

Genomic insights into the modifications of spike morphology traits during wheat breeding

Yangyang Liu¹ | Rui Yu² | Liping Shen^{1,3} | Mengjing Sun⁴ | Yanchun Peng⁵ |
Qingdong Zeng⁶ | Kuocheng Shen^{1,7} | Xuchang Yu^{1,7} | He Wu^{1,7} |
Botao Ye^{1,7} | Ziyang Wang^{1,7} | Zhiweng Sun^{1,7} | Danning Liu^{1,7} | Xiaohui Sun⁸ |
Zhiliang Zhang^{7,9} | Jiayu Dong^{7,9} | Jing Dong⁵ | Dejun Han⁶ | Zhonghu He^{4,10} |
Yuanfeng Hao⁴ | Jianhui Wu²  | Zifeng Guo^{1,3,7}

¹Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing, China

²State Key Laboratory of Crop Stress Resistance and High-Efficiency Production, Northwest A&F University, Yangling, Shaanxi, China

³China National Botanical Garden, Beijing, China

⁴Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (CAAS), Beijing, China

⁵Key Laboratory of Crop Molecular Breeding, Ministry of Agriculture and Rural Affairs, Institute of Food Crops, Hubei Academy of Agricultural Sciences, Wuhan, China

⁶State Key Laboratory of Crop Stress Biology for Arid Areas, Northwest A&F University, Yangling, Shaanxi, China

⁷University of Chinese Academy of Sciences, Beijing, China

⁸Yantai Academy of Agricultural Sciences, Yantai, China

⁹State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Innovative Academy of Seed Design, Chinese Academy of Sciences, Beijing, China

¹⁰International Maize and Wheat Improvement Center (CIMMYT) China Office, c/o CAAS, Beijing, China

Correspondence

Yuanfeng Hao, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (CAAS), Beijing, 100081, China.
Email: haoyuanfeng@caas.cn

Jianhui Wu, State Key Laboratory of Crop Stress Resistance and High-Efficiency Production, Northwest A&F University, Yangling, Shaanxi 712100, China.
Email: wujh@nwfau.edu.cn

Zifeng Guo, Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China.
Email: guozifeng@ibcas.ac.cn

Funding information

Strategic Priority Research Program of Chinese Academy of Sciences, Grant/Award Number: XDA24010104-2; National Natural Science Foundation of China, Grant/Award Number:

Abstract

Over the past century, environmental changes have significantly impacted wheat spike morphology, crucial for adaptation and grain yield. However, the changes in wheat spike modifications during this period remain largely unknown. This study examines 16 spike morphology traits in 830 accessions released from 1900 to 2020. It finds that spike weight, grain number per spike (GN), and thousand kernel weight have significantly increased, while spike length has no significant change. The increase in fertile spikelets is due to fewer degenerated spikelets, resulting in a higher GN. Genome-wide association studies identified 49,994 significant SNPs, grouped into 293 genomic regions. The accumulation of favorable alleles in these genomic regions indicates the genetic basis for modification in spike morphology traits. Genetic network analysis of these genomic regions reveals the genetic basis for phenotypic correlations among spike morphology traits. The haplotypes of the

Yangyang Liu, Rui Yu, Liping Shen, Mengjing Sun, and Yanchun Peng authors contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Plant, Cell & Environment* published by John Wiley & Sons Ltd.

32272122; Post-doctoral Fellowship Program
of CPSF under Grant Number,
Grant/Award Number: GZB20240820

identified genomic regions display obvious geographical differentiation in global accessions and environmental adaptation over the past 120 years. In summary, we reveal the genetic basis of adaptive evolution and the interactions of spike morphology, offering valuable resources for the genetic improvement of spike morphology to enhance environmental adaptation.

KEYWORDS

adaptive evolution, genetic network, grain yield, GWAS

1 | INTRODUCTION

Bread wheat (*Triticum aestivum*. L) is a staple food for over 35% of the world's population, playing a crucial role in global food security (Nigus et al., 2022; Shiferaw et al., 2013). Climate change and population growth have increased food demand (Rezaei et al., 2023), widening the gap between demand and production (van Dijk et al., 2021). To meet the rising food demand due to population growth and environmental challenges, global wheat production needs to remarkably increase (van Dijk et al., 2021). Increasing grain yield remains a primary goal of wheat breeding. Wheat yield is primarily determined by three factors: spike number per area, grain number per spike (GN), and TKW (Cao et al., 2020; Lázaro et al., 2010; Wang et al., 2023). Wheat breeders face the challenge of developing high-efficiency breeding strategy for varieties with higher grain number and size (Kong et al., 2022; Kuzay et al., 2019; Wang et al., 2019; Wittern et al., 2022; Yao et al., 2021; Zhang et al., 2018) since these traits often exhibit a trade-off and tend to change together (Liu et al., 2023c). This requires a better understanding of the genetic architecture and networks underlying multiple traits, as focusing on a single trait is insufficient for increasing grain yield.

Spike is the key component for determining grain yield in wheat. Spike morphology is closely linked to GN and grain size (Guo et al., 2018). For instance, the GN is determined by grain number per spikelet and fertile spikelet number per spike (FSN), spikelet number (SN) per spike (Zhou et al., 2021a). Spike length and SN determine spike compactness. Spike morphology traits are primarily quantitative, controlled by multiple quantitative trait loci (QTLs) and influenced by environmental factors (Alqudah et al., 2020; Katz et al., 2022; Lin et al., 2021; Liu et al., 2023c). Therefore, a key task for accelerating wheat variety development is the genetic dissection of architecture and networks governing spike morphology traits.

Genome-wide association studies (GWAS) have been successfully used to dissect the genetic architecture and network of agronomic traits (Fang et al., 2017; Liu et al., 2023b). To capture the available genetic variability for the targeted traits, the GWAS population must have high genetic diversity. This includes varieties released at different breeding stages that are adapted to various geographical environments and adverse stresses (Hao et al., 2020; Li et al., 2022a; Sansaloni et al., 2020). The varieties development during the Chinese wheat breeding process exhibits high genetic diversity due to significant changes in climatic conditions (Hao et al., 2020; Li et al., 2022a; Walkowiak et al., 2020).

In wheat breeding, continuous domestication and selection for spike morphology traits can cause fluctuation in allele frequency, leading to a loss of genetic diversity (Haudry et al., 2007; Li et al., 2022b; Wang, Wang, et al., 2022; Zhao et al., 2023; Zhou et al., 2021b; Zhou et al., 2020). Chinese wheat breeding has successfully improved grain yield over the past decades (Hao et al., 2021; Qin et al., 2015). Simultaneously, spike morphology has undergone significant changes for adaptive evolution (He et al., 2021). To our knowledge, no systematic analyses of spike morphology traits in Chinese wheat varieties over the past decades have been conducted. Additionally, the genetic basis for the modifications in spike morphology remains largely unknown.

In this study, we collected 830 wheat accessions released or planted from 1900 to 2020 and assessed 16 spike morphology traits. Comprehensive GWAS analyses allowed us to identify novel genomic regions and genetic networks related to spike morphology traits. Furthermore, we uncovered the genetic basis for the modifications in spike morphology for environmental adaptation.

2 | RESULTS

2.1 | The variations of spike morphology traits across wheat breeding process

We analyzed 16 spike morphology traits in 830 Chinese wheat accessions released from 1900 to 2020 (Table S1). These traits included 12 measured traits: spike length, spike weight, SN, GN, grain number per spikelet, grain weight per spike, grain weight per spikelet, chaff weight per spike, TKW, fertile SN per spike, the number of sterile spikelet at top of individual spike (SSN-T), the number of sterile spikelet at bottom of individual spike (SSN-B). Additionally, four ratios or indirect traits were analyzed: SN/FSN, grain number per spike/chaff weight (GN/CW), grain weight/chaff weight (GW/CW), and grain weight/spike weight (GW/SW) (Table S2). Most of the 16 spike traits were highly correlated (Supporting Information S1: Figure S1a). Principal component analysis (PCA) grouped the 16 traits into three categories: (1) GN/CW; (2) TKW; (3) GN, chaff weight, FSN, FSN/SN, grain number per spikelet, grain weight per spike, grain weight per spikelet, GW/CW, GW/SW, spike length, SSN-B, SSN-T, and spike weight (Supporting Information S1: Figure S1b). Their

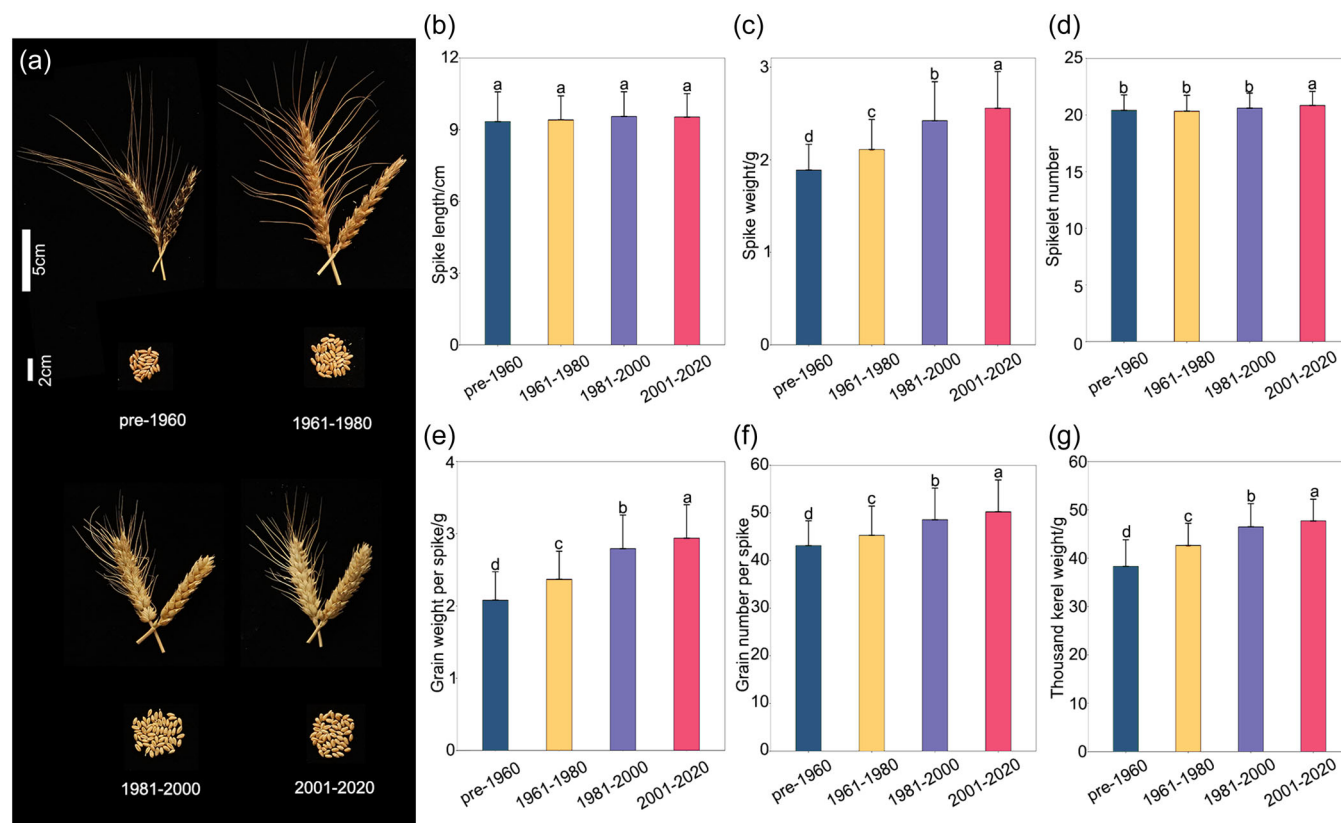


FIGURE 1 The modifications of spike morphology traits across Chinese breeding process. (a) The spikes at different phases of Chinese breeding. (b–g), The differences of spike length (b), spike weight (c), spikelet number per spike (d), grain weight per spike (e), grain number per spike (f), thousand kernel weight (g) among the different phases of Chinese breeding. Different letters above the bars indicate significant differences at $p < 0.05$ (one-way analysis of variance followed by LSD test).

broad-sense heritabilities of those traits ranged from 0.43 to 0.89 (Supporting Information S1: Figure S1c).

We evaluated the dynamics in the phenotypic values of 16 spike morphology traits throughout the Chinese breeding process (Table S3). The average values of most traits increased during breeding process (Figure 1b–g, Table S3). However, nonsignificant changes in spike length were observed in this study (Figure 1b). Comparing to accessions from pre-1960, accessions from 2001–2020 showed significant increases on average values: spike weight by 35.44%, SN by 2.11%, grain weight per spike by 41.34%, GN by 16.44% and TKW by 24.53% (Figure 1b–g, Table S3). Additionally, the numbers of SSN-T and SSN-B gradually decreased, while the fertile SN per spike increased throughout the breeding process. Compared to pre-1960 accessions, SSN-T and SSN-B decreased by 17.06% and 26.87%, respectively, while the number of fertile spikelet per spike increased by 5.40% in the accessions of 2001–2020 (Table S3). These results indicated that the modern breeding increased the number of fertile spikelets by reducing spikelet abortion at the top and bottom of the spike. This increase in GN has implications for improving wheat grain yield (Table S3).

These findings suggest that spike morphology traits have undergone significant changes, resulting in a fuller appearance. These structural characteristics may play a crucial role in improving yield. It

is intriguing to consider that spike morphology traits might have been extensively selected during the wheat breeding process in China.

2.2 | Accumulation of favorable alleles linked to spike morphological changes

To understand the genetic basis of the spike morphology trait modifications during the Chinese wheat breeding process, we conducted GWAS on 16 spike morphology traits using 21,279,862 SNPs from 830 Chinese accessions. We identified a total of 49,994 SNPs ($-\log_{10}(P \text{ value}) > 5.0$) (Figure 2a), and 40.53% (20,263/49,994) associated with at least two traits (Figure 2a, Table S4). To evaluate the relationship between the accumulation of favorable alleles and the changes in spike morphology, we determined the allele frequency of significant SNPs throughout the Chinese breeding process. An increased frequency of favorable alleles was associated with increases in spike weight, grain weight per spike, grain weight per spikelet, and TKW (Figure 2b). These traits exhibited a consistent trend of improvement (Figures 1c,e,g and 2c–e, Tables S2 and S6). Furthermore, the numbers of accumulated favorable alleles for spike weight, grain weight per spike, and TKW showed close connections with the phenotypic changes over the breeding history (Figure 2f–h). For spikelet degeneration (SSN-T, SSN-B),

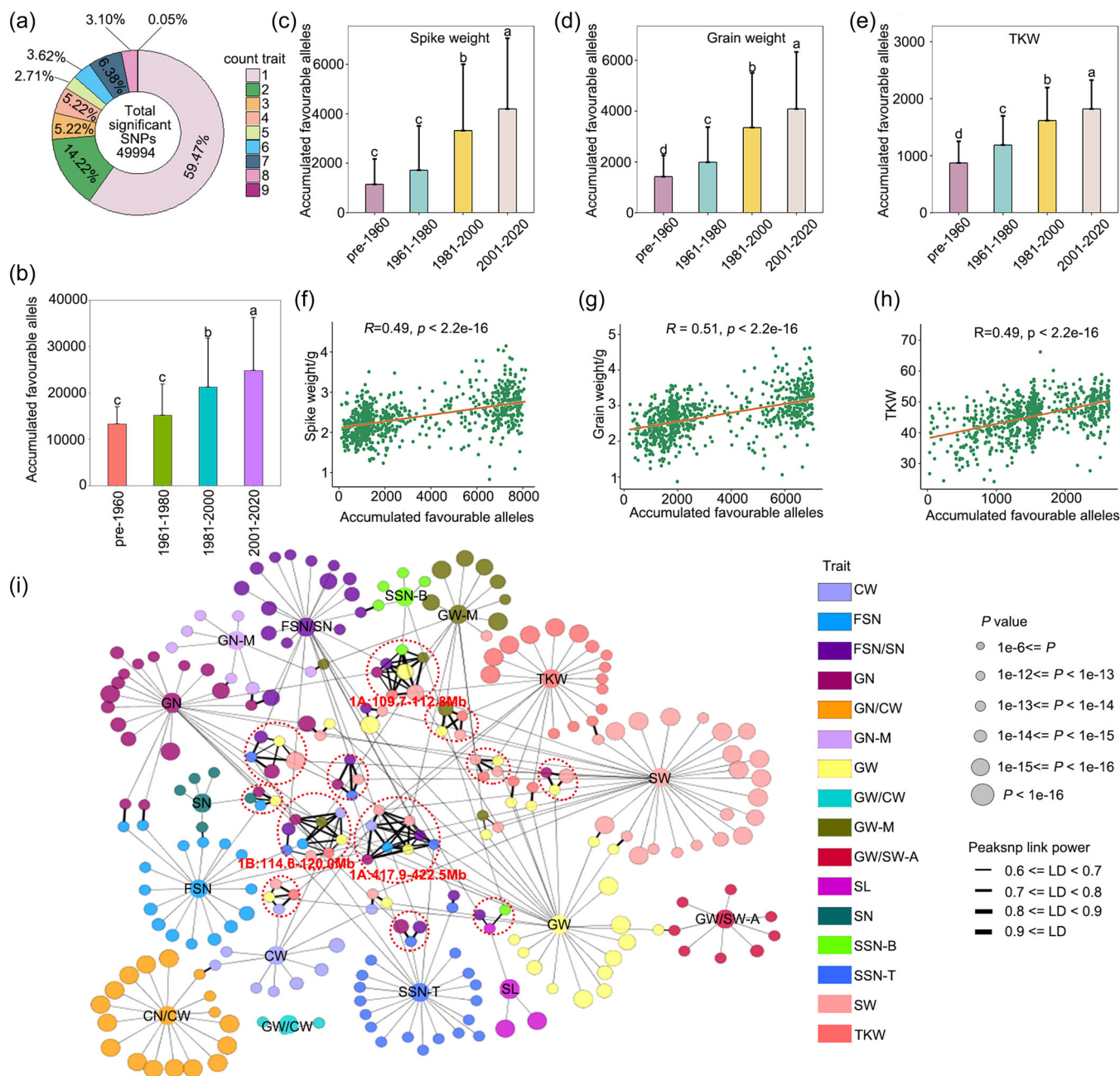


FIGURE 2 Accumulation of favorable alleles for the spike weight (SW), grain weight per spike (GW) and thousand kernel weight (TKW) during China wheat breeding and association networks across different spike traits. (a) Summary of 16 traits significant SNPs (p value $< 1e-5$). (b), The number of favorable alleles related to spike morphology in wheat at different breeding stages. (c–e), The number of favorable alleles related to spike weight (c), GW (d) and TGW, (e) at different breeding stages. (f, g) Correlations between favorable alleles and spike morphology traits: SW (f), GW, (g) and TKW (h). (i) The nodes represent wheat spike architecture traits and their associated loci. The sixteen traits are indicated by different colors. The edges between the loci from different traits are linked by their peak SNP LD. Only the edges with an average LD ≥ 0.6 are displayed. Red dashed circles indicate loci specifically associated with different traits. The size of the circle indicates the P -value of the correlation between the SNP within the loci and the phenotype; smaller P -values indicate higher correlations. SW, spike weight.

the frequency of unfavorable alleles exhibited decreased trend (Tables S5 and S6). For SN per spike and fertile spikelet number per spike, favorable alleles were abundant in pre-1960 accessions and increased from 1960 to 1980 to 2001–2020 (Tables S5 and S6). In summary, these results indicate that the accumulation of favorable alleles provide the genetic basis for modifications of spike morphology during wheat breeding.

Allele frequency of identified SNPs were assessed across historical periods using 145 wheat accessions, comprising modern Chinese cultivars, Chinese landraces, and introduced modern cultivars (IMC) (Hao et al., 2020). We quantified the prevalence of significant SNPs for spike weight, grain weight per spike, TKW and GN in each accessions, identifying the top 10 accessions for each trait (Table S7).

Analysis showed that the IMCs like St 2422/464 had higher frequencies of favorable alleles associated with spike weight, grain weight per spike, and TKW (Table S7). Likewise, modern Chinese cultivars including Yannong 19, Shaan 213, Xinkehan 9, Jinan 16, Zhoumai 16, and Wuyimai exhibited abundant favorable alleles for spike weight, grain weight per spike, and TKW (Table S7). Moreover, Chinese landraces such as Paozimai, Chinese Spring, and Baihuamai showed numerous favorable alleles for grain number per spike. Conversely, introduced accessions like Villa Glori and Neuzucht had higher frequencies of favorable alleles for GN and grain number per spikelet (Table S7). These findings offer valuable insights into the genetic diversity and variation in spike morphology traits among introduced and locally improved wheat varieties in China.

Utilizing linkage disequilibrium (LD) and marker connections, we identified 293 genomic regions associated with the 16 traits (Table S8). These regions included 141 for the A genome, 95 for the B genome, and 57 for the D genome, and 55.63% (163/293) associated with multiple traits (Table S8). Additionally, 28.67% of identified genomic regions overlapped with previously reported QTLs (Table S9). Known genes such as *TaNAC77-7B* (Luo et al., 2021), *TaPIN1-6D* (Yao et al., 2021), and *TaVRN2-5A* (Yao et al., 2021) were among those identified (Supporting Information S1: Figure S2, Table S9). We developed a genetic network (Figure 2i) to explore the coregulation of the 293 genomic regions and dissect their genetic architecture across various spike morphology traits. Most identified genomic regions showed correlations across traits (Figure 2i), except for GW/CW and GW/SW (Figure 2i, Table S8). Specifically, genomic regions associated with spike weight, grain weight per spike, grain weight per spikelet, chaff weight per spike, and GN exhibited close associations (Figure 2i, Table S8). Certain genomic regions, like chr1A: 109.7–112.8 Mb, chr1A: 417.9–422.5 Mb, and chr1B: 114.8–120.0 Mb, were pivotal nodes in genetic network (Figure 2i, Table S8). Notably, chr1B: 114.8–120.0 Mb, associated with SN (qSN1B.1) and grain setting (qGS1B.2) (Pang et al., 2020), influenced traits including spike weight, grain weight per spike, TKW, fertile SN per spike and GN (Figure 2i, Table S8).

2.3 | Identification of genomic regions related to important agronomic traits

To evaluate the biological functions of the genomic regions with the coregulated network, we focused on two specific loci chr1A: 109.7–112.8 Mb and chr1B: 114.8–120.0 Mb (Table S8). XP-CLR and *Fst* analyses demonstrated that the genomic region (chr1A: 109.7–112.8 Mb) underwent selection and exhibited population differentiation during the Chinese breeding process (Figure 3a,b). This region was closely associated with spike weight, grain weight per spikelet, TKW, SSN-B, FSN/SN, GN, and grain number per spikelet (Figure 3c, Table S8).

We identified a total of 12 haplotypes (Hap-A1–12) within this region, with Hap-A2 being the most prevalent and dominant across different breeding phases. The proportions of Hap-A3 and Hap-A4

increased gradually over time, while Hap-A6 and Hap-A10 decreased (Figure 3d, Tables S10 and S11). Hap-A3 and Hap-A4 exhibited relatively higher average values for spike weight, grain weight per spike, TKW, GN, FSN/SN, but lower SSN-B compared to other haplotypes. The frequency of Hap-A3 and Hap-A4 notably increased during the breeding process (Figure 3e, Table S12).

Understanding the genetic composition of wheat backbone parents is crucial for optimizing their use in high-yield breeding programs. Within the chr1A: 109.7–112.8 Mb interval, we identified three haplotypes among Chinese wheat backbone parents, encompassing the 145 wheat accessions analyzed previously (Figure 3f, Table S13). Notably, Wuyimai possessed a unique haplotype distinct from those found in other backbone parents, which were further divided into two distinct haplotypes (Figure 3f, Table S13).

We examined the haplotype frequency differences for the interval chr1A: 109.7–112.8 Mb in Chinese and American cultivars and landraces based on previous data set (Niu et al., 2023). In landraces, six haplotypes were identified, while only four were found in Chinese cultivars and one in American cultivars, respectively. Hap-cu-A4 was the predominant haplotype in both Chinese and American accessions. Conversely, Hap-cu-A2 and Hap-cu-A3 were exclusive to landraces, indicating a selection preference by Chinese and American breeders for this genomic region (Figure 3g, Table S14).

XP-CLR and *Fst* analyses indicated that the genomic region chr1B: 114.8–120.0 Mb was subject to selected and showed population differentiation during wheat breeding process in China (Figure 4a,b). The region was significantly associated with nine traits: spike weight, grain weight per spike, chaff weight per spike, TKW, SSN-T, fertile spike number per spike, FSN/SN and GN, and grain number per spikelet (Figure 4c, Table S8). Eight haplotypes (Hap-B1–8, number greater than 5) were detected for this genomic region (Table S15). Hap-B1 emerged as the dominant haplotype, showing an increasing frequency trend, whereas Hap-B3 to Hap-B6 showed a decreasing trend. Notably, Hap-B5's frequency significantly declined from 17% to 4% (Table S16). Hap-B1 exhibited higher values of spike weight, grain weight per spike, GN, and TKW compared to Hap-B5 (Figure 4d,e, Table S17). These results suggest that Chinese breeders have selected high-yield haplotypes for this genomic region. Two haplotypes were identified in Chinese wheat backbone parents for this interval, with Aimengniu and Xiaoyan6 showing identical haplotypes to the reference genome (Chinese Spring) (Figure 4f, Table S18). When analyzing haplotype distribution in Chinese and American cultivars and landraces using a previous data set (Niu et al., 2023), seven haplotypes were identified in landraces, while only three were found in Chinese cultivars and two in American cultivars. Hap-cu-B1 was predominantly found in Chinese cultivars (44.68%), whereas most American cultivars (96.97%) carried Hap-cu-B7 (Figure 4g, Table S19). Hap-cu-B7 had the highest frequency among the haplotypes (Figure 4g, Table S19). The selection process in this genomic region by Chinese and American breeders has led to reduced haplotype diversity in the cultivars (Figure 4g). These findings indicate that these two colocalized regions have been under

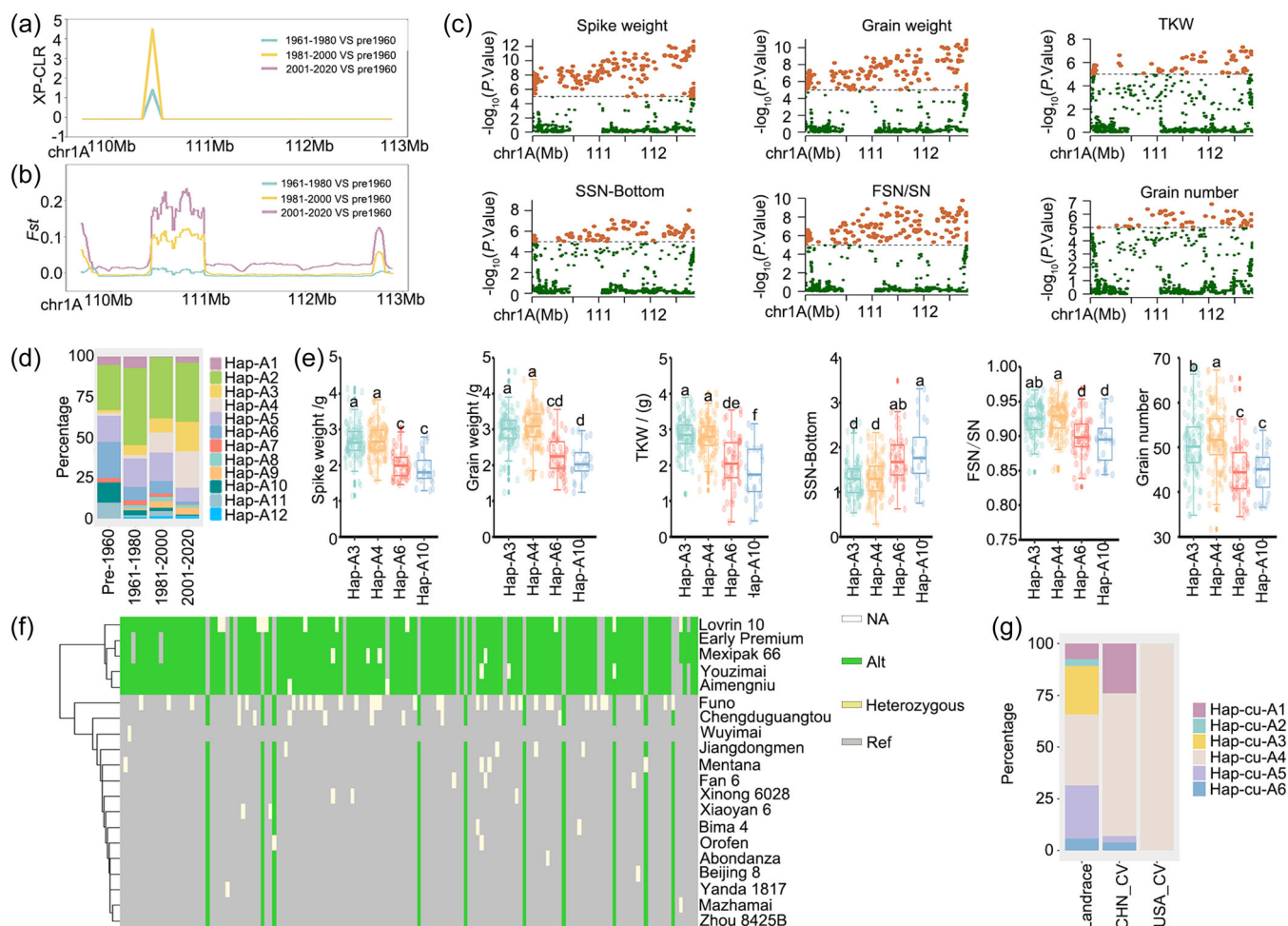


FIGURE 3 Identification and breeding selection of the haplotype blocks on chromosome chr1A: 109706033-112856267. (a) Genome-wide selection signals (XP-CLR scores) of different wheat breeding eras in China. (b) F_{st} of different wheat breeding eras in China. (c) Manhattan plots for spike weight (SW), grain weight per spike (GW), thousand kernel weight (TKW), the number of sterile spikelet at bottom of individual spike (SSN-B), FSN/SN, grain number per spike (GN) on chromosome chr1A: 109706033-112856267. The orange dots represent the significantly associated signals around the identified interval at the threshold of $P < 1.0 \times 10^{-5}$. (d) Percentages of different haplotypes over Chinese breeding process. (e) The differences of SW, GW, TKW, the number of sterile spikelet at SSN-B, FSN/SN and GN among the four haplotypes. In each plot, the central line indicates the median, the box limits indicate the upper and lower quartiles, the whiskers indicate $1.5 \times$ the interquartile range. The significance of the difference was analyzed using Fisher's least significant difference in 12 haplotypes. (f) Distribution of haplotype blocks of genomic region (chr1A: 109706033-112856267) in the wheat backbone parents. The genome that is the same as the Chinese Spring is indicated in gray (Ref), the heterozygous sites (Heterozygous) are indicated in yellow, and the homozygous change sites (Alt) are indicated in green. The missing sites (NA) are indicated in white. (g) Percentages of different haplotypes in landraces, China and America cultivars. FSN, fertile spikelet number per spike.

artificial selection for an extended period, focusing on increasing grain yield and environmental adaptation.

2.4 | Identification of a favorable allele for the increase of spike weight, grain weight per spike, and grain number per spike

We identified a major peak at chr1B: 114.8-120.0 Mb, which is associated with spike weight, grain weight per spike, GN, and TKW (Figure 5a-d, Tables S4 and S8). Further analysis of linkage disequilibrium in this region revealed that the SNPs are within a single

block (Figure 5e). This block encompasses 14 genes, which *TraesCS1B02G104500* showing the highest expression in grain (Table S20). Thus, *TraesCS1B02G104500* is likely the candidate gene influencing grain weight per spike, GN and spike weight in wheat. We detected four SNPs in the association population (+2T > C, +2,265 G/C, +2,394 G/C, +2,819 T/C), with two being nonsynonymous (+2 Met/Trp, +2,394 Gly/Ala). Using these SNPs, we classified all the accessions into five haplotypes (Hap1-5). Hap1 (402 accessions) and Hap2 (298 accessions) are the two major haplotypes, with Hap1 having a sequence similar to the reference genome (Chinese Spring) (Figure 5g). Hap1 exhibited higher average values for spike weight, grain weight per spike, GN and TKW (Figure 5h-k).

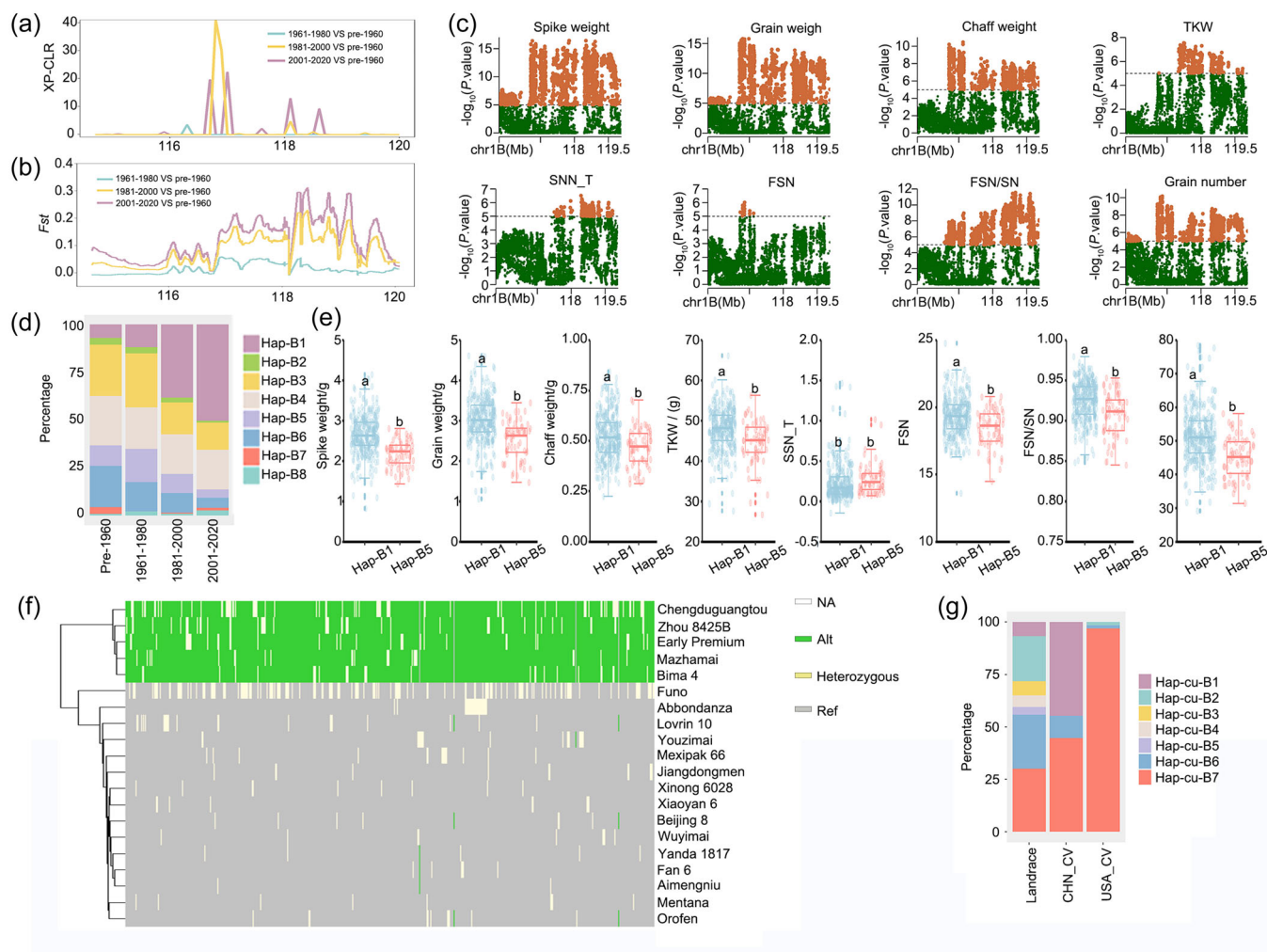


FIGURE 4 Identification and breeding selection of the haplotype blocks on chromosome 1B: 114665493-120067499. (a) Genome-wide selection signals (XP-CLR scores) of different wheat breeding eras in China. (b) F_{st} of different wheat breeding eras in China. (c) Manhattan plots for spike weight (SW), grain weight per spike (GW), chaff weight per spike (CW), thousand kernel weight (TKW), the number of sterile spikelet at top of individual spike (SSN-T), fertile spikelet number per spike (FSN), FSN/SN, grain number per spike (GN) on chromosome chr1B: 114665493-120067499. The orange dots represent the significantly associated signals around the identified interval at the threshold of $P < 1.0 \times 10^{-5}$. (d) Percentages of different haplotypes over Chinese breeding process. (e), The differences of SW, GW, CW, TKW, SSN-T, FSN, FSN/SN, GN between the two haplotypes. In each plot, the central line indicates the median, the box limits indicate the upper and lower quartiles, the whiskers indicate $1.5 \times$ the interquartile range. The significance of the difference was analyzed using Fisher's least significant difference in 8 haplotypes. (f), Distribution of haplotype blocks of locus (chr1B: 114665493-120067499) in the wheat backbone parent. The genome that is the same as the Chinese Spring is indicated in gray (Ref), the heterozygous sites (Heterozygous) are indicated in yellow, and the homozygous change sites (Alt) are indicated in green. The missing sites (NA) are indicated in white. (g), Percentages of different haplotypes in landraces, Chinese and American cultivars. CW, chaff weight.

We identified a SNP in *TraesCS1B02G10450* (G/C at +2,394 bp in the gene), resulting in an amino acid substitution from Glycine to Alanine. To understand the impact of this SNP on spike morphology traits, we examined recombinant inbred lines (RILs) derived from the parents Bainong (carrying the G allele) and Jingshuang (carrying the C allele). On average, RILs with the G allele exhibited higher values for spike weight, GN and grain weight per spike (Table S19). We further constructed near-isogenic lines (NILs) that carried the G allele from Bainong through repeated backcrossing into Jingshuang. Compared to the parental line Jingshuang with the C allele (NIL^C), the improved Jingshuang

(NIL^G) showed an increase in spike weight and grain weight (Figure 5i,m, Tables S21 and S22).

3 | DISCUSSION

Numerous studies have extensively examined wheat spike morphology traits, including rachis (Avni et al., 2017; Guo et al., 2020), SN (Pu-Yang et al., 2022; Kong et al., 2022; Kuzay et al., 2019; Yao et al., 2021), GN (Glenn et al., 2022; Sakuma et al., 2019; Wang et al., 2023; Yao et al., 2021), TKW (Li et al., 2021; Liu et al., 2019;

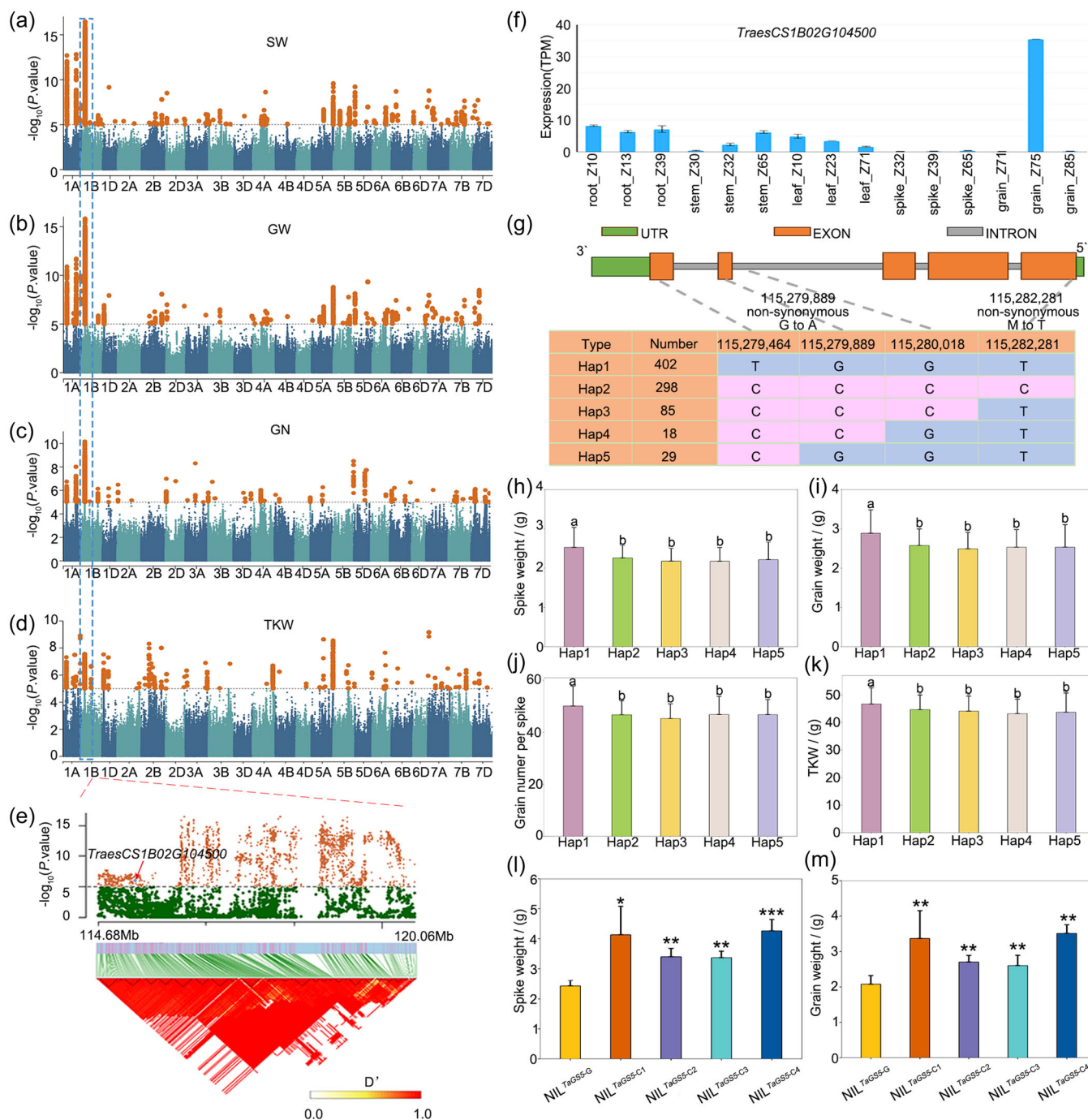


FIGURE 5 GWAS identification of candidate gene for spikelet traits variation and functional validation of superior allele. (a–d), Manhattan plot of spike weight (SW), grain weight per spike (GW), grain number per spike (GN), and thousand kernel weight (TKW). (e) Local Manhattan plot of SW in the region of chr1B: 114.8–120.0 Mb (top), and linkage disequilibrium (LD) heat map (bottom). (f) The gene expression of the candidate gene in different tissues. (g) Gene structure and haplotype analyses of *TraesCS1B02G104500*, the blue and pink colors indicate the reference and the alternative sequence, respectively. (h–k), the differences of SW, GW, GN and TKW among the five haplotypes. (i–m), The differences of spike weight and grain weight among near-isogenic lines (NILs). The data are shown as means $\pm$ SD. Significant differences were determined by Student's *t* test (two sided, * p < 0.05, ** p < 0.01, *** p < 0.001).

Sun et al., 2020). However, our study not only included these traits, but also focused on other traits like grain number per spikelet, grain weight per spikelet, and the number of degenerated spikelets at both the top and the bottom of the spike, which have not been thoroughly analyzed. The goal of crop breeding is to combine desirable traits

from different parental lines to enhance grain yield and environmental adaptation (Scott et al., 2020; Zeng et al., 2017). Identifying favorable alleles aids breeders in selecting suitable parents and guiding breeding efforts for high grain yield and environmental adaptability. In this study, we performed GWAS of 16 spike

morphology traits using 830 Chinese wheat accessions released from 1900 to 2020. We identified significant SNPs associated with multiple traits. This genetic analysis of spike morphology traits will help to uncover potential genomic signals contributing to high wheat yield and environmental adaptation. Additionally, we comprehensively evaluation the frequency of potential favorable alleles throughout the breeding process. Our findings demonstrated a significant positive correlation among the number of favorable alleles and phenotypic modifications, indicating that selecting favorable loci positively impact spike morphology traits during breeding process. These results provide valuable references for parent selection in wheat breeding programs.

Breeders need to consider the potential trade-offs between agronomic traits due to their close correlations (Chen and Lübberstedt, 2010; Fang et al., 2017). To better understand the interactions between spike morphology traits, we constructed a genetic network of these traits. This network offers valuable insights for developing effective strategies that avoid trade-offs in molecular design of spike morphology for enhance environmental adaptation.

The 293 identified genomic regions provided valuable resources for a deeper understanding of the genetic structure and co-localization of wheat spike traits. Notably, we identified several previously studied genes, such as *TaZIM-A1* (Liu et al., 2019), *TaPIN-D1* (Yao et al., 2021), and *TaNAC-B77* (Luo et al., 2021; Yao et al., 2021). The genomic region (chr1B: 114.8–120.0 Mb) overlaps with an interval associated with spike number as determined by a previous study (Pang et al., 2020). This region shows selection preferences in both Chinese and American varieties. Interestingly, we did not identify previously reported genes related to spike traits, such as *WHEAT ORTHOLOG OF APO1* (*WAP01*) (Kuzay et al., 2019), *FRIZZY PANICLE* (*FZP*) (Du et al., 2021), and *VERNALIZATION1* (*VRN1*) (Yan et al., 2003). To clarify the reasons of the undiscovered these genes, we carried out further analyses. Firstly, we retrieved SNPs from the imputation genotype and found no SNPs within these genes. Although these genes significantly influence spike traits, the absence of SNPs in these regions may have hindered their identification. Secondly, GWAS was conducted based on the BLUP values of phenotypic data in ten environments. This may affect the identification of the genes, which exhibit obvious phenotype in a given condition. For instance, *VRN1* (Yan et al., 2003), a wheat vernalization gene, might require specific environmental conditions to identify gene-environment interactions. Thirdly, the wheat spike traits are determined by different genes with weak influences. The effects of well-known genes may be diluted among the numerous contributing genes, making their identification challenging in this population.

The genotype-phenotype map and genetic network of wheat spike morphological traits offer valuable resources for breeding superior wheat varieties with environmental stability. They also provide crucial insights into the genetic mechanisms underlying these traits. Different environments can trigger unique gene expressions, resulting in varied spike phenotypes. Future GWAS analyses should investigate the genetic basis of traits across different environmental conditions, identifying the variability and conservation of genes

related to spike traits. This approach will reveal the mechanisms of wheat's environmental adaptation and inform the selection of improved genes for breeding regionally adaptable varieties.

4 | METHODS

4.1 | Plant materials and growth conditions

The 830 Chinese wheat accessions in this study were categorized into four groups (pre-1960, 1961–1980, 1981–2000, 2001–2020) on the basis of their release year (Table S1). These accessions were cultivated under field conditions in 2020 and 2021 at ten locations: Xiangshan Experimental Station in Beijing (40.06°N, 116.11°E) (location 1), Zhaoxian (37.76°N, 114.78°E) (location 2), Qingdao (36.07°N, 120.38°E) (location 3), Linfen (34.7°N, 112.5°E) (location 4), Xinxiang (35.37°N, 113.90°E) (location 5), Luoyang (34.7°N, 112.5°E) (location 6), Zhengzhou (34.72°N, 113.64°E) (location 7), Nangyang (33.03°N, 112.5°E) (location 8), Suqian (33.86°N, 118.27°E) (location 9), and Yangling (34.31°N, 108.10°E) (location 10). The experiment was conducted using an augmented experimental design by ACBD software with 16 blocks and five controls repeated in each block. Each accession was planted in a 3 m six-row plot with 20 cm between rows (1.2 × 3 m), and sowing rate was decided based on germination rate and 2.7 million/hectare basic seeding. Each field applied about 750 kg of compound fertilizer (N:P:K = 15%:15%:15%)/hectare and 150 kg urea (46%N)/hectare before sowing. A winter irrigation was carried out for each experimental plot from November to next year January except for Linfen and Xinxiang. Other field management followed local agronomic practices.

4.2 | Phenotyping of wheat spike traits

We measured the 16 spike morphology traits across 10 locations, which include 12 direct measurements and four ratios/undirected traits. The 12 traits consist of spike weight (SW), spike length, chaff weight (CW), grain number per spike (GN), grain weight (GW), thousand kernel weight (TKW), SN, fertile spikelet number (FSN), grain number per spikelet (GN-M), grain weight per spikelet (GW-M), the number of sterile spikelets at top of individual spike (SSN-T), the number of sterile spikelets at bottom of individual spike (SSN-B). The four ratios/undirected traits comprised GN/CW, GW/CW, FSN/SN, and GW/SW (Table S2). To determine these traits, five spikes were selected from the main shoots of different plants.

4.3 | Phenotype analysis

The BLUP for each trait in the 10 environments (10 locations × 1 year) was estimated with the 'lme4' package in R (Bates et al., 2015) using the liner mixed model:

$$Y_{ij} = \mu + \text{gen}_i + \text{env}_j + (\text{genXenv})_{ij} + \text{error}_{ij}$$

Where Y_{ij} is the trait of the i_{th} genotype in j_{th} environment, μ means genotype, env means environment, error was set to random. The BLUP value of each inbred line was used for basic statistical analysis of the phenotypes, correlation analysis, cluster analysis and PCA. Correlation among the 16 spike traits was estimated using Pearson correlations between pairs of traits with the 'cor' function in R. The formula used was:

$$\rho_{X,Y} = \text{cov}(X, Y) / (\sigma_X \times \sigma_Y)$$

where $\text{cov}(X,Y)$ is the co-variance between two spike traits, and σ_X and σ_Y represent the standard deviation (s.d.) of the two traits and the cluster analysis of 16 spike morphology traits was performed using the 'ggcorrplot' function in R 'ggcorrplot' package (Kassambara, 2018) with the 'ward.D' method. The 'principal' functions of the 'prcomp' function in R package were used to perform PCA (R version R-3.5). The broad-sense heritability (h^2) was calculated using the following formula:

$$H^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_{GE}^2/2 + \sigma_e^2/2r)$$

Where σ_G^2 is the genotypic variance, σ_{GE}^2 is the genotype by the environment effect, σ_e^2 is the residual error, and r is the average number of replications ($r = 5$).

4.4 | Genotype imputation and SNP filter

Genotype imputation (van Leeuwen et al., 2015) estimates missing genotypes from the reference panel and enhance the power of genome-wide association studies and conducting subsequent fine-mapping. For this study, we utilized wheat 660 K chip sequencing to generate genotype data for 830 wheat accessions, and SNPs for 830 wheat accessions were obtained throughout genotype imputation based on the SNPs from 306 worldwide wheat accessions (Zhou et al., 2020a). Genotype imputations followed the methodology of previous studied (Guo et al., 2022; He et al., 2008; Hou et al., 2014; Zhou et al., 2020a). We used the Beagle program (Browning et al., 2018; Browning et al., 2021) for genotype imputation, applying parameters of 'overlap = 500, window = 5000, ne = 12,000' (He et al., 2019). The accuracy of imputation was estimated to be over 90% in both GT and GL modes (He et al., 2019).

4.5 | Identification of selective sweeps for wheat breeding improvement

The XP-CLR test was utilized to detect selective sweeps, identifying potential improvement-related sweeps (Chen & Patterson, & Reich, 2010). We used 116 accessions from pre-1960 as a reference and compared them with 98 accessions for 1961–1980, 198 accessions for 1981–2000, 418 accessions for 2001–2020 to identify potential selection-related sweeps. The XP-CLR score between two

wheat populations (pre-1960 vs 1961–1980, pre-1960 vs 1981–2000, pre-1960 vs 2001–2020) was calculated using the parameters "--rate 1e-8 --maxsnps 200 --size 500000 --step 100000". To detect the selection of the identified genomic regions, selection sweeps were ranked based on decreasing XP-CLR scores, with the top 5% regions as selective sweeps.

4.6 | Identification of differentiated regions during breeding area

We estimated F_{st} in 100 kb sliding windows with a step size of 1 kb to quantify genomic differentiation between the pre-1960 accession pool and the 1961–1980, 1981–2000, 2001–2020 accession pools using VCFtools (v.0.1.15) (Danecek et al., 2011). Initially, sliding windows with the top 5% highest F_{st} values were selected. Outlier windows less than 20 kb were then merged. Differentiated regions were identified as those in the top 5% of both statistics.

4.7 | GWAS of spike traits and candidate gene annotation

A large-scale GWAS was conducted on 16 spike morphology traits using 21,279,862 high-quality SNPs (MAF > 0.05; missing rate < 20%; missing genotype rate < 10%). Genome-wide association analysis was performed in GEMMA (version 0.98.4) (Zhou and Stephens, 2012) using the mixed linear model (MLM). The kinship (K) matrix was build using simple matching coefficients. The top three principal components (PCs) from principal component analysis (PCA) were used for population structure correction in Plink (Purcell et al., 2007) with parameters set to "--pca 10". Genetic relationship between accessions were modeled as a random effect using the K matrix. We conservatively chose $-\log_{10}(P\text{-value}) = 5.0$ as the threshold for significant association to control genome-wide type I errors (Hao et al., 2020; Liu et al., 2023a, 2023c). Significant SNPs (above the threshold) in LD blocks with $r^2 \geq 0.2$ were considered candidate genomic region. Gene within the candidate genomic regions were selected as candidate genes for the GWAS associations.

4.8 | Linkage disequilibrium analysis

The software LDBlockShow (Dong et al., 2021) was utilized to perform pairwise LD analysis of the associated genomic region during the chr1B: 114.8–120.0 Mb.

4.9 | Construction of trait-candidate genomic regions association networks

The analysis of association network was performed using Cytoscape (Version: 3.2.1) (Shannon et al., 2003). The network illustrated the

connections between the traits and their candidate genomic regions, as well as the links between different traits candidate genomic regions (average $r^2 \geq 0.6$). First, all 16 traits and their corresponding candidate genomic regions were designated as nodes. Links between individual traits and candidate genomic regions, as well as links within consensus candidate genomic regions, were assigned as edges. The edges' effective scores between trait and trait-relative candidate genomic regions were represented by most significant P -values of the correlation between SNP in candidate genomic regions and trait, while edges' effective scores between pairs of candidate regions was represented as peak SNP LD.

4.10 | Analysis of favorable alleles by GWAS significant SNPs

The significant SNP sets generated by GWAS and BLUP values for 16 traits were utilized to explore the changes in the frequency of favorable alleles during wheat breeding process and improvement in China. Multiple comparisons of phenotypic traits corresponding to different SNPs were performed to identify SNPs with significant phenotypic differences, and the phenotypes with large values of spike weight, grain weight, grain number and SN were called favorable allele, while the phenotypes with small values of aborted spike number in spike top position (SSN-T), aborted spike number in spike bottom position (SSN-B) were called favorable allele. Finally, the ratio of favorable allele was calculated for each of the 16 traits in 830 wheat accessions.

4.11 | Identification of haplotypes in genomic regions and candidate gene

Initially, VCFtools (Danecek et al., 2011) was employed to extract SNPs from candidate intervals or genes into separate VCF files. Subsequently, the VCF files were imported into R using the vcfR package (Knaus and Grunwald, 2017). SNPs matching the reference genome (Chinese Spring) were coded as 0, mutated sites as 2, missing sites as 3, and heterozygous sites as the R package pheatmap (<https://CRAN.Rproject.org/package=pheatmap>) was then utilized to visualize interval haplotypes and conduct cluster analysis. Principal Component Analysis (PCA) was applied to examine the clustering results. Categories containing more than 5 accessions were selected for haplotype statistical analysis.

ACKNOWLEDGEMENTS

This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDA24010104-2), the National Natural Science Foundation of China (32272122), and Supported by the Post-doctoral Fellowship Program of CPSF under Grant Number GZB20240820.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Jianhui Wu  <https://orcid.org/0000-0001-8154-1199>

REFERENCES

- Alqudah, A.M., Haile, J.K., Alomari, D.Z., Pozniak, C.J., Kobijlski, B. & Börner, A. (2020) Genome-wide and SNP network analyses reveal genetic control of spikelet sterility and yield-related traits in wheat. *Scientific Reports*, 10, 2098.
- Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S.O., Gundlach, H. et al. (2017) Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, 357, 93–97.
- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Browning, B.L., Tian, X., Zhou, Y. & Browning, S.R. (2021) Fast two-stage phasing of large-scale sequence data. *The American Journal of Human Genetics*, 108, 1880–1890.
- Browning, B.L., Zhou, Y. & Browning, S.R. (2018) A one-penny imputed genome from next-generation reference panels. *The American Journal of Human Genetics*, 103, 338–348.
- Cao, S., Xu, D., Hanif, M., Xia, X. & He, Z. (2020) Genetic architecture underpinning yield component traits in wheat. *Theoretical and Applied Genetics*, 133, 1811–1823.
- Chen, H., Patterson, N. & Reich, D. (2010) Population differentiation as a test for selective sweeps. *Genome Research*, 20, 393–402.
- Chen, Y. & Lübberstedt, T. (2010) Molecular basis of trait correlations. *Trends in Plant Science*, 15, 454–461.
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A. et al. (2011) The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158.
- van Dijk, M., Morley, T., Rau, M.L. & Saghai, Y. (2021) A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nature Food*, 2, 494–501.
- Dong, S.S., He, W.M., Ji, J.J., Zhang, C., Guo, Y. & Yang, T.L. (2021) LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Briefings in Bioinformatics*, 22, 1–6.
- Du, D., Zhang, D., Yuan, J., Feng, M., Li, Z., Wang, Z. et al. (2021) FRIZZY PANICLE defines a regulatory hub for simultaneously controlling spikelet formation and awn elongation in bread wheat. *New Phytologist*, 231, 814–833.
- Fang, C., Ma, Y., Wu, S., Liu, Z., Wang, Z., Yang, R. et al. (2017) Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biology*, 18, 161.
- Glenn, P., Zhang, J., Brown-Guedira, G., DeWitt, N., Cook, J.P., Li, K. et al. (2022) Identification and characterization of a natural polymorphism in FT-A2 associated with increased number of grains per spike in wheat. *Theoretical and Applied Genetics*, 135, 679–692.
- Guo, L., Ma, M., Wu, L., Zhou, M., Li, M., Wu, B. et al. (2022) Modified expression of *TaCYP78A5* enhances grain weight with yield potential by accumulating auxin in wheat (*Triticum aestivum* L.). *Plant Biotechnology Journal*, 20, 168–182.
- Guo, W., Xin, M., Wang, Z., Yao, Y., Hu, Z., Song, W. et al. (2020) Origin and adaptation to high altitude of Tibetan semi-wild wheat. *Nature Communications*, 11, 5085.

- Guo, Z., Zhao, Y., Röder, M.S., Reif, J.C., Ganal, M.W., Chen, D. et al. (2018) Manipulation and prediction of spike morphology traits for the improvement of grain yield in wheat. *Scientific Reports*, 8, 14435.
- Hao, B., Ma, J., Chen, P., Jiang, L., Wang, X., Li, C. et al. (2021) Wheat breeding in China over the past 80 years has increased grain zinc but decreased grain iron concentration. *Field Crops Research*, 271, 108253.
- Hao, C., Jiao, C., Hou, J., Li, T., Liu, H., Wang, Y. et al. (2020) Resequencing of 145 landmark cultivars reveals asymmetric sub-genome selection and strong founder genotype effects on wheat breeding in China. *Molecular Plant*, 13, 1733–1751.
- Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D. & Poncet, C. (2007) Grinding up wheat: a massive loss of nucleotide diversity since domestication. *Molecular Biology and Evolution*, 24, 1506–1517.
- He, F., Pasam, R., Shi, F., Kant, S., Keeble-Gagnere, G., Kay, P. et al. (2019) Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nature Genetics*, 51, 896–904.
- He, G., Zhang, Y., Liu, P., Jing, Y., Zhang, L., Zhu, Y. et al. (2021) The transcription factor *TaLAX1* interacts with *Q* to antagonistically regulate grain threshability and spike morphogenesis in bread wheat. *New Phytologist*, 230, 988–1002.
- He, X.Y., Zhang, Y.L., He, Z.H., Wu, Y.P., Xiao, Y.G., Ma, C.X. et al. (2008) Characterization of phytoene synthase 1 gene (*Psy1*) located on common wheat chromosome 7A and development of a functional marker. *Theoretical and Applied Genetics*, 116, 213–221.
- 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.
- Kassambara, A. (2018). ggcorrplot: Visualization of a Correlation Matrix using 'ggplot2'. R package version 0.1.2. <https://CRAN.R-project.org/package=ggcorrplot>
- Katz, A., Byrne, P., Reid, S., Bratschun, S., Haley, S. & Pearce, S. (2022) Identification and validation of a QTL for spikelet number on chromosome arm 6BL of common wheat (*Triticum aestivum* L.). *Molecular Breeding*, 42, 17.
- Knaus, B.J. & Grünwald, N.J. (2017) vcfr: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, 17, 44–53.
- Kong, X., Wang, F., Geng, S., Guan, J., Tao, S., Jia, M. et al. (2022) The wheat *AGL6*-like MADS-box gene is a master regulator for floral organ identity and a target for spikelet meristem development manipulation. *Plant Biotechnology Journal*, 20, 75–88.
- Kuzay, S., Xu, Y., Zhang, J., Katz, A., Pearce, S., Su, Z. et al. (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.
- Lázaro, L., Abbate, P.E., Cogliatti, D.H. & Andrade, F.H. (2010) Relationship between yield, growth and spike weight in wheat under phosphorus deficiency and shading. *The Journal of Agricultural Science*, 148, 83–93.
- van Leeuwen, E.M., Kanterakis, A., Deelen, P., Kattenberg, M.V., Slagboom, P.E., de Bakker, P.I.W. et al. (2015) Population-specific genotype imputations using minimac or IMPUTE2. *Nature Protocols*, 10, 1285–1296.
- Li, A.L., Hao, C.Y., Wang, Z.Y., Geng, S.F., Jia, M.L., Wang, F. et al. (2022a) Wheat breeding history reveals synergistic selection of pleiotropic genomic sites for plant architecture and grain yield. *Molecular Plant*, 15, 504–519.
- Li, H., Nie, F., Zhu, L.L., Mu, M.H., Fan, R.X., Li, J.Y. et al. (2022b) New insights into the dispersion history and adaptive evolution of taxon in China. *Journal of Genetics and Genomics*, 49, 185–194.
- Li, J., Xie, L., Tian, X., Liu, S., Xu, D., Jin, H. et al. (2021) *Ta*NAC100 acts as an integrator of seed protein and starch synthesis exerting pleiotropic effects on agronomic traits in wheat. *The Plant Journal*, 108, 829–840.
- Lin, Y., Jiang, X., Hu, H., Zhou, K., Wang, Q., Yu, S. et al. (2021) QTL mapping for grain number per spikelet in wheat using a high-density genetic map. *The Crop Journal*, 9, 1108–1114.
- Liu, H., Li, T., Wang, Y., Zheng, J., Li, H., Hao, C. et al. (2019) *Ta*ZIM-A1 negatively regulates flowering time in common wheat (*Triticum aestivum* L.). *Journal of Integrative Plant Biology*, 61, 359–376.
- Liu, Y., Shen, K., Yin, C., Xu, X., Yu, X., Ye, B. et al. (2023a) Genetic basis of geographical differentiation and breeding selection for wheat plant architecture traits. *Genome Biology*, 24, 114.
- Liu, Y.Y., Chen, J., Yin, C.B., Wang, Z.Y., Wu, H., Shen, K.C. et al. (2023c) A high-resolution genotype–phenotype map identifies the *Ta*SPL17 controlling grain number and size in wheat. *Genome Biology*, 24, 1–28.
- Liu, Y.Y., Shen, K.C., Yin, C.B., Xu, X.W., Yu, X.C., Ye, B.T. et al. (2023b) Genetic basis of geographical differentiation and breeding selection for wheat plant architecture traits. *Genome Biology*, 24, 1–25.
- Luo, G., Shen, L., Zhao, S., Li, R., Song, Y., Song, S. et al. (2021) Genome-wide identification of seed storage protein gene regulators in wheat through coexpression analysis. *The Plant Journal*, 108, 1704–1720.
- Nigus, M., Shimelis, H., Mathew, I. & Abady, S. (2022) Wheat production in the highlands of eastern Ethiopia: opportunities, challenges and coping strategies of rust diseases. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, 72, 563–575.
- Niu, J.Q., Ma, S.W., Zheng, S.S., Zhang, C., Lu, Y.R., Si, Y.Q. et al. (2023) Whole-genome sequencing of diverse wheat accessions uncovers genetic changes during modern breeding in China and the United States. *The Plant Cell*, 35, 4199–4216.
- Pang, Y., Liu, C., Wang, D., St. Amand, P., Bernardo, A. & Li, W. (2020) High-resolution genome-wide association study identifies genomic regions and candidate genes for important agronomic traits in wheat. *Molecular Plant*, 13, 1311–1327.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R. & Bender, D. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81, 559–575.
- Pu-Yang, D., Zi-Qiang, M., Hua-Ping, T., Yang, M., Mei, D., Qian-Tao, J. et al. (2022) A major and stable QTL for wheat spikelet number per spike validated in different genetic backgrounds. *Journal of Integrative Agriculture*, 21, 1551–1562.
- Qin, X., Zhang, F., Liu, C., Yu, H., Cao, B., Tian, S. et al. (2015) Wheat yield improvements in China: past trends and future directions. *Field Crops Research*, 177, 117–124.
- Rezaei, E.E., Webber, H., Asseng, S., Boote, K., Durand, J.L. & Ewert, F. (2023) Climate change impacts on crop yields. *Nature Reviews Earth & Environment*, 4, 831–846.
- Sakuma, S., Golan, G., Guo, Z., Ogawa, T., Tagiri, A., Sugimoto, K. et al. (2019) Unleashing floret fertility in wheat through the mutation of a homeobox gene. *Proceedings of the National Academy of Sciences*, 116, 5182–5187.
- Sansaloni, C., Franco, J., Santos, B., Percival-Alwyn, L., Singh, S., Petroli, C. et al. (2020) Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. *Nature Communications*, 11, 4572.
- Scott, M.F., Ladejobi, O., Amer, S., Bentley, A.R., Biernaskie, J., Boden, S.A. et al. (2020) Multi-parent populations in crops: a toolbox integrating genomics and genetic mapping with breeding. *Heredity*, 125, 396–416.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T. & Ramage, D. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13, 2498–2504.
- Shiferaw, B., Smale, M., Braun, H.J., Duveiller, E., Reynolds, M. & Muricho, G. (2013) Crops that feed the world 10. Past successes and

- future challenges to the role played by wheat in global food security. *Food Security*, 5, 291–317.
- Sun, H., Zhang, W., Wu, Y., Gao, L., Cui, F. & Zhao, C. (2020) The circadian clock gene, *TaPRR1*, is associated with yield-related traits in wheat (*Triticum aestivum* L.). *Frontiers in Plant Science*, 11, 1–16.
- Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M.T., Brinton, J. et al. (2020) Multiple wheat genomes reveal global variation in modern breeding. *Nature*, 588, 277–283.
- Wang, W., Pan, Q., Tian, B., He, F., Chen, Y., Bai, G. et al. (2019) Gene editing of the wheat homologs of TONNEAU1-recruiting motif encoding gene affects grain shape and weight in wheat. *The Plant Journal*, 100, 251–264.
- Wang, Y., Du, F., Wang, J., Wang, K., Tian, C., Qi, X. et al. (2022) Improving bread wheat yield through modulating an unselected AP2/ERF gene. *Nature Plants*, 8, 930–939.
- Wang, Z., Wang, W., Xie, X., Wang, Y., Yang, Z., Peng, H. et al. (2022) Dispersed emergence and protracted domestication of polyploid wheat uncovered by mosaic ancestral haploblock inference. *Nature Communications*, 13, 3891.
- Wittern, L.M., Barrero, J.M., Bovill, W.D., Verbyla, K.L., Hughes, T., Swain, S.M. et al. (2022) Overexpression of the WAPO-A1 gene increases the number of spikelets per spike in bread wheat. *Scientific Reports*, 12, 14229.
- 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.
- Yao, F.Q., Li, X.H., Wang, H., Song, Y.N., Li, Z.Q., Li, X.G. et al. (2021) Down-expression of *TaPIN1s* increases the tiller number and grain yield in wheat. *BMC Plant Biology*, 21, 443–453.
- Zeng, D., Tian, Z., Rao, Y., Dong, G., Yang, Y., Huang, L. et al. (2017) Rational design of high-yield and superior-quality rice. *Nature Plants*, 3, 17031.
- Zhang, Y., Li, D., Zhang, D., Zhao, X., Cao, X., Dong, L. et al. (2018) Analysis of the functions of TaGW2homoeologs in wheat grain weight and protein content traits. *The Plant Journal*, 94, 857–866.
- Zhao, X., Guo, Y., Kang, L., Yin, C., Bi, A. & Xu, D. (2023) Population genomics unravels the Holocene history of bread wheat and its relatives. *Nature Plants*, 9, 403–419.
- Zhou, H., Riche, A.B., Hawkesford, M.J., Whalley, W.R., Atkinson, B.S. & Sturrock, C.J. (2021a) Determination of wheat spike and spikelet architecture and grain traits using X-Ray Computed Tomography imaging. *Plant Methods*, 17, 1–9.
- Zhou, X. & Stephens, M. (2012) Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, 44, 821–824.
- Zhou, Y., Bai, S.L., Li, H., Sun, G.L., Zhang, D.L., Ma, F.F. et al. (2021b) Introgressing the genome into wheat as a basis for cereal improvement. *Nature Plants*, 7, 774–786.
- Zhou, Y., Zhao, X.B., Li, Y.W., Xu, J., Bi, A.Y., Kang, L.P. et al. (2020) *Triticum* population sequencing provides insights into wheat adaptation. *Nature Genetics*, 52, 1412–1422.

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: Liu, Y., Yu, R., Shen, L., Sun, M., Peng, Y., Zeng, Q. et al. (2024) Genomic insights into the modifications of spike morphology traits during wheat breeding. *Plant, Cell & Environment*, 47, 5470–5482. <https://doi.org/10.1111/pce.15117>