Uncovering genomic regions controlling plant architectural traits in hexaploid wheat using GWAS

Lingqiang Wang (✉ lqwang@mail.hzau.edu.cn)

Huazhong Agricultural University College of Plant Science and Technology

Ali Muhammad

Huazhong Agricultural University College of Plant Science and Technology

Weicheng Hu

Huazhong Agricultural University College of Plant Science and Technology

Jinsheng Yu

Zhejiang A and F University

Shahid Ullah Khan

Huazhong Agricultural University College of Plant Science and Technology

Guosheng Xie

Huazhong Agricultural University College of Plant Science and Technology

Jibin Wang

Guangxi University Guangxi Agriculture College

Research article

Keywords: Wheat; Plant architectural traits; Single-nucleotide polymorphisms (SNP); Genome-wide association studies (GWAS)

Posted Date: February 11th, 2020

DOI: <https://doi.org/10.21203/rs.2.23079/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Wheat is a staple food crop worldwide. Plant height is a key factor in plant architecture as it plays a crucial role in lodging and thus affects yield and quality. Genome-wide studies are mostly applied in crop plants, due to its advanced genotyping technologies, identification of novel loci, and improved statistical approaches.

Results: In this study, the population was genotyped by using Illumina iSelect 90K single nucleotide polymorphism (SNP) assay and finally 22,905 high-quality SNPs were used to perform a genome-wide association study (GWAS) for plant architectural traits employing four multi-locus GWAS (ML-GWAS) and three single-locus GWAS (SL-GWAS) models. As a result, 174 and 97 significant SNPs controlling plant architectural traits were detected by four ML-GWAS and three SL-GWAS methods, respectively. Among these SNP makers, 43 SNPs were commonly detected, including seven across multiple environments and thirty-six across multiple methods. Interestingly, five most stable SNPs (Kukri_c34553_89, RAC875_c8121_1490, wsnp_Ex_rep_c66315_64480362, Ku_c5191_340, and tpb0049a09_1302) consistently detected across multiple environments and methods, possibly played a role in modulating plant height and flag leaf length. When comparing ML-GWAS methods, pLARmEB was the most powerful and accountable for the detection of 49 significant SNPs that mostly contributed to plant height (36 SNPs). However, in SL-GWAS the FarmCPU model detected most of the significant SNPs. Moreover, a total of 152 candidate genes were found that are likely to be involved in plant growth and development which may provide insightful information related to plant architectural traits.

Conclusion: Altogether, our results reveal 174 and 97 significant SNPs controlling plant architectural traits using four ML-GWAS and three SL-GWAS methods, respectively. The detection of the stable loci across multiple environments and methods, possibly play a role in modulating plant architectural traits in hexaploid wheat, and finally will contribute to the discovery of valuable SNP loci for marker-assisted selection (MAS) in wheat molecular breeding.

Background

Wheat (*Triticum aestivum* L) is the most important food crop cultivated worldwide on about 200 million hectares and provides 20% of the total needs of the world population [1–3]. Wheat is among the three major cereal crops ranked its position next to maize and rice annually produced [4]. For the first time in human history, wheat cultivation has produced a sufficient amount of food, which in turn, supported the development of cities and contributed to the rise of human civilization in Middle Eastern empires of Babylon and Egypt [5]. High grain yield is one of the critical breeding goals in developing cultivars. Wheat has been grown in China for more than 4,000 years and is currently cultivated in 10 agricultural zones based on its different kinds, cultivation season, effects of temperature and photoperiod [6–8]. Crop yield is a complex trait significantly associated with thousand grain weight, grains number per spike and fertile spikes/m². Moreover, kernel shape, plant height, effective tillers per plant, flag leaf length and flag leaf width traits can also affect yield through effects on photosynthetic ability, grain filling duration and dry matter translocation [9–12]. Larger grains are the potential indicator of grain yield, apart from this increased grain size is positively associated with seeding growth and

early maturity [13]. Using semi-dwarf cultivars contained dwarfing genes (Rht1 and Rht2) were the main reason for the success of the Green Revolution [11].

Prior studies have identified chromosomal regions associated with yield-related traits, either by QTL mapping using linkage maps [14–17] or by genome-wide association study [18–21]. However, the standard QTL mapping using bi-parental crosses only detect information on two alleles at a given locus with low resolution, and unable to detect QTLs in complex and genetically diverse traits [22]. In addition, the evaluation of the mapping population is a time consuming and costly process. As a complement to conventional QTL mapping, GWAS analysis provides an opportunity to analyze the diverse genetic background of complex traits with more accuracy [23]. GWAS has three obvious advantages compared to QTL mapping: (1) it shortens the time of constructing populations; (2) it can be used to evaluate a wider range of germplasms for phenotyping many traits with one cycle of genotyping; (3) it offers much higher resolution for fine mapping [20, 24]. Currently research have been accomplished through GWAS in wheat [25], maize [26], rice [27], soybean [28] and cotton [29]. Comparatively hexploid wheat has larger genome (≈ 15.8 GB) than rice (≈ 400 Mb) and maize (≈ 2.32 Gb) [25]. Over the past ten years, the absence of fully sequenced reference genome has limited the gene discovery of wheat and recent advances in functional genomics have provided breeders with a new impetus to achieve their goals [30].

However, substantially more work is required because of the experimental bottleneck emerging from the absence of inconsistency among studies, and the utilization of low-density marker platforms in gene mapping studies. To achieve better understanding with high resolution, larger number of molecular markers are preferred for mapping population. Currently, the best alternate choice for wheat breeder is the utilization of single nucleotide polymorphism (SNP) to completely detect candidate genes controlling complex traits [20]. Comparatively, the high-density wheat Illumina iSelect array comprising about 90,000 SNPs has been extensively adapted than SSR marker due to their efficiency, high density, stable inheritance, low cost, representativeness and capability of automatic detection [25, 31].

To address these issues the present study was designed to use genome-wide association study (GWAS) approach employing high-density wheat iSelect 90K SNP array. Our main objectives were to 1) investigate marker-trait associations (MTAs) for plant architecture-related traits. 2) detect candidate genes responsible for corresponding morphological traits. Indeed, this study will provide useful insights by integrating three single-locus and four newly developed multi-locus GWAS methods, and further establish a regulatory network in wheat breeding programs.

Results

Statistical analysis of phenotypic traits

In the present study, we evaluated wheat germplasm collections for plant architectural traits, including plant height, flag leaf length, flag leaf width and the number of tillers per plant. The phenotypic characteristics for plant height across four environments and flag leaf length, flag leaf width and number of tillers per plant across two environments as shown in Additional file 1 and Fig. 1. All the traits exhibited the normal distribution pattern each year, indicating the quantitative nature of these traits (Fig. 1). Descriptive statistics revealed large phenotypic variations for all the traits as given in Additional file S1. PH ranged from 48.40 to 124.82 cm with

coefficient of variations (CVs) ranged from 11.09–16.11%. The FLL varied from 16.06 to 31.57 cm, FLW varied from 1.22 to 2.75 cm and for tillers, the number of tillers per plant ranged from 8.10 to 14.67. Analysis of variance indicated highly significant differences ($P < 0.001$) for all the studied traits (Additional file 2).

Broad sense heritability was also estimated for PH, FLL, FLW and number of tillers with values ranged from 0.569 (FLW) to 0.794 (PH), suggesting the stability of these traits. The correlation analysis revealed significant correlation between different environments for each of the four traits, indicating the consistency of these traits across various environments (Fig. 2). Furthermore, PH was significantly and positively correlated with FLL in almost all the environments. However, they correlated significantly but negatively with FLW in most of the environments indicating competition of these traits to assimilate at the plant growth stage. Finally, the relatively weak correlation of tillers with other traits suggested the independency of this trait.

Population Structure Analysis

Population structure is important due to the large number of diverse genotypes used in the study may produce false associations between the phenotypic values and unlinked markers. Therefore, a comprehensive analysis of population structure is prerequisite for evaluating successful association mapping. The number of subpopulations were calculated by the rate of change in the log probability of data between successive K-values. ΔK was calculated for increasing the number of K value determined by STRUCTURE analysis according to the procedure of Evanno [32]. At $k = 2$, a break in the slope was observed followed by flattening of the curve (Fig. 3a). Hence, the most likely number of subpopulations was two ($k = 2$) (Fig. 3b). Moreover, this result was confirmed by PCA based on standardized covariance of genetic distances of SNP markers (Fig. 3c).

Gwas For Plant Architectural Traits

Four multi-locus GWAS (ML-GWAS) models: FASTmrMLM, FASTmrEMMA, mrMLM, and pLARmEB, and three single-locus GWAS (SL-GWAS) models: FarmCPU, MLM, and MLMM implemented by Genomic Association and Prediction Integrated Tool (GAPIT) in R were used to detect the association signals. In ML-GWAS the SNPs with the critical logarithm of odds (LOD) score of 3 or greater than 3 were regarded as significant SNPs for associated traits. In SL-GWAS the screening criteria for detecting significant SNPs were $P = 1/m$ (m is the number of total SNPs). To obtain more reliable results, the SNPs that were simultaneously detected in at least two years or by at least two methods were considered as most stable SNPs. A total of 113 and 62 significant SNPs were identified by ML-GWAS and SL-GWAS methods, respectively (Fig. 4a).

Gwas Using Four Multi-locus Models

ML-GWAS models i.e. FASTmrMLM, FASTmrEMMA, mrMLM, and pLARmEB screened 47, 32, 46, and 49, significant SNPs, respectively for PH (117 SNPs) across four environments and for FLL (24), FLW (15) and Tillers (18) across two environments (Fig. 4b, Table 1–3). Among these SNPs, 30, 8, 5, and 4 were detected by FASTmrMLM for PH, FLL, FLW, and Tillers, respectively (Fig. 4b, Table 1). By using FASTmrEMMA the number of significant SNPs identified for the above-mentioned traits were 24, 4, 1, and 3, respectively. Also 27, 6, 9, and 4 significant SNPs were detected by mrMLM approach for the said traits, respectively. Finally, the pLARmEB model

identified 36, 6, 0, and 7 significant association signals with PH, FLL, FLW, and Tillers, respectively (Fig. 4b, Table 1).

Table 1

Common SNPs detected through different ML-GWAS methods associated with plant architectural traits

Trait	Years	Methods (1–4)	SNP	Chr	Position (bp)	LOD > 3	r ² (%)
PH	2015	1,2,4	BobWhite_24364_73	3B	1551850	3.84–5.19	2.41 ~ 5.93
	2015	2,4	BS00089076_51	5A	2414392	3.47–4.94	1.49–3.53
	2015	2,4	D_contig10675_778	2D	958016	3.1–3.58	1.31–2.35
	2015	1,3	RAC875_c34971_137	7A	3550660	5.21–9.59	4.03–9.99
	2015	1,2	wsnp_Ex_c39592_46849607	5A	2274396	3.20–3.94	1.35–3.70
	2015	1,2	wsnp_Ex_rep_c67296_65839761	4D	2143869	3.38–7.15	5.21–6.17
	2016	2,3,4	Excalibur_03094_523	7D	4302086	4.27–7.19	2.22–5.60
	2016	1,2,4	Ku_c5191_340	6B	3397322	4.54–7.23	4.23–8.44
	2016	2,4	Kukri_04598_614	2A	87562	4.26–6.13	2.25–2.38
	2016	1,4	Kukri_c34553_89	2B	887912	3.22–6.43	2.32–3.42
	2016	1,2,3,4	RAC875_c8121_1490	3A	1272572	4.98–9.25	5.67–9.34
	2016	1,4	wsnp_Ex_rep_c66315_64480362	6B	3350805	3.12–6.78	3.10–5.73
	2017	2,3,4	GENE_4469_430	7A	3542226	3.69–5.51	2.01–3.09
	2017	1,4	Ku_c5191_340	6B	3397322	7.92–9.30	4.10–12.99

Methods (1–4): Corresponding four multi-locus GWAS methods i.e. mrMLM, FASTmrMLM, FASTmrEMMA, and pLARM, respectively. r² (%) represents proportion of total phenotypic variation explained by each SNP. PH (Plant height); FLL (Flag leaf length); FLW (Flag leaf width); TILL (Number of tillers per plant).

To validate the findings, we further compared the results across multiple environments and 6 and 1 SNPs were co-identified in at least two of the environments for PH and FLL, respectively (Fig. 4c, Table 2). These environment-stable SNPs were located on different chromosomes. For PH there was one SNP on chromosome 2B, one on 3A, one located on 3B, one on 6A and two SNPs located on chromosome 6B. The LOD scores ranged from 3.05 to 7.92. One stable SNP across two environments located on chromosome 5A associated with FLL with LOD value ranging from 3.88 to 6.55 (Table 2). Comparing the results across different methods, we found 36 common SNPs were co-detected simultaneously by at least two approaches (Fig. 4c, Table 1). Among these, four significant SNPs (RAC875_c8121_1490, Ku_27771_508, Tdurum_contig42962_2138, BS00022127_51) were detected by all four methods (Table 1).

Trait	Years	Methods (1–4)	SNP	Chr	Position (bp)	LOD > 3	r ² (%)
	2017	1,2	Kukri_c34553_89	2B	887912	3.40–4.60	1.95–3.04
	2017	1,2	RAC875_232_1895	5A	2256162	3.88–4.22	1.53–2.29
	2017	1,2,4	RAC875_c33937_211	1B	1022	3.46–4.25	1.26–3.56
	2017	1,2,3	RAC875_c8121_1490	3A	1272572	7.07–7.81	7.44–13.88
	2018	1,2,4	Excalibur_01167_1207	5A	2153585	3.00–4.62	0.54–3.24
	2018	1,2,3,4	Ku_27771_508	3B	1402891	4.21–8.20	1.42–8.38
	2018	3,4	Kukri_2596_146	3A	1287688	3.03–3.14	0.43–1.53
	2018	1,2,4	RAC875_c6911_412	7A	3921772	4.29–6.33	0.82–4.45
	2018	1,2,3,4	Tdurum_contig42962_2138	5A	2426777	4.96–6.13	1.51–7.82

Methods (1–4): Corresponding four multi-locus GWAS methods i.e. mrMLM, FASTmrMLM, FASTmrEMMA, and pLARmEB, respectively. r² (%) represents proportion of total phenotypic variation explained by each SNP. PH (Plant height); FLL (Flag leaf length); FLW (Flag leaf width); TILL (Number of tillers per plant).

To validate the findings, we further compared the results across multiple environments and 6 and 1 SNPs were co-identified in at least two of the environments for PH and FLL, respectively (Fig. 4c, Table 2). These environment-stable SNPs were located on different chromosomes. For PH there was one SNP on chromosome 2B, one on 3A, one located on 3B, one on 6A and two SNPs located on chromosome 6B. The LOD scores ranged from 3.05 to 7.92. One stable SNP across two environments located on chromosome 5A associated with FLL with LOD value ranging from 3.88 to 6.55 (Table 2). Comparing the results across different methods, we found 36 common SNPs were co-detected simultaneously by at least two approaches (Fig. 4c, Table 1). Among these, four significant SNPs (RAC875_c8121_1490, Ku_27771_508, Tdurum_contig42962_2138, BS00022127_51) were detected by all four methods (Table 1).

Trait	Years	Methods (1–4)	SNP	Chr	Position (bp)	LOD > 3	r ² (%)
	2018	3,4	wsnp_Ex_25573_34834321	3B	1536039	3.98–5.66	0.97–2.04
FLL	2017	1,2	BobWhite_c5694_1201	4B	2087004	3.73–5.13	4.29–5.65
	2017	2,4	BS00021881_51	6D	3496747	4.384.59	1.98–2.97
	2017	1,2	tplb0049a09_1302	5A	2453837	4.96–7.17	6.61–13.14
FLW	2017	1,2	IAAV8465	3D	1660873	3.70–4.19	5.48–5.77
	2017	1,2	RAC875_01969_384	7A	3547319	3.05–6.53	5.13–7.53
	2018	1,3	BS00068592_51	5B	2485486	3.79–5.65	4.95–5.37
TILL	2017	1,2,3,4	BS00022127_51	7B	3924027	3.11–5.15	2.60–5.56
	2017	2,4	D_contig07330_330	7D	4292747	3.20–3.75	1.16–2.06
	2017	1,4	Kukri_c4750_452	6A	2998535	3.23–3.46	1.59–6.30
	2017	1,4	TA001732_0977	2B	580007	3.10–3.96	1.15–5.85
	2017	2,3	wsnp_BE499835B-Ta_2_5	5B	2481742	3.81–4.51	3.64–3.99
	2018	1,3	BobWhite_07783_174	2A	359541	4.11–4.27	4.18–5.93
Methods (1–4): Corresponding four multi-locus GWAS methods i.e. mrMLM, FASTmrMLM, FASTmrEMMA, and pLARmEB, respectively. r ² (%) represents proportion of total phenotypic variation explained by each SNP. PH (Plant height); FLL (Flag leaf length); FLW (Flag leaf width); TILL (Number of tillers per plant).							
To validate the findings, we further compared the results across multiple environments and 6 and 1 SNPs were co-identified in at least two of the environments for PH and FLL, respectively (Fig. 4c, Table 2). These environment-stable SNPs were located on different chromosomes. For PH there was one SNP on chromosome 2B, one on 3A, one located on 3B, one on 6A and two SNPs located on chromosome 6B. The LOD scores ranged from 3.05 to 7.92. One stable SNP across two environments located on chromosome 5A associated with FLL with LOD value ranging from 3.88 to 6.55 (Table 2). Comparing the results across different methods, we found 36 common SNPs were co-detected simultaneously by at least two approaches (Fig. 4c, Table 1). Among these, four significant SNPs (RAC875_c8121_1490, Ku_27771_508, Tdurum_contig42962_2138, BS00022127_51) were detected by all four methods (Table 1).							

Table 2

Stable SNPs co-detected in multiple environments associated with plant height and flag leaf length

Trait	Years	SNP	Chr	Position (bp)	LOD > 3	r ² (%)
PH	2016, 2017	Kukri_c34553_89	2B	887912	4.60–6.43	1.95–2.32
	2016, 2017	RAC875_c8121_1490	3A	1272572	5.05–7.07	5.76–13.88
	2015, 2017, 2018	IACX3190	3B	1429597	5.05–6.35	4.29–8.51
	2017, 2018	GENE_3659_104	6A	2991495	3.44–3.86	1.50–2.54
	2016, 2017	wsnp_Ex_rep_c66315_64480362	6B	3350804	3.67–6.78	3.10–4.63
	2016, 2017, 2018	Ku_c5191_340	6B	3397322	3.05–7.92	0.72–4.39
FLL	2017, 2018	tplb0049a09_1302	5A	2453837	3.88–7.17	5.04–13.14
PH (Plant height); FLL (Flag leaf length). r ² (%) represents proportion of total phenotypic variation explained by each SNP.						
We further checked the co-detected common SNPs simultaneously in multiple environments and different methods and detect five most stable SNPs (Kukri_c34553_89, RAC875_c8121_1490, wsnp_Ex_rep_c66315_64480362, Ku_c5191_340, and tplb0049a09_1302). Among these, one SNP was associated with FLL and the rest of four were identified for PH across multiple environments and different years (Fig. 4c, Table 3). Comparatively, the four ML-GWAS models i.e. FASTmrMLM, FASTmrEMMA, mrMLM, and pLARMEB, pLARMEB was the most powerful and useful approach, that detected 49 significant association signals, mostly associated with PH (36 SNPs), however FASTmrEMMA was the least effective approach accountable for 32 significant SNPs (Fig. 3b). The Manhattan and QQ plots of the above four ML-GWAS methods for plant height, flag leaf length, flag leaf width and number of tillers are presented in Additional file 7 (a-e).						

Table 3

Stable SNPs identified simultaneously in different environments and different methods associated with plant height and flag leaf length

Trait	Years	Methods (1–4)	SNP	Chr	Position (bp)	LOD > 3	r^2 (%)
PH	2016, 2017	2016 (1,4), 2017 (1,2)	Kukri_c34553_89	2B	887912	3.22–6.43	1.95–3.42
	2016, 2017	2016 (1,2,3,4), 2017 (1,2,3)	RAC875_c8121_1490	3A	1272572	4.98–9.25	5.67–13.88
	2016, 2017	2016 (1,4)	wsnp_Ex_rep_c66315_64480362	6B	3350804	3.12–6.78	3.10–5.73
	2016, 2017, 2018	2016 (1,2,4), 2017 (1,4)	Ku_c5191_340	6B	3397322	3.05–9.30	0.72–12.99
FLL	2017, 2018	2017 (1,2)	tplb0049a09_1302	5A	2453837	3.88–7.17	5.04–13.14
Methods (1–4): Corresponding four multi-locus GWAS methods i.e. mrMLM, FASTmrMLM, FASTmrEMMA, and pLARM, respectively. r^2 (%) represents proportion of total phenotypic variation explained by each SNP. PH (Plant height); FLL (Flag leaf length).							

Gwas Using Three Single-locus Models

Three single-locus GWAS (SL-GWAS) methods i.e. FarmCPU, MLM, and MLMM were used to further analyze the results of the same plant architectural traits. A total of 62 significant SNPs were detected by the three SL-GWAS methods for the above mentioned traits across multiple environments (Fig. 4d, Additional file 3). In the three SL-GWAS methods, FarmCPU detected 56 significant SNPs, MLM detected 19, and MLMM detected 22 significant SNPs associated with different traits in multiple environments (Fig. 4d, Additional file 3). We further checked the common SNPs detected by all three SL-GWAS methods across multiple environments and methods and found three most stable SNPs i.e. RAC875_c8121_1490, BS00049008_51, and tplb0049a09_1302 were repeated consistently by all methods in most of the environments (Additional file 3). By comparing the results of all three SL-GWAS methods, the FarmCPU is the most effective and reliable approach accountable for 56 SNPs for all the traits, while MLM was the least effective method responsible for 19 SNPs detection.

Traits Having Common Associations

SNPs associated with more than one trait are very useful for marker-assisted selection. A total of five SNPs were detected associated with more than one trait across multiple environments (Additional file 4). Among these, one SNP on chromosome 5A (Excalibur_01167_1207) associated with PH and FLW. The rest of four pleiotropic

SNPs were associated with PH and FLL across multiple environments. The presence of pleiotropic effects of these SNPs controlling plant height and flag leaf length were confirmed by the correlation analysis (Fig. 2). These pleiotropic SNPs i.e. Jagger_c6772_80, RAC875_c8121_1490, BS00089954_51, and Ku_c5191_340 were located on chromosome 1A, 3A, 3B, and 6B. Moreover, these pleiotropic associations suggest that the aforementioned SNPs have multifaceted role in plant architectural traits and highlight the significance of flag leaf length and width to plant height.

Candidate Genes Identification

To further understand the genetic basis of plant architectural traits, we detected several candidate genes that were surrounding the peak SNPs. Interestingly, several major candidate genes that were directly associated with the consensus SNPs had exact same annotations (Fig. 5 and Additional file 6). For instance, several putative candidate genes for PH and FLL, annotated as Laccase which is used for lignin polymerization to help in a variety of functions in plant development [33]. Similarly, the putative genes responsible for PH and FLL annotated as Cysteine proteinase inhibitor, has a function in plant growth and defense [34]. Seven genes annotated as Glutathione S-transferase responsible for resistance against biotic stress [35]. Additionally, six putative candidate genes surrounding significant SNPs associated with number of tillers have annotations as F-box family protein, involved in plant vegetative and reproductive growth [36]. Further examples are given in Fig. 5 and Additional file 6. These results will provide useful information for future work and suggest that the likely gene families are important for plant architectural traits.

Discussion

Plant architecture related traits

In this study, we employed four multi-locus models: FASTmrMLM, FASTmrEMMA, mrMLM, and pLARmEB and three single-locus models i.e. FarmCPU, MLM, and MLMM to identify significant association with plant height, flag leaf length, flag leaf width, and number of tillers across multiple environments in hexaploid wheat. Plant height is a key factor in crop breeding as it plays a crucial role in reshaping plant architecture. In wheat, it greatly affects lodging and thus grain yield and quality [37]. Consequently, the identification of major dwarfing genes in green revolution to reduce plant height and improve yield was a major component in wheat breeding [38]. In this study, we identified significant SNPs across multiple years by different ML-GWAS approaches (Table 1–3, supplementary Fig. 1a-e). By integrated ML-GWAS results, a total of 24 stable SNPs consistently detected by most of the ML-GWAS methods for PH (Table 1 and Fig. 4c). Most of these SNPs located on chromosome 5A and 7A. Earlier studies reported significant association signals with PH located on chromosome 5A and 7A [39–41]. It is revealed that chromosome 5A harboured the highest number of significant SNPs associated with plant architectural traits (Table 1). Several previous studies confirmed that chromosome 5A is the most useful and reproducible in wheat genome [42–45]. S Sukumaran, S Dreisigacker, M Lopes, P Chavez and MP Reynolds [22] evaluated spring wheat population for yield related traits and reported that most of the significant SNPs identified on chromosome 5A and 6A. Similarly, the SNPs on chromosomes 5A and 6A are most likely the MTAs reported previously [43, 46].

Among the four stable SNPs simultaneously detected across multiple environments and methods for PH (Table 3), two SNPs were located on chromosome 6B, revealing the significance of chromosome 6B which are consistent with the findings of [47–49] most of their significant SNPs located on chromosome 6B controlling FLW. Another SNP (Kukri_c34553_89) located on chromosome 2B with LOD ranging from 3.22–6.43, simultaneously detected in two environments by three ML-GWAS methods i.e. mrMLM, FASTmrMLM, and pLARmEB. J Chen, F Zhang, C Zhao, G Lv, C Sun, Y Pan, X Guo and F Chen [50] evaluated six quality related traits in Chinese wheat and detected Kukri_c34553_89 among the environment-stable SNPs and revealed the positive effects of the aforementioned SNP on harvest index across five environments. Furthermore, two consensus SNPs (RAC875_c8121_1490 and Ku_c5191_340), located on chromosome 3A and 6B respectively, significantly associated with plant height across multiple environments and methods. L Gao, G Zhao, D Huang and J Jia [51] constructed a selection map for domestication and improvement in wheat and reported both RAC875_c8121_1490 and Ku_c5191_340 in their study. Another candidate SNP (tplb0049a09_1302) located on chromosome 5A, simultaneously detected by mrMLM and FASTmrMLM approaches and consistently repeated in two environments (Table 3). Similarly, tplb0049a09_1302 is reported earlier by Q-u Ain, A Rasheed, A Anwar, T Mahmood, M Imtiaz, Z He, X Xia and UM Quraishi [39], using 90K array to identify several genomic regions associated with yield related traits in historical wheat genotypes of Pakistan. Another SNP (BobWhite_c5694_1201) located on chromosome 4B is likely same to the QTL identified by J Zou, K Semagn, M Iqbal, H Chen, M Asif, A N'Diaye, A Navabi, E Perez-Lara, C Pozniak and R-C Yang [52], investigating the effect of 90K SNP array and QTL detection in a spring wheat population. For FLL three consensus SNPs were detected on chromosome 4B, 5A, and 6D (Table 1). F Li, W Wen, J Liu, Y Zhang, S Cao, Z He, A Rasheed, H Jin, C Zhang and J Yan [20] also reported significant SNPs associated with FLL on chromosome 5A and 6D. The SNP (BS00021881_51) associated with FLL simultaneously detected via two ML-GWAS approaches i.e. FASTmrMLM and pLARmEB was reported earlier in QTL mapping [53]. Two stable SNPs (BS00022127_51 and wsnp_BE499835B-Ta_2_5) associated with the number of tillers per plant corresponded to the previously reported SNPs in wheat [54, 55]. Among the stably detected SNPs for number of tillers, BS00022127_51 located on chromosome 7B, simultaneously detected by all four ML-GWAS methods (Table 1). F Li, W Wen, J Liu, Y Zhang, S Cao, Z He, A Rasheed, H Jin, C Zhang and J Yan [20] reported significant SNPs associated with number of tillers located on the same chromosomes 7B and 7D.

Significance Of Gwas Using High-density Genotyping

Sequencing larger data with new technology will provide the base to use high-density genotyping approach for quicker and cost-effective operations. Compared to tradition QTL mapping, GWAS has three major advantages: high resolution power for QTL identification, ability to detect more alleles, and less estimation time [56]. GWAS is mostly applicable in crop plants, due to its modern genotyping technologies, identification of novel alleles, and improved statistical methods [56]. Crop genotyping has been a common approach since 1990s, but recently several improvements have been occurred in different types of polymorphisms and genotyping platforms [57]. Previously used methods depend on patterns of DNA digestion restriction and hybridization, randomly amplified PCR fragments, the advent of next generation sequencing (NGS) technologies enable genotyping in depth investigation with higher resolution [58, 59]. The most widely used NGS is single nucleotide polymorphisms (SNPs) which allow better detection power for markers associated with agronomic traits [60]. Previously genotyping efforts were not very effective in screening complex traits because the associated loci for desirable

traits were not completely detected due to their weak individual effects. The discovery of high-throughput genotyping and advanced bioinformatics tools now solve this issue by increasing the accuracy, while reducing the cost of genotyping [61, 62]. The application of Single Nucleotide Polymorphism (SNP) as a molecular marker provides better understanding of variation in an organism or individual part. Using SNPs provide the base for high throughput genotyping. Molecular markers are mostly used segregation in analysis, forensic examination, genetic mapping and diagnosis, and numerous biological applications [63–65]. Up to now, different genetic markers are being used in crop breeding but most of these are limited in their applications due to their unavailability and high cost of operation. SNPs are the most widely used markers having a wide range of applications in genome analysis [66, 67]

Combining strategy of SL and ML-GWASs can improve the power of GWAS

With the advent of molecular markers and genome studies, several association mapping approaches have been developed to reveal the genetic architecture of complex traits in crops [68, 69]. In earlier studies, mostly SL-GWAS methods were adopted to dissect complex traits, but only few SNPs for each trait have been identified due to its procedural limitations. GLM has an obvious shortcoming of high false positive rate due to the absence of kinship among materials as covariate [70]. In MLM, due to the setting of very high threshold, many small-effect loci are missed [71]. To make up for the limitations of these methods, some multi-locus methodologies have recently been implemented, such as FASTmrMLM [72], FASTmr EMMA [73], mrMLM [71], and pLARM EB [74] are more effective approaches, which were used in this study. Using these models can improve the accuracy of SNPs with high detection power and less stringent criteria, which can effectively overcome the above issues. The most obvious advantage of these multi-locus models is that no Bonferroni multiple test correction is needed ([71, 73]. To study quantitative traits with the complex genetic background as the number of molecular markers is comparatively larger than sample sizes, it is recommended to simultaneously use multiple GWAS methods.

In the past ten years, several GWAS approaches have been used to identify significant SNPs, especially evaluating agronomic traits in common wheat (*T. aestivum* L.) traits using a single locus and two multi-locus approaches [75]. Conclusively, V Jaiswal, V Gahlaut, PK Meher, RR Mir, JP Jaiswal, AR Rao, HS Balyan and PK Gupta [75] verified that ML-GWAS has more detection power than SL-GWAS by revealing ten Marker Trait Associations (MTAs) through SL-GWAS while, 22 MTAs through multi locus mixed model (MLMM) and 58 MTAs through multi-trait mixed model (MTMM). A more recent study of Ward and his co-researchers utilized a conventional mixed linear model and recently developed FarmCPU approach to dissect the genetic architecture of yield related traits in winter wheat. Comparatively, FarmCPU detected 74 significant associations while, the single locus model only screened nine significant associations for different yield-related traits. Furthermore, FarmCPU approach is more complicated and less obvious to user than mix linear model, and hence more care should be taken during using FarmCPU algorithm [18]. Y Zhang, P Liu, X Zhang, Q Zheng, M Chen, F Ge, Z Li, W Sun, Z Guan and T Liang [76] evaluated maize lines through a series of multi locus GWAS approaches to detect some novel loci responsible for lodging resistance. By comparing four multi-locus methods i.e. FASTmrEMMA, mrMLM, pLARM EB, and ISIS EM-BLASSO methods, it was confirmed that ISIS EM-BLASSO was the most effective approach for QTL identification [76]. The combination of two SL-GWAS and ML-GWAS methods contributes efficiently to detection of significant loci associated with pre-harvest sprouting tolerance in wheat

[77]. SU Khan, J Yangmiao, S Liu, K Zhang, MHU Khan, Y Zhai, A Olalekan, C Fan and Y Zhou [68] combined SL and ML-GWAS approaches and revealed that ML-GWAS methods are more effective with high robustness and power of QTN detection than SL-GWAS in the genetic dissection of yield related traits of rapeseed genotypes. Li and his co-workers integrated the results of three single locus GWAS and three multi locus methods for fiber quality traits in upland cotton [78]. A total of 342 significant QTNs were detected of which 72 were consistently detected by at least two approaches or in at least two environments. According to C Li, Y Fu, R Sun, Y Wang and Q Wang [78], the power of QTN detection in association analysis can be improved by combining single locus and multi-locus GWASs.

In this study, we detected a total of 113 and 62 significant SNPs by ML-GWAS and SL-GWAS approaches, respectively. Furthermore, 19 SNPs co-detected by using ML-GWAS and SL-GWAS methods together (Additional file 5). A comparison of the four ML-GWAS methods revealed that pLARmEB was more powerful and robust [68], than the other three models in the detection significant SNPs for plant architectural traits. Through integrating the results of four ML-GWAS and three SL-GWAS methods led to the verification of the significance of ML-GWAS models. However, some recent findings revealed the reliability of association studies can be improved by combining single-locus and multi-locus GWAS approaches [78–81].

Conclusion

We applied four multi-locus and three single-locus GWAS models to investigate the complex genetic background of plant architectural traits (PH, FLL, FLW, and TILL) in hexaploid wheat. A total of 271 significant SNPs associated with plant architectural traits across multiple environments and in different methods were detected by employing four multi-locus and three single-locus methods. Comparatively, 174 and 97 significant association signals were detected by ML-GWAS and SL-GWAS models, respectively which signifies the importance of ML-GWAS over SL-GWAS approaches. Taken together, the results of ML-GWAS revealed five most stable SNPs i.e. Kukri_c34553_89, RAC875_c8121_1490, wsnp_Ex_rep_c66315_64480362, Ku_c5191_340, and tplb0049a09_1302 which were consistently detected across multiple environments and methods. In total, 152 putative candidate genes for three traits were identified, which will provide useful information for future work and highlight the role of the likely gene families for plant architectural traits.

Materials And Methods

Plant materials and phenotyping

A total of 319 wheat germplasm accessions from the collection at the Hubei Agricultural Science Institute in Hubei Province, China, which represent a wheat gene pool adapted to central China and the Yangzi River region, were performed on the experimental farm of Huazhong Agricultural University, Wuhan, China for consecutive four years (2015–2018). Twenty individuals from each variety (line) were grown in two rows with a distance of 15 cm between plants in each row and 20 cm between rows. Field management essentially followed normal local wheat cropping practices. The lines were harvested individually at maturity to prevent seed contamination among lines. Four phenotypic traits were evaluated, including plant height (PH) across four environments and the rest of three traits i.e. flag leaf length (FLL), flag leaf width (FLW), and the number of tiller per plant (TILL) across two environments. Plant height was measured after physiological maturity by measuring the distance

between the stem base and the top of the spike excluding awns. Flag leaf length was measured as the distance from the base to the tip of the leaf. Flag leaf width as the width of the widest section of the leaf. Number of tiller was recorded by counting the total number of fertile tillers per plant.

Genotyping

A total of 319 wheat accessions were genotyped using the Illumina iSelect 90K SNP array [31]. Genotyping was performed by Beijing Compass Technology & investment Co. Ltd (<http://www.bjcompass.com/>), using the 90K wheat genotyping assay [31], following the manufacturer's protocol. A quality preprocessing of genotyping data was done for sample call rate, SNP call rate, minor allele frequency (MAF) and Hardy–Weinberg equilibrium (HWE). This preprocessing was implemented in PLINK software (<http://pngu.mgh.harvard.edu/purcell/plink/>) [82].

Statistical Analysis

Descriptive analysis, ANOVA, correlation analysis and heritability estimates were conducted in the R statistical package [83]. The broad sense heritability for the traits was estimated by the formula $H^2 = VG/(VG + VE)$ where VG and VE represent estimates of genetic and environmental variance, respectively [84].

Population Structure And Kinship Analysis

Population structure using a Bayesian cluster analysis was estimated by STRUCTURE HARVESTER software [85]. A putative number of subpopulations ranging from $k = 1$ to 7 was assessed using 100,000 burn-in iterations followed by 500,000 recorded Markov-Chain iterations. To estimate the sampling variance (robustness) of inferred population structure, 10 independent runs were carried out for each k . K was estimated using an ad-hoc statistic ΔK based on the rate of change in log probability of data between successive values [32]. Principle component analysis (PCA) was calculated by R software for evaluating the population structure and compared to the result of STRUCTURE [25]. Linkage disequilibrium (LD) among markers was calculated using observed vs. expected allele frequencies of the markers in TASSEL v.5.0 [39].

Genome-wide Association Studies

In this study, we used mrMLM software for four ML-GWAS (FASTmrMLM, FASTmrEMMA, mrMLM, and pLARmEB) and three SL-GWAS (FarmCPU, MLM, and MLM) implemented by Genomic Association and Prediction Integrated Tool (GAPIT) in R [86]. Previously, SL-GWAS methods were mostly applied such as GLM and MLM. However, single-locus approaches have some limitations such as GLM leads to high false-positive rates (FPRs), while MLM utilizes Bonferroni corrections for loci detection to reduce the FPRs [87]. Though, this procedure is so stringent that results in missing significant SNPs [71]. Therefore, multi-locus GWAS approaches are the best alternatives. The stringent Bonferroni multiple test correction in the SL-GWAS analysis is substituted by a flexible selection criterion in multi-locus GWAS analysis, that reduces the possibility of missing out significant loci [71, 73]. The four ML-GWAS methods were performed with default parameters, and the screening criteria for significance were set with LOD scores 3 or > 3 [71, 73, 74, 88, 89]. However, for SL-GWAS models, the

threshold for P-value was calculated based on the number of the markers ($P = 1/n$, n = total SNP used) according to the method of [90]. Significant markers were visualized with a Manhattan plot using Haploview 4.2 software [91]. Important p-value distributions (expected vs. observed p-values on a $-\log^{10}$ scale) were shown with a quantile-quantile plot.

Candidate Gene Analysis

Candidate gene sites were aligned and downloaded from the ViroBLAST database (<https://urgi.versailles.inra.fr/blast/docs/aboutviroblast.html>). The R Package Pathway Association Study Tool (PAST) version 1.0.1 was used to identify genes around the peak SNPs with a window size of 200 kb. To find candidate genes or putative related proteins of SNP flanking-regions, BLASTx search was conducted for significant marker-trait associations (MTAs) against recently released genome sequence IWGSC RefSeq v1.0 [92].

Declarations

Abbreviations

GWAS: Genome-Wide Association Study; ML: Multi-locus; SL: Single-locus; SNP: Single Nucleotide Polymorphism; MAS: Marker-Assisted Selection; QTL: Quantitative Trait Loci; MTA: Marker-Trait Associations; PH: Plant Height; FLL: Flag Leaf Length; FLW: Flag Leaf Width; TILL: Tillers; CV: Coefficient of Variation; GAPIT: Genomic Association and Prediction Integrated Tool; NGS: Next Generation Sequencing; MAF: Minor Allele Frequency; MTAs: Marker-trait associations; HWE: Hardy–Weinberg Equilibrium; ANOVA: Analysis of Variance; LD: Linkage disequilibrium.

Acknowledgements

Not applicable.

Authors' contributions

JW and LW conceived and designed the project, AM and WH performed the experiment, JY and SUK analyzed and interpreted the data, AM and GX drafted and revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported in part by grants from the National Natural Science Foundation of China (31771775 and 31171524), the Programme of Introducing Talents of Discipline to Guangxi University (A3310051010 and EE101701).

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

References

1. FAOSTAT F: **Available online: [http://www.fao.org/faostat/en/# data](http://www.fao.org/faostat/en/#data). QC (accessed on January 2018) 2017.**
2. Schulte D, Close TJ, Graner A, Langridge P, Matsumoto T, Muehlbauer G, Sato K, Schulman AH, Waugh R, Wise RP: **The international barley sequencing consortium—at the threshold of efficient access to the barley genome.** *Plant Physiol.* 2009, 149(1):142-7. DOI: 10.1104/pp.108.128967.
3. Tester M, Langridge P: **Breeding technologies to increase crop production in a changing world.** *Science.* 2010, 12;327(5967):818-22. DOI: 10.1126/science.1183700.
4. Graybosch RA, Peterson CJ: **Specific adaptation and genetic progress for grain yield in Great Plains hard winter wheats from 1987 to 2010.** *Crop Sci.* 2012, 52(2):631-643; DOI: 10.2135/cropsci2011.08.0412
5. Curtis T, Halford N: **Food security: the challenge of increasing wheat yield and the importance of not compromising food safety.** *Ann Appl Biol.* 2014, 164(3):354-372; DOI: 10.1111/aab.12108
6. He Z: **A history of wheat breeding in China:** Cimmyt; 2001.
7. Zhuang Q: **Chinese wheat improvement and pedigree analysis.** Chinese) China Agriculture Press, Beijing, China 2003.
8. Hao C, Dong Y, Wang L, You G, Zhang H, Ge H, Jia J, Zhang X: **Genetic diversity and construction of core collection in Chinese wheat genetic resources.** *Chinese Science Bulletin* 2008, 53(10):1518-1526; DOI: 10.1007/s11434-008-0212-x
9. Xiao Y, Qian Z, Wu K, Liu J, Xia X, Ji W, He Z: **Genetic gains in grain yield and physiological traits of winter wheat in Shandong Province, China, from 1969 to 2006.** *Crop Sci.* 2012, 52(1):44-56; DOI: 10.2135/cropsci2011.05.0246
10. 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(2):760-773; DOI: 10.2135/cropsci2016.05.0362
11. Ellis M, Rebetzke G, Azanza F, Richards R, Spielmeyer W: **Molecular mapping of gibberellin-responsive dwarfing genes in bread wheat.** *Theoretical and Applied Genetics* 2005, 111(3):423-430.
12. Zhou Y, He Z, Sui X, Xia X, Zhang X, Zhang G: **Genetic improvement of grain yield and associated traits in the northern China winter wheat region from 1960 to 2000.** *Crop Sci.* 2007, 47(1):245-253; DOI: doi:10.2135/cropsci2006.03.0175

13. Peng J, Ronin Y, Fahima T, Röder MS, Li Y, Nevo E, Korol A: **Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat.** PNAS. 2003, **100**(5):2489-2494; DOI: 10.1073/pnas.252763199
14. Fan X, Cui F, Ji J, Zhang W, Zhao X, Liu J, Meng D, Tong Y, Wang T, Li J: **Dissection of pleiotropic QTL regions controlling wheat spike characteristics under different nitrogen treatments using traditional and conditional QTL mapping.** Front Plant Sci. 2019, **10**; DOI: 10.3389/fpls.2019.00187
15. Fu X, Xu J, Zhou M, Chen M, Shen L, Li T, Zhu Y, Wang J, Hu J, Zhu L: **Enhanced Expression of QTL *qLL9/DEP1* Facilitates the Improvement of Leaf Morphology and Grain Yield in Rice.** Int J Mol Sci. 2019, **20**(4):866; DOI: 10.3390/ijms20040866
16. Hu Z, Cao L, Sun X, Zhu Y, Zhang T, Jiang L, Liu Y, Dong S, Sun D, Yang J: **Fine mapping of a major quantitative trait locus, *qgnp7 (t)*, controlling grain number per panicle in African rice (*Oryza glaberrima* S.).** Breed Sci. 2018, **68**(5):606-613; DOI: 10.1270/jsbbs.18084
17. Segami S, Yamamoto T, Oki K, Noda T, Kanamori H, Sasaki H, Mori S, Ashikari M, Kitano H, Katayose Y: **Detection of novel QTLs regulating grain size in extra-large grain rice (*Oryza sativa* L.) lines.** Rice. 2016, **9**(1):34; DOI: 10.1186/s12284-016-0109-2
18. Ward BP, Brown-Guedira G, Kolb FL, Van Sanford DA, Tyagi P, Sneller CH, Griffey CA: **Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia.** PloS one. 2019, **14**, e0208217; DOI: 10.1371/journal.pone.0208217.
19. Garcia M, Eckermann P, Haefele S, Satija S, Sznajder B, Timmins A, Baumann U, Wolters P, Mather DE, Fleury D: **Genome-wide association mapping of grain yield in a diverse collection of spring wheat (*Triticum aestivum* L.) evaluated in southern Australia.** PloS one. 2019, **14**(2):e0211730; DOI: 10.1371/journal.pone.0211730.
20. Li F, Wen W, Liu J, Zhang Y, Cao S, He Z, Rasheed A, Jin H, Zhang C, Yan J: **Genetic architecture of grain yield in bread wheat based on genome-wide association studies.** BMC Plant Biol. 2019, **19**, 168; DOI:10.1186/s12870-019-1781-3.
21. Li L, Mao X, Wang J, Chang X, Reynolds M, Jing R: **Genetic dissection of drought and heat-responsive agronomic traits in wheat.** Plant Cell Environ. 2019, **42**(9):2540-2553 DOI: 10.1111/pce.13577.
22. Sukumaran S, Dreisigacker S, Lopes M, Chavez P, Reynolds MP: **Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments.** Theor Appl Genet. 2015, **128**, 353-363; DOI: 10.1007/s00122-014-2435-3.
23. Nordborg M, Weigel D: **Next-generation genetics in plants.** Nature. 2008, **456**(7223):720; DOI: 10.1038/nature07629.
24. Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT: **Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines.** Nature. 2010, **465**(7298):627; DOI: 10.1038/nature08800.
25. Sun C, Zhang F, Yan X, Zhang X, Dong Z, Cui D, Chen F: **Genome-wide association study for 13 agronomic traits reveals distribution of superior alleles in bread wheat from the Yellow and Huai Valley of China.** Plant Biotechnol. J. 2017, **15**, 953-969; DOI: 10.1111/pbi.12690.
26. Samayoa L, Cao A, Santiago R, Malvar R, Butrón A: **Genome-wide association analysis for fumonisin content in maize kernels.** BMC Plant Biol. 2019, **19**(1):166; DOI: 10.1186/s12870-019-1759-1.

27. Chen J, Zhou H, Xie W, Xia D, Gao G, Zhang Q, Wang G, Lian X, Xiao J, He Y: **Genome-wide association analyses reveal the genetic basis of combining ability in rice.** *Plant Biotechnol J.* 2019, 17(11):2211-2222; DOI: 10.1111/pbi.13134
28. Do TD, Vuong TD, Dunn D, Clubb M, Valliyodan B, Patil G, Chen P, Xu D, Nguyen HT, Shannon JG: **Identification of new loci for salt tolerance in soybean by high-resolution genome-wide association mapping.** *BMC genomics.* 2019, 20(1):318; DOI: 10.1186/s12864-019-5662-9.
29. Dilnur T, Peng Z, Pan Z, Palanga KK, Jia Y, Gong W, Du X: **Association analysis of salt tolerance in Asiatic cotton (*Gossypium arboreum*) with SNP markers.** *Int J Mol Sci.* 2019, 20(9):2168; DOI: 10.3390/ijms20092168.
30. Pingault L, Choulet F, Alberti A, Glover N, Wincker P, Feuillet C, Paux E: **Deep transcriptome sequencing provides new insights into the structural and functional organization of the wheat genome.** *Genome Biol.* 2015, 16(1):29; DOI: 10.1186/s13059-015-0601-9.
31. Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L: **Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array.** *Plant Biotechnol. J.* 2014, 12, 787-796; DOI: 10.1111/pbi.12183.
32. Evanno G, Regnaut S, Goudet J: **Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study.** *Mol Ecol.* 2005, 14, 2611-2620; DOI: 10.1111/j.1365-294X.2005.02553.x.
33. Berthet S, Thevenin J, Baratiny D, Demont-Caulet N, Debeaujon I, Bidzinski P, Leple J-C, Huis R, Hawkins S, Gomez L-D: **Role of plant laccases in lignin polymerization.** *Adv Bot Res.* 2012, 61, 145-172; DOI: 10.1016/B978-0-12-416023-1.00005-7.
34. Li R, Wang W, Wang W, Li F, Wang Q, Xu Y, Wang S: **Overexpression of a cysteine proteinase inhibitor gene from *Jatropha curcas* confers enhanced tolerance to salinity stress.** *Electron J Biotechnol.* 2015, 18, 368-375; DOI: 10.1016/j.ejbt.2015.08.002.
35. Gullner G, Komives T, Király L, Schröder P: **Glutathione S-transferase enzymes in plant-pathogen interactions.** *Front Plant Sci.* 2018, 9, 1836; Doi: 10.3389/fpls.2018.01836.
36. Peng J, Yu D, Wang L, Xie M, Yuan C, Wang Y, Tang D, Zhao X, Liu X: **Arabidopsis F-box gene *FOA1* involved in ABA signaling.** *Sci China: Life Sci.* 2012, 55, 497-506; DOI: 10.1007/s11427-012-4332-9.
37. Griffiths S, Simmonds J, Leverington M, Wang Y, Fish L, Sayers L, Alibert L, Orford S, Wingen L, Snape J: **Meta-QTL analysis of the genetic control of crop height in elite European winter wheat germplasm.** *Mol Breed.* 2012, 29, 159-171; DOI: 10.1007/s11032-010-9534-x.
38. Hedden P: **The genes of the Green Revolution.** *Trends Genet.* 2003, 19, 5-9; DOI: 10.1016/S0168-9525(02)00009-4.
39. Ain Q-u, Rasheed A, Anwar A, Mahmood T, Imtiaz M, He Z, Xia X, Quraishi UM: **Genome-wide association for grain yield under rainfed conditions in historical wheat cultivars from Pakistan.** *Front Plant Sci.* 2015, 6, 743; DOI: 10.3389/fpls.2015.00743.
40. Sheoran S, Jaiswal S, Kumar D, Raghav N, Sharma R, Pawar S, Paul S, Iquebal MA, RAI A, Kumar D: **Uncovering genomic regions associated with 36 agro-morphological traits in Indian spring wheat using GWAS.** *Front Plant Sci.* 2019, 10, 527; DOI: 10.3389/fpls.2019.00527.
41. Gao F, Wen W, Liu J, Rasheed A, Yin G, Xia X, Wu X, He Z: **Genome-wide linkage mapping of QTL for yield components, plant height and yield-related physiological traits in the Chinese wheat cross *Zhou***

8425B/Chinese Spring. Front Plant Sci. **2015**, 6, 1099; DOI: 10.3389/fpls.2015.01099.

42. Cuthbert JL, Somers DJ, Brûlé-Babel AL, Brown PD, Crow GH: **Molecular mapping of quantitative trait loci for yield and yield components in spring wheat (*Triticum aestivum* L.)**. Theor Appl Genet. **2008**, 117, 595-608; DOI: 10.1007/s00122-008-0804-5.
43. Huang X, Kempf H, Ganai M, Röder M: **Advanced backcross QTL analysis in progenies derived from a cross between a German elite winter wheat variety and a synthetic wheat (*Triticum aestivum* L.)**. Theor Appl Genet. **2004**, 109, 933-943; DOI: 10.1007/s00122-004-1708-7.
44. Marza F, Bai G-H, Carver B, Zhou W-C: **Quantitative trait loci for yield and related traits in the wheat population Ning7840× Clark**. Theor Appl Genet. **2006**, 112, 688-698; DOI: 10.1007/s00122-005-0172-3.
45. Quarrie S, Steed A, Calestani C, Semikhodskii A, Lebreton C, Chinoy C, Steele N, Pljevljakusić D, Waterman E, Weyen J: **A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring× SQ1 and its use to compare QTLs for grain yield across a range of environments**. Theor Appl Genet. **2005**, 110, 865-880; DOI: 10.1007/s00122-004-1902-7.
46. Börner A, Schumann E, Fürste A, Cöster H, Leithold B, Röder M, Weber W: **Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.)**. Theor Appl Genet. **2002**, 105, 921-936; DOI: 10.1007/s00122-002-0994-1.
47. Wu Q, Chen Y, Fu L, Zhou S, Chen J, Zhao X, Zhang D, Ouyang S, Wang Z, Li D: **QTL mapping of flag leaf traits in common wheat using an integrated high-density SSR and SNP genetic linkage map**. Euphytica. **2016**, 208, 337-351; DOI: 10.1007/s10681-015-1603-0.
48. Yang D, Liu Y, Cheng H, Chang L, Chen J, Chai S, Li M: **Genetic dissection of flag leaf morphology in wheat (*Triticum aestivum* L.) under diverse water regimes**. BMC Genet. **2016**, 17, 94; DOI: 10.1007/s10681-015-1603-0.
49. Fan X, Cui F, Zhao C, Zhang W, Yang L, Zhao X, Han J, Su Q, Ji J, Zhao Z: **QTLs for flag leaf size and their influence on yield-related traits in wheat (*Triticum aestivum* L.)**. Mol Breed. **2015**, 35, 24; DOI: 10.1007/s11032-015-0205-9.
50. Chen J, Zhang F, Zhao C, Lv G, Sun C, Pan Y, Guo X, Chen F: **Genome-wide association study of six quality traits reveals the association of the *TaRPP13L1* gene with flour colour in Chinese bread wheat**. Plant Biotechnol J. **2019**, 17, 2106-2122; DOI: 10.1111/pbi.13126.
51. Gao L, Zhao G, Huang D, Jia J: **Candidate loci involved in domestication and improvement detected by a published 90K wheat SNP array**. Sci Rep. **2017**, 7, 44530; DOI: 10.1038/srep44530 (2017).
52. Zou J, Semagn K, Iqbal M, Chen H, Asif M, N'Diaye A, Navabi A, Perez-Lara E, Pozniak C, Yang R-C: **QTLs associated with agronomic traits in the *Attila*× *CDC* Go spring wheat population evaluated under conventional management**. PloS one. **2017**, 12, e0171528; DOI: 10.1371/journal.pone.0171528.
53. He X, Lillemo M, Shi J, Wu J, Bjørnstad Å, Belova T, Dreisigacker S, Duveiller E, Singh P: **QTL characterization of Fusarium head blight resistance in CIMMYT bread wheat line *Soru# 1***. PLoS One. **2016**, 11, e0158052; DOI: 10.1371/journal.pone.0158052.
54. Downie RC, Bouvet L, Furuki E, Gosman N, Gardner KA, Mackay IJ, Campos Mantello C, Mellers G, Phan HT, Rose GA: **Assessing European wheat sensitivities to *Parastagonospora nodorum* necrotrophic effectors and fine-mapping the Snn3-B1 locus conferring sensitivity to the effector *SnTox3***. Front Plant Sci. **2018**, 9, 881; DOI: 10.3389/fpls.2018.00881.

55. Naraghi SM, Simsek S, Kumar A, Al Rabbi SH, Alamri MS, Elias EM, Mergoum M: **Deciphering the Genetics of Major End-Use Quality Traits in Wheat**. *G3: Genes Genomes Genet.* **2019**, 9, 1405-1427; DOI: 10.1534/g3.119.400050.
56. Zhu C, Gore M, Buckler ES, Yu J: **Status and prospects of association mapping in plants**. *The plant genome.* **2008**, 1, 5-20; DOI:10.3835/plantgenome2008.02.0089.
57. Van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C: **Ten years of next-generation sequencing technology**. *Trends Genet.* **2014**, 30, 418-426; DOI: 10.1016/j.tig.2014.07.001.
58. Gupta P, Rustgi S, Mir R: **Array-based high-throughput DNA markers for crop improvement**. *Heredity.* **2008**, 101, 5; DOI: 10.1038/hdy.2008.35.
59. Voss-Fels K, Snowdon RJ: **Understanding and utilizing crop genome diversity via high-resolution genotyping**. *Plant Biotechnol J.* **2016**, 14, 1086-1094; DOI: 10.1111/pbi.12456.
60. He J, Zhao X, Laroche A, Lu Z-X, Liu H, Li Z: **Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding**. *Front Plant Sci.* **2014**, 5, 484; DOI: 10.3389/fpls.2014.00484.
61. Ruperao P, Edwards D: **Bioinformatics: identification of markers from next-generation sequence data**. In: *Plant Genotyping*. Springer; 2015: 29-47.
62. Varshney RK, Singh VK, Hickey JM, Xun X, Marshall DF, Wang J, Edwards D, Ribaut J-M: **Analytical and decision support tools for genomics-assisted breeding**. *Trends Plant Sci.* **2016**, 21, 354-363; DOI: 10.1016/j.tplants.2015.10.018.
63. Lam H-M, Xu X, Liu X, Chen W, Yang G, Wong F-L, Li M-W, He W, Qin N, Wang B: **Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection**. *Nat Genet.* **2010**, 42, 1053; DOI:10.1038/ng.715.
64. Singh H, Deshmukh RK, Singh A, Singh AK, Gaikwad K, Sharma TR, Mohapatra T, Singh NK: **Highly variable SSR markers suitable for rice genotyping using agarose gels**. *Mol Breed.* **2010**, 25, 359-364; DOI: 10.1007/s11032-009-9328-1.
65. Sonah H, Deshmukh RK, Sharma A, Singh VP, Gupta DK, Gacche RN, Rana JC, Singh NK, Sharma TR: **Genome-wide distribution and organization of microsatellites in plants: an insight into marker development in Brachypodium**. *Plos one* **2011**, 6, e21298; DOI: 10.1371/journal.pone.0021298.
66. Agarwal M, Shrivastava N, Padh H: **Advances in molecular marker techniques and their applications in plant sciences**. *Plant Cell Rep.* **2008**, 27, 617-631; DOI: 10.1007/s00299-008-0507-z.
67. Sonah H, Deshmukh RK, Singh VP, Gupta DK, Singh NK, Sharma TR: **Genomic resources in horticultural crops: status, utility and challenges**. *Biotechnol Adv.* **2011**, 29, 199-209; DOI: 10.1016/j.biotechadv.2010.11.002.
68. Khan SU, Yangmiao J, Liu S, Zhang K, Khan MHU, Zhai Y, Olalekan A, Fan C, Zhou Y: **Genome-wide association studies in the genetic dissection of ovule number, seed number, and seed weight in Brassica napus L.** *Ind Crops Prod.* **2019**, 142, 111877; DOI: 10.1016/j.indcrop.2019.111877.
69. Feng J, Wen Y, Zhang J, Zhang Y: **Advances on methodologies for genome-wide association studies in plants**. *Acta Agron Sin.* **2016**, 42, 945-956; DOI: 10.3724/SP.J.1006.2016.00945.
70. Pace J, Yu X, Lübberstedt T: **Genomic prediction of seedling root length in maize (*Zea mays* L.)**. *The Plant J.* **2015**, 83, 903-912; DOI: 10.1111/tpj.12937.

71. Wang S-B, Feng J-Y, Ren W-L, Huang B, Zhou L, Wen Y-J, Zhang J, Dunwell JM, Xu S, Zhang Y-M: **Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology.** Sci Rep. **2016**, 6, 19444; DOI: 10.1038/srep19444.
72. Zhang Y-M, Tamba CL: **A fast mrMLM algorithm for multi-locus genome-wide association studies.** bioRxiv. **2018**, 341784; DOI: 10.1101/341784.
73. Wen Y-J, Zhang H, Ni Y-L, Huang B, Zhang J, Feng J-Y, Wang S-B, Dunwell JM, Zhang Y-M, Wu R: **Methodological implementation of mixed linear models in multi-locus genome-wide association studies.** Briefings Bioinf. **2017**, 19, 700-712; DOI: 10.1093/bib/bbw145.
74. Zhang J, Feng J, Ni Y, Wen Y, Niu Y, Tamba C, Yue C, Song Q, Zhang Y: **pLARmEB: integration of least angle regression with empirical Bayes for multilocus genome-wide association studies.** Heredity. **2017**, 118, 517; DOI:10.1038/hdy.2017.8.
75. Jaiswal V, Gahlaut V, Meher PK, Mir RR, Jaiswal JP, Rao AR, Balyan HS, Gupta PK: **Genome wide single locus single trait, multi-locus and multi-trait association mapping for some important agronomic traits in common wheat (*T. aestivum* L.).** PloS one. **2016**, 11, e0159343; DOI: 10.1371/journal.pone.0159343.
76. Zhang Y, Liu P, Zhang X, Zheng Q, Chen M, Ge F, Li Z, Sun W, Guan Z, Liang T: **Multi-locus genome-wide association study reveals the genetic architecture of stalk lodging resistance-related traits in maize.** Front Plant Sci. **2018**, 9, 611; DOI: 10.3389/fpls.2018.00611.
77. Zhu Y, Wang S, Wei W, Xie H, Liu K, Zhang C, Wu Z, Jiang H, Cao J, Zhao L: **Genome-wide association study of pre-harvest sprouting tolerance using a 90K SNP array in common wheat (*Triticum aestivum* L.).** Theor Appl Genet. **2019**, 132, 2947-2963; DOI: 10.1007/s00122-019-03398-x.
78. Li C, Fu Y, Sun R, Wang Y, Wang Q: **Single-locus and multi-locus genome-wide association studies in the genetic dissection of fiber quality traits in upland cotton (*Gossypium hirsutum* L.).** Front Plant Sci. **2018**, 9, 1083; DOI: 10.3389/fpls.2018.01083.
79. Chang F, Guo C, Sun F, Zhang J, Wang Z, Kong J, He Q, Sharmin RA, Zhao T: **Genome-wide association studies for dynamic plant height and number of nodes on the Main stem in summer sowing soybeans.** Front Plant Sci. **2018**, 9, 1184; DOI: 10.3389/fpls.2018.01184.
80. Xu Y, Yang T, Zhou Y, Yin S, Li P, Liu J, Xu S, Yang Z, Xu C: **Genome-wide association mapping of starch pasting properties in maize using single-locus and multi-locus models.** Front Plant Sci. **2018**, 9, 1311; DOI: 10.3389/fpls.2018.01311.
81. He L, Xiao J, Rashid KY, Yao Z, Li P, Jia G, Wang X, Cloutier S, You FM: **Genome-wide association studies for pasmo resistance in flax (*Linum usitatissimum* L.).** Front Plant Sci. **2018**, 9, 1982; DOI: 10.3389/fpls.2018.01982.
82. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ: **PLINK: a tool set for whole-genome association and population-based linkage analyses.** Am J Hum Genet. **2007**, 81, 559-575; DOI: 10.1086/519795.
83. Team RC: **R: A language and environment for statistical computing.** 2013.
84. Arora S, Singh N, Kaur S, Bains NS, Uauy C, Poland J, Chhuneja P: **Genome-wide association study of grain architecture in wild wheat *Aegilops tauschii*.** Front Plant Sci. **2017**, 8, 886; DOI: 10.3389/fpls.2017.00886.
85. Earl DA: **STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.** Conservation Genet Resour. **2012**, 4, 359-361; DOI: 10.1007/s12686-011-

9548-7.

86. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z: **GAPIT: genome association and prediction integrated tool**. *Bioinformatics* **2012**, 28, 2397-2399; DOI:10.1093/bioinformatics/bts444.
87. Ma L, Liu M, Yan Y, Qing C, Zhang X, Zhang Y, Long Y, Wang L, Pan L, Zou C: **Genetic dissection of maize embryonic callus regenerative capacity using multi-locus genome-wide association studies**. *Front Plant Sci.* **2018**, 9, 561; DOI: 3389/fpls.2018.00561.
88. Ren W-L, Wen Y-J, Dunwell JM, Zhang Y-M: **pKWmEB: integration of Kruskal–Wallis test with empirical Bayes under polygenic background control for multi-locus genome-wide association study**. *Heredity*. 2018, **120**(3):208-218; DOI: [10.1038/s41437-017-0007-4](https://doi.org/10.1038/s41437-017-0007-4)
89. Tamba CL, Ni Y-L, Zhang Y-M: **Iterative sure independence screening EM-Bayesian LASSO algorithm for multi-locus genome-wide association studies**. *PLoS Comput Biol.* 2017, **13**(1):e1005357; DOI: [10.1371/journal.pcbi.1005357](https://doi.org/10.1371/journal.pcbi.1005357)
90. Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N: **Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels**. *Nat Genet.* **2013**, 45, 43; DOI:10.1038/ng.2484.
91. Barrett JC, Fry B, Maller J, Daly MJ: **Haploview: analysis and visualization of LD and haplotype maps**. *Bioinformatics.* **2004**, 21, 263-265; DOI: 10.1093/bioinformatics/bth457.
92. Consortium IWGS: **Shifting the limits in wheat research and breeding using a fully annotated reference genome**. *Science.* **2018**, 361, eaar7191; DOI: 10.1126/science. aar7191.

Supplementary Files Legend

Additional file 1. Descriptive statistics of plant architectural traits across multiple environments.

Additional file 2. Analysis of Variance (ANOVA) and heritability for plant height, flag leaf length, flag leaf width and no of tillers per plant.

Additional file 3. Significant SNPs detected through SL-GWAS method across different environments and methods.

Additional file 4. List of pleiotropic SNPs associated with more than one trait.

Additional file 5. Common SNPS detected through ML-GAWS and SL-GWAS methods.

Additional file 6. Summary of potential candidate genes identified for plant architectural traits.

Additional file 7(a-e). Manhattan and quantile-quantile (Q-Q) plots for plant architectural traits across different environments; **(a & b)** Manhattan and Q-Q plots for Plant height (2015-2018); **(c)** Manhattan and Q-Q plots for flag leaf length (2017-2018); **(d)** Manhattan and Q-Q plots for flag leaf width (2017-2018); **(e)** Manhattan and Q-Q plots for number of tillers (2017-2018).

Figures

Figure 1

Phenotype distributions for plant architectural traits (a) Plant height across four environments (2015, 2016, 2017, and 2018); (b) Flag leaf length (2017 and 2018); (c) Flag leaf width (2017 and 2018); (d) Number of tillers per plant (2017 and 2018).

Figure 2

Correlation among plant architectural traits across multiple environments. PH-Plant height (2015-2018), FLL-Flag leaf length (2017-2018), FLW-Flag leaf width (2017-2018) and TILL-Number of tillers per plant (2017-2018). * and ** indicate significant correlation at $P < 0.05$ and 0.01 , respectively.

Figure 3

Population structure analysis of 319 wheat varieties based on unlinked SNP markers. (a) plot of ΔK value with putative k ranging from 2 to 6. (b) two major subpopulations $k=2$, indicated by two colors (red and green) represent the proportion of each subpopulation. (c) analysis of principal component analysis of the genotypic data.

Figure 4

Significant SNPs detected through different GWAS methods in multiple environments. The traits include PH (Plant height); FLL (Flag leaf length); FLW (Flag leaf width) and TILL (Number of tillers per plant). (a) Significant SNPs detected via ML-GWAS and SL-GWAS methods, (b) Significant SNPs detected through four ML-GWAS methods i.e. FASTmrMLM, FASTmrEMMA, mrMLM, and pLARMmEB, (c) Significant SNPs detected across multiple years and different ML-GWAS methods, (d) Significant SNPs detected through three SL-GWAS methods i.e. FarmCPU, MLM, and MLMM.

Figure 5

Putative candidate genes responsible for important functions associated with plant architectural traits. PH (Plant height); FLL (Flag leaf length); TILL (Number of tillers per plant).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile15.docx](#)
- [Additionalfile7ae.ManhattanandQQplots.pptx](#)
- [Additionalfile6.Candidategenes.xlsx](#)