), 426–438. doi:10.1094/MPMI-11-20-0304-R

Shuai, B., Reynaga-Pena, C. G., and Springer, P. S. (2002). The lateral organ boundaries gene defines a novel, plant-specific gene family. *Plant Physiol.* 129 (2), 747–761. doi:10.1104/pp.010926

Singh, K., Batra, R., Sharma, S., Saripalli, G., Gautam, T., Singh, R., et al. (2021). WheatQTLdb: A QTL database for wheat. *Mol. Genet. Genomics.* 296 (5), 1051–1056. doi:10.1007/s00438-021-01796-9

Sukumaran, S., Dreisigacker, S., Lopes, M., Chavez, P., and Reynolds, M. P. (2015). Genome wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theor. Appl. Genet.* 128 (2), 353–363. doi:10.1007/s00122-014-2435-3

Sukumaran, S., Reynolds, M. P., and Sansaloni, C. (2018). Genome-wide association analyses identify QTL hotspots for yield and component traits in durum wheat grown under yield potential, drought, and heat stress environments. *Front. Plant Sci.* 9, 81. doi:10.3389/fpls.2018.00081

Suliman, S., Alemu, A., Abdelmula, A. A., Badawi, G. H., Al-Abdallat, A., and Tadesse, W. (2021). Genome-wide association analysis uncovers stable QTLs for yield and quality traits of spring bread wheat (*Triticum aestivum*) across contrasting environments. *Plant gene.* 25, 100269. doi:10.1016/j.plgene.2020.100269

Sun, C., Zhang, F., Yan, X., Zhang, X., Dong, Z., Cui, D., et al. (2017). Genome wide association study for 13 agronomic traits reveals distribution of superior alleles in bread wheat from the Yellow and Huai Valley of China. *Plant Biotechnol. J.* 15 (8), 953–969. doi:10.1111/pbi.12690

Ullah, M. I., Mahpara, S., Bibi, R., Shah, R. U., Ullah, R., Abbas, S., et al. (2021). Grain yield and correlated traits of bread wheat lines: Implications for yield improvement. *Saudi J. Biol. Sci.* 28 (10), 5714–5719. doi:10.1016/j.sjbs.2021.06.006

Ur Rehman, S., Wang, J., Chang, X., Zhang, X., Mao, X., and Jing, R. (2019). A wheat protein kinase gene TaSnRK2.9-5A associated with yield contributing traits. *Theor. Appl. Genet.* 132 (4), 907–919. doi:10.1007/s00122-018-3247-7

Vikas, V. K., Pradhan, A. K., Budhlakoti, N., Mishra, D. C., Chandra, T., Bhardwaj, S. C., et al. (2022). Multi-locus genome-wide association studies (ML-GWAS) reveal novel genomic regions associated with seedling and adult plant stage leaf rust resistance in bread wheat (*Triticum aestivum* L.). *Heredity* 128, 434–449. doi:10.1038/s41437-022-00525-1

Vos, P. G., Paulo, M. J., Voorrips, R. E., Visser, R. G., van Eck, H. J., and van Eeuwijk, F. A. (2017). Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. *Theor. Appl. Genet.* 130 (1), 123–135. doi:10.1007/s00122-016-2798-8

Wang, J., and Zhang, Z. (2021). GAPIT version 3: Boosting power and accuracy for genomic association and prediction. *Genomics Proteomics Bioinforma.* 19 (4), 629–640. doi:10.1016/j.gpb.2021.08.005

Wang, L., Shen, R., Chen, L. T., and Liu, Y. G. (2014). Characterization of a novel DUF1618 gene family in rice. *J. Integr. Plant Biol.* 56 (2), 151–158. doi:10.1111/jipb.12130

Wang, R., Ma, J., Zhang, Q., Wu, C., Zhao, H., Wu, Y., et al. (2019). Genome-wide identification and expression profiling of glutathione transferase gene family under multiple stresses and hormone treatments in wheat (*Triticum aestivum* L.). *BMC genomics* 20 (1), 986–1015. doi:10.1186/s12864-019-6374-x

Wang, S. X., Zhu, Y. L., Zhang, D. X., Shao, H., Liu, P., Hu, J. B., et al. (2017). Genome-wide association study for grain yield and related traits in elite wheat varieties and advanced lines using SNP markers. *PloS ONE* 12 (11), e0188662. doi:10.1371/journal.pone.0188662

Wang, Z., Li, P., Yang, Y., Chi, Y., Fan, B., and Chen, Z. (2016). Expression and functional analysis of a novel group of legume-specific WRKY and Exo70 protein variants from soybean. *Sci. Rep.* 6, 32090. doi:10.1038/srep32090

Wang, Z., Zhang, R., Cheng, Y., Lei, P., Song, W., Zheng, W., et al. (2021). Genome-wide identification, evolution, and expression analysis of LBD transcription factor family in bread wheat (*Triticum aestivum* L.). *Front. Plant Sci.* 12, 721253. doi:10.3389/fpls.2021.721253

Ward, B. P., Brown-Guedira, G., Kolb, F. L., Van Sanford, D. A., Tyagi, P., Sneller, C. H., et al. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PLoS ONE* 14 (2), e0208217. doi:10.1371/journal.pone.0208217

Wheat Initiative (2013). An international vision for wheat improvement. Available at: https://www.wheatinitiative.org/vision-paper-1 (Accessed 25 June, 2022).

Yan, Y. S., Chen, X. Y., Yang, K., Sun, Z. X., Fu, Y. P., Zhang, Y. M., et al. (2011). Overexpression of an F-box protein gene reduces abiotic stress tolerance and promotes root growth in rice. *Mol. Plant* 4 (1), 190–197. doi:10.1093/mp/ssq066

Yu, H., Deng, Z., Xiang, C., and Tian, J. (2014). Analysis of diversity and linkage disequilibrium mapping of agronomic traits on B-genome of wheat. *J. Genomics* 2, 20–30. doi:10.7150/jgen.4089

Zhang, H., Chen, J., Li, R., Deng, Z., Zhang, K., Liu, B., et al. (2016). Conditional QTL mapping of three yield components in common wheat (*Triticum aestivum* L.). *Crop J.* 4 (3), 220–228. doi:10.1016/j.cj.2016.01.007

Zhang, J., Yao, Q., Li, R., Lu, Y., Zhou, S., Han, H., et al. (2022). Identification of genetic loci on chromosome 4B for improving the grain number per spike in pre-breeding lines of wheat. *Agronomy* 12, 171. doi:10.3390/agronomy12010171

Zhang, Y. M., Jia, Z., and Dunwell, J. M. (2019). Editorial: The applications of New multi-locus GWAS methodologies in the genetic dissection of complex traits. *Front. Plant Sci.* 10, 100. doi:10.3389/fpls.2019.00100

Zhao, J., Zhang, X., Wan, W., Zhang, H., Liu, J., Li, M., et al. (2018). Identification and characterization of the *EXO70* gene family in polyploid wheat and related species. *Int. J. Mol. Sci.* 20 (1), 60. doi:10.3390/ijms20010060