

GWAS analysis revealed wheat growth characteristics and yield-related genes in the Yellow River Basin

WeiJun Zhang^{1†}, Jinshang He^{1†}, Xiaoliang Wang¹, Ling Kang¹, Yan Dong¹, Lina Cao¹, Fuguo Zhang², Dongsheng Chen^{1*} 

Crop Breeding and Applied Biotechnology
25(3): e51572539, 2025
Brazilian Society of Plant Breeding.
Printed in Brazil
<http://dx.doi.org/10.1590/1984-70332025v25n3a39>

Abstract: *Wheat is the main food crop in the Yellow River Basin of Northwest China. To explore the genetic mechanisms of wheat yield traits and identify related candidate genes, 250 varieties were studied. Genotyping was performed with the wheat 90K SNP chip, and genome-wide association analysis was carried out through multiple models. A total of 335 SNPs were significantly associated with yield-related traits, and 75 markers were simultaneously detected in at least 2 models. Of the 73 candidate genes screened, 4 had been reported in previous studies, and 69 were new. Among the significantly associated loci, 53 loci were significantly associated with multiple wheat traits, and a 5A_B1_Hap2 dominant haplotype was identified that can simultaneously regulate spikelet number, spikelet weight, thousand-grain weight, and effective tiller. This study provides effective loci for molecular marker-assisted selection breeding of wheat and cloning of yield-related trait genes.*

Keywords: *Wheat, yield-related traits, GWAS, haplotype blocks, candidate genes*

INTRODUCTION

Wheat (*Triticum aestivum* L.) is an important food crop in China, with approximately 17% of the world's wheat production over the past few years (<https://www.fao.org/faostat/zh/#compare>). With the adjustment of planting structure, the wheat planting area in northwest China, especially in the upper region of the Yellow River, has recently had a downward trend (Sinclair and Bai 1997, Li et al. 2019). The wheat industry in the upper reaches of the Yellow River should be revitalized by breeding new varieties. To achieve high yields, continuous improvement of wheat varieties is the main goal of wheat breeding (Kuzay et al. 2019). Wheat yield is influenced by multiple traits, such as effective tiller number per plant (ETN), kernel number per spike (KNS), kernel weight per spike (KWS), and thousand-grain weight (TKW). Its stability depends on adaptive factors like winter-spring characteristics (WS) and growth stage. Thus, analyzing yield-related agronomic traits and understanding the genetic structure of wheat are crucial for future wheat breeding programs (Pang et al. 2020, Alonso et al. 2021, Franco et al. 2022).

Recent advances in DNA sequencing technology, SNP detection, genotyping, and assembling genomes can be conducted in a high-throughput and cost-effective manner, which have greatly facilitated understanding of large crop genomes,



***Corresponding author:**
E-mail: cdsnky@163.com

Received: 11 February 2025
Accepted: 08 June 2025
Published: 26 August 2025

¹ Crop Research Institute, Ningxia Academy of Agriculture and Forestry Sciences, Yinchuan, 750002, Ningxia, China

² Ningxia Academy of Agriculture and Forestry Sciences, Yinchuan, 750002, Ningxia, China

especially of wheat, including *Triticum urartu*, wild emmer, durum wheat, and multiple varieties of hexaploid wheat (Xiao et al. 2022). Led by the International Wheat Genome Sequencing Consortium (IWGSC), scientists completed the world's first complete genome map of the hexaploid wheat "Chinese Spring" in 13 years (IWGSC 2018). With the continuous improvement of reference genomes, mapping technologies, such as genotyping and SNP array analysis, genome-wide association studies (GWAS), and linkage analysis continue to make breakthroughs (Li et al. 2019, Xiao et al. 2022). These advancements have facilitated the cloning of important wheat trait genes. Although about 150 functional markers have been developed, they still cannot meet the requirements of high-throughput genotyping and molecular identification (Rasheed et al. 2016, Scherlosky et al. 2018). GWAS serve as a foundation experiment in gene mapping of wheat, are an excellent tool to link trait analysis to its underlying genetics (Korte and Farlow 2013), and have played a pivotal role in the related research of wheat functional genomics in recent years (Li et al. 2019, Xiao et al. 2022).

Natural wheat varieties in populations show rich genetic diversity, which facilitates finding genes linked to yield and adaptability. Using multiple models can reduce false results in gene localization. This study conducted GWAS on 250 natural wheat germplasms with a 90K breeding chip. The objectives were to find loci related to yield traits and predict candidate genes. Detecting key SNP loci for yield could help boost wheat yield in the Yellow River Basin and clarify the molecular basis of yield formation.

MATERIAL AND METHODS

Plant materials and phenotyping

A total of 250 varieties consisting of wheat breeding accessions and foreign introduced accessions were used as test materials (Table S1). All materials were planted in Yongning County, Yinchuan City, China, from 2019 to 2020. Winter wheat was sown in October and harvested in June the following year. Spring wheat was planted in February and harvested in July. A random block design was adopted, with two rows, two repetitions, a length of 1 m, a row spacing of 20 cm, and 135 plants per row. Field management was carried out according to local conventional methods. Ten individual plants of each material were randomly selected to investigate their plant height (PLH), kernel number per spike (KNS), tiller number per plant (TIN), and effective tiller number per plant (ETN). Once dried naturally after threshing, thousand kernel weight (TKW) and kernel weight per spike (KWS) were measured. The wheat growing time (WGT) was calculated by subtracting the sowing date from the maturity date. The tested materials were divided into winter or spring (WS) types of wheat based on their vernalization requirement.

Genotyping

The DNA of seeds was extracted using the CTAB method (Wei et al. 2021). The DNA quality was assessed by 1.2% agarose gel electrophoresis. The DNA concentration was measured with a NanoDrop™ND-2000 spectrophotometer (Wang et al. 2014), and 250 wheat materials were scanned using the KPS Wheat 90K genotyping chip at the Beijing Compsom Biotechnology Co. Beijing, China. Affymetrix Power Tools were used for the original SNP typing of the samples. Markers were a filtration deletion rate of more than 10% and a minimum allele frequency of less than 5%. High-quality SNP markers were retained for subsequent analysis.

Population structure and estimation of linkage disequilibrium

The principal component analysis and Neighbor-joining evolutionary tree were estimated for the population structure through GCTA v1.94 (Yang et al. 2011), and finally plotted through the Matplotlib package for Python. The squared correlation coefficient between loci (r^2) was used as a parameter to measure the linkage disequilibrium between two polymorphic loci of populations, and r^2 was mainly calculated using Pop LD decay v3.41 (Zhang et al. 2019).

Genome-wide association analysis and haplotype block analysis

The GWAS was carried out by the GAPIT package in the R software (Lipka et al. 2012). To minimize false positives, a mixed linear (MLM) (Yu et al. 2006), the FarmCPU (Liu et al. 2016), CMLM (Zhang et al. 2010), and BLINK (Huang et al. 2019) models were employed. Because LD (Linkage disequilibrium) was extensive in this population, assuming that all tested markers were statistically independent and deriving a Bonferroni-corrected significance threshold would have

been too conservative. Therefore, we estimated the effective number of independent markers (Me) (Li et al. 2012) and derived a less-conservative threshold following $0.05/\text{Me}$ (1.0×10^{-4} equivalent to $-\log_{10}(P) = 4$). The SNPs were visualized using a Manhattan map and quantile–quantile (QQ) map in R (v4.0.4). Haplotype analyses of the significant SNPs were performed to delineate the candidate region utilizing Haploview software (Barrett et al. 2005).

Candidate gene identification

Candidate genes were identified according to their physical positions and functions based on the IWGSC RefSeq v1.0 (<http://www.wheatgenome.org/>), and the significant SNPs were annotated to the corresponding genes using ANNOVAR (Wang et al. 2010). GO enrichment was conducted to classify the function of candidate genes.

RESULTS AND DISCUSSION

Phenotypic data analysis

The descriptive statistics for the seven yield traits (WGT, PLH, TIN, ETN, KNS, TKW, and KWS) were listed in Table S2. Wide phenotypic differences were observed across all traits, with the coefficient of variation ranging from 14% to 35%. The yield-related traits of wheat exhibited nearly normal distribution (Figure S1a). The correlation analysis results (Table S3, Figure S1b) indicated a significant correlation between KNS and KWS ($r = 0.899$) along with TIN and ETN ($r = 0.844$). The WGT was negatively correlated with KNS and KWS.

Population structure and linkage disequilibrium

The 90 K wheat SNP array with 77,024 SNPs was used for genotyping. After quality control, 73,156 SNP markers were used for the association mapping. According to the results of PCA analysis (Figure S2a) and evolutionary tree analysis (Figure S2b), wheat could be divided into two subgroups according to winter wheat and spring wheat. Some of the samples were more complex, indicating that the gene exchange between the natural populations was relatively frequent.

The LD decay distance was calculated using PopLDdecay software. The LD decay diagram (Figure S2c, d) illustrates that the average LD decay distance in the population is 5 Mb. The LD attenuation distance was greater than the average distance between markers (0.28 Mb).

Genome-wide association analysis

Based on four different models (MLM, FarmCPU, BLINK, and CMLM), a total of 335 significantly associated loci with traits were detected, distributed across 21 chromosomes. The phenotypic variance (PVE) ranged from 3.58% to 14.96% (Table S4, Figure 1). Notably, 75 significant loci were detected by more than two models, and 53 loci were pleiotropic (significantly associated with multiple traits). For subsequent candidate gene prediction, 88 stable associated loci were identified as reliable trait-linked loci. The results from the four models mutually validated each other, enhancing the reliability of the analysis.

For plant height, nine significant loci were mapped, explaining 6.72% to 8.78% of phenotypic variance. Three significantly associated loci were obtained in the 16.93–18.54 Mb interval on chromosome 4D, which were significant in all four models. Twenty-four SNPs were significantly associated with KNS, explaining 3.58% to 7.79% of phenotypic variance, with 21 significant loci clustered in the 585.48–589.24 Mb interval on chromosome 5A. For KWS, nine significant loci were identified, explaining 4.99% to 8.37% of phenotypic variance. Among them, loci *Affx-109116213* and *Affx-110331741* were also significantly associated with KNS, indicating they are pleiotropic. Six significant loci for TKW were detected, explaining 5.11% to 6.95% of phenotypic variance, distributed on chromosomes 1B, 3A, 7A, 6B, and 7B, respectively.

Six SNPs were significantly associated with the tiller number, explaining 5.22% to 7.37% of phenotypic variance, with five significant loci within an 805 kb region on chromosome 5A. Fifty-four SNPs were significantly associated with WGT, explaining 6.00% to 14.20% of phenotypic variance, while 60 SNPs were associated with WS, explaining 6.00% to 14.96%. Due to the strong correlation between WS and WGT, 61 loci were significantly associated with both traits, including 33 loci on chromosome 5A distributed in the 470.24–691.83 Mb interval.

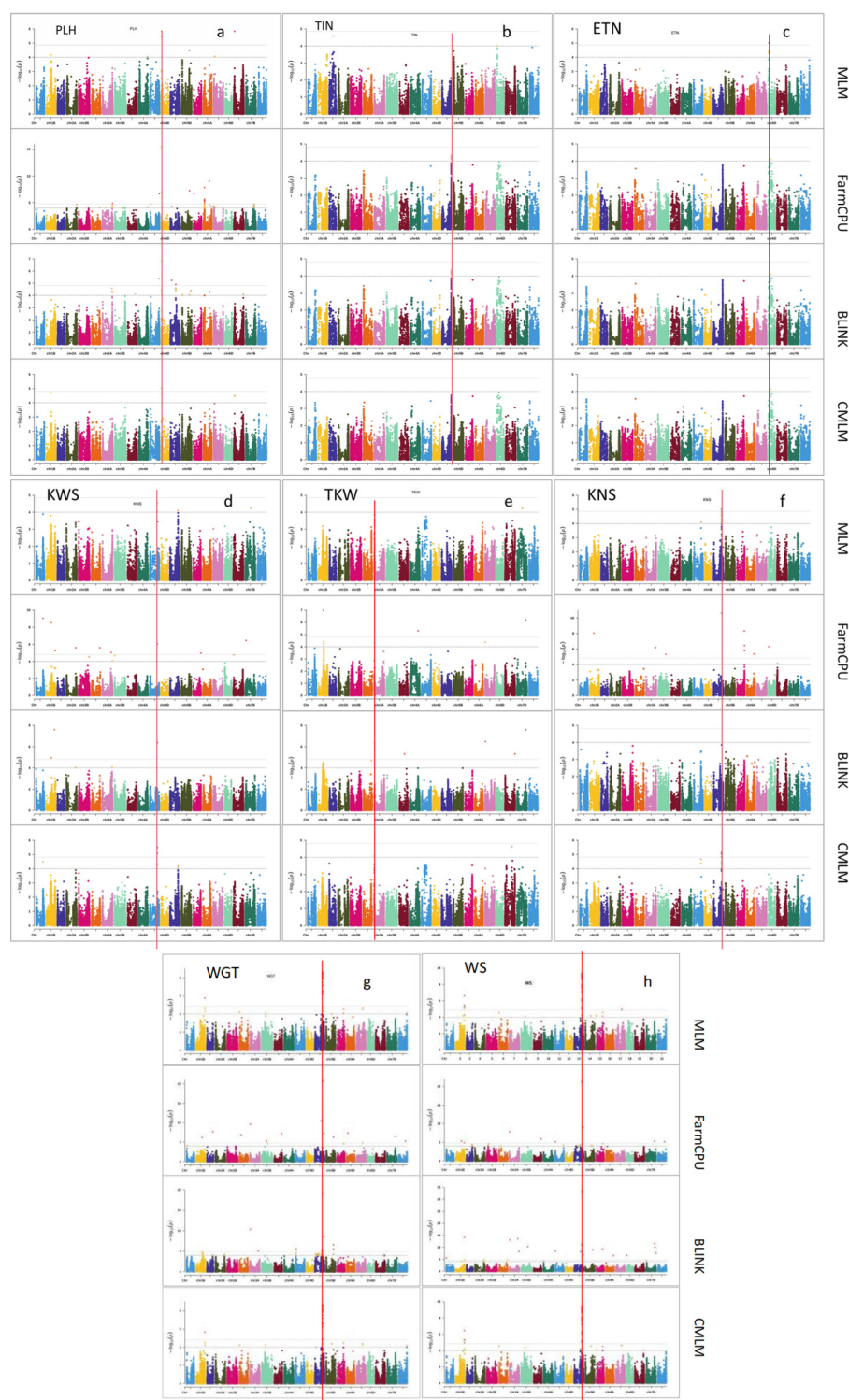


Figure 1. Manhattan plots of association analysis for eight traits using the MLM, FarmCPU, BLINK, and CMLM models. **a.** Plant height (PLH), **b.** Tiller number (TIN), **c.** Effective tiller number (ETN), **d.** Kernel weight per spike (KWS), **e.** Thousand-grain weight (TKW), **f.** Kernel number per spike (KNS), **g.** Whole growing stage (WGT), **h.** Winter-spring type (WS). X- and Y-axes represent the SNP markers along each chromosome of wheat and the $-\log_{10}(P\text{-value})$, respectively. The dotted lines designate $-\log_{10}(10^{-4})$, which was regarded as the threshold for significant associations.

Linkage disequilibrium haplotype block analysis

Haplotype analysis was conducted for KNS (Figure 2), ETN (Figure S3), TKW (Figure S4), and KWS (Figure S5). Regarding KNS, five SNP markers were found near *Affx-109116213* in the region of 586.14–586.60 Mb on chromosome 5A span LD Block1 (Figure 2a, b). This Block has three haplotypes (*5A_B1_Hap1*, *5A_B1_Hap2*, and *5A_B1_Hap3*). The distribution frequencies were 0.524, 0.45, and 0.012, respectively (Figure 2c). The mean values of KNS for each haplotype were 28.98, 38.97, and 32.19, respectively. The KNS with the *5A_B1_Hap2* haplotype was significantly higher than those with the *5A_B1_Hap1* and *5A_B1_Hap3* haplotypes (Figure 2d), while the WGT of the *5A_B1_Hap2* haplotype was significantly lower than that of the *5A_B1_Hap1* and *5A_B1_Hap3* haplotypes (Figure 2e).

Among the six SNPs significantly associated with ETN, five SNPs spanning 0.26 Mb on chromosome 6D form a LD block (Figure S3a, b). Two haplotypes, *6D_B1_Hap1* and *6D_B1_Hap2*, were detected with the frequency distribution of 0.912 and 0.088, respectively (Figure S3c). The number of effective tillers with haplotype *6D_B1_Hap1* was significantly higher than that with haplotype *6D_B1_Hap2* (Figure S3d).

On chromosome 3A, near the *Affx-109174053* marker significantly associated with TKW, a 51 kb region formed two LD blocks (Figure S4a, b). Block1 contains the four haplotypes *3A-B1-Hap1*, *3A-B1-Hap2*, *3A-B1-Hap3*, and *3A-B1-Hap4*, with distribution frequencies of 0.402, 0.377, 0.187, and 0.026, respectively (Figure S4c). The TKW means for these haplotypes are 45.15, 44.94, 48.17, and 46.34, respectively, with *3A-B1-Hap3* showing significantly higher TKW than the others (Figure S4d). Block2 also has four haplotypes: *3A-B2-Hap1*, *3A-B2-Hap2*, *3A-B2-Hap3*, and *3A-B2-Hap4*, with frequencies of 0.591, 0.307, 0.013, and 0.081 (Figure S4c). Their mean TKWs are 46.11, 44.67, 44.65, and 44.59, respectively, and *3A-B2-Hap1* shows better thousand kernel weight than the other haplotypes (Figure S4d).

Haplotype analysis was conducted for KWS. A 138-kb LD block was formed near the *Affx-109347856* marker on chromosome 4B (Figure S5a, b). The 3 haplotypes detected in this block were *4B-B1-Hap1*, *4B-B1-Hap2*, and *4B-B1-Hap3*, with distribution frequencies of 0.640, 0.243, and 0.071 (Figure S5c), and their corresponding grain weights per spike were 1.46, 1.58, and 2.00, respectively. Haplotype *4B-B1-Hap3* showed significantly higher grain weight per spike than the other two haplotypes (Figure S5d).

A haplotype takes two or more SNP alleles as one genetic unit and largely improves the reliability of association analysis, which has become one of the important means of molecular marker-assisted breeding application (Liu and Zhang 2016, Chaves et al. 2020). In this study, haplotype analysis was performed on the KNS, ETN, TKW, and KWS. The number of effective tillers and distribution frequency in *6D_B1_Hap1* were significantly higher than those in *6D_B1_Hap2*, indicating that the haplotype of *6D_B1_Hap1* was subjected to positive selection by breeders and was widely used in wheat breeding. In addition, three haplotypes were formed for KNS on chromosome 5A. The number of grains per spike of the haplotype *5A_B1_Hap2* with a significantly higher KNS and shorter WGT than that of *5A_B1_Hap1* and *5A_B1_Hap3* indicated that haplotype *5A_B1_Hap2* had a large utility for the genetic improvement of grain number per spike and the whole growth period. Haplotype *4B_B1_Hap3* had better grain weight per spike, and *3A_B1_Hap3* had significantly higher TKW than other haplotypes. However, their current distribution frequencies are low, so focusing on breeding these superior haplotypes is recommended.

Candidate gene prediction

In this study, 73 potential candidate genes associated with multiple wheat yield traits were identified (Table S5). The four genes are *TaTB1-4D*, *TaVRN1-5A*, *PhyC*, and *Pinb*, and the remaining 69 are novel genes. The four candidate genes related to wheat tillering were *TraesCS6D01G166500*, *TraesCS5A01G508000*, *TraesCS5A01G506900*, and *TraesCS5A01G506800*. *TraesCS6D01G166500*, which are involved in processes related to hexose transferase activity. Nine candidate genes associated with wheat plant height were identified (Table S5). The proteins encoded by these genes include extra-large GTP-binding protein 3, Acetyl-CoA carboxylase, cotton fiber protein, and Acetolactate synthase. Many of these enzymes have catalytic activities and are involved in various metabolic pathways during plant growth. *TaTB1-4D* was located in the region of 18.461–18.465 Mb on chromosome 4D and was very closed to *Rht1* (Peng et al. 1999); thus, it can be considered the same as *Rht1*. This study also identified 12 candidate genes significantly associated with thousand-grain weight and KWS. Their encoded products cover F-box/RNI-class superfamily proteins, glyceraldehyde-3-phosphate dehydrogenase, 30S ribosomal protein S1, ethylene-responsive transcription factors, etc.

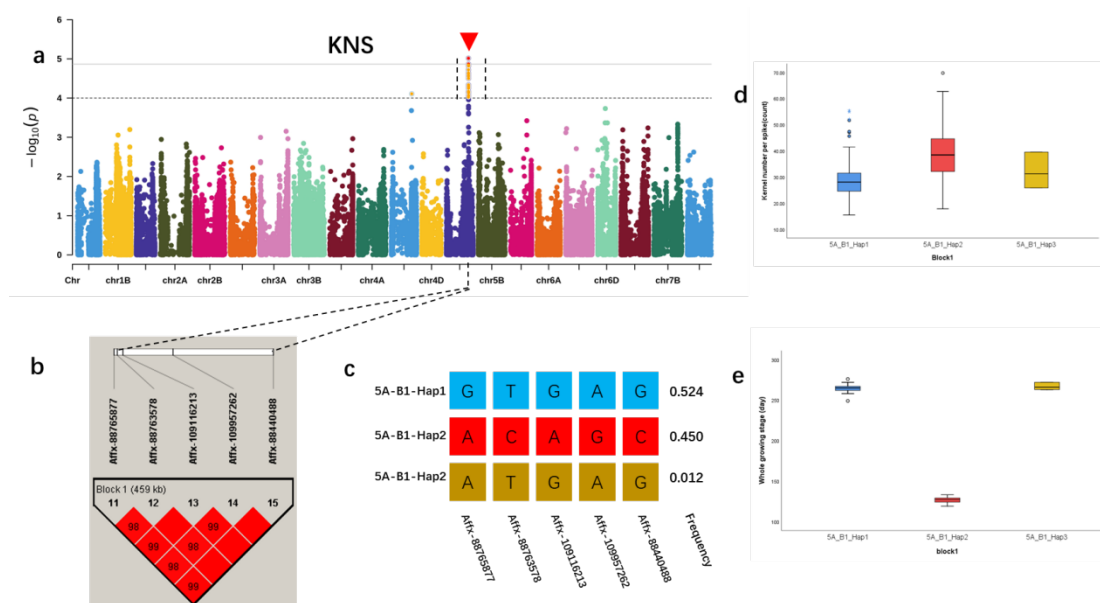


Figure 2. Genome-wide association study results for KNS and the analysis of the haplotype of KNS on chromosome 5A. **a.** Manhattan plot for Kernel number per spike (KNS). The horizontal line represents the significance threshold ($-\log_{10}(P\text{-value}) = 4$). The arrows indicate the location of the main peaks studied. **b.** Genomic locations of four SNP loci and the LD based on D' values between SNPs on chromosome 5A. **c.** Haplotypes detected in 250 accessions based on the four SNPs. **d.** Differences in the KNS among three haplotypes. **e.** Differences in the wheat growing stage among three haplotypes.

Additionally, 49 candidate genes associated with the winter-spring characteristics and the entire growth period of wheat were identified. The proteins encoded by these genes include dynamin-like proteins, peroxisomal membrane protein 2, DUF2431 domain-containing proteins, ABC transporter B family proteins, and mitochondrial carrier family proteins. Notably, the candidate gene *TraesCS5A01G391300* is a cloned gene of *PhyC* (Mizuno et al. 2012) related to heading date. The candidate gene *TraesCS5A01G391700*, annotated by significant marker *Affx-110331741*, was the vernalization gene *TaVRN1-5A* (Yan et al. 2003). The candidate gene *TraesCS5D01G004300* annotated by the marker *Affx-88590952* is grain hardness gene *Pinb* (Bhave and Morris 2008), suggesting that vernalization traits affect the grain material accumulation pattern through genetic linkage and developmental process regulation. Twenty candidate genes associated with the number of grains per spike were identified, 18 of which were located on chromosome 5A and simultaneously related to vernalization and the whole growing stage. The other two candidate genes -*TraesCS4B01G268100* and *TraesCS7A01G072400*- were also significantly associated with KWS. *TraesCS7A01G072400* is located in different regions of the same chromosome as the reported grain number genes *WAP01* (674.081–674.082Mb) (Kuzay et al. 2019) and *TaPIN1-7A* (148.414–148.418 Mb) (Yao et al. 2021), suggesting it is a novel gene related to spike traits. *TraesCS4B01G268100* (541.910–541.912 Mb) is close to the reported grain number gene *SVP3-4B* (589.180–589.195 Mb) (Li et al. 2020) and may have similar gene functions.

Reliability of multi-model GWAS in mapping complex agronomic traits

A total of 335 significantly associated loci for 8 yield-related traits through joint analysis were using four models. Among them, 88 loci showed stable pleiotropy across multiple models, confirming the effectiveness of multi-model cross-validation in controlling false positives in GWAS. Compared with single-model analysis, multi-model integration enhances mapping accuracy. For example, the *Affx-110331741* locus was significantly associated with WGT, WS, and KNS in both CMLM and MLM, and its annotated *TraesCS5A01G391700* gene is the known vernalization and flowering gene *TaVRN1-5A* (Yan et al. 2003), validated as a key locus through multi-model integration.

Genetic mechanism interpretation of yield trait-associated loci

In the 585.48–589.24 Mb interval on chromosome 5A, this study mapped 21 SNPs significantly associated with KNS, which partially overlaps with the SC/SL QTL cluster (381.76–533.29 Mb) (Li et al. 2021), suggesting conserved expression of the same QTL under different genetic backgrounds. Notably, the 585.5–595.4 Mb region on chromosome 5A enriches 27 pleiotropic SNPs related to WS, WGT, and KWS, located only 11.3 Mb from the winter-spring trait gene *TaPIL1* (574.194–574.199 Mb) (Zhang et al. 2022), implying that this region may regulate multi-trait co-expression through photoperiod response. Five tillering inhibitory genes (*Tin1*, *Tin2*, *Tin3*, *Tin4*, and *Ftin*) have been reported (Spielmeyer and Richards 2004, Kuraparthi et al. 2007, Zhang et al. 2013, Wang et al. 2022). However, the tillering-related region (chr6D:23.15–23.42Mb) identified here does not overlap with *TaPIN1-6D* (chr6D:397.082–397.084 Mb) (Yao et al. 2021), suggesting a novel QTL. Additionally, the close linkage of the plant height-associated locus *Affx-108864084* with *Rht1* (Peng et al. 1999) validates the selection signature of dwarf genes in modern variety improvement.

Research limitations and future directions

Although multi-model analysis improved locus reliability, some minor-effect QTLs (e.g. PVE < 5%) might remain undetected due to sample size limitations. Follow-up studies could integrate multi-year phenotypic data from different ecological zones in Ningxia to efficiently enhance detection of weak-effect loci via Meta-GWAS. Furthermore, fine mapping and transgenic validation are required to resolve causal variations in the yield gene cluster on chromosome 7B and the pleiotropic region on chromosome 5A. Stress-resistant genes embedded in Ningxia's germplasm resources should be further explored for application in stress-resistant breeding through integration of association analysis and phenomics.

ACKNOWLEDGMENTS

This work was supported by the Key Research and Development Program of Ningxia (2023BCF01011, 2023BSB03059, 2025BBF01010); Ningxia natural science fund project (2023AAC05050); National Key Research and Development Program of China (2022YFD1200900); Innovative Team for Precision Evaluation and Gene Mining of Crop Germplasm Resources.

DATA AVAILABILITY

The datasets generated and/or analyzed in this study, as well as the supplementary tables and figures, are available from the corresponding author upon reasonable request.

REFERENCES

- Alonso MP, Vanzetti LS, Crescente JM, Mirabella NE, Panelo JS and Pontaroli AC (2021) QTL mapping of spike fertility index in bread wheat. **Crop Breeding and Applied Biotechnology** 21: e34402113.
- Barrett JC, Fry B, Maller J and Daly MJ (2005) Haploview: Analysis and visualization of LD and haplotype maps. **Bioinformatics** 21: 263–265.
- Bhave M and Morris CF (2008) Molecular genetics of puroindolines and related genes: Allelic diversity in wheat and other grasses. **Plant Molecular Biology** 66: 205–219.
- Chaves MS, Silva GBP, Caierão E, Federizzi LC and Martinelli JA (2020) A century of wheat breeding in Brazil: The origin and inheritance of the *Lr34* locus in wheat varieties released from 1922 to 2016. **Crop Breeding and Applied Biotechnology** 20: e27952027.
- Franco MF, Lori GA, Cendoya MG, Panelo J, Alonso MP, Malbrán I and Pontaroli AC (2022) QTL mapping for Type II resistance to Fusarium head blight and spike architecture traits in bread wheat. **Crop Breeding and Applied Biotechnology** 22: e38242229.
- Huang M, Liu X, Zhou Y, Summers RM 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: giy154.
- International Wheat Genome Sequencing Consortium - IWGSC (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. **Science** 361: eaar7191.
- Korte A and Farlow A (2013) The advantages and limitations of trait analysis with GWAS: A review. **Plant Methods** 9: 29.
- Kuraparthi V, Sood S, Dhaliwal HS, Chhuneja P and Gill BS (2007) Identification and mapping of a tiller inhibition gene (*tin3*) in wheat. **Theoretical and Applied Genetics** 114: 285–294.
- Kuzay S, Xu Y, Zhang J, Katz A, Pearce S, Su Z, Fraser M, Anderson JA, Brown-Guedira G, DeWitt N, Haugrud AP, Faris JD, Akhunov E, Bai G and 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.

- Li H, Zhou Y, Xin W, Wei Y, Zhang J and Guo L (2019) Wheat breeding in northern China: Achievements and technical advances. **The Crop Journal** 7: 718-729.
- Li K, Debernardi JM, Li C, Lin H, Zhang C, Jernstedt J, Korff MV, Zhong J and Dubcovsky J (2020) Interactions between SQUAMOSA and SVP MADS-box proteins regulate meristem transitions during wheat spike development. **Plant Cell** 33:3621-3644.
- Li MX, Yeung JMY, Cherny SS and Sham PC (2012) Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. **Human Genetics** 131: 747-756.
- 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 and 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.
- Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES and Zhang Z (2012) GAPIT: genome association and prediction integrated tool. **Bioinformatics** 28: 2397-2399.
- Liu G and Zhang D (2016) The application of the functional molecular marker in wheat breeding. **Biotechnology Bulletin** 32: 18.
- Liu X, Huang M, Fan B, Buckler ES and Zhang Z (2016) Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. **PLOS Genetics** 12: e1005767.
- Mizuno N, Nitta M, Sato K and Nasuda S (2012) A wheat homologue of PHYTOCLOCK 1 is a candidate gene conferring the early heading phenotype to einkorn wheat. **Genes & Genetic Systems** 87: 357-367.
- 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, Wu Y, Li A, Wang H, Kong L, Bai G and 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.
- Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD and Harberd NP (1999) "Green revolution" genes encode mutant gibberellin response modulators. **Nature** 400: 256-261.
- Rasheed A, Wen W, Gao F, Zhai S, Jin H, Liu J, Guo Q, Zhang Y, Dreisigacker S, Xia X and He Z (2016) Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. **Theoretical and Applied Genetics** 129: 1843-1860.
- Scherlosky A, Marchioro VS, Franco FA, Braccini AL and Schuster I (2018) Genetic variability of Brazilian wheat germplasm obtained by high-density SNP genotyping. **Crop Breeding and Applied Biotechnology** 18: 399-408.
- Sinclair TR and Bai Q (1997) Analysis of high wheat yields in Northwest China. **Agricultural Systems** 53: 373-385.
- Spielmeyer W and Richards RA (2004) Comparative mapping of wheat chromosome 1AS which contains the tiller inhibition gene (tin) with rice chromosome 5S. **Theoretical and Applied Genetics** 109: 1303-1310.
- Wang K, Li M and Hakonarson H (2010) ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. **Nucleic Acids Research** 38: e164.
- Wang S, Zhu Y, Zhang H, Chang C and Ma C (2014) Analysis of genetic diversity and relationship among wheat breeding parents by SSR markers. **Journal of Triticeae Crops** 34: 621-627.
- Wang Z, Wu F, Chen X, Zhou W, Shi H, Lin Y, Hou S, Yu S, Zhou H, Li C and Liu Y (2022) Fine mapping of the tiller inhibition gene TIN4 contributing to ideal plant architecture in common wheat. **Theoretical and Applied Genetics** 135: 527-535.
- Wei MZ, Deng TS, Zhang Q, Cheng Z and Li S (2021) Seed-Mediated Synthesis of Gold Nanorods at Low Concentrations of CTAB. **ACS Omega** 6: 9188-9195.
- Xiao J, Liu B, Yao Y, Guo Z, Jia H, Kong L, Zhang A, Ma W, Ni Z, Xu S, Lu F, Jiao Y, Yang W, Lin X, Sun S, Lu Z, Gao L, Zhao G, Cao S, Chen Q, Zhang K, Wang M, Wang M, Hu Z, Guo W, Li G, Ma X, Li J, Han F, Fu X, Ma Z, Wang D, Zhang X, Ling HQ, Xia G, Tong Y, Liu Z, He Z, Jia J and Chong K (2022) Wheat genomic study for genetic improvement of traits in China. **Science China Life Sciences** 65: 1718-1775.
- Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T and Dubcovsky J (2003) Positional cloning of the wheat vernalization gene VRN1. **Proceedings of the National Academy of Sciences** 100: 6263-6268.
- Yang J, Lee SH, Goddard ME and Visscher PM (2011) GCTA: A tool for genome-wide complex trait analysis. **American Journal of Human Genetics** 88: 76-82.
- Yao FQ, Li XH, Wang H, Song YN, Li ZQ, Li XG, Gao XQ, Zhang XS and Bie XM (2021) Down-expression of *TaPIN1s* increases the tiller number and grain yield in wheat. **BMC Plant Biology** 21: 443.
- Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S and Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. **Nature Genetics** 38: 203-208.
- Zhang C, Dong SS, Xu JY, He WM and Yang TL (2019) PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. **Bioinformatics** 35: 1786-1788.
- Zhang J, Bernard D, Biddulph B, Brockman F, Walker E, Khan N, Wong D, Hayden M and Apples R (2013) Wild-type alleles of *Rht-B1* and *Rht-D1* as independent determinants of thousand-grain weight and kernel number per spike in wheat. **Molecular Breeding** 32: 771-783.
- Zhang L, He G, Li Y, Yang Z, Liu T, Xie X, Kong X and Sun J (2022) PIL transcription factors directly interact with SPLs and repress tillering/

GWAS analysis revealed wheat growth characteristics and yield-related genes in the Yellow River Basin

branching in plants. **The New Phytologist** **233**: 1414-1425.

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordoñas JM and Buckler ES (2010) Mixed linear

model approach adapted for genome-wide association studies. **Nature Genetics** **42**: 355-360.

 This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.