

## ORIGINAL ARTICLE

# Genome-wide association analysis of spike and kernel traits in the U.S. hard winter wheat

Jyotirmoy Halder<sup>1</sup> | Harsimardeep S. Gill<sup>1</sup> | Jinfeng Zhang<sup>1</sup> | Rami Altameemi<sup>1</sup> | Eric Olson<sup>2</sup> | Brent Turnipseed<sup>1</sup> | Sunish K. Sehgal<sup>1</sup> 

<sup>1</sup>Dep. of Agronomy, Horticulture & Plant Science, South Dakota State Univ., Brookings, SD 57007, USA

<sup>2</sup>Dep. of Plant, Soil and Microbial Sciences, Michigan State Univ., East Lansing, MI 48824, USA

## Correspondence

Sunish K. Sehgal, Dep. of Agronomy, Horticulture & Plant Science, South Dakota State Univ., Brookings, SD 57007, USA.  
Email: [sunish.sehgal@sdstate.edu](mailto:sunish.sehgal@sdstate.edu)

Assigned to Associate Editor Alexander Lipka.

## Funding information

National Institute of Food and Agriculture, Grant/Award Number: 2022-68013-36439; South Dakota Wheat Commission, Grant/Award Number: 3X1340; U.S. Department of Agriculture, Grant/Award Number: SD00H695-20

## Abstract

A better understanding of the genetic control of spike and kernel traits that have higher heritability can help in the development of high-yielding wheat varieties. Here, we identified the marker-trait associations (MTAs) for various spike- and kernel-related traits in winter wheat (*Triticum aestivum* L.) through genome-wide association studies (GWAS). An association mapping panel comprising 297 hard winter wheat accessions from the U.S. Great Plains was evaluated for eight spike- and kernel-related traits in three different environments. A GWAS using 15,590 single-nucleotide polymorphisms (SNPs) identified a total of 53 MTAs for seven spike- and kernel-related traits, where the highest number of MTAs were identified for spike length (16) followed by the number of spikelets per spike (15) and spikelet density (11). Out of 53 MTAs, 14 were considered to represent stable quantitative trait loci (QTL) as they were identified in multiple environments. Five multi-trait MTAs were identified for various traits including the number of spikelets per spike (NSPS), spikelet density (SD), kernel width (KW), and kernel area (KA) that could facilitate the pyramiding of yield-contributing traits. Further, a significant additive effect of accumulated favorable alleles on the phenotype of four spike-related traits suggested that breeding lines and cultivars with a higher number of favorable alleles could be a valuable resource for breeders to improve yield-related traits. This study improves the understanding of the genetic basis of yield-related traits in hard winter wheat and provides reliable molecular markers that will facilitate marker-assisted selection (MAS) in wheat breeding programs.

**Abbreviations:** BLUE, best linear unbiased estimate; CEnv, combined environment; FarmCPU, fixed and random model circulating probability unification; GWAS, genome-wide association studies; HWWAMP, hard winter wheat association mapping panel; KA, kernel area; KL, kernel length; KW, kernel width; LD, linkage disequilibrium; MAS, marker-assisted selection; MTA, marker-trait association; NKPS, number of kernels per spike; NSPS, number of spikelets per spike; SD, spikelet density; SL, spike length; SNP, single-nucleotide polymorphism; TKW, 1,000-kernel weight.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *The Plant Genome* published by Wiley Periodicals LLC on behalf of Crop Science Society of America.

## 1 | INTRODUCTION

Bread wheat (*Triticum aestivum* L.) is one of the most important and widely grown food crops, playing a crucial role in feeding 35% of the world's population by providing 19% of calories and 21% of proteins (Tadesse et al., 2019). To meet the growing demand, global food production needs to double by 2050; however, the average annual yield increase of wheat has been reported to be 1%, which is far less than the required rate (1.7%) (Pal et al., 2022; Tadesse et al., 2019). Therefore, to feed an ever-increasing world population with a gradual decrease in farmland, yield improvement remains the primary focus for wheat breeding programs globally.

Wheat yield is a complex phenomenon influenced by various factors, such as morphological characteristics, physiological indices, grain-related traits, and different environmental conditions, making yield enhancement increasingly challenging (Liu et al., 2018b; Nadolska-Orczyk et al., 2017). However, grain yield can be dissected into different component traits such as the number of spikelets per spike (NSPS), spike length (SL), spike number per unit area, number of kernels per spike (NKPS), kernel size, and 1,000-kernel weight (TKW), which are less sensitive to the environment and have higher heritability than that of grain yield itself (Hai et al., 2008; Kato et al., 2000). Three major components that collectively determine the final yield of wheat are the number of spikes per unit area, NKPS, and TKW (Liu et al., 2018b). In addition, other spike-related traits, like spike number per plant, SL, and spikelet number, also play a significant role in wheat yield improvement (Guo et al., 2017). Therefore, the identification and further deployment of genomic regions that control different yield-related traits are essential in dissecting the genetic basis of yield and the overall improvement of wheat.

Wheat yield and yield-contributing traits exhibit polygenic inheritance and are usually controlled by multiple genes and quantitative trait loci (QTL) (Anuarbek et al., 2020; Jin et al., 2020). Linkage mapping and genome-wide association studies (GWAS) have been intensively used to dissect the genetic basis of complex traits (Liu et al., 2018a). Further, the effectiveness of GWAS has already been established to capture genetic factors for various traits in wheat including agronomic (Sukumaran et al., 2015; Sun et al., 2017), disease resistance (Arruda et al., 2016; Halder et al., 2019; Juliana et al., 2018; Zhang et al., 2022), and end-use quality (Chen et al., 2019). Likewise, these approaches have been used to identify significant QTL for several yield-related traits in wheat (Singh et al., 2022). Quantitative trait loci and genes associated with spike- (Cui et al., 2014; Gao et al., 2015; Guo et al., 2017; Huang et al., 2006; Liu et al., 2017b; Naruoka et al., 2011; Pang et al., 2020; Ward et al., 2019) and kernel-related traits (Chen et al., 2016; Jaiswal et al., 2015; Liu et al., 2018b; Zhang et al., 2012) have been identified on all the wheat chromosomes. Numer-

### Core Ideas

- Final grain yield in wheat is primarily determined by various kernel and spike traits.
- Fifty-three MTAs were identified for spike and kernel traits in hard winter wheat germplasm.
- Fourteen MTAs for different traits were consistent over multiple environments.
- Promising germplasm and SNPs associated with yield-contributing traits will facilitate future wheat breeding.

ous previous studies reported major effect genes and QTL, such as *TaGS5*, *TaSus1*, *TaSus2*, and *TaGW2*, significantly associated with kernel size, kernel weight, spike, peduncle length, and grain weight (Bednarek et al., 2012; Hou et al., 2014; Ma et al., 2016; Wang et al., 2015; Zhang et al., 2018). Two major genes responsible for modern wheat spike morphology (*Q* and *C*) mapped on chromosomes 5A and 2D were also found associated with grain size, shape, grain number, grain yield, and TKW (Johnson et al., 2008; Xie et al., 2018). Genes such as *TaMOC1-7A*, *TaTEF-7A*, and *WAP0-A1* were found to be associated with spikelet or grain number per spike (Kuzay et al., 2019, 2022; Zhang et al., 2015; Zheng et al., 2014), whereas *TaCKX6-D1* and *TaSAP1* have been reported to have significant effect on TKW and NKPS (Chang et al., 2013; Zhang et al., 2012).

Even though a considerable number of studies dissected the genetics of yield-related traits in wheat, relatively few studies used winter wheat germplasm (Dhakal et al., 2021; Gill et al., 2022; Lozada et al., 2017; Ward et al., 2019; Zanke et al., 2015). As winter and spring wheat possess considerable genetic differences with respect to growth habits and adaptation, the genetic exchange between these two major wheat classes is limited. The winter wheat breeding programs mostly breed within the winter class and hybridizations among these two classes are made only to introgress specific traits like disease resistance or other marker-linked traits of agronomic importance. This necessitates the identification of genomic regions related to yield and associated traits within winter wheat for direct use. Further, hard winter wheat germplasm likely harbors many new QTL affecting the yield and contributing traits that are yet to be explored.

The present study used an association mapping panel comprising hard winter wheat accessions from the U.S. Great Plains breeding programs, which has been successfully used to dissect various important traits (Altameemi et al., 2021; Navrotskyi et al., 2020; Sidhu et al., 2020). We sought to perform GWAS using the hard winter wheat association mapping panel (HWWAMP) to identify the genomic regions associated

with various spike- and kernel-related traits and explore the putative candidate genes harbored in these regions using the wheat reference genome.

## 2 | MATERIALS AND METHODS

### 2.1 | Plant materials and field trials

In the current study, we used 297 accessions of the HWWAMP developed under the USDA–TCAP project (Guttieri et al., 2015). This HWWAMP comprises advanced breeding lines and released varieties since the 1940s from the Great Plains region of the United States including Montana, North Dakota, South Dakota, Nebraska, Kansas, Colorado, Oklahoma, and Texas. The association mapping panel was evaluated at three experimental stations of South Dakota State University at the Aurora Farm (E1), Brookings Farm (E2), and Felt Farm (E3) in South Dakota during the 2019–2020 growing season using a randomized complete block design with two replications. Each accession was planted in a 1.25 m long two-row plot with an interrow spacing of 20 cm. The trials were managed using recommended agronomic practices for proper growth and development.

### 2.2 | Phenotypic trait evaluation and statistical analysis

Eight morphological traits were evaluated including NSPS, SL, spikelet density (SD), NKPS, TKW, kernel length (KL), kernel width (KW), and KA. Ten random spikes from each accession per replication were manually harvested at physiological maturity. Spike length was measured from the base of the rachis to the topmost spikelet excluding the awns. The NSPS were counted from the basal sterile spikelet to the top fertile spikelet. The SL and NSPS were the means of 10 randomly selected spikes in each replication. Spikelet density was calculated as the ratio of NSPS to SL. The TKW was measured by weighing 1,000 kernels from each accession in two replicates. Three kernel-related traits (KL, KW, and KA) were recorded following an automatic grain analyzer Vibe QM3 (Vibe Imaging Analytics).

META-R (Multi Environment Trial Analysis with R) v6.04 (Alvarado et al., 2020) was used for estimating the best linear unbiased estimates (BLUEs) for all the traits. The BLUEs of two replicates for individual locations were estimated using the following model:

$$y_{ij} = \mu + G_i + R_j + e_{ij}$$

where  $y_{ij}$  is the trait of interest,  $\mu$  is the overall mean,  $G_i$  is the fixed effect of the  $i$ th genotype or accession,  $R_j$  is the ran-

dom effect of the  $j$ th replicate, and  $e_{ij}$  is the random error term associated with the  $i$ th genotype and  $j$ th replicate. For multilocation analysis, BLUEs over three environments were obtained using the following model:

$$y_{ijk} = \mu + G_i + E_j + R_{k(j)} + GE_{ij} + e_{ijk}$$

where  $y_{ijk}$  is the trait of interest,  $\mu$  is the overall mean,  $G_i$  is the fixed effect of the  $i$ th genotype,  $E_j$  is the random effect of the  $j$ th environment,  $R_{k(j)}$  is the random effect of the  $k$ th replicate nested within the  $j$ th environment,  $GE_{ij}$  is the random effect of the genotype  $\times$  environment (G  $\times$  E) interaction, and  $e_{ijk}$  is the random error term. To estimate the broad-sense heritability ( $H^2$ ) across environments, variance components were estimated by fitting genotype ( $G_i$ ) effects as random. The heritability was calculated as follows:

$$H^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_{GE}^2/nLoc + \hat{\sigma}_e^2/(nLoc \times nRep)}$$

where  $\hat{\sigma}_g^2$  genotype variance component, and  $\hat{\sigma}_{GE}^2$  is G  $\times$  E interaction variance component, and  $\hat{\sigma}_e^2$  is the error variance components, respectively. The  $nRep$  term represents the number of replications within an environment and  $nLoc$  term represents the number of environments in the analysis.

### 2.3 | Genotyping and SNP discovery

The HWWAMP was genotyped using the wheat Infinium 90K iSelect array (Illumina Inc.) under the USDA–TCAP (Cavanagh et al., 2013). The genotypic data was obtained from the T3 Toolbox (<https://wheat.triticeaetoolbox.org/>). The dataset was then filtered to remove the single-nucleotide polymorphisms (SNPs) with any heterozygosity, >10% missing data, and minor allele frequency of <5%. The filtered genotypic dataset consisted of a total of 15,590 high-quality SNPs, which were then imputed in TASSEL v5.0 (Bradbury et al., 2007) for further analysis. The genetic positions of the wheat Infinium 90K iSelect SNP markers were obtained from the consensus genetic map of 46,977 SNPs (Wang et al., 2014). The physical positions of the SNPs associated with various spike- and kernel-related traits were obtained by BLAST searching the flanking sequences of respective SNPs to wheat Chinese Spring RefSeq v1.1 (IWGSC et al., 2018).

### 2.4 | Population structure and linkage disequilibrium

Population structure among the 297 hard winter wheat accessions was assessed using a Bayesian model-based clustering

program STRUCTURE v2.3.4 assuming an admixture model (Pritchard et al., 2000). The most likely number of subgroups was inferred based on an ad hoc statistic ( $\Delta K$ ) based on the rate of change in the log probability between runs using successive  $K$  values (Evanno et al., 2005) using STRUCTURE HARVESTER (Earl & vonHoldt, 2012). We used 10 subgroups ( $K = 1-10$ ) with five independent runs for each subgroup using a burn-in period of 10,000 iterations followed by 10,000 Monte Carlo iterations. The HWWAMP was analyzed for linkage disequilibrium (LD) decay pattern in one of our previous studies (Ayana et al., 2018) using a set of 1,842 SNPs in TASSEL v5.0 (Bradbury et al., 2007). The LD decay was estimated as the distance where the LD value ( $r^2$ ) falls below 0.1 or half strength of  $D'$  ( $D' = 0.5$ ). For the whole genome, the LD dropped to 0.5 at 4.5 cM. The LD decay was found similar for both A (3.4 cM) and B (3.6 cM) subgenomes; however, it was much higher for the D subgenome (14.2 cM).

## 2.5 | Marker–trait associations

Genome-wide associations were analyzed using two different algorithms, namely the mixed linear model (Yu et al., 2006) and fixed and random model circulating probability unification (FarmCPU) (Liu et al., 2016). Both the models were implemented in the Genomic Association and Prediction Integrated Tool (GAPIT) (Lipka et al., 2012). In brief, the mixed linear model incorporates kinship and population structure as covariates to minimize the confounding effects and control false positives. However, it leads to false negatives because of the confounding between testing markers and cofactors simultaneously. The FarmCPU model is an improved multiple-locus model that controls false positives by fitting the associated markers detected from the iterations as cofactors to perform marker tests within a fixed-effect model. The quantile–quantile plots suggested that the FarmCPU performed better than the mixed linear model for most traits. Therefore, we employed FarmCPU to report the marker–trait associations (MTAs) for all the spike- and kernel-related traits. Though we found several associations to be significant based on the Bonferroni correction for multiple testing, most of these associations were limited to one or two environments. Thus, we used an arbitrary threshold of  $-\log_{10} P \geq 3.0$  to report MTAs following previous studies (Dong et al., 2016; Liu et al., 2017a; Sidhu et al., 2020; Wang et al., 2017), and only those MTAs were reported that surpassed this threshold in the combined environment analysis (CEnv) and at least two individual environments.

Further, we sought to analyze the effect of accumulated favorable alleles on the trait's phenotype. The accessions from the mapping panel were grouped based on the alleles of identified MTAs for all the spike-related traits. These

groups were compared using Tukey's HSD (honestly significant difference) test at a 5% level of significance to verify the additive effect of the favorable alleles on the phenotype of the traits.

## 2.6 | Identification of candidate genes

For candidate gene analysis, only stable MTAs for three spike traits were selected. The candidate regions were demarcated within  $\pm 1$  Mb of the most significant SNP for each region to identify the candidate genes. The high-confidence genes in the selected region were retrieved from the wheat genome assembly IWGSC RefSeq v1.1 and gene ontology annotation information of these genes was extracted from IWGSC Functional Annotation v1.0 (IWGSC et al., 2018). The gene expression browser was used to exclude unlikely candidates and shorten the candidate list to fewer genes for each selected MTA (Ramírez-González et al., 2018).

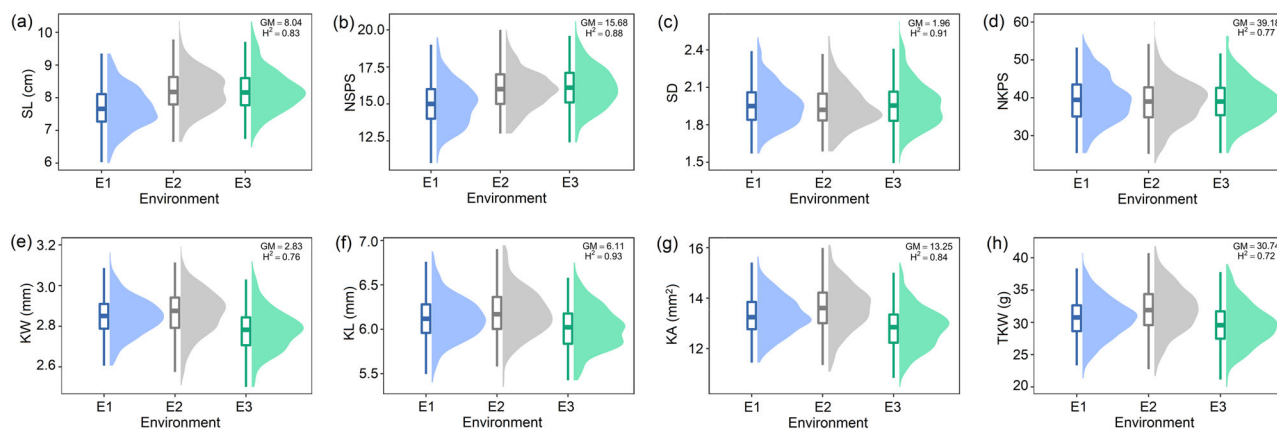
## 3 | RESULTS

### 3.1 | Phenotypic variation and correlations

A wide variation was observed in the HWWAMP for all the spike- and kernel-related traits across three environments (Supplemental Table S1). Overall, phenotypic variabilities of traits appeared to be normally distributed with noticeable variation among three environments (Figure 1). Correlation coefficients among three different environments and CEnv values for each trait were generally high (Supplemental Figures S1 and S2). The broad-sense heritability estimates for eight traits ranged from .72 (TKW) to .93 (KL). Pearson's correlation among eight traits was estimated using across-environment BLUE values (Figure 2). A moderate to high positive correlation was observed between NSPS and other spike-related traits such as SL ( $r = 0.42$ ), SD ( $r = 0.64$ ), and NKPS ( $r = 0.32$ ). However, NSPS was negatively correlated with all four kernel-related traits (KW, KL, KA, and TKW). On the other hand, SL showed significant positive correlations with all the studied traits except SD ( $r = -0.43$ ) (Figure 2). As expected, we observed a strong positive correlation within four kernel-related traits (KW, KL, KA, and TKW), and kernel-related traits showed a negative correlation with spike-related traits except for SL (Figure 2).

### 3.2 | Genotyping and population structure

The genotype data was filtered to obtain a total of 15,590 high-quality SNPs that were distributed across all 21 wheat



**FIGURE 1** Distributions of eight spike- and kernel-related traits in three different environments (E1, E2, and E3). (a) spike length (SL); (b) number of spikelets per spike (NSPS); (c) spikelet density (SD); (d) number of kernels per spike (NKPS); (e) kernel width (KW); (f) kernel length (KL); (g) kernel area (KA); and (h) thousand-kernel weight (TKW). GM, grand mean across three environments;  $H^2$ , broad-sense heritability

chromosomes (Supplemental Table S2). We investigated the population structure to assess if the association-mapping panel is structured based on the breeding programs and origin and to figure out any relationship of population structure with the spike and kernel traits. Four subpopulations (P1, P2, P3, and P4) were identified in the HWWAMP, where P1, P2, P3, and P4 consist of 120, 34, 33, and 110 genotypes. We found a significant difference in means among the four subpopulations (P1, P2, P3, and P4) for all the traits except SL (Supplemental Table S3).

### 3.3 | GWAS on the spike- and kernel-related traits

A GWAS was performed on the spike- and kernel-related traits using BLUES from across-environment analysis (CEnv) and individual environments (Supplemental Figure S3).

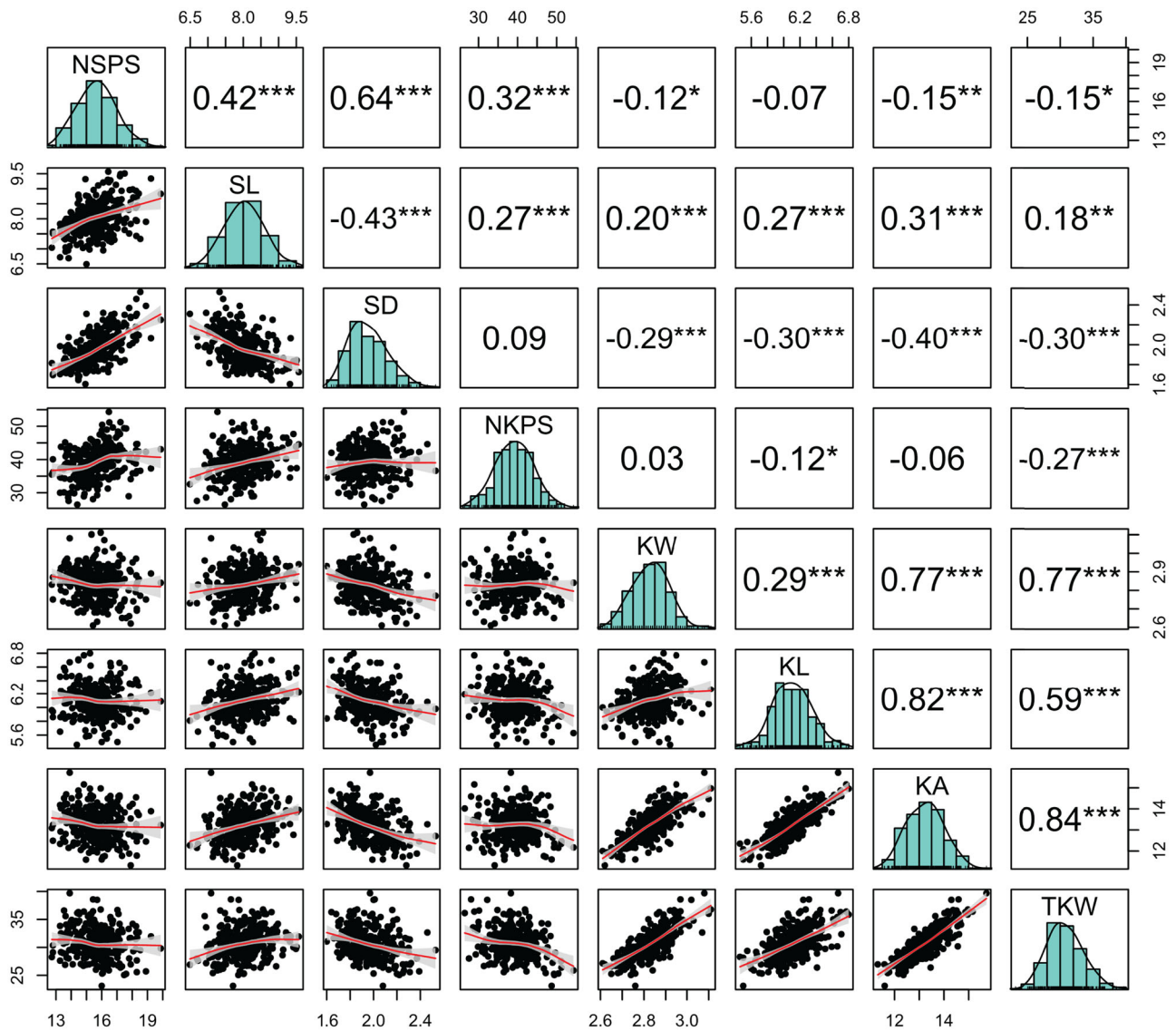
The GWAS identified a total of 53 MTAs ( $P < .001$ ) in combined environment analysis (CEnv) for seven traits out of eight studied traits, as no MTA was observed for KL (Supplemental Table S4). The MTAs were distributed across all the chromosomes except 2D and 5D (Figure 3). Most of the MTAs were identified on the A subgenome (28), followed by the B (19) and D (6) subgenomes. The highest number of MTAs were located on chromosomes 2B and 4A (six MTAs each), followed by 1A, 2A, and 3B (five MTAs each) (Figure 3; Supplemental Table S4). Of the 53 MTAs from the CEnv analysis, we emphasized only on 14 MTAs that were identified in CEnv as well as at least two individual environments (Table 1). Thus, these 14 MTAs were considered to represent stable QTL for spike or kernel traits in U.S. hard winter wheat breeding germplasm, whereas the 39 MTA identified in CEnv could be considered putative QTL (Supplemental Table S4).

#### 3.3.1 | Spike-related traits

Of the 15 MTAs identified for NSPS in CEnv (Supplemental Table S4), four MTAs, namely *Excalibur\_c97022\_396* (37.4 Mb), *RFL\_Contig3175\_1217* (604.9 Mb), *IWA4455* (462.6 Mb), and *IWA5913* (674.3 Mb) on chromosomes 6A, 6A, 6D, and 7A, respectively, were identified in multiple environments (Table 1). For SL, a total of 16 MTAs (Supplemental Table S4) were identified of which only three SNPs, namely *Kukri\_c10860\_1283* (87.9 Mb), *Tdurum\_contig82393\_484* (730.6 Mb), and *IWA3639* (610.9 Mb) on chromosomes 2A, 2B, and 7A, were identified in multiple environments (Table 1). Out of the 11 MTAs for SD in CEnv analysis, three MTAs, including *IWA2519* (371.6 Mb), *IWA5913* (674.3 Mb), and *IWA1902* (530 Mb) on chromosomes 3A, 7A, and 7D, were identified in multiple environments. Similarly, we identified five MTAs for NKPS on chromosomes 1D, 2B, and 5B, of which *BS00021959\_51* (110.8 Mb) and *Excalibur\_c1921\_1191* (427.7 Mb) on chromosomes 2B and 5B were identified in multiple environments (Supplemental Table S4; Figure 3).

#### 3.3.2 | Kernel-related traits

The GWAS identified only six MTAs for three (KW, KA, and TKW) kernel-related traits (Supplemental Table S4, Supplemental Figure S3). In total, three MTAs were detected for KW on chromosomes 1A, 2A, and 4A. Out of the three MTAs, only *BS00044274\_51* (47.8 Mb) on chromosome 2A was significant in CEnv and at least two environments (Table 1). Further, two MTAs were identified for KA on chromosomes 3D and 4A; however, only *Kukri\_c74409\_199* (37.8 Mb) on chromosome 4A was found in multiple environments and CEnv. We identified only one MTA for TKW located on chromosome 2A (Supplemental Table S4).



**FIGURE 2** Pearson's correlation matrix among spike- and kernel-related traits based on their best linear unbiased estimates. Values inside the rectangle represent the correlation coefficient and three symbols \*, \*\*, and \*\*\*, represent correlation coefficient significance levels at  $\alpha < .05$ ,  $.01$ , and  $.001$ , respectively. Bivariate scatter plots with fitted lines are on the left side of the diagonally arranged phenotypic distribution plots. NSPS, number of spikelets per spike; SL, spike length; SD, spikelet density; NKPS, number of kernels per spike; KW, kernel width; KL, kernel length; KA, kernel area; TKW, thousand-kernel weight

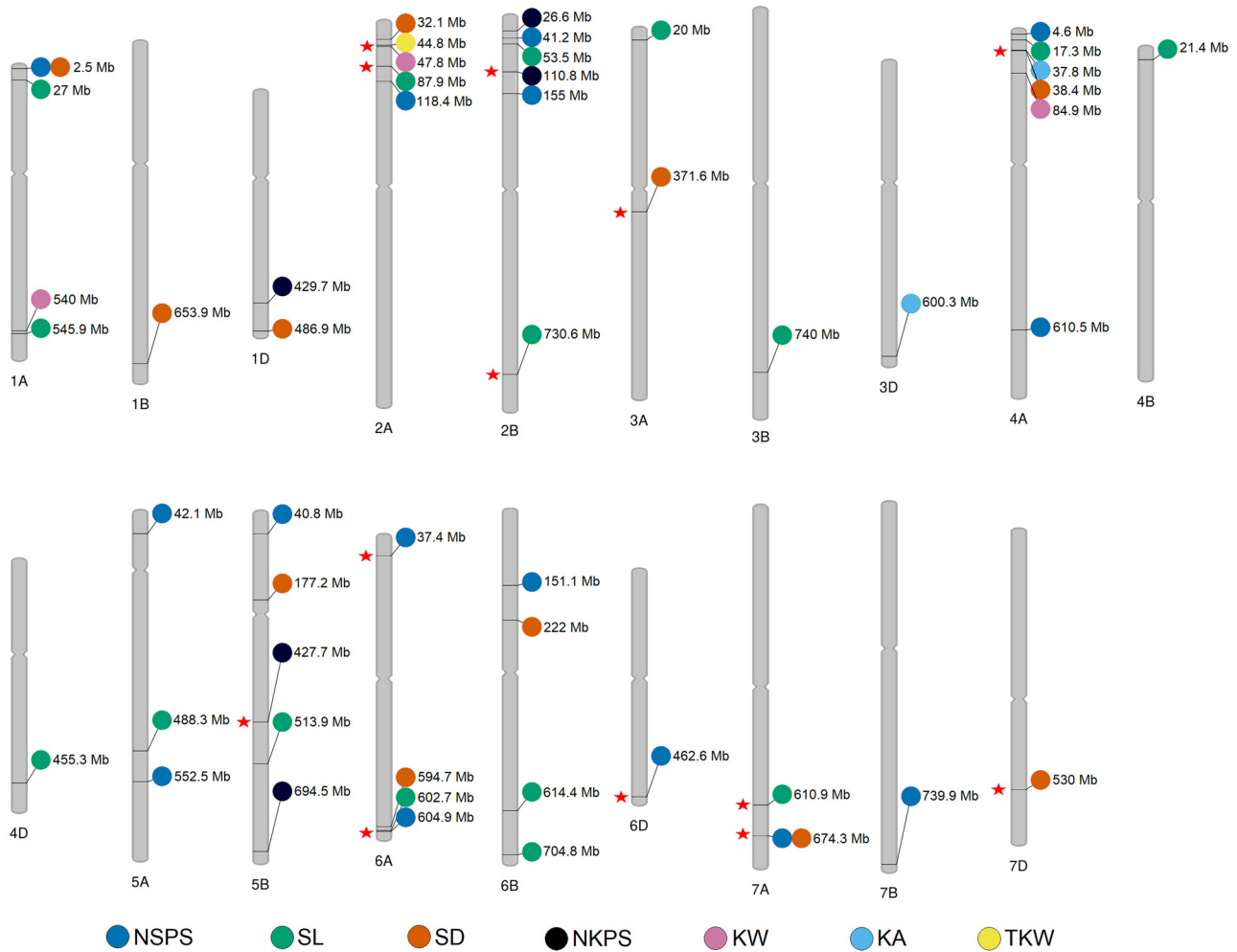
### 3.4 | Genomic regions or SNPs associated with multiple traits

We identified a total of five MTAs that were associated with more than one trait (Supplemental Table S4). Two MTAs, namely *Excalibur\_c35316\_154* (2.5 Mb) and *IWA5913* (674.2 Mb) located on chromosomes 1A and 7A, were associated with both NSPS and SD. Similarly, *BS00044274\_51* (47.8 Mb) on chromosome 2A was associated with KA and KW, whereas *IWA6659* (84.9 Mb) on chromosome 4A was associated with several spike- and kernel-related traits including KW, KA, TKW, and SL. Another MTA, *IWA6485* (600.2 Mb) on chromosome 3D, was

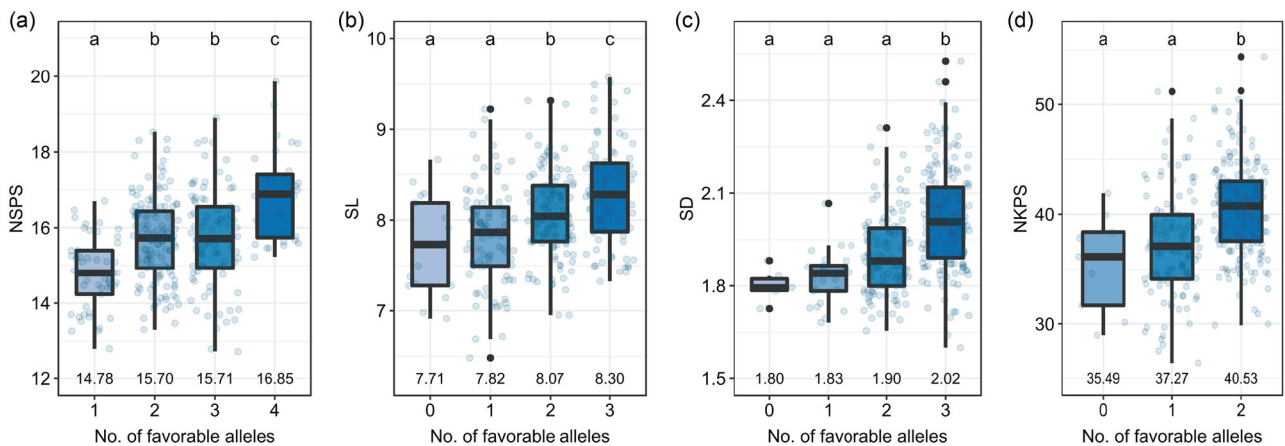
significantly associated with KA and SL (Supplemental Table S4).

### 3.5 | Phenotypic effects of favorable alleles of detected MTAs

To identify the accumulative effect of favorable alleles for various traits, we grouped the HWWAMP based on the number of favorable alleles carried by the accessions for four spike-related traits (Figure 4). We used only stable MTAs identified for spike-related traits in this analysis (Table 1). For NSPS, accessions with one, two, three, and four favorable alleles



**FIGURE 3** Physical position (Mb) of the marker–trait associations (MTAs) identified for seven spike- and kernel-related traits in the combined analysis (CEnv). The physical positions are based on Chinese Spring RefSeq 1.1 (IWGSC et al., 2018). The consistent MTAs for respective traits identified over multiple environments are highlighted using a red star on left side of the chromosome bars. NSPS, number of spikelets per spike; SL, spike length; SD, spikelet density; NKPS, number of kernels per spike; KW, kernel width; KA, kernel area; TKW, thousand-kernel weight



**FIGURE 4** Effect of accumulation of favorable alleles of the identified stable associations (marker–trait associations identified in multiple environments) for four spike-related traits: (a) number of spikelets per spike (count); (b) spike length (cm); (c) spikelet density; and (d) kernel per spike (count). The x axis represents the favorable allele count and y axis represents the respective phenotypic value. Significantly different groups are represented with different letters (on top of the boxplots) based on Tukey's HSD test at .05 level of significance. NSPS, number of spikelets per spike; SL, spike length; SD, spikelet density; NKPS, number of kernels per spike

**TABLE 1** Stable marker–trait associations identified in multiple environments (CEnv and at least two individual environments) for various spike- and kernel-related traits in the U.S. hard winter wheat

Trait	SNP <sup>a</sup>	Chr	Alleles <sup>b</sup>	Pos <sup>c</sup>	–Log <sub>10</sub> <i>P</i>	Environments <sup>d</sup>	Effect
NSPS	<i>Excalibur_c97022_396</i>	6A	<u>C</u> /T	37,415,157	4.0	E1, E3, CEnv	–0.266
	<i>RFL_Contig3175_1217</i>	6A	T/ <u>C</u>	604,877,158	3.5	E2, E3, CEnv	–0.168
	<i>IWA4455</i>	6D	A/ <u>G</u>	462,631,946	3.8	E1, E3, CEnv	0.061
	<i>IWA5913</i>	7A	<u>A</u> /G	674,276,906	14.0	E1, E2, E3, CEnv	–0.500
SL	<i>Kukri_c10860_1283</i>	2A	<u>G</u> /A	87,857,405	4.7	E1, E2, CEnv	0.163
	<i>Tdurum_contig82393_484</i>	2B	<u>C</u> /A	730,562,664	8.4	E1, E3, CEnv	0.147
	<i>IWA3639</i>	7A	<u>G</u> /A	610,934,198	7.6	E1, E3, CEnv	–0.138
SD	<i>IWA2519</i>	3A	<u>C</u> /T	371,628,644	3.3	E1, E2, CEnv	–0.026
	<i>IWA5913</i>	7A	<u>A</u> /G	674,276,906	9.4	E1, E2, E3, CEnv	–0.055
	<i>IWA1902</i>	7D	<u>A</u> /G	530,035,575	4.1	E1, E2, E3, CEnv	–0.041
NKPS	<i>BS00021959_51</i>	2B	<u>C</u> /T	110,818,850	3.2	E1, E3, CEnv	–1.862
	<i>Excalibur_c1921_1191</i>	5B	<u>G</u> /A	427,650,909	4.0	E1, E2, CEnv	1.390
KW	<i>BS00044274_51</i>	2A	<u>T</u> /G	47,826,702	3.3	E2, E3, CEnv	0.024
KA	<i>Kukri_c74409_199</i>	4A	<u>G</u> /A	37,773,890	3.4	E1, E3, CEnv	0.250

Note. Chr, chromosome; Pos, physical position in base pair (based on IWGSC RefSeq); NSPS, number of spikelets per spike; SL, spike length; SD, spikelet density; NKPS, number of kernels per spike; KW, kernel width; KA, kernel area.

<sup>a</sup>Single-nucleotide polymorphisms (SNPs) were significant in multiple environments.

<sup>b</sup>Favorable allele is underlined.

<sup>c</sup>Position in base pair (bp) based on Chinese Spring RefSeq 1.1.

<sup>d</sup>Environment 1 (E1), Environment 2 (E2), Environment 3 (E3), combined best linear unbiased estimates (CEnv).

had a mean value of 14.78, 15.70, 15.71, and 16.85 spikelets per spike, respectively. The accessions with three or more favorable alleles had significantly higher NSPS than lines with only one favorable allele (Figure 4). We also observed a significant difference in SL with the increment of favorable alleles. The HWWAMP lines that carried none of the identified favorable alleles for SL had a mean of 7.7 cm, whereas lines with one, two, and three favorable alleles had a mean SL of 7.82, 8.06, and 8.30 cm, respectively. A similar additive effect was observed for SD, as three favorable alleles significantly increased the compactness of the spike (SD) compared with lines with a lower number of favorable alleles. Similarly, the trait mean for NKPS was 35.49, 37.27, and 40.53 when the accessions had zero, one, and two favorable alleles, respectively (Figure 4).

### 3.6 | Candidate gene analysis for stable MTAs

Candidate gene analysis was performed for five stable MTAs having the  $-\log_{10} P \geq 4.0$ . For each MTA, a window of  $\pm 1$  Mb around the most significant SNP was demarcated to identify the candidate genes. A total of 120 high-confidence genes were retrieved by using IWGSC RefSeq v1.0 annotation. However, we excluded the unlikely candidates using publicly available RNA sequencing expression data from the Wheat

Expression Browser and a thorough review of related literature. Finally, 14 putative candidate genes were selected for three different traits based on their relatively high expression in the shoot, spike, and grain development at the vegetative and reproductive stages of wheat (Table 2). The selected genes encode for various proteins including peptidyl-prolyl cis-trans isomerase, cold shock protein, NADPH-cytochrome P450 reductase, glycoprotein membrane GPI-anchored, ubiquitin-like protein, BZIP transcription factor, and ATP-dependent RNA helicase (Table 2).

## 4 | DISCUSSION

Wheat yield is the most important and complex target trait for wheat breeding programs. Grain yield mainly depends on the accumulative effect of the different yield-contributing traits including spike- and kernel-related traits. Breeders have relied upon the identification and deployment of novel genes and QTL governing these crucial traits. Nevertheless, relatively few studies (Dhakal et al., 2021; Lozada et al., 2017; Ward et al., 2019; Zanke et al., 2015) have been reported on winter wheat. In this study, a diverse panel of varieties and advanced breeding lines obtained from various hard winter wheat breeding programs from the U.S. Great Plains was used to dissect the genetics of spike and kernel traits.



**TABLE 2** Putative candidate genes within the identified regions associated with wheat spike-related traits

Trait	Significant SNP	Chr	Gene ID	Protein
NSPS	<i>Excalibur_c97022_396</i>	6A	<i>TraesCS6A01G068300</i>	Gamma-glutamylcyclotransferase
			<i>TraesCS6A01G068900</i>	Peptidyl-prolyl cis-trans isomerase
			<i>TraesCS6A01G069500</i>	Cold shock protein
SL	<i>Tdurum_contig82393_484</i>	2B	<i>TraesCS2B01G533500</i>	Sulfhydryl oxidase 1
			<i>TraesCS2B01G534200</i>	NADPH-cytochrome P450 reductase
			<i>TraesCS2B01G534700</i>	UPF0136 membrane protein
	<i>Kukri_c10860_1283</i>	2A	<i>TraesCS2A01G141800</i>	Transcription-associated protein 1
			<i>TraesCS2A01G141900</i>	Glycoprotein membrane GPI-anchored (TaMs1)
			<i>TraesCS2A01G142700</i>	Ubiquitin-like protein
<i>IWA3639</i>	7A	<i>TraesCS2A01G142800</i>	BZIP transcription factor	
		<i>TraesCS7A01G419100</i>	60S acidic ribosomal protein P3	
		<i>TraesCS7A01G419400</i>	ATP-dependent RNA helicase	
SD	<i>IWA1902</i>	7D	<i>TraesCS7D01G411600</i>	60S acidic ribosomal protein P3
			<i>TraesCS7D01G412200</i>	ATP-dependent RNA helicase

Note. NSPS, number spikelets of per spike; SL, spike length; SD, spikelet density, Chr, chromosome.

#### 4.1 | Phenotypic correlations among traits and heritability

A wide range of phenotypic variation was observed among the germplasm for all the traits studied, making this panel suitable for GWAS (Supplemental Table S1). High broad-sense heritability was observed for all the traits ranging from .72 for TKW to .93 for KL, indicating less environmental influence on the observed traits (Figure 1). The high heritability estimates were in line with several previous studies (Alqudah et al., 2020; Garcia et al., 2019; Muhammad et al., 2020; Wang et al., 2017). We observed a significant positive correlation between NSPS and other spike-related traits (SL, SD, and NKPS), whereas a strong negative correlation was observed between SL and SD (Figure 2). In this study, we observed a moderately positive correlation (0.32) between NSPS and NKPS, which was lower than expected. All of our experiments were conducted in dryland (without irrigation) to mimic the South Dakota and U.S. Great Plains environments. Two of the three environments were drier during flowering and grain-filling stages likely resulting in several spikelets being sterile, which is in line with previous studies that showed abiotic environmental factors such as drought impairs wheat growth and leads to spikelet sterility (Gol et al., 2021; Ji et al., 2010). In our study, we counted the total number of spikelets (both fertile and sterile). Thus, the lower-than-expected correlation between NSPS and NKPS observed in this study might be related to the increased number of sterile spikelets. Spikelet density showed a strong positive relation with NSPS but was negatively correlated with SL, owing to its mathematical dependency on these traits. These results also indicate that the loci controlling these three traits (NSPS, SL, and SD) are often not independent and could be in an

inverse relationship. Therefore, to accelerate the breeding process, it is essential to determine relationships among different interdependent traits. On the other hand, all the kernel-related traits were positively correlated among themselves and with SL, showing that increased SL positively affects kernel traits (Figure 2). Several previous studies also reported positive associations among TKW, KL, and KW (Brescghello & Sorrells, 2007; Chen et al., 2016; Muhammad et al., 2020; Ramya et al., 2010; Rasheed et al., 2014). We also observed that an increase in NSPS, NKPS, and SD causes a decrease in kernel-related traits (KW, KL, KA, and TKW), which is consistent with the previous reports (Muhammad et al., 2020; Qu et al., 2021; You et al., 2021). Further, two yield-component traits, grain number and grain weight or KA, are known to be negatively correlated (Bustos et al., 2013; García et al., 2013; Sadras, 2007), and this might be a consequence of trade-offs between these traits. Further, the competition for assimilates between spikelets leads to an unbalanced distribution of grain number along with the spike (Guo et al., 2017; Molero et al., 2019).

#### 4.2 | Marker-trait associations and comparison with previous studies

We used the FarmCPU algorithm to identify the MTAs that use both fixed and random effect models iteratively to control the false positives arising from population structure and familial relatedness. In this study, a total of 53 MTAs were identified for the spike- and kernel-related traits using combined BLUEs across three environments (CEnv) (Supplemental Table S4). However, we mainly focused on the MTAs identified in multiple environments and MTAs

associated with multiple traits (Table 1). We identified 47 MTAs for spike-related traits (Supplemental Table S4), while only a few associations were detected for kernel-related traits. Several MTAs identified in this study were likely to be novel, whereas many of the MTAs colocalized with QTLs that have been reported to be reliable in various previous studies. Spike-related traits are generally complex and most of the wheat chromosomes harbor genetic factors affecting these traits (Liu et al., 2018b). Therefore, optimization of wheat inflorescence traits, such as head length or SL, spikelet number, and SD can effectively enhance the integrated sink capacity and ultimate yield potential of wheat (Fan et al., 2019; Yao et al., 2019; You et al., 2021).

The MTAs for NSPS were distributed on the chromosomes 1A, 2A, 2B, 4A, 5A, 5B, 6A, 6B, 6D, 7A, and 7B (Supplemental Table S4). Similar to our findings, many previous studies reported several important chromosomes, such as 1A, 5A, 7A, and 7B, harbor significant stable QTL for NSPS (Cui et al., 2012; Hu et al., 2020; Kato et al., 2000; Ma et al., 2007; Liu et al., 2018a). In our study, four stable MTAs, namely *Excalibur\_c97022\_396* (37.4 Mb), *RFL\_Contig3175\_1217* (604.9 Mb), *IWA4455* (462.6 Mb), and *IWA5913* (674.3 Mb) on chromosomes 6AS, 6AL, 6DL, and 7AL, respectively (Table 1), were found to be significantly associated with NSPS. Of the four stable associations for NSPS, one MTA (*IWA5913*) identified on chromosome 7A overlapped with *WHEAT ORTHOLOG OF APO1 (WAPO1)* that was cloned from the wheat chromosome arm 7AL (~674 Mb) in a recent study as a gene regulating NSPS (Kuzay et al., 2019, 2022). Further, Sun et al. (2017) identified an SNP (*Kukri\_c264\_539*) for NSPS on 6A (29.5 Mb) in four environments, which is a similar region as the stable MTA *Excalibur\_c97022\_396* (37.4 Mb) detected in this study. We did not find any QTL reported for NSPS in regions *RFL\_Contig3175\_1217* (6A) and *IWA4455* (6D), therefore these two stable MTA could be novel. In addition to these stable MTAs, we also identified two other MTAs for NSPS that were only identified in CEnv analysis. One MTA on chromosome 7B (*GENE-4848\_95*; 740 Mb) was colocalized in a similar region with a stable QTL reported by Liu et al. (2018a). Moreover, one significant marker (*BS00011235\_51*; 552 Mb) for NSPS was detected on chromosome 5A, which is located ~35 Mb away from the region containing an important gene controlling vernalization (*Vrn-A1*) (Yan et al., 2003). This *Vrn-A1* gene has been previously reported to be involved in increasing the NSPS (Whitechurch & Snape, 2003).

In our study, 16 MTAs were identified on chromosomes 1A, 2A, 2B, 3A, 3B, 4A, 4B, 4D, 5A, 5B, 6A, 6B, and 7A for SL. Out of those MTAs, three associations, namely *Kukri\_c10860\_1283* (87.9 Mb), *Tdurum\_contig82393\_484* (730.6 Mb), and *IWA3639* (610.9 Mb) on chromosomes 2A, 2B, and 7A, were found to be stable over multiple environments (Table 1). However, MTAs *Kukri\_c10860\_1283*

and *Tdurum\_contig82393\_484* for SL have been previously reported in various studies (Liu et al., 2017b, 2018a; Mwadzingeni et al., 2017), but we did not find any previous reports on SL QTL in proximity to *IWA3639* (610.9 Mb) on chromosome 7A. Another MTA (*Excalibur\_rep\_c68588\_1196*) identified in the CEnv analysis on chromosome 4BS (21 Mb) for SL, corresponds to the physical location of the semidwarfing gene *Rht-B1* (~30 Mb). Other than reducing plant height, *Rht-B1* was reported to exhibit pleiotropic effects on grain yield and yield components including SL (Guan et al., 2020; Okada et al., 2019). Several other MTAs identified on chromosomes 5A, 6A, and 6B for SL in the CEnv analysis in this study were found to be colocalized with MTAs reported by previous studies (Liu et al., 2018a; Mwadzingeni et al., 2017; Yu et al., 2014) (Supplemental Table S4).

For SD, 11 MTAs distributed on chromosomes 1A, 1B, 1D, 2A, 3A, 4A, 5B, 6A, 6B, 7A, and 7D were identified; however, MTAs on chromosomes 3A, 7A, and 7D were stable and significant in multiple environments (Table 1; Supplemental Table S4). Similar to our finding, several previous studies reported MTAs for SD on chromosomes 1A, 1B, 3A, 4A, 6A, and 6B (Jantauriyarat et al., 2004; Liu et al., 2018b). The QTL identified for SD, such as *qSc-7A* (Fan et al., 2019) and *Qsd.sau-2SY-7A.2* (You et al., 2021), were found in the vicinity of stable association (*IWA5913*) detected in this study on chromosome 7A (674 Mb), the region harboring major NSPS-associated gene *WAPO1*. Another association for SD (*IWA1902*) at 530 Mb on chromosome 7D was consistently identified in all three individual environments and the CEnv analysis, indicating the importance of this region in spike morphology. A minor QTL for SD on chromosome 7D was also identified in a multiple-environment joint analysis by Liu et al. (2020) in a similar genomic region as detected in our study. The chromosome 7D MTA could be further studied and validated in different genetic backgrounds and multiple environments to facilitate its application in wheat breeding.

For NKPS, we observed five MTAs that were distributed on chromosomes 1D, 2B, and 5B (Supplemental Table S4). One stable association on chromosome 2B (*BS00021959\_51*) was detected at 110.8 Mb (90.9 cM) for NKPS (Table 1). Shi et al. (2017) also reported an association for NKPS at 122.3 Mb and suggested that this marker may be linked to the photoperiod-insensitive gene *Ppd-B1* in common wheat. A similar association was detected by Gao et al. (2015) at 92 cM for the same trait on chromosome 2B in a biparental mapping population in two environments. It would be interesting to know if this QTL for NKPS is *Ppd-B1*. Further, we identified another stable MTA (*Excalibur\_c1921\_1191*) for NKPS on chromosome 5B at 427.6 Mb (Table 1). Tang et al. (2011) mapped three QTL for NKPS spanned by markers *Xgwm499* and *Xgwm213* (418.8–477.5 Mb) overlapping the physical position identified in our study. Two other

MTAs were identified using the CEnv analysis for NKPS on chromosomes 2B (*CAP8\_c5108\_139*) and 5B (*IWA4329*), which colocalized with the previously reported MTAs in several studies (Guo et al., 2017; Liu et al., 2018a; Neumann et al., 2011) (Supplemental Table S4).

A total of three MTAs were identified for KW on chromosomes 1A, 2A, and 4A. Several studies have reported associations in similar physical locations on chromosomes 1A and 2A for KW (Li et al., 2019; Wu et al., 2015). Further, we also identified MTAs for KA on chromosomes 3D and 4A and for TKW on 2A; however, previous studies also identified significant association for KA and TKW on chromosomes 4A and 2A, respectively (Liu et al., 2018a; Su et al., 2018) (Table 1; Supplemental Table S4). Thus, the present study also validated many of the QTL reported in previous studies from different backgrounds in addition to the identification of several novel QTL for spike and kernel traits.

### 4.3 | Multi-trait QTL regions

We identified several genomic regions that were associated with more than one trait and this phenomenon could likely be due to the complex relationships among these traits (Figure 3). Our study identified two MTAs (*Excalibur\_c35316\_154* and *IWA5913*) responsible for NSPS, which were also found to be associated with the SD (Supplemental Table S4). In our study, these could be likely a result of the mathematical dependency of SD on NSPS, as SD is expressed as the ratio of NSPS to SL. However several previous studies also reported numerous QTL clusters simultaneously affecting various spike-related traits and involving major genes (Cui et al., 2012; Fan et al., 2019; Heidari et al., 2011; Zhai et al., 2016). In the case of kernel-related traits, MTA for KW (*IWA6659*) was found to have a pleiotropic relationship with other kernel and spike traits including TKW and SL (Supplemental Table S4), and multi-trait QTL have also been previously reported for grain yield, TKW, KW, NSPS, SL, and NKPS in wheat (Chen et al., 2016; Patil et al., 2013; Prashant et al., 2012; Wu et al., 2015). Pleiotropic effects may also partially be explained by correlations between agronomic traits (Chen et al., 2016; Kumar et al., 2016; Liu et al., 2017b), as observed in our study (Figure 2).

### 4.4 | Additive effect of favorable alleles for wheat yield enhancement

Integrating multiple favorable loci into the same genetic background can significantly optimize plant traits (Fan et al., 2019; Li et al., 2021; You et al., 2021), which could be an effective strategy for modern wheat improvement. In our study, favorable alleles showed significant additive effects on the phenotype of four spike-related traits (Figure 4). The lines car-

rying a combination of favorable alleles showed significantly greater trait performance than other lines without the alleles, indicating the importance of combining favorable alleles to improve the performance of yield-contributing traits in wheat. Our study showed that the germplasm with three or more favorable alleles of identified QTL significantly improved the NSPS count compared with germplasm with only one favorable allele (Figure 4). Several of the desirable and undesirable lines, based on the highest and lowest phenotypic value for spike-related traits, were compared to examine the cumulative allele effect on the individual germplasm phenotypes (Table 3). The line with the highest NSPS in our HWWAMP was MT9982 (19.87) followed by OK05108 (19.23), which had all four favorable alleles from four stable MTAs, whereas lines with only one favorable allele, such as TAM400 (12.80), TAM109 (13.27), and HV906-865 (13.27), were among the worst (Table 3). Similarly, an additive effect of favorable alleles was observed for other traits (SL, SD, and NKPS) as well, such as a line with the highest mean SL in all the environments was OK1067274 (9.57 cm), which had all the three favorable alleles, whereas line MT85200 (6.49 cm), with the shortest SL, had only one favorable allele (Table 3). In our study, we also observed that many released varieties and advanced breeding lines from various breeding programs have a better combination of favorable alleles from stable MTAs detected for multiple spike-related traits. For instance, one advanced line OK05108 was among the best for NSPS and SL, with all the seven favorable alleles from all stable MTAs detected for these two traits (Table 3).

In addition, there are many released cultivars from both before ('Hail', 'Rose', and 'Cossack') and in the early 21st century ('Neosho' 'RonL', and 'SmokyHill') that showed good phenotypic performance and were found to have a good combination of favorable alleles for multiple spike-related traits (Table 3; Supplemental Table S1). Moreover, we observed that released cultivars and advanced breeding lines from recent decades have accumulated favorable alleles for various traits and this germplasm exhibited higher trait mean values than most of the cultivars released prior 1980s (Table 3). These findings suggest that continuous efforts by wheat breeders have resulted in the development of lines with a higher number of favorable alleles over time resulting in increased trait means. Such breeding lines and cultivars with a higher number of favorable alleles for multiple traits could serve as a valuable resource for the development of new high-yielding varieties in the hard winter wheat breeding programs.

### 4.5 | Spatial distribution of favorable alleles of stable MTAs

The spatial distribution of favorable alleles of stable MTAs was investigated by grouping 297 HWWAMP accessions

**TABLE 3** List of undesirable and desirable lines carrying low and high number of favorable alleles, respectively, for spike-related traits

Trait	Entry name	No. of favorable alleles	Trait mean	Year of release	Origin <sup>a</sup>	Trait	Entry name	No. of favorable alleles	Trait mean	Year of release	Origin
NSPS	CUTTER	1	13.33	2002	KS	SD	SAGE	0	1.73	1973	KS
	TAM400	1	12.8	2001	TX		CO04025	0	1.78	–	CO
	TAMW-101	1	13.53	1971	TX		LARNED	0	1.79	1976	KS
	TAM109	1	13.27	1991	TX		EAGLE	0	1.79	1970	KS
	HV906-865	1	13.27	–	KS		NORKAN	0	1.83	1986	KS
	TX86A6880	4	18.43	–	TX		RAWHIDE	3	2.32	1990	NE
	MT9904	2	18.53	–	MT		ROSE	3	2.37	1979	SD
	MT9513	3	18.9	–	MT		MT9513	3	2.39	–	MT
	OK05108	4	19.23	–	OK		COSSACK	3	2.46	1998	KS
	MT9982	4	19.87	–	MT		MT9904	3	2.53	–	MT
SL	MT85200	1	6.49	–	MT	NKPS	TANDEM	1	26.45	1997	SD
	HOMESTEAD	1	6.69	1973	NE		HOMESTEAD	1	27.55	1973	NE
	HV906-865	1	6.72	–	KS		G1878	1	28.08	–	KS
	OVERLEY	0	6.91	2004	KS		TRIUMPH64	1	28.6	1964	OK
	TRIUMPH64	2	6.95	1964	OK		AGATE	1	28.92	1979	NE
	OK05108	3	9.34	–	OK		RONL	2	49.65	2007	KS
	NEOSHO	3	9.42	2006	KS		TX06V7266	2	50.45	–	TX
	GUYMON	3	9.49	2005	OK		SMOKYHILL	1	51.17	2006	KS
	HAIL	3	9.51	1982	CO		TX05V7269	2	51.23	–	TX
	OK1067274	3	9.57	–	OK		HV9W06-504	2	54.33	–	KS

Note. NSPS, number of spikelets per spike; SL, spike length; SD, spikelet density; NKPS, number of kernels per spike.

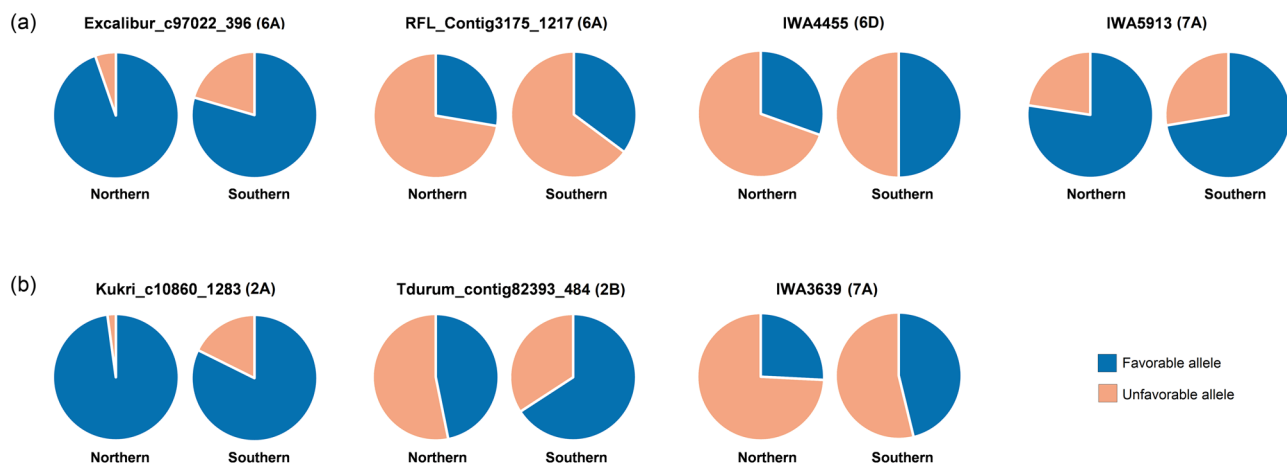
<sup>a</sup>Origin represents the breeding program

based on their region of origin. We grouped the accessions into two broad groups: Northern and Southern Great Plains. The Northern group comprised accessions originating from the northern states including Montana, Nebraska, North Dakota, and South Dakota, whereas the Southern group included accessions from Colorado, Kansas, Oklahoma, and Texas. Out of the four stable genomic regions identified for NSPS, the MTA represented by *Excalibur\_c97022\_396* was observed to be fixed in germplasm from the Northern region and also exhibited high frequency in germplasm from the Southern region. This genomic region has been previously found to be associated with NSPS in multiple environments (Sun et al., 2017), suggesting a presence of stable QTL for this trait that might be effective in a broad range of environments. Apart from this, favorable allele for another NSPS-associated MTA corresponding to the location of *WAP01* was found to be accumulated in the majority of accessions from Northern and Southern regions. However, there is still an opportunity to use *WAP01* in the Great Plains region, as about 25% of germplasm did not have the favorable allele (Figure 5). In addition to these two MTAs for NSPS, another genomic region represented by *IWA4456* was observed to be better distributed in germplasm from the Southern region, suggesting its role in adaptation to the Southern region.

Out of three stable MTAs for SL, the region represented by *Kukri\_c10860\_1283* was observed to be almost fixed in the Northern germplasm along with a high frequency of the favorable allele in the Southern germplasm, suggesting the importance of this genomic region in the Great Plains. In contrast, MTAs represented by *Tdurum\_contig82393\_484* and *IWA3639* were observed to be more frequent in germplasm from the Southern region (Figure 5).

#### 4.6 | Putative candidate genes for several important MTAs

Fourteen putative candidate genes were identified in the regions surrounding five MTAs ( $-\log_{10} P \geq 4.0$ ) detected for different spike-related traits (NSPS, SL, and SD) based on their potential involvement and higher expression in the shoot, spike, and kernel development at vegetative and reproductive stages of wheat (Table 2). Gene *TraesCS6A01G068900* was identified in the vicinity of an MTA for NSPS on chromosome 6A that is annotated to encode for the protein peptidyl-prolyl *cis-trans* isomerase (PPIase), which has a significant role in the deposition of storage proteins in wheat (Dutta et al., 2011). It was also found that PPIase activity



**FIGURE 5** Pie charts showing the frequency distribution of favorable alleles of stable marker–trait associations (MTAs) in the germplasm from two broad wheat growing regions, Northern and Southern Great Plains for the two spike-related traits: (a) number of spikelets per spike (NSPS, count) and (b) spike length (SL, cm). The ID of the associated SNP is given on the top of the pie charts with respective chromosome numbers in parenthesis

in the wheat was regulated by the developmental stages and was also cultivar dependent (Dutta et al., 2011). Similarly, *TraesCS6A01G069500* encodes for cold shock protein (CSP), a family known for its role in cold acclimation, gene expression regulation, and developmental processes such as flower and seed development, among others (Behl et al., 2020). A field experiment conducted by Yu et al. (2017) found that the cold shock protein *SeCspA* significantly improved TKW and grain yield in wheat.

Several putative candidate genes were identified for SL encoding for proteins such as NADPH-cytochrome P450 reductase, glycoprotein membrane glycosylphosphatidylinositol-anchored, and others (Table 2), which might play a significant role in multiple processes of plant growth and development (Pandian et al., 2020). For instance, *TaMsl1* (glycosylphosphatidylinositol-anchored lipid transfer protein) is a wheat fertility gene, was found to express in pollen development and encodes a GPI-LTP targeted to the plasma membrane (Kouidri et al., 2018). Protein ubiquitination is a major posttranslational modification that occurs in eukaryotes and regulates diverse biological processes such as lysine ubiquitination in common wheat, which mediates diverse cellular processes including carbohydrate metabolism, signal transduction, and photosynthesis (Zhang et al., 2017). Therefore, these putative genes may provide importance insights for understanding the underlying genetic mechanism of spike development in common wheat.

In conclusion, the current study showed that the HWWAMP is a valuable source for exploiting genetic variation for various yield-contributing traits. The negative relationships between spike- and kernel-related traits indicate that the loci controlling these traits are typically in an inverse relationship, leading to a moderate physiological trade-off between primary grain yield components. Genome-wide

association studies effectively identified both stable and environment-specific QTL for various spike- and kernel-related traits. The SNP markers linked to these stable QTL identified in this study could be useful for marker-assisted breeding in the hard winter wheat breeding programs and could also be incorporated into multivariate genomic selection models for the enhancement of yield-related traits in wheat.

#### DATA AVAILABILITY STATEMENT

The genotyping dataset analyzed during the current study is available online (<https://wheat.triticeaetoolbox.org/breeders/trial/2772>). The data generated during this study are included in the Supplemental information file.

#### ACKNOWLEDGMENTS

The authors would like to thank the South Dakota Agriculture Experimental Station (Brookings, SD, USA) for providing the resources to conduct the experiments. The authors are also thankful to Cody Hall and Navreet Brar for their efforts in planting and harvesting the experimental trials. This project was collectively funded by the USDA hatch projects SD00H695-20, and the USDA Agriculture and Food Research Initiative Competitive Grants 2022-68013-36439 (Wheat-CAP) from the USDA National Institute of Food and Agriculture and South Dakota Wheat Commission grant 3×1340. The funders had no role in the study design, data collection, analysis, decision to publish, or manuscript preparation.

#### AUTHOR CONTRIBUTIONS

Jyotirmoy Halder: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review &

editing. Harsimardeep Gill: Data curation, Formal analysis, Investigation, Software, Visualization, Writing – review & editing. Jinfeng Zhang: Data curation, Investigation, Writing – review & editing. Rami Altameemi: Investigation. Eric Olson: Writing – review & editing. Brent Turnipseed: Writing – review & editing. Sunish Sehgal: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

## 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.

## ORCID

Sunish K. Sehgal  <https://orcid.org/0000-0001-5644-7547>

## REFERENCES

- Alqudah, A. M., Sallam, A., Stephen Baenziger, P., & Börner, A. (2020). GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: Lessons from barley – A review. *Journal of Advanced Research*, 22, 119–135. <https://doi.org/10.1016/j.jare.2019.10.013>
- Altameemi, R., Gill, H. S., Ali, S., Ayana, G., Halder, J., Sidhu, J. S., Gill, U. S., Turnipseed, B., Hernandez, J. L. G., & Sehgal, S. K. (2021). Genome-wide association analysis permits characterization of *Stagonospora nodorum* blotch (SNB) resistance in hard winter wheat. *Scientific Reports*, 11, 12570. <https://doi.org/10.1038/s41598-021-91515-6>
- Alvarado, G., Rodríguez, F. M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., Pérez-Rodríguez, P., & Lopez-Cruz, M. A. (2020). META-R: A software to analyze data from multi-environment plant breeding trials. *Crop Journal*, 8, 745–756. <https://doi.org/10.1016/j.cj.2020.03.010>
- Anuarbek, S., Abugalieva, S., Pecchioni, N., Laidā<sup>2</sup>, G., Maccaferri, M., Tuberosa, R., & Turuspekov, Y. (2020). Quantitative trait loci for agronomic traits in tetraploid wheat for enhancing grain yield in Kazakhstan environments. *PLoS ONE*, 15, e0234863. <https://doi.org/10.1371/JOURNAL.PONE.0234863>
- Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., Thurber, C., Merrill, K. R., Foresman, B. J., & Kolb, F. L. (2016). Genome-wide association mapping of *Fusarium* head blight resistance in wheat using genotyping-by-sequencing. *The Plant Genome*, 9. <https://doi.org/10.3835/plantgenome2015.04.0028>
- Ayana, G. T., Ali, S., Sidhu, J. S., Gonzalez Hernandez, J. L., Turnipseed, B., & Sehgal, S. K. (2018). Genome-wide association study for spot blotch resistance in hard winter wheat. *Frontiers in Plant Science*, 9, 926. <https://doi.org/10.3389/fpls.2018.00926>
- Bednarek, J., Boulaflous, A., Girousse, C., Ravel, C., Tassy, C., Barret, P., Bouzidi, M. F., & Mouzeyar, S. (2012). Down-regulation of the *TaGW2* gene by RNA interference results in decreased grain size and weight in wheat. *Journal of Experimental Botany*, 63, 5945–5955. <https://doi.org/10.1093/jxb/ers249>
- Behl, A., Kumar, V., Shevtsov, M., & Singh, S. (2020). Pleiotropic roles of cold shock proteins with special emphasis on unexplored cold shock protein member of *Plasmodium falciparum*. *Malaria Journal*, 19, 382. <https://doi.org/10.1186/s12936-020-03448-6>
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23, 2633–2635. <https://doi.org/10.1093/bioinformatics/btm308>
- Breseghele, F., & Sorrells, M. E. (2007). QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. *Field Crops Research*, 101, 172–179. <https://doi.org/10.1016/j.fcr.2006.11.008>
- Bustos, D. V., Hasan, A. K., Reynolds, M. P., & Calderini, D. F. (2013). Combining high grain number and weight through a DH-population to improve grain yield potential of wheat in high-yielding environments. *Field Crops Research*, 145, 106–115. <https://doi.org/10.1016/j.fcr.2013.01.015>
- Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G. L., Akhunova, A., See, D., Bai, G., Pumphrey, M., Tomar, L., Wong, D., Kong, S., Reynolds, M., Da Silva, M. L., Bockelman, H., ... Akhunov, E. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proceedings of the National Academy of Sciences*, 110, 8057–8062. <https://doi.org/10.1073/pnas.1217133110>
- Chang, J., Zhang, J., Mao, X., Li, A., Jia, J., & Jing, R. (2013). Polymorphism of *TaSAP1-A1* and its association with agronomic traits in wheat. *Planta*, 237, 1495–1508. <https://doi.org/10.1007/s00425-013-1860-x>
- Chen, G., Zhang, H., Deng, Z., Wu, R., Li, D., Wang, M., & Tian, J. (2016). Genome-wide association study for kernel weight-related traits using SNPs in a Chinese winter wheat population. *Euphytica*, 212, 173–185. <https://doi.org/10.1007/s10681-016-1750-y>
- Chen, J., Zhang, F., Zhao, C., Lv, G., Sun, C., Pan, Y., Guo, X., & Chen, F. (2019). Genome-wide association study of six quality traits reveals the association of the *TaRPP13L1* gene with flour colour in Chinese bread wheat. *Plant Biotechnology Journal*, 17, 2106–2122. <https://doi.org/10.1111/pbi.13126>
- Cui, F., Ding, A., Li, J., Zhao, C., Wang, L., Wang, X., Qi, X., Li, X., Li, G., Gao, J., & Wang, H. (2012). QTL detection of seven spike-related traits and their genetic correlations in wheat using two related RIL populations. *Euphytica*, 186, 177–192. <https://doi.org/10.1007/s10681-011-0550-7>
- Cui, F., Zhao, C., Ding, A., Li, J., Wang, L., Li, X., Bao, Y., Li, J., & Wang, H. (2014). Construction of an integrative linkage map and QTL mapping of grain yield-related traits using three related wheat RIL populations. *Theoretical and Applied Genetics*, 127, 659–675. <https://doi.org/10.1007/s00122-013-2249-8>
- Dhakar, S., Liu, X., Chu, C., Yang, Y., Rudd, J. C., & Ibrahim, A. M. H. (2021). Genome-wide QTL mapping of yield and agronomic traits in two widely adapted winter wheat cultivars from multiple mega-environments. *PeerJ*, 9, e12350. <https://doi.org/10.7717/PEERJ.12350/SUPP-2>
- Dong, Y., Liu, J., Zhang, Y., Geng, H., Rasheed, A., Xiao, Y., Cao, S., Fu, L., Yan, J., Wen, W., Zhang, Y., Jing, R., Xia, X., & He, Z. (2016). Genome-wide association of stem water soluble carbohydrates in bread wheat. *PLoS ONE*, 11, e0164293. <https://doi.org/10.1371/JOURNAL.PONE.0164293>
- Dutta, T., Kaur, H., Singh, S., Mishra, A., Tripathi, J. K., Singh, N., Pareek, A., & Singh, P. (2011). Developmental changes in storage proteins and peptidyl prolyl cis-trans isomerase activity in grains of different wheat cultivars. *Food Chemistry*, 128, 450–457. <https://doi.org/10.1016/j.foodchem.2011.03.052>
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and

- implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–20. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fan, X., Cui, F., Ji, J., Zhang, W., Zhao, X., Liu, J., Meng, D., Tong, Y., Wang, T., & Li, J. (2019). Dissection of pleiotropic QTL regions controlling wheat spike characteristics under different nitrogen treatments using traditional and conditional QTL mapping. *Frontiers in Plant Science*, 10, 187. <https://doi.org/10.3389/fpls.2019.00187>
- Gao, F., Wen, W., Liu, J., Rasheed, A., Yin, G., Xia, X., Wu, X., & He, Z. (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. *Frontiers in Plant Science*, 6, 1099. <https://doi.org/10.3389/fpls.2015.01099>
- García, G. A., Hasan, A. K., Puhl, L. E., Reynolds, M. P., Calderini, D. F., & Miralles, D. J. (2013). Grain yield potential strategies in an elite wheat double-haploid population grown in contrasting environments. *Crop Science*, 53, 2577–2587. <https://doi.org/10.2135/cropsci2012.11.0669>
- Garcia, M., Eckermann, P., Haefele, S., Satija, S., Sznajder, B., Timmins, A., Baumann, U., Wolters, P., Mather, D. E., & Fleury, D. (2019). Genome-wide association mapping of grain yield in a diverse collection of spring wheat (*Triticum aestivum* L.) evaluated in southern Australia. *PLoS ONE*, 14, e0211730. <https://doi.org/10.1371/journal.pone.0211730>
- Gill, H. S., Halder, J., Zhang, J., Rana, A., Kleinjan, J., Amand, P. St., Bernardo, A., Bai, G., & Sehgal, S. K. (2022). Whole-genome analysis of hard winter wheat germplasm identifies genomic regions associated with spike and kernel traits. *Theoretical and Applied Genetics*, 135, 2953–2967. <https://doi.org/10.1007/s00122-022-04160-6>
- Gol, L., Haraldsson, E. B., & Von Korff, M. (2021). Ppd-H1 integrates drought stress signals to control spike development and flowering time in barley. *Journal of Experimental Botany*, 72, 122–136. <https://doi.org/10.1093/jxb/eraa261>
- Guan, P., Shen, X., Mu, Q., Wang, Y., Wang, X., Chen, Y., Zhao, Y., Chen, X., Zhao, A., Mao, W., Guo, Y., Xin, M., Hu, Z., Yao, Y., Ni, Z., Sun, Q., & Peng, H. (2020). Dissection and validation of a QTL cluster linked to *Rht-B1* locus controlling grain weight in common wheat (*Triticum aestivum* L.) using near-isogenic lines. *Theoretical and Applied Genetics*, 133, 2639–2653. <https://doi.org/10.1007/s00122-020-03622-z>
- Guo, Z., Chen, D., Alqudah, A. M., Röder, M. S., Ganai, M. W., & Schnurbusch, T. (2017). Genome-wide association analyses of 54 traits identified multiple loci for the determination of floret fertility in wheat. *New Phytologist*, 214, 257–270. <https://doi.org/10.1111/nph.14342>
- Guttieri, M. J., Baenziger, P. S., Frels, K., Carver, B., Arnall, B., & Waters, B. M. (2015). Variation for grain mineral concentration in a diversity panel of current and historical great plains hard winter wheat germplasm. *Crop Science*, 55, 1035–1052. <https://doi.org/10.2135/cropsci2014.07.0506>
- Hai, L., Guo, H., Wagner, C., Xiao, S., & Friedt, W. (2008). Plant science genomic regions for yield and yield parameters in Chinese winter wheat (*Triticum aestivum* L.) genotypes tested under varying environments correspond to QTL in widely different wheat materials. *Plant Science*, 175, 226–232. <https://doi.org/10.1016/j.plantsci.2008.03.006>
- Halder, J., Zhang, J., Ali, S., Sidhu, J. S., Gill, H. S., Talukder, S. K., Kleinjan, J., Turnipseed, B., & Sehgal, S. K. (2019). Mining and genomic characterization of resistance to tan spot, *Stagonospora nodorum* blotch (SNB), and Fusarium head blight in Watkins core collection of wheat landraces. *BMC Plant Biology*, 19, 480. <https://doi.org/10.1186/s12870-019-2093-3>
- Heidari, B., Sayed-Tabatabaei, B. E., Saeidi, G., Kearsey, M., & Suenaga, K. (2011). Mapping QTL for grain yield, yield components, and spike features in a doubled haploid population of bread wheat. *Genome*, 54, 517–527. <https://doi.org/10.1139/g11-017>
- Hou, J., Jiang, Q., Hao, C., Wang, Y., Zhang, H., & Zhang, X. (2014). Global selection on sucrose synthase haplotypes during a century of wheat breeding. *Plant Physiology*, 164, 1918–1929. <https://doi.org/10.1104/pp.113.232454>
- Hu, J., Wang, X., Zhang, G., Jiang, P., Chen, W., Hao, Y., Ma, X., Xu, S., Jia, J., Kong, L., & Wang, H. (2020). QTL mapping for yield-related traits in wheat based on four RIL populations. *Theoretical and Applied Genetics*, 133, 917–933. <https://doi.org/10.1007/s00122-019-03515-w>
- Huang, X. Q., Cloutier, S., Lycar, L., Radovanovic, N., Humphreys, D. G., Noll, J. S., Somers, D. J., & Brown, P. D. (2006). Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 113, 753–766. <https://doi.org/10.1007/s00122-006-0346-7>
- Jaiswal, V., Gahlaut, V., Mathur, S., Agarwal, P., Khandelwal, M. K., Khurana, J. P., Tyagi, A. K., Balyan, H. S., & Gupta, P. K. (2015). Identification of novel SNP in promoter sequence of *TaGW2-6A* associated with grain weight and other agronomic traits in wheat (*Triticum aestivum* L.). *PLoS ONE*, 10, e0129400. <https://doi.org/10.1371/journal.pone.0129400>
- Jantauriyarat, C., Vales, M. I., Watson, C. J. W., & Riera-Lizarazu, O. (2004). Identification and mapping of genetic loci affecting the free-threshing habit and spike compactness in wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 108, 261–273. <https://doi.org/10.1007/s00122-003-1432-8>
- Ji, X., Shiran, B., Wan, J., Lewis, D. C., Jenkins, C. L. D., Condon, A. G., Richards, R. A., & Dolferus, R. (2010). Importance of pre-anthesis anther sink strength for maintenance of grain number during reproductive stage water stress in wheat. *Plant, Cell & Environment*, 33, 926–942. <https://doi.org/10.1111/j.1365-3040.2010.02130.x>
- Jin, J., Liu, D., Qi, Y., Ma, J., & Zhen, W. (2020). Major QTL for seven yield-related traits in common wheat (*Triticum aestivum* L.). *Frontiers in Genetics*, 11, 1012. <https://doi.org/10.3389/fgene.2020.01012>
- Johnson, E. B., Nalam, V. J., Zemetra, R. S., & Riera-Lizarazu, O. (2008). Mapping the compactum locus in wheat (*Triticum aestivum* L.) and its relationship to other spike morphology genes of the Triticeae. *Euphytica*, 163, 193–201. <https://doi.org/10.1007/s10681-007-9628-7>
- Juliana, P., Singh, R. P., Singh, P. K., Poland, J. A., Bergstrom, G. C., Huerta-Espino, J., Bhavani, S., Crossa, J., & Sorrells, M. E. (2018). Genome-wide association mapping for resistance to leaf rust, stripe

- rust and tan spot in wheat reveals potential candidate genes. *Theoretical and Applied Genetics*, *131*, 1405–1422. <https://doi.org/10.1007/s00122-018-3086-6>
- Kato, K., Miura, H., & Sawada, S. (2000). Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. *Theoretical and Applied Genetics*, *101*, 1114–1121. <https://doi.org/10.1007/s001220051587>
- Koudiri, A., Baumann, U., Okada, T., Baes, M., Tucker, E. J., & Whitford, R. (2018). Wheat TaMs1 is a glycosylphosphatidylinositol-anchored lipid transfer protein necessary for pollen development. *BMC Plant Biology*, *18*, 332. <https://doi.org/10.1186/s12870-018-1557-1>
- Kumar, A., Mantovani, E. E., Seetan, R., Soltani, A., Echeverryâ Solarte, M., Jain, S., Simsek, S., Doehlert, D., Alamri, M. S., Elias, E. M., Kianian, S. F., & Mergoum, M. (2016). Dissection of genetic factors underlying wheat kernel shape and size in an elite × nonadapted cross using a high density SNP linkage map. *The Plant Genome*, *9*, plantgenome2015.09.0081. <https://doi.org/10.3835/plantgenome2015.09.0081>
- Kuzay, S., Lin, H., Li, C., Chen, S., Woods, D. P., Zhang, J., Lan, T., Von Korff, M., & Dubcovsky, J. (2022). WAPO-A1 is the causal gene of the 7AL QTL for spikelet number per spike in wheat. *PLoS Genetics*, *18*, e1009747. <https://doi.org/10.1371/JOURNAL.PGEN.1009747>
- Kuzay, S., Xu, Y., Zhang, J., Katz, A., Pearce, S., Su, Z., Fraser, M., Anderson, J. A., Brown-Guedira, G., Dewitt, N., Peters Haugrud, A., Faris, J. D., Akhunov, E., Bai, G., & Dubcovsky, J. (2019). Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theoretical and Applied Genetics*, *132*, 2689–2705. <https://doi.org/10.1007/s00122-019-03382-5>
- Li, F., Wen, W., Liu, J., Zhang, Y., Cao, S., He, Z., Rasheed, A., Jin, H., Zhang, C., Yan, J., Zhang, P., Wan, Y., & Xia, X. (2019). Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biology*, *19*, 168. <https://doi.org/10.1186/s12870-019-1781-3>
- Li, T., Deng, G., Su, Y., Yang, Z., Tang, Y., Wang, J., Qiu, X., Pu, X., Li, J., Liu, Z., Zhang, H., Liang, J., Yang, W., Yu, M., Wei, Y., & Long, H. (2021). Identification and validation of two major QTLs for spike compactness and length in bread wheat (*Triticum aestivum* L.) showing pleiotropic effects on yield-related traits. *Theoretical and Applied Genetics*, *134*, 3625–3641. <https://doi.org/10.1007/s00122-021-03918-8>
- Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., & Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics*, *28*, 2397–2399. <https://doi.org/10.1093/bioinformatics/bts444>
- Liu, H., Ma, J., Tu, Y., Zhu, J., Ding, P., Liu, J., Li, T., Zou, Y., Habib, A., Mu, Y., Tang, H., Jiang, Q., Liu, Y., Chen, G., Zheng, Y., Wei, Y., & Lan, X. (2020). Several stably expressed QTL for spike density of common wheat (*Triticum aestivum*) in multiple environments. *Plant Breeding*, *139*, 284–294. <https://doi.org/10.1111/PBR.12782>
- Liu, J., He, Z., Rasheed, A., Wen, W., Yan, J., & Zhang, P. (2017a). Genome-wide association mapping of black point reaction in common wheat (*Triticum aestivum* L.). *BMC Plant Biology*, *17*, 220. <https://doi.org/10.1186/s12870-017-1167-3>
- Liu, J., Xu, Z., Fan, X., Zhou, Q., Cao, J., Wang, F., Ji, G., Yang, L., Feng, B., & Wang, T. (2018a). A genome-wide association study of wheat spike related traits in China. *Frontiers in Plant Science*, *871*, 1584. <https://doi.org/10.3389/fpls.2018.01584>
- Liu, K., Sun, X., Ning, T., Duan, X., Wang, Q., Liu, T., An, Y., Guan, X., Tian, J., & Chen, J. (2018b). Genetic dissection of wheat panicle traits using linkage analysis and a genome-wide association study. *Theoretical and Applied Genetics*, *131*, 1073–1090. <https://doi.org/10.1007/s00122-018-3059-9>
- Liu, X., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genetics*, *12*, e1005767. <https://doi.org/10.1371/journal.pgen.1005767>
- Liu, Y., Lin, Y., Gao, S., Li, Z., Ma, J., Deng, M., Chen, G., Wei, Y., & Zheng, Y. (2017b). A genome-wide association study of 23 agronomic traits in Chinese wheat landraces. *Plant Journal*, *91*, 861–873. <https://doi.org/10.1111/tpj.13614>
- Lozada, D. N., Mason, R. E., Babar, M. A., Carver, B. F., Guedira, G. B., & Merrill, K. (2017). Association mapping reveals loci associated with multiple traits that affect grain yield and adaptation in soft winter wheat. *Euphytica*, *213*, 222. <https://doi.org/10.1007/S10681-017-2005-2>
- Ma, L., Li, T., Hao, C., Wang, Y., Chen, X., & Zhang, X. (2016). *TaGS5-3A*, a grain size gene selected during wheat improvement for larger kernel and yield. *Plant Biotechnology Journal*, *14*, 1269–1280. <https://doi.org/10.1111/pbi.12492>
- Ma, Z., Zhao, D., Zhang, C., Zhang, Z., Xue, S., Lin, F., Kong, Z., Tian, D., & Luo, Q. (2007). Molecular genetic analysis of five spike-related traits in wheat using RIL and immortalized F<sub>2</sub> populations. *Molecular Genetics and Genomics*, *277*, 31–42. <https://doi.org/10.1007/s00438-006-0166-0>
- Molero, G., Joynson, R., Pineraâ Chavez, F. J., Gardiner, L. J., Riveraâ Amado, C., Hall, A., & Reynolds, M. P. (2019). Elucidating the genetic basis of biomass accumulation and radiation use efficiency in spring wheat and its role in yield potential. *Plant Biotechnology Journal*, *17*, 1276–1288. <https://doi.org/10.1111/pbi.13052>
- Muhammad, A., Hu, W., Li, Z., Li, J., Xie, G., Wang, J., & Wang, L. (2020). Appraising the genetic architecture of kernel traits in hexaploid wheat using GWAS. *International Journal of Molecular Sciences*, *21*, 5649. <https://doi.org/10.3390/ijms21165649>
- Mwadingeni, L., Shimelis, H., Rees, D. J. G., & Tsilo, T. J. (2017). Genome-wide association analysis of agronomic traits in wheat under drought-stressed and non-stressed conditions. *PLoS ONE*, *12*, e0171692. <https://doi.org/10.1371/journal.pone.0171692>
- Nadolska-Orczyk, A., Rajchel, I. K., Orczyk, W., & Gasparis, S. (2017). Major genes determining yield-related traits in wheat and barley. *Theoretical and Applied Genetics*, *130*, 1081–1098. <https://doi.org/10.1007/s00122-017-2880-x>
- Naruoka, Y., Talbert, L. E., Lanning, S. P., Blake, N. K., Martin, J. M., & Sherman, J. D. (2011). Identification of quantitative trait loci for productive tiller number and its relationship to agronomic traits in spring wheat. *Theoretical and Applied Genetics*, *123*, 1043–1053. <https://doi.org/10.1007/s00122-011-1646-0>
- Navrotskyi, S., Belamkar, V., Baenziger, P. S., & Rose, D. J. (2020). Insights into the genetic architecture of bran friability and water retention capacity, two important traits for whole grain end-use quality in winter wheat. *Genes*, *11*, 838. <https://doi.org/10.3390/GENES11080838>
- Neumann, K., Kobiljski, B., Dena, S., Varshney, R. K., & B rner, A. (2011). Genome-wide association mapping: A case study in bread wheat (*Triticum aestivum* L.). *Molecular Breeding*, *27*, 37–58. <https://doi.org/10.1007/s11032-010-9411-7>



- Okada, T., Jayasinghe, J. E. A. R. M., Eckermann, P., Watson-Haigh, N. S., Warner, P., Hendrikse, Y., Baes, M., Tucker, E. J., Laga, H., Kato, K., Albertsen, M., Wolters, P., Fleury, D., Baumann, U., & Whitford, R. (2019). Effects of *Rht-B1* and *Ppd-D1* loci on pollinator traits in wheat. *Theoretical and Applied Genetics*, *132*, 1965–1979. <https://doi.org/10.1007/s00122-019-03329-w>
- Pal, N., Saini, D. K., Kumar, S., Pal, N., Saini, D. K., & Kumar, S. (2022). Breaking yield ceiling in wheat: Progress and future prospects. In M.-R. Ansari (Ed.), *Wheat*. <https://doi.org/10.5772/INTECHOPEN.102919>
- Pandian, B. A., Sathishraj, R., Djanaguiraman, M., Prasad, P. V. V., & Jugulam, M. (2020). Role of cytochrome P450 enzymes in plant stress response. *Antioxidants*, *9*, 454. <https://doi.org/10.3390/antiox9050454>
- Pang, Y., Liu, C., Wang, D., St Amand, P., Bernardo, A., Li, W., He, F., Li, L., Wang, L., Yuan, X., Dong, L., Su, Y., Zhang, H., Zhao, M., Liang, Y., Jia, H., Shen, X., Lu, Y., Jiang, H., ... Liu, S. (2020). High-resolution genome-wide association study identifies genomic regions and candidate genes for important agronomic traits in wheat. *Molecular Plant*, *13*, 1311–1327. <https://doi.org/10.1016/j.molp.2020.07.008>
- Patil, R. M., Tamhankar, S. A., Oak, M. D., Raut, A. L., Honrao, B. K., Rao, V. S., & Misra, S. C. (2013). Mapping of QTL for agronomic traits and kernel characters in durum wheat (*Triticum durum* Desf.). *Euphytica*, *190*, 117–129. <https://doi.org/10.1007/s10681-012-0785-y>
- Prashant, R., Kadoo, N., Desale, C., Kore, P., Dhaliwal, H. S., Chhuneja, P., & Gupta, V. (2012). Kernel morphometric traits in hexaploid wheat (*Triticum aestivum* L.) are modulated by intricate QTL × QTL and genotype × environment interactions. *Journal of Cereal Science*, *56*, 432–439. <https://doi.org/10.1016/j.jcs.2012.05.010>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Qu, X., Liu, J., Xie, X., Xu, Q., Tang, H., Mu, Y., Pu, Z., Li, Y., Ma, J., Gao, Y., Jiang, Q., Liu, Y., Chen, G., Wang, J., Qi, P., Habib, A., Wei, Y., Zheng, Y., Lan, X., & Ma, J. (2021). Genetic mapping and validation of loci for kernel-related traits in wheat (*Triticum aestivum* L.). *Frontiers in Plant Science*, *12*, 667493. <https://doi.org/10.3389/fpls.2021.667493>
- Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., & Venturini, L. (2018). The transcriptional landscape of polyploid wheat. *Science*, *361*, 6403. <https://doi.org/10.1126/SCIENCE.AAR6089>
- Ramya, P., Chaubal, A., Kulkarni, K., Gupta, L., Kadoo, N., Dhaliwal, H. S., Chhuneja, P., Lagu, M., & Gupta, V. (2010). QTL mapping of 1000-kernel weight, kernel length, and kernel width in bread wheat (*Triticum aestivum* L.). *Journal of Applied Genetics*, *51*, 421–429. <https://doi.org/10.1007/BF03208872>
- Rasheed, A., Xia, X., Ogonnaya, F., Mahmood, T., Zhang, Z., Mujeeb-Kazi, A., & He, Z. (2014). Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. *BMC Plant Biology*, *14*, 128. <https://doi.org/10.1186/1471-2229-14-128>
- Sadras, V. O. (2007). Evolutionary aspects of the trade-off between seed size and number in crops. *Field Crops Research*, *100*, 125–138. <https://doi.org/10.1016/j.fcr.2006.07.004>
- Shi, W., Hao, C., Zhang, Y., Cheng, J., Zhang, Z., Liu, J., Yi, X., Cheng, X., Sun, D., Xu, Y., Zhang, X., Cheng, S., Guo, P., & Guo, J. (2017). A combined association mapping and linkage analysis of kernel number per spike in common wheat (*Triticum aestivum* L.). *Frontiers in Plant Science*, *8*, 1412. <https://doi.org/10.3389/fpls.2017.01412>
- Sidhu, J. S., Singh, D., Gill, H. S., Brar, N. K., Qiu, Y., Halder, J., Al Tameemi, R., Turnipseed, B., & Sehgal, S. K. (2020). Genome-wide association study uncovers novel genomic regions associated with coleoptile length in hard winter wheat. *Frontiers in Genetics*, *10*, 1345. <https://doi.org/10.3389/fgene.2019.01345>
- Singh, K., Saini, D. K., Saripalli, G., Batra, R., Gautam, T., & Singh, R. (2022). WheatQTLdb V2.0: A supplement to the database for wheat QTL. *Molecular Breeding*, *42*, 56. <https://doi.org/10.1007/S11032-022-01329-1>
- Su, Q., Zhang, X., Zhang, W., Zhang, N., Song, L., Liu, L., Xue, X., Liu, G., Liu, J., Meng, D., Zhi, L., Ji, J., Zhao, X., Yang, C., Tong, Y., Liu, Z., & Li, J. (2018). QTL detection for kernel size and weight in bread wheat (*Triticum aestivum* L.) using a high-density SNP and SSR-based linkage map. *Frontiers in Plant Science*, *9*, 1484. <https://doi.org/10.3389/fpls.2018.01484>
- Sukumar, S., Dreisigacker, S., Lopes, M., Chavez, P., & 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. *Theoretical and Applied Genetics*, *128*, 353–363. <https://doi.org/10.1007/s00122-014-2435-3>
- Sun, C., Zhang, F., Yan, X., Zhang, X., Dong, Z., Cui, D., & Chen, F. (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 Biotechnology Journal*, *15*, 953–969. <https://doi.org/10.1111/pbi.12690>
- Tadesse, W., Sanchez-Garcia, M., Assefa, S. G., Amri, A., Bishaw, Z., & Ogonnaya, F. C. (2019). Genetic gains in wheat breeding and its role in feeding the world. *Crop Breeding, Genetics and Genomics*, *1*, e190005. <http://doi.org/10.20900/cbgs20190005>
- Tang, Y.-L., Li, J., Wu, Y.-Q., Wei, H.-T., Li, C.-S., Yang, W.-Y., & Chen, F. (2011). Identification of QTLs for yield-related traits in the recombinant inbred line population derived from the cross between a synthetic hexaploid wheat-derived variety Chuanmai 42 and a Chinese elite variety Chuannong 16. *Agricultural Sciences in China*, *10*, 1665–1680. [https://doi.org/10.1016/S1671-2927\(11\)60165-X](https://doi.org/10.1016/S1671-2927(11)60165-X)
- The International Wheat Genome Sequencing Consortium (IWGSC). Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., Pozniak, C. J., Choulet, F., Distelfeld, A., Poland, J., Ronen, G., Sharpe, A. G., Barad, O., Baruch, K., Keeble-Gagnère, G., Mascher, M., Ben-Zvi, G., Josselin, A.-A., ... Wang, L. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, *361*, eaar7191. <https://doi.org/10.1126/science.aar7191>
- Wang, S.-X., Zhu, Y.-L., Zhang, D.-X., Shao, H., Liu, P., Hu, J.-B., Zhang, H., Zhang, H.-P., Chang, C., Lu, J., Xia, X.-C., Sun, G.-L., & Ma, C.-X. (2017). Genome-wide association study for grain yield and related traits in elite wheat varieties and advanced lines using SNP markers. *PLoS ONE*, *12*, e0188662. <https://doi.org/10.1371/journal.pone.0188662>
- Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., Maccaferri, M., Salvi, S., Milner, S. G., Cattivelli, L., Mastrangelo, A. M., Whan, A., Stephen, S., Barker, G., Wieseke, R., Plieske, J., Lillmo, M., Mather, D., Appels, R., ... Akhunov, E. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnology Journal*, *12*, 787–796. <https://doi.org/10.1111/pbi.12183>

- Wang, S., Zhang, X., Chen, F., & Cui, D. (2015). A single-nucleotide polymorphism of *TaGS5* gene revealed its association with kernel weight in Chinese bread wheat. *Frontiers in Plant Science*, *6*, 1166. <https://doi.org/10.3389/fpls.2015.01166>
- Ward, B. P., Brown-Guedira, G., Kolb, F. L., Van Sanford, D. A., Tyagi, P., Sneller, C. H., & Griffey, C. A. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PLoS ONE*, *14*, e208217. <https://doi.org/10.1371/journal.pone.0208217>
- Whitechurch, E. M., & Snape, J. W. (2003). Developmental responses to vernalization in wheat deletion lines for chromosomes 5A and 5D. *Plant Breeding*, *122*, 35–39. <https://doi.org/10.1046/j.1439-0523.2003.00749.x>
- Wu, Q.-H., Chen, Y.-X., Zhou, S.-H., Fu, L., Chen, J.-J., Xiao, Y., Zhang, D., Ouyang, S.-H., Zhao, X.-J., Cui, Y., Zhang, D.-Y., Liang, Y., Wang, Z.-Z., Xie, J.-Z., Qin, J.-X., Wang, G.-X., Li, D.-L., Huang, Y.-L., Yu, M.-H., ... Liu, Z.-Y. (2015). High-density genetic linkage map construction and QTL mapping of grain shape and size in the wheat population Yanda1817 × Beinong6. *PLoS ONE*, *10*, e0118144. <https://doi.org/10.1371/journal.pone.0118144>
- Xie, Q., Li, N., Yang, Y., Lv, Y., Yao, H., Wei, R., Sparkes, D. L., & Ma, Z. (2018). Pleiotropic effects of the wheat domestication gene *Q* on yield and grain morphology. *Planta*, *247*, 1089–1098. <https://doi.org/10.1007/s00425-018-2847-4>
- Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., & Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene *VRN1*. *Proceedings of the National Academy of Sciences*, *100*, 6263–6268. <https://doi.org/10.1073/pnas.0937399100>
- Yao, H., Xie, Q., Xue, S., Luo, J., Lu, J., Kong, Z., Wang, Y., Zhai, W., Lu, N., Wei, R., Yang, Y., Han, Y., Zhang, Y., Jia, H., & Ma, Z. (2019). *HL2* on chromosome 7D of wheat (*Triticum aestivum* L.) regulates both head length and spikelet number. *Theoretical and Applied Genetics*, *132*, 1789–1797. <https://doi.org/10.1007/S00122-019-03315-2>
- You, J., Liu, H., Wang, S., Luo, W., Gou, L., Tang, H., Mu, Y., Deng, M., Jiang, Q., Chen, G., Qi, P., Peng, Y., Tang, L., Habib, A., Wei, Y., Zheng, Y., Lan, X., & Ma, J. (2021). Spike density quantitative trait loci detection and analysis in tetraploid and hexaploid wheat recombinant inbred line populations. *Frontiers in Plant Science*, *12*, 3389. <https://doi.org/10.3389/fpls.2021.796397>
- Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., & Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, *38*, 203–208. <https://doi.org/10.1038/ng1702>
- Yu, M., Mao, S.-L., Chen, G.-Y., Pu, Z.-E., Wei, Y.-M., & Zheng, Y.-L. (2014). QTLs for uppermost internode and spike length in two wheat RIL populations and their effect upon plant height at an individual QTL level. *Euphytica*, *200*, 95–108. <https://doi.org/10.1007/s10681-014-1156-7>
- Yu, T.-F., Xu, Z.-S., Guo, J.-K., Wang, Y.-X., Abernathy, B., Fu, J.-D., Chen, X., Zhou, Y.-B., Chen, M., Ye, X.-G., & Ma, Y.-Z. (2017). Improved drought tolerance in wheat plants overexpressing a synthetic bacterial cold shock protein gene *SeCspA*. *Scientific Reports*, *7*, 44050. <https://doi.org/10.1038/srep44050>
- Zanke, C. D., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Korzun, V., Argillier, O., Stiewe, G., Hinze, M., Neumann, F., Eichhorn, A., Polley, A., Jaenecke, C., Ganal, M. W., & RöDer, M. S. (2015). Analysis of main effect QTL for thousand grain weight in European winter wheat (*Triticum aestivum* L.) by genome-wide association mapping. *Frontiers in Plant Science*, *6*, 644. <https://doi.org/10.3389/fpls.2015.00644>
- Zhai, H., Feng, Z., Li, J., Liu, X., Xiao, S., Ni, Z., & Sun, Q. (2016). QTL analysis of spike morphological traits and plant height in winter wheat (*Triticum aestivum* L.) using a high-density SNP and SSR-based linkage map. *Frontiers in Plant Science*, *7*, 1617. <https://doi.org/10.3389/fpls.2016.01617>
- Zhang, B., Liu, X., Xu, W., Chang, J., Li, A., Mao, X., Zhang, X., & Jing, R. (2015). Novel function of a putative MOC1 ortholog associated with spikelet number per spike in common wheat. *Scientific Reports*, *5*, 12211. <https://doi.org/10.1038/srep12211>
- Zhang, J., Gill, H S., Halder, J., Brar, N K., Ali, S., Bernardo, A., Amand, P St., Bai, G., Turnipseed, B., & Sehgal, S K. (2022). Multi-locus genome-wide association studies to characterize Fusarium head blight (FHB) resistance in hard winter wheat. *Frontiers in Plant Science*, *13*, 946700. <https://doi.org/10.3389/FPLS.2022.946700>
- Zhang, L., Zhao, Y. L., Gao, L. F., Zhao, G. Y., Zhou, R. H., Zhang, B. S., & Jia, J. Z. (2012). *TaCKX6-D1*, the ortholog of rice *OsCKX2*, is associated with grain weight in hexaploid wheat. *New Phytologist*, *195*, 574–584. <https://doi.org/10.1111/j.1469-8137.2012.04194.x>
- Zhang, N., Zhang, L., Shi, C., Tian, Q., Lv, G., Wang, Y., Cui, D., & Chen, F. (2017). Comprehensive profiling of lysine ubiquitome reveals diverse functions of lysine ubiquitination in common wheat. *Scientific Reports*, *7*, 13601. <https://doi.org/10.1038/s41598-017-13992-y>
- Zhang, Y., Li, D., Zhang, D., Zhao, X., Cao, X., Dong, L., Liu, J., Chen, K., Zhang, H., Gao, C., & Wang, D. (2018). Analysis of the functions of *TaGW2* homoeologs in wheat grain weight and protein content traits. *Plant Journal*, *94*, 857–866. <https://doi.org/10.1111/TPJ.13903>
- Zheng, J., Liu, H., Wang, Y., Wang, L., Chang, X., Jing, R., Hao, C., & Zhang, X. (2014). *TEF-7A*, a transcript elongation factor gene, influences yield-related traits in bread wheat (*Triticum aestivum* L.). *Journal of Experimental Botany*, *65*, 5251–5265. <https://doi.org/10.1093/jxb/eru306>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Halder, J., Gill, H. S., Zhang, J., Altameemi, R., Olson, E., Turnipseed, B., & Sehgal, S. K. (2023). Genome-wide association analysis of spike and kernel traits in the U.S. hard winter wheat. *The Plant Genome*, *16*, e20300. <https://doi.org/10.1002/tpg2.20300>