Identification of Genetic Loci on Chromosome 4B for Improving the Grain Number per Spike in Pre-Breeding Lines of Wheat

Jinpeng Zhang 1, Qifu Yao 2, Ruixin Li 1, Yuqing Lu 1, Shenghui Zhou 1, Haiming Han 1, Weihua Liu 1, Xiuquan Li 1, Xining Yang 1 and Lihui Li 1,*

1 Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China; zhangjinpeng@caas.cn (J.Z.); liruixin2021@126.com (R.L.); luyuqing@caas.cn (Y.L.); zhoushenghui@caas.cn (S.Z.); hanhaiming@caas.cn (H.H.); liuweihua@caas.cn (W.L.); lixiuquan@caas.cn (X.L.); yangxining@caas.cn (X.Y.)
2 College of Agro-Forestry Engineering & Planning, Tongren University, Tongren 554300, China; yaoqifu@126.com
* Correspondence: lilihui@caas.cn

Abstract: The grain number per spike (GNPS) is an important yield component, and much attention is given to the increase in GNPS for current yield improvement of common wheat. Here, a panel of 259 pre-breeding lines and elite commercial varieties were collected for the investigation of 12 agronomic traits, especially for spike-related traits, with 2-year replicates. The high correlation between GNPS and kernel number per spikelet (KNS) suggested that the high GNPS trait in our pre-breeding lines was mainly controlled by grain set number per spikelet. Genome-wide association studies (GWAS) using the 660K SNP genotyping assay suggested that a major locus on chromosomes 4BS contributed to the high GNPS trait, which contributed to 33% and 48% of the variation in KNS and GNPS, respectively. A good diagnostic KASP marker AX-109286577 flanking the 4BS locus was developed for easy selection of the large spike trait. Taken together, the results suggested that untapped rare allele variation in our pre-breeding lines can be used for improvement of the yield component of set grain number per spike.

Keywords: wheat; GWAS; grain number per spike; kernel number per spikelet; pre-breeding; large spike

1. Introduction

Common wheat (Triticum aestivum, allohexaploid, AABBDD genome) is the most widely grown crop globally, providing approximately 20% of the calories and protein in human diets [1]. Increasing wheat yield is a global challenge for food security, and six complementary approaches have been proposed, including optimizing developmental patterns to maximize spike fertility and grain number, and improving potential grain size and grain filling [2]. Wheat yield components are determined by three main factors, namely the spike number per unit area, grain number per spike (GNPS), and thousand-kernel weight (TKW). Breeding practices have indicated that improvement of the spike number per unit area and TKW have increased to very high levels in China and other countries [3–5], while the increase in GNPS will therefore be of widespread concern for future yield improvements [2,6].

Natural variation in wheat genetic stocks constitutes an important gene source for wheat improvement. A number of quantitative trait loci (QTLs) affecting grain number per spike have been mapped in common wheat. Some major QTLs were identified for GNPS on chromosomes 1A, 2B, 4B, 5B [7–10], for spikelet number per spike (SNS) on chromosome 2D, 4A, 5A, 6A, 7AL [11–14], and for kernel number per spikelet (KNS) on 2B [15] based on a biparental cross population. Other loci related to GNPS were detected on 2A, 2B, 4B and 5A by using association mapping of materials with natural variation [16,17].
A few genes have been cloned to increase the GNPS, SNS and KNS. The GNI1 gene located on chromosome 2A was identified as responsible for increased floret fertility and grain number [18]. The physical region of the 7A-QTL revealed that TaAPO-A1 positively controls the spikelet number on the spikes [19,20]. An associative transcriptomic approach identified three candidate genes, TaVRS1-2B, TaTFL1-2D, and TaPAP2-5A, that affect inflorescence development in wheat and increase wheat grain number [21].

New sources of germplasm with maximized grain number per spike are necessary for increasing the yield potential, particularly in regard to spikelet fertility, i.e., the number of kernels per spikelet [22]. Our team continued to pre-breed large spike germplasms and reported two high GNPS wheat germplasm lines Pubing 3228 and Pubing 3504 for increasing grain number [23,24]. By crossing the high GNPS wheat germplasm lines with parents collected worldwide, we produced a panel of pre-breeding lines with diversified traits, including large spike, high grain weight, earliness and dwarf characteristics. Among these agronomic traits, spike-related traits are the most prominent. The panel of pre-breeding lines show the wide phenotypic variation, which is suitable for genome-wide association studies (GWASs). The aim of this study was mainly to evaluate yield component-related trait variations and identify the genetic loci of natural variation by completing a 2-year phenotypic investigation and whole-genome SNP genotyping. GWAS and population genetics analysis were employed to identify the genomic region for increasing grain number per spike, and the study provides us with new genetic resources and genetic loci for improvement of yield potential in wheat.

2. Materials and Methods

2.1. Plant Materials

A panel of 255 derivative lines from the cross between wheat and *Agropyron cristatum* were collected. These derivative lines were developed by Professor Lihui Li of the Chinese academy of agricultural sciences since 1998 [25] and show wide genetic diversity. These pre-breeding lines carry the genetic component from the worldwide parents. Among these parents, common wheat-*A. cristatum* derivatives include FC_big_spike, Pubing 4201, Pubing 4363, 924144-2, Zhi 7730, Pubing 2011, Pubing 4844, Zi 08-69. The introduced parents include Jagger and Hi-Line from USA, CMH 83.605 from CIMMYT, and Fukuho from Japan. The commercial varieties in China include Jimai 22, Zhoumai 18, Xiaoyan 22 and Goacheng 8901. Combined with four elite commercial varieties (Zhoumai 18, Zhoumai 27, Jimai 22, Fukuho), a total of 259 lines were used for population structure and association mapping analysis. All the lines were sown at Xinxiang station of CAAS (Henan Province, China) during the two winter wheat growing seasons of 2016–2017 and 2017–2018. The plant density was 25 cm by 10 cm with a 2 m length. The field was arranged in a randomized complete block design with two replications. The pedigree information is provided in Table S1.

2.2. Genotyping

A wheat 660K SNP array designed by the Chinese Academy of Agricultural Sciences and synthesized by Affymetrix was applied to genotype all 259 collected wheat accesses (http://wheat.pw.usda.gov/ggpages/topics/Wheat660_SNP_array_developed_by_CAAS.pdf, accessed on 30 November 2021). The wheat 660K SNP array contained wide genetic variation from hexaploid and tetraploid wheat, emmer wheat and *Aegilops tauschii*. All genotypes were filtered by PLINK v1.90p 64-bit software (https://www.cog-genomics.org/plink/1.9, accessed on 30 November 2021). After removing nucleotide variations with missing rates ≥ 0.1 and minor allele frequency (<0.05), 310,097 SNPs were used for the population genetics analysis and GWAS (Table S2).

2.3. Population Structure and Phylogenetic Tree Analysis

Population structure was assessed using Admixture 1.3 software [26]. LD decay was analyzed by the PopLDdecay software v3.40 [27]. K values were set from K = 2 to
K = 18. The minimum CV error value appeared at K = 16. The phylogenetic tree was constructed with the neighbor joining (NJ) method with the program PHYLIP package 3.697 [https://evolution.genetics.washington.edu/phylip.html, accessed on 30 November 2021] and was visualized using the online tool of the ITOL website (http://itol.embl.de, accessed on 30 November 2021). Principal component analysis (PCA) using whole-genome SNPs was performed with Plink 1.9 software.

2.4. Phenotype Evaluation and Correlation Analysis

After harvest, the plant height (PH), fertile tiller numbers, grain number per spike (GNPS), kernel number per spikelet (KNS), total spikelet number per spike, spike length, and fertile spikelet number per spike were investigated. The heading date (Hd) was recorded by counting days from the sowing date to the date when ears of approximately half of the genotypes were fully visible. After threshing, the grain length (GL), grain width (GW) and thousand-kernel weight (TKW) were measured by the image analysis method provided with SC-E software (Hangzhou Wanshen Detection Technology Co., Ltd., Hangzhou, China). We measured five individual plants for each line except for the heading date trait; the average value of individuals was calculated for trait evaluation. All phenotype data are provided in Table S3. The broad heritability and BLUPs for traits were estimated using the R package lme4 [28]. Broad-sense heritability ($H^2$) was calculated using the following formula: $H^2 = \text{var (LINE)}/[\text{var (LINE)} + \text{var (LINE:YEAR)}/2 + \text{var (RESIDUAL)}]/4$. Visualization of a correlation matrix was performed with the R package ggcorrplot (http://www.sthda.com/english/wiki/ggcorrplot-visualization-of-a-correlation-matrix-using-ggplot2, accessed on 30 November 2021). Boxplot figures were drawn by using the R package ggpubr (https://cran.r-project.org/web/packages/ggpubr/index.html, accessed on 30 November 2021).

2.5. Genome-Wide Association Studies (GWAS)

GWAS was performed with the R package Genomic Association and Prediction Integrated Tool (GAPIT, version 3) using the mixed linear model (MLM) [29]. Principal component analysis (PCA = 4) was used to account for population structure and fitted as fixed effect [30]. The markers with an adjusted $-\log_{10}(P) \geq 6.0$ were regarded as significant for each trait. A high-quality drawing tool was used to create Manhattan plots of the genomic analysis with the R package CMplot (https://github.com/YinLiLin/R-CMplot, accessed on 30 November 2021).

2.6. KASP Markers Development and Validation

All SNP probe sequences on the wheat 660K SNP Array were used in a BLASTN search (e-value cut-off of $1 \times 10^{-5}$) against the wheat reference sequence (IWGSC RefSeq v1), and the flanking 100 bp sequences near the SNP position were extracted for primer design. The KASP™ primer set containing two allele-specific forward primers and one common reverse primer (Table S4) was designed by LGC Biosearch Technologies, Hoddesdon, UK. The primer set was synthesized with a PAGE purification method by Sangon Biotech (Shanghai, China). KASP assays were performed with 384-plate format 5 µL reactions (2.5 µL of template DNA (25 ng), 2.5 µL of 2 × Kaspar mix, and 0.07 µL of primer mix). A FluOstar Omega Microplate Reader was used to read the fluorescence signals.

2.7. Candidate Gene Analysis

The target region gene lists on wheat chromosome 4BS were extracted and annotated using the tools of the Triticeae Multiomics Center website (http://202.194.139.32/tools/intervalTools.html, accessed on 30 November 2021). The expression data of wheat genes were collected from the WheatExp database (https://wheat.pw.usda.gov/WheatExp/, accessed on 30 November 2021).
3. Results

3.1. Phenotype Variation and Trait Correlation Analysis

A panel of 255 pre-breeding lines derived from the cross between common wheat and A.cristatum and 4 elite commercial varieties were collected for agronomic trait investigation. To avoid the effects of environmental error on phenotype traits, 2-year replicates were conducted. Among the 12 investigated traits, PH, GNPS, total spikelet number, spike length, spikelet density, KNS and TKW showed wide variation ranges, whereas Hd showed the smallest variation. The maximum coefficient of variation (CV) was the 22.72% for the GNPS trait. Broad-sense heritability ($H^2$) of plant height showed the highest value, and the $H^2$ of GNPS was up to 0.78, ranking third (Table 1). The panel of the collected pre-breeding lines contained the line Pubing 3228 with prominently high GNPS features [23]. The correlation of 12 traits showed that GNPS is highly correlated with fertile spikelet numbers and KNS (Figure 1A), which suggests that the high GNPS trait in our pre-breeding lines was controlled both by kernel number per spikelet and spikelet number. The distribution figure shows that the GNPS trait of 259 lines varied widely in both 2-year replicates and exhibited a skewed distribution (Figure 1B), which suggests that major genes are responsible for the high GNPS trait and the panel is suitable for GWAS.

Table 1. Investigation of 12 agronomic traits of 259 lines in 2-year replicates.

| Trait                  | 2016–2017 Variation Range | CV † | Median | Variation Range | CV † | Median | $H^2$ ‡ |
|------------------------|---------------------------|------|--------|-----------------|------|--------|---------|
| PH(cm)                 | 65–115                    | 9.46%| 83     | 57–93           | 9.34%| 75     | 0.91    |
| Spike_number           | 5.9–17.6                  | 18.77%| 9.9  | 5.2–17.3       | 15.22%| 10.8  | 0.56    |
| GNPS                   | 37.2–105.6               | 22.72%| 9.9  | 34.0–96.9      | 17.14%| 54.2  | 0.78    |
| Total_spikelet_number  | 16.8–26.4                | 6.75%| 22.3  | 16.5–24.0      | 6.38%| 20.0  | 0.64    |
| Fertile_spikelet_number| 13.6–24.8                | 11.87%| 19.0 | 14.8–23.9      | 7.82%| 19.0  | 0.52    |
| Spike_length(cm)       | 7.7–14.5                 | 9.73%| 10.7  | 7.1–11.5       | 8.61%| 9.3   | 0.74    |
| Spikelet_density       | 16.2–26.7                | 10.27%| 21.0 | 16.8–27.7      | 9.39%| 21.9  | 0.82    |
| KNS                    | 2.7–6.0                  | 17.99%| 3.5  | 3.1–6.1        | 10.67%| 3.9   | 0.68    |
| TKW(g)                 | 32.9–60.8                | 11.20%| 49.0 | 26.1–68.9      | 14.40%| 47.2  | 0.74    |
| GW(mm)                 | 2.6–3.9                  | 8.19%| 3.5  | 3.0–3.9        | 5.06%| 3.6   | 0.66    |
| GL(mm)                 | 4.8–7.6                  | 7.29%| 6.6  | 5.9–7.9        | 4.05%| 6.9   | 0.65    |
| Hd                     | 180–188                  | 0.99%| 182.0 | 166–179      | 1.24%| 175.0 | 0.63    |

† CV represents coefficient of variation. ‡ $H^2$ represents broad-sense heritability. GNPS and KNS represent grain number per spike and kernel number per spikelet, respectively.

3.2. Estimation of Population Structure

To reveal the genetic relationship of all pre-breeding lines, estimate the population structure and identify the number of subgroups, two methods, including phylogenetic tree and population structure analyses, were used. An SNP-based phylogenetic tree (with 310,097 SNP markers distributed throughout the whole genome of wheat) divided all accessions into 10 major groups, including a breeding line group and 9 other pre-breeding lines groups (Figure 2A). The admixture structure showed identical results to the phylogenetic tree (Figure 2B). The principal component analysis (PCA) of all accessions was consistent with the phylogeny and the population structure (Figure 2C). We found that the population structure of 259 pre-breeding panels showed the perfect fit with the pedigree relationship. In our pre-breeding lines, the Pubing01 group was the largest group, containing 80 lines, in which the commercial varieties Zhoumai 18 and Zhoumai 27 belonged to the same group. The Pubing01 group can be further divided into eight subgroups. Pubing02 and Pubing04 were the second largest groups, which contained 48 lines. Pubing03, Pubing05, Pubing06, Pubing07, Pubing08 and Pubing09 were divided into relatively independent groups. The commercial variety Jimai 22 is the same group as Pubing06. Ten breeding lines (Kenong 2011, PZW 9, Pubing 176, Pubing 143, Pubing 9946, Pubing 3228, Pubing 3882, Jinmai 80, Pubing 701 and Pubing 717) and Fukuho introduced...
from Japan (receipt cultivar of a wide cross between wheat and A. cristatum) showed rich genetic variations, which means that the breeding lines carry more complex genetic backgrounds than pre-breeding lines. According to the pedigree, the Pubing07, Pubing08 and Pubing09 group lines belonging to a relatively separate clade are early generation pre-breeding lines, and their spike traits show prominently high GNPS. As a result, genetic relationships revealed that our pre-breeding lines show wide variation and their genetic backgrounds are also diversified for association mapping studies.

3.3. GWAS Identify Genetic Loci Associated with Spike-Related Traits

To identify the genetic loci controlling the GNPS trait, a GWAS was completed using 259 accessions and whole-genome SNP markers combined with 2 years of phenotypic data. After filtering out low-quality SNP markers, a total of 310,097 SNPs were retained. These SNPs were distributed over each of the 21 chromosomes (Figure S1). Marker density varied among chromosomes with a minimum of 4.53 markers per Mb on chromosome 4D and a maximum of 41.02 markers per Mb on chromosome 5B. Principal component analysis (PCA) was performed with GAPIT software. We conducted a Bayesian information criterion (BIC)-based model selection to find the optimal number of PCs for the GWAS. The scree plot shows that the eigenvalues start to form a relatively straight line after the fourth principal component (Figure S2). Therefore, we selected PCA = 4 for optimum population structure estimation in the GWAS. In total, 477 marker-trait associations (MTAs) were detected for 12 traits across all the environments (Table S5). Finally, we identified 3 significant QTL regions on wheat chromosomes 1AL, 4BS and 5AL (Figures 3 and S3). SNP markers AX-109302049 and AX-108839971 flanking the 1AL locus in the physical region from 583.3 M to 584.7 M were tightly associated with both GNPS and KNS with $p$-values of $1.13 \times 10^{-7}$ and $3.93 \times 10^{-6}$, respectively. Another major locus on chromosome 4BS delimited in the region between markers AX-109286577 and AX-94438527 was associated with GNPS and KNS traits, which suggests that the genomic region exhibits pleiotropic action. Moreover, we identified a locus on chromosome 5AL flanked by the marker AX-111514840 at the physical position 537.133 Mb, which has pleiotropic effects on spike length, spikelet number, and spikelet density. The reported Q gene on chromosome 5AL was an
APETALA2 (AP2)-like gene controlling floral homeotic gene expression and free-threshing character, but its position was located at the position of 650.127 Mb [31]. Hence, the spikelet density locus on 5AL in this study is considered to be an independent locus different from Q gene.
region of Chinese Spring Refv1.0 chr4B [30], we analyzed the possible relationship between our identified locus and Rht-B1. The phenotypic correlation indicated that a weak positive correlation coefficient (r = 0.19) was found between the PH and GNPS trait (Figure 1). The average plant height of high GNPS groups is significantly higher than that of low GNPS groups and no significant PH allele were associated with the 2.76 Mb region on chromosome 4BS from 26.49 Mb to 29.25 Mb (Table S5), which means the locus confer grain number improvement is independent of the Rth-B1 gene.

Figure 3. Manhattan plots for spike-related traits including grain number per spike, i.e., GNPS (A) and kernel number per spikelet, i.e., KNS (B) from the GWAS. The threshold to determine the significant marker-trait associations was set at $1.0 \times 10^{-6}$. The arrowhead indicates the peak signal containing the candidate pleiotropic genes.

To trace and utilize the genetic loci near the AX-109286577 SNP, we developed the KASP diagnostic marker for marker-assisted selection breeding. As shown in Figure 4C, we successfully developed a high-resolution KASP marker that can distinguish the allele AX-109286577 SNP into two obvious groups, i.e., the homozygous allele AA type and TT type. Using the AX-109286577 KASP marker, we detected other materials of our pre-breeding lines, including early generation parents and derivative breeding lines. We found that the AX-109286577 allele A with high GNPS and KNS genetic effects could only be detected in our early generation parent lines, including Pubing3228 and Pubing3504, whereas the AX-109286577 allele T existed in our present breeding lines and commercial varieties (Table S6). The results suggested that the locus with high GNPS and KNS alleles on 4BS is an untapped rare locus for grain number improvement programs.
Zhoumai18 (Figure 4D). Hence, the 2.76 Mb region on chromosome 4BS from 26.49 Mb to 29.25 Mb was verified to carry the genes for increasing KNS and GNPS of wheat. Our results suggest that the AX-109286577-A SNP haplotype is a good allele type for large spike traits. In view of the fact that the Rth-B1 gene TraesCS4B043100 is located on the 30.861–30.863 Mb region of Chinese Spring Refv1.0 chr4B [30], we analyzed the possible relationship between our identified locus and Rht-B1. The phenotypic correlation indicated that a weak positive correlation coefficient (r = 0.19) was found between the PH and GNPS trait (Figure 1). The average plant height of high GNPS groups is significantly higher than that of low GNPS groups and no significant PH allele were associated with the 2.76 Mb region on chromosome 4BS from 26.49 Mb to 29.25 Mb (Table S5), which means the locus confer grain number improvement is independent of the Rth-B1 gene.

Figure 4. The SNP allele AX-109286577 in the chromosome 4BS genomic region was highly related to the grain number per spikelet trait. Box plot for the SNP allele AX-109286577 related highly to the grain number per spike, i.e., GNPS (A) and kernel number per spikelet, i.e., KNS (B) traits, respectively. (C), the KASP diagnostic marker SNP AX-109286577 was developed, blue circles represent high GNPS materials and red circles represent low GNPS materials. (D), Spike morphology comparison of commercial varieties and pre-breeding line L711 (Table S5). Scale bar = 2 cm.

3.5. Identification of Candidate Genes and Their Expression

To identify the candidate genes in the 4BS chromosome region for high GNPS and KNS traits, we searched the annotated genes based on Chinese Spring IWGSC RefSeq v1.1 reference genome annotation [32]. From the 26.49 Mb to 28.95 Mb physical regions of chromosome 4B, a total of 56 high-confidence genes were annotated in the Chinese Spring reference genome (Table S7). Combined with the expression data of the published transcriptome, we found that two genes near the Manhattan plot p-value peak (26.49 Mb position, p-value = 6.06 × 10^{-7}) were highly expressed in spike tissues. The two genes are TraesCS4B02G037000 (zinc finger CCCH domain-containing protein, i.e., TraesCS4B03G0080800 in Chinese Spring IWGSC RefSeq v2.1) and TraesCS4B02G037200 (WD-repeat protein, i.e., TraesCS4B03G0081100 in Chinese Spring IWGSC RefSeq v2.1). Their expression in different tissues is shown in Figure S4. According to the WheatExp database, we found that the TraesCS4B02G037200 gene was highly expressed in spike tissue.
and showed the highest expression level on chromosome 4B compared to chromosomes 4A and 4D. The detailed functional annotation suggests that the TraesCS4B02G037200 gene encodes a WD-repeat protein with high homology to the maize cell division cycle protein cdt2 and the barley protein HORVU4Hr1G005810.7, which suggests that the TraesCS4B01G037200 gene is a candidate gene for increasing the GNPS and KNS traits.

4. Discussion

The increase in GNPS, an important yield component, will be of widespread concern in future yield improvements. It has been reported that spikelet fertility has increased during recent decades, especially in the European lines, and spikelet fertility is suggested as an essential driver of grain yield progress in wheat [22]. However, studies on the increase in floret set number are limited. Here, we reported that a panel of pre-breeding derivatives derived from the cross between high GNPS germplasm and diverse elite varieties carry prominent GNPS and KNS phenotypes. The correlation coefficient was as high as 0.81 between GNPS and KNS, which means that grain number per spike in our pre-breeding lines is mainly contributed by grain set number per spikelet and floret fertility. Our study showed the same feature as our previous results, in which GNPS showed significant positive correlations with KNS in the Pubing 3504 bипarent population [33]. Additionally, GNPS in the Pubing 3228 and Pubing 3504 germplasms is independently inherited and freely recombined with the expected fertile tiller number and grain weight, which suggests that our high GNPS pre-breeding lines are suitable as breeding parents for the improvement of yield potential [23]. Guo et al. (2016) found that the post-anthesis process of grain set/abortion was important in determining genotypic variation in KNS, and an increase in KNS was mainly associated with improved grain survival [34]. Based on our observation of the developmental dynamic anatomy, good floret fertility after the anthesis process can also explain the high KNS phenotype in our pre-breeding lines. As a result, this study provides us with many useful pre-breeding lines for increasing the GNPS yield component.

Elucidation of the genetic loci controlling the GNPS and KNS in our pre-breeding lines is necessary. This study reported that a major locus on 4BS controlled both GNPS and KNS traits in our pre-breeding lines. According to the Chinese spring v1.1 genome reference, the approximately 2.76 Mb region on chromosome 4BS from 26.49 Mb to 29.25 Mb contains 56 high-confidence genes. To accurately detect the elite haplotype of the GNPS-related genetic loci, we further developed a diagnostic KASP marker in the target genomic region of chromosome 4BS. The marker AX-109286577 will be useful for genomic selection breeding of the KNS trait in early offspring. Finally, we narrowed down the gene TraesCS4B02G037200, which was highly expressed in the spike tissues as a candidate gene. Moreover, we found that the elite haplotype on chromosome 4BS was only detected in our pre-breeding lines and not in commercial varieties, which means that the loci are still rare for a large number of wheat varieties. In the distal region of the short arm of chromosome 4B, another team also conducted an association mapping study of spike-related traits by using a panel of 768 wheat cultivars and genotyping-by-sequencing technology. They identified a genomic region on 4BS spanning from 28.7 to 44.4 Mb near the Rht-B1 locus and two candidate genes, TraesCS4B02G049100 encoding RNA-binding protein 34 and TraesCS4B02G049300 encoding a GPN-loop GTPase-like protein [17]. It is reported that the semi-dwarfing Rht-B1b gene reduced plant height and increased grain number and yield [35,36]. Our result showed that the locus for GNPS identified in our large-spike pre-breeding lines was independent of the dwarfing Rht-B1b gene. The difference between the two loci on 4BS suggests that the different genes take part in the process of grain number per spike. Overall, our results proved that the rare allele variation in our pre-breeding lines can be used for wheat improvement of yield potential, especially in the yield component of grain number per spike.
**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy12010171/s1. Figure S1: The distribution of SNPs with minor allele frequencies larger than 0.05 and missing rates ≤ 90%. Figure S2: The screen plot of principal components was used to find the optimal number of PCs for the GWAS. The eigenvalues start to form a relatively straight line after the fourth principal component. Figure S3: Manhattan plots for spike-related traits including spikelet density (A), spike length (B) and total spikelet number per spike (C) from the GWAS. The arrowhead indicates the peak signal containing the candidate pleiotropic genes. Figure S4: The expression profiles of TraesCS4B02G037000 and TraesCS4B02G037200 in different tissues and organs of wheat. The expression data of wheat genes were collected from the WheatExp database (https://wheat.pw.usda.gov/WheatExp/, accessed on 30 November 2021). Table S1: Pedigree information of a panel of 255 derivative lines and 4 elite commercial varieties. Table S2: The VCF file of 310,097 SNPs in all materials after removing nucleotide variations with missing rates ≥ 0.1 and minor allele frequency (<0.05). Table S3: All phenotype data were provided including 2-year replicates. Table S4: KASP primer sequences used in this study. Table S5: The list of marker-trait associations (MTAs) for 12 investigated traits across all the environments. Table S6: Other pre-breeding lines, including early generation lines and derivative breeding lines, were chosen for detecting the AX-109286577 SNP genotype on chromosome 4B. Table S7: High-confidence genes were annotated in the Chinese Spring reference genome from the 26.49 Mb to 28.95 Mb physical regions of chromosome 4B.

**Author Contributions:** Conceptualization, L.L. and J.Z.; methodology, J.Z.; investigation, Q.Y., R.L., Y.L., S.Z., H.H., W.L., X.L., X.Y.; resources, L.L.; data curation, L.L., J.Z.; writing—original draft preparation, J.Z.; writing—review and editing, L.L., J.Z.; supervision, L.L.; project administration, L.L., J.Z.; funding acquisition, L.L., J.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Natural Science Foundation of China, grant number NSFC No. 31971879 and the Chinese Agriculture Research System, grant number CARS-03.

**Data Availability Statement:** The data presented in this study are available in Supplementary Materials.

**Acknowledgments:** We would like to thank Feng Chen at Henan Agricultural University for providing technical help with the GWAS analysis.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Shiferaw, B.; Smale, M.; Braun, H.-J.; Duveiller, E.; Reynolds, M.; Muricho, G. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Secur.* 2013, 5, 291–317. [CrossRef]
2. Foulkes, M.J.; Slafé, G.A.; Davies, W.J.; Berry, P.M.; Sylvester-Bradley, R.; Martre, P.; Calderini, D.F.; Griffiths, S.; Reynolds, M.P. Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. *J. Exp. Bot.* 2010, 62, 469–486. [CrossRef]
3. Fischer, R. The importance of grain or kernel number in wheat: A reply to Sinclair and Jamieson. *Field Crop Res.* 2008, 105, 15–21. [CrossRef]
4. Gao, F.; Ma, D.; Yin, G.; Rasheed, A.; Dong, Y.; Xiao, Y.; Xia, X.; Wu, X.; He, Z. Genetic Progress in Grain Yield and Physiological Traits in Chinese Wheat Cultivars of Southern Yellow and Huai Valley since 1950. *Crop Sci.* 2017, 57, 760–773. [CrossRef]
5. Shearman, V.J.; Sylvester-Bradley, R.; Scott, R.K.; Foulkes, M.J. Physiological processes associated with wheat yield progress in the UK. *Crop Sci.* 2005, 45, 175–185. [CrossRef]
6. Xiao, S.H. Yield potential improvement in wheat. In *Chinese Wheat Improvement and Pedigree Analysis*; Zhuang, Q.S., Ed.; Chinese Agriculture Press: Beijing, China, 2003; pp. 497–519. (In Chinese)
7. Deng, Z.; Cui, Y.; Han, Q.; Fang, W.; Li, J.; Tian, J. Discovery of Consistent QTLs of Wheat Spike-Related Traits under Nitrogen Treatment at Different Development Stages. *Front. Plant Sci.* 2017, 8, 2120. [CrossRef] [PubMed]
8. Heidari, B.; Seyed-Tabatabaei, B.E.; Saedidi, G.; Kearsey, M.; Suenaga, K. Mapping QTL for grain yield, yield components, and spike features in a doubled haploid population of bread wheat. *Genome* 2011, 54, 517–527. [CrossRef]
9. Wang, J.; Liao, X.; Li, Y.; Zhou, R.; Yang, X.; Gao, L.; Jia, J. Fine mapping a domestication-related QTL for spike-related traits in a synthetic wheat. *Genome* 2010, 53, 798–804. [CrossRef] [PubMed]
10. Yang, L.; Zhao, D.; Meng, Z.; Xu, K.; Yan, J.; Xia, X.; Cao, S.; Tian, Y.; He, Z.; Zhang, Y. QTL mapping for grain yield-related traits in bread wheat via SNP-based selective genotyping. *Theor. Appl. Genet.* 2019, 133, 857–872. [CrossRef]
11. Deng, M.; Wu, F.; Zhou, W.; Li, J.; Shi, H.; Wang, Z.; Lin, Y.; Yang, X.; Wei, Y.; Zheng, Y.; et al. Mapping of QTL for total spikelet number per spike on chromosome 2D in wheat using a high-density genetic map. *Genet. Mol. Biol.* **2019**, *42*, 603–610. [CrossRef] [PubMed]

12. Ma, Z.; Zhao, D.; Zhang, C.; Zhang, Z.; Xue, S.; Lin, F.; Kong, Z.; Tian, D.; Luo, Q. Molecular genetic analysis of five spike-related traits in wheat using RIL and immortalized F2 populations. *Mol. Genet. Genom.* **2006**, *277*, 31–42. [CrossRef] [PubMed]

13. Ma, J.; Ding, P.; Liu, J.; Li, T.; Zou, Y.; Habib, A.; Mu, Y.; Tang, H.; Jiang, Q.; Liu, Y.; et al. Identification and validation of a major and stably expressed QTL for spikelet number per spike in bread wheat. *Theor. Appl. Genet.* **2019**, *132*, 3155–3167. [CrossRef] [PubMed]

14. Wang, R.; Liu, Y.; Isham, K.; Zhao, W.; Wheeler, J.; Klassen, N.; Hu, Y.; Bonman, J.M.; Chen, J. QTL identification and KASP marker development for productive tiller and fertile spikelet numbers in two high-yielding hard white spring wheat cultivars. *Mol. Breed.* **2018**, *38*, 135. [CrossRef]

15. Zhou, Y.; Conway, B.; Miller, D.; Marshall, D.; Cooper, A.; Murphy, P.; Chao, S.; Brown-Guedira, G.; Costa, J. Quantitative Trait Loci Mapping for Spike Characteristics in Hexaploid Wheat. *Plant Genome* **2017**, *10*, plantgenome2016-10. [CrossRef]

16. Liu, J.; Xu, Z.; Fan, X.; Zhou, Q.; Cao, J.; Wang, F.; Ji, G.; Yang, L.; Feng, B.; Wang, T. A Genome-Wide Association Study of Wheat Spike Related Traits in China. *Front. Plant Sci.* **2018**, *9*, 1584. [CrossRef] [PubMed]

17. Pang, Y.; Liu, C.; Wang, D.; Amand, P.S.; Bernardo, A.; Li, W.; He, F.; Li, L.; Wang, L.; Yuan, X.; et al. High-Resolution Genome-Wide Association Study Identifies Genomic Regions and Candidate Genes for Important Agronomic Traits in Wheat. *Mol. Plant* **2020**, *13*, 1311–1327. [CrossRef]

18. Sakuma, S.; Golan, G.; Guo, Z.; Ogawa, T.; Tagiri, A.; Sugimoto, K.; Bernhardt, N.; Brassac, J.; Mascher, M.; Hersel, G.; et al. Unleashing floret fertility in wheat through the mutation of a homeobox gene. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 5182–5187. [CrossRef] [PubMed]

19. Kuzay, S.; Xu, Y.; Zhang, J.; Katz, A.; Pearce, S.; Su, Z.; Fraser, M.; Anderson, J.A.; Brown-Guedira, G.; DeWitt, N.; et al. Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theor. Appl. Genet.* **2019**, *132*, 2689–2705. [CrossRef]

20. Muqaddasi, Q.H.; Brassac, J.; Koppolu, R.; Plieske, J.; Canal, M.W.; Röder, M.S. TaAPO-A1, an ortholog of rice aberrant panicle initiative 1, is associated with total spikelet number per spike in elite European hexaploid winter wheat (*Triticum aestivum* L.) varieties. *Sci. Rep.* **2019**, *9*, 13853. [CrossRef] [PubMed]

21. Wang, Y.; Yu, H.; Tian, C.; Sajjad, M.; Gao, C.; Tong, Y.; Wang, X.; Jiao, Y. Transcriptome Association Identifies Regulators of Wheat Grain Number Per Spike in Two Wheat Germplasm Lines. *Plant Genome* **2018**, *12*, 171. [CrossRef]

22. Würschum, T.; Leiser, W.L.; Langer, S.M.; Tucker, M.R.; Longin, C.F.H. Phenotypic and genetic analysis of spike and kernel characteristics in wheat reveals long-term genetic trends of grain yield components. *Theor. Appl. Genet.* **2018**, *131*, 2071–2084. [CrossRef]

23. Chen, D.; Zhang, J.-P.; Wang, J.-S.; Yang, X.-M.; Liu, W.-H.; Gao, A.-N.; Li, X.-Q.; Li, L.-H. Inheritance and Availability of High Grain Number Per Spike in Two Wheat Germplasm Lines. *J. Integr. Agric.* **2018**, *17*, 1786–1788. [CrossRef]

24. Wang, J.; Liu, W.; Wang, H.; Li, L.; Wu, J.; Yang, X.; Li, X.; Gao, A. QTL mapping of yield-related traits in the wheat germplasm 3504 with high grain number per spike under varying environments. *Front. Plant Sci.* **2016**, *7*, 1409–1416. [CrossRef]

25. Li, L.H.; Yang, X.M.; Li, X.Q.; Dong, Y.C.; Chen, X.M. Introduction of desirable genes from *Triticum aestivum* L.) unleft and stably expressed QTL for spikelet number per spike in bread wheat (*Triticum aestivum* L.) varieties. *Sci. Rep.* **2019**, *9*, 13853. [CrossRef] [PubMed]

26. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [CrossRef] [PubMed]

27. Zhang, C.; Dong, S.-S.; Xu, J.-Y.; He, W.-M.; Yang, T.-L. PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **2019**, *35*, 1786–1788. [CrossRef]

28. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **2015**, *67*, 1–48. [CrossRef]

29. Lipka, A.E.; Tian, F.; Wang, Q.; Peiffer, J.; Li, M.; Bradbury, P.J.; Gore, M.; Buckler, E.; Zhang, Z. GAPT: Genome association and prediction integrated tool. *Bioinformatics* **2012**, *28*, 2397–2399. [CrossRef]

30. Tang, Y.; Liu, X.; Wang, J.; Li, M.; Wang, Q.; Tian, F.; Su, Z.; Fan, Y.; Liu, D.; Lipka, A.E.; et al. GAPT Version 2: An Enhanced Integrated Tool for Genomic Association and Prediction. *Plant Genome* **2016**, *9*, plntgenome2015-11. [CrossRef]

31. Faris, J.D.; Bellers, J.P.; Brooks, S.A.; Gill, B.S. A Bacterial Artificial Chromosome Contig Spanning the Major Domestication Locus Q in Wheat and Identification of a Candidate Gene. *Genetics* **2003**, *164*, 311–321. [CrossRef] [PubMed]

32. Ma, S.; Wang, M.; Wu, J.; Guo, W.; Chen, Y.; Li, G.; Wang, Y.; Shi, W.; Xia, G.; Fu, D.; et al. WheatOmics: A platform combining multiple omics data to accelerate functional genomics studies in wheat. *Mol. Plant* **2021**, *14*, 1965–1968. [CrossRef]

33. Chen, D.; Wu, X.-Y.; Wu, K.; Zhang, J.-P.; Liu, W.-H.; Yang, X.-M.; Li, X.-Q.; Lu, Y.-Q.; Li, L.-H. Novel and favorable genomic regions for spike related traits in a wheat germplasm Pubing 3504 with high grain number per spike under varying environments. *J. Integr. Agric.* **2017**, *16*, 2386–2401. [CrossRef]

34. Guo, Z.; Slafert, G.A.; Schnurbusch, T. Genotypic variation in spike fertility traits and ovary size as determinants of floret and grain survival rate in wheat. *J. Exp. Bot.* **2016**, *67*, 4221–4230. [CrossRef] [PubMed]
35. Liu, Y.; Zhang, J.; Hu, Y.-G.; Chen, J. Dwarfing genes Rht4 and Rht-B1b affect plant height and key agronomic traits in common wheat under two water regimes. *Field Crops Res.* **2017**, *204*, 242–248. [CrossRef]

36. Rebetzke, G.; Ellis, M.; Bonnett, D.; Mickelson, B.; Condon, A.; Richards, R. Height reduction and agronomic performance for selected gibberellin-responsive dwarfing genes in bread wheat (*Triticum aestivum* L.). *Field Crops Res.* **2012**, *126*, 87–96. [CrossRef]