ORIGINAL ARTICLE

Genome-wide association analysis of stem water-soluble carbohydrate content in bread wheat

Luping Fu[1] · Jingchun Wu[1] · Shurong Yang[2] · Yirong Jin[3] · Jindong Liu[4] · Mengjiao Yang[1] · Awais Rasheed[1,5] ·
Yong Zhang[1] · Xianchun Xia[1] · Ruilian Jing[1] · Zhonghu He[1,6] · Yonggui Xiao[1]

Abstract

Key message GWAS identified 36 potentially new loci for wheat stem water-soluble carbohydrate (WSC) contents
and 13 pleiotropic loci affecting WSC and thousand-kernel weight. Five KASP markers were developed and validated.
Abstract Water-soluble carbohydrates (WSC) reserved in stems contribute significantly to grain yield (GY) in wheat. How-
ever, knowledge of the genetic architecture underlying stem WSC content (SWSCC) is limited. In the present study, 166
diverse wheat accessions from the Yellow and Huai Valleys Winter Wheat Zone of China and five other countries were grown
in four well-watered environments. SWSCC at 10 days post-anthesis (10DPA), 20DPA and 30DPA, referred as WSC10,
WSC20 and WSC30, respectively, and thousand-kernel weight (TKW) were assessed. Correlation analysis showed that TKW
was significantly and positively correlated with WSC10 and WSC20. Genome-wide association study was performed on
SWSCC and TKW with 373,106 markers from the wheat 660 K and 90 K SNP arrays. Totally, 62 stable loci were detected
for SWSCC, with 36, 24 and 19 loci for WSC10, WSC20 and WSC30, respectively; among these, 36 are potentially new, 16
affected SWSCC at two or three time-points, and 13 showed pleiotropic effects on both SWSCC and TKW. Linear regression
showed clear cumulative effects of favorable alleles for increasing SWSCC and TKW. Genetic gain analyses indicated that
pyramiding favorable alleles of SWSCC had simultaneously improved TKW. Kompetitive allele-specific PCR markers for
five pleiotropic loci associated with both SWSCC and TKW were developed and validated. This study provided a genome-
wide landscape of the genetic architecture of SWSCC, gave a perspective for understanding the relationship between WSC
and GY and explored the theoretical basis for co-improvement of WSC and GY. It also provided valuable loci and markers
for future breeding.

Abbreviations

BLUE	Best linear unbiased estimation
FarmCPU	Fixed and random model circulating prob-ability unification

GWAS	Genome-wide association study
GY	Grain yield
H^2	Broad-sense heritability
KASP	Kompetitive allele-specific PCR
LD	Linkage disequilibrium
MAF	Minor allele frequency
MAS	Marker-assisted selection
MTA	Marker-trait association
NIRS	Near-infrared spectroscopy

✉ Zhonghu He
zhhecaas@163.com

✉ Yonggui Xiao
xiaoyonggui@caas.cn

[1] Institute of Crop Sciences, National Wheat Improvement
Center, Chinese Academy of Agricultural Sciences (CAAS),
Beijing 100081, China

[2] College of Agronomy, Gansu Agricultural University,
Lanzhou 730000, Gansu, China

[3] Dezhou Institute of Agricultural Sciences, Dezhou 253000,
Shandong, China

[4] Agricultural Genomics Institute at Shenzhen,
Chinese Academy of Agricultural Sciences (CAAS),
Shenzhen 518000, China

[5] Department of Plant Sciences, Quaid-i-Azam University,
Islamabad 45320, Pakistan

[6] International Maize and Wheat Improvement Center
(CIMMYT) China Office, Beijing 100081, China

PIC   Polymorphism information content
QTL   Quantitative trait locus/loci
SWSCC   Stem water-soluble carbohydrate content
TKW   Thousand-kernel weight
WSC   Water-soluble carbohydrate
WSC10   Stem WSC content at 10 days post-anthesis
WSC20   Stem WSC content at 20 days post-anthesis
WSC30   Stem WSC content at 30 days post-anthesis
YHVWWZ   Yellow and Huai Valleys Winter Wheat
             Zone

Introduction

Bread wheat (*Triticum aestivum* L.) is among the most important food crops worldwide. It was estimated that a genetic gain of 50% in yield, or an annual gain of ~2%, is essential to meet predicted global requirements over the next 20 years (Lopes et al. 2012). However, annual gains in yield were only 0.6–0.7% in past decades (Sharma et al. 2012; Gao et al. 2017), mainly achieved by conventional breeding. Therefore, it is urgent to improve grain yield (GY) potential with better dissecting the genetic basis of yield and related traits in wheat.

Grain filling in wheat relies on two major carbon sources, namely direct photosynthetic assimilation from green leaves and reserved carbohydrates in stems and leaf sheaths (Ehdaie et al. 2008). When the photosynthetic source is depressed either by leaf senescence or by drought/heat stresses, grain filling becomes more dependent on mobilized resources (Bidinger et al. 1977; Kobata et al. 1992; Blum et al. 1994). Water-soluble carbohydrates (WSC) are stored in stems and leaf sheaths during vegetative and early reproductive stages, and they are remobilized and transported to grains at the later grain filling stages (Pheloung and Siddique 1991; Wardlaw and Willenbrink 2000). Stem-reserved WSC could account for 10–20% and 30–50% of the wheat GY under well-watered and terminal drought conditions, respectively (Aggarwal and Sinha 1984; Wardlaw and Willenbrink 2000; Foulkes et al. 2010; Ovenden et al. 2017). Therefore, improvement in SWSCC can be a valuable approach to improve GY (Shearman et al. 2005; Ruuska et al. 2008; Sadras and Lawson 2011; Xiao et al. 2012; Gao et al. 2017).

In addition to improving GY, WSC also play important roles in coping with abiotic stresses caused by water deficiency and/or high temperature (Livingston et al. 2009). SWSCC are higher in drought-tolerant cultivars than in sensitive ones (Foulkes et al. 2002; Goggin and Setter 2004), and increasing the genetic capacity for WSC accumulation was considered an approach to improve drought tolerance in wheat (Ovenden et al. 2017). In addition, the components of WSC, e.g., fructose, glucose and sucrose, are involved in plant immunity as signaling molecules for regulation of defense genes on biotic stress (Bolouri-Moghaddam and van den Ende 2013; Trouvelot et al. 2014). It is obvious that WSC are involved in a complex system of plant growth, development and diverse biotic and abiotic stress responses (Rolland et al. 2006; Trouvelot et al. 2014). And researchers indicated selection for higher SWSCC has potential in breeding for improved adaptation across a range of environmental stresses (Rebetzke et al. 2008).

Many studies have shown that genotypic differences in SWSCC are repeatable across diverse environments with high broad-sense heritability (H^2) of 0.7–0.9 (Zhang et al. 2014; Dong et al. 2016a, b). This indicates that variation in SWSCC is largely genetically determined; however, SWSCC is also significantly affected by environmental factors like drought and heat stresses (Ovenden et al. 2017). Quantitative trait loci (QTL) mapping researches of SWSCC have been reported in barley (Teulat et al. 2001), rice (Nagata et al. 2002; Wang et al. 2017; Phung et al. 2019), maize (Thévenot et al. 2005; Bian et al. 2015), perennial ryegrass (Turner et al. 2006), and sorghum (Brenton et al. 2016). In wheat, diverse bi-parental populations were used in identification of QTL or genomic regions associated with SWSCC (Snape et al. 2007; Yang et al. 2007; Rebetzke et al. 2008; Dong et al. 2016b). Nevertheless, these studies didn't provide a genome-wide landscape of the complex genetic architecture, as family-based genetic populations have limited diversity. In addition, the amount of recombination places a limit on mapping resolution in family-based QTL mapping (Korte and Farlow 2013). It is also possible that QTL with moderate or small effects may be missed in QTL mapping, and the Beavis effect can cause a biased estimation of QTL effects especially when the population size is small (Xu 2003). Moreover, the relatively long genetic distances between linked markers and causal genes limit their use in marker-assisted selection (MAS, Platten et al. 2019).

Genome-wide association studies (GWAS) on SWSCC were performed at flowering, mid-grain filling and maturity stages (Zhang et al. 2014; Li et al. 2015), but the detailed genetic architecture was not revealed due to use of only 209 SSR markers. Dong et al. (2016a) conducted GWAS on SWSCC at 14 days post-anthesis (DPA) using 18,207 markers from the wheat 90 K SNP array (Wang et al. 2014), but there were still large gaps in the genetic map, and marker coverage for the D genome was particularly sparse (Liu et al. 2017). Therefore, it is necessary for a more precise dissection of the genetic architecture underlying the complexity of stem carbohydrate metabolism. In addition, genes controlling SWSCC express dynamically at different growth stages (Veenstra et al. 2017; Yáñez et al. 2017; Hou et al. 2018), and knowledge of basis of SWSCC over time would provide more valuable information for breeding. In maize, QTL analyses of stalk sugar contents at different growth stages were conducted by Bian et al. (2015). Although QTL mapping

works on wheat SWSCC have been performed at different stages (Snape et al. 2007; Yang et al. 2007; Rebetzke et al. 2008; Zhang et al. 2014; Li et al. 2015; Dong et al. 2016a, b) knowledge of the dynamic expression patterns of genes associated with WSC remains limited. It is therefore important to gain deep insights into the genetic expression patterns related to wheat SWSCC at different stages.

In the present study, a diversity panel of 166 winter wheat accessions was planted in four well-watered environments, and SWSCC at three time-points at the grain filling stage and thousand-kernel weight (TKW) were investigated. GWAS was performed on SWSCC and TKW using a high-density physical map constructed with the markers from wheat 660 K and 90 K SNP arrays. The objectives were to (1) evaluate the relationship between SWSCC and TKW, (2) identify loci and candidate genes associated with SWSCC and TKW, providing insights into the genetic basis of SWSCC in wheat and (3) develop high-throughput kompetitive allele-specific PCR (KASP) markers for MAS targeting WSC and TKW improvement.

Materials and methods

Plant materials and field trials

A diversity panel of 166 representative wheat accessions chosen from more than 400 cultivars was used for GWAS on SWSCC and TKW, including 144 accessions collected from the Yellow and Huai Valleys Winter Wheat Zone (YHVWWZ) of China, and 22 from five other countries (Liu et al. 2017; Zhai et al. 2018; Li et al. 2019; Table S1). Among them, 130 Chinese wheat cultivars released from 1947 to 2016 (Table S1) were used to investigate the genetic progress in improvement in SWSCC and TKW. These cultivars were divided into five groups, i.e., 9 cultivars released during 1947–1979, 13 in the 1980s, 36 in the 1990s, 59 in the 2000s and 13 released in the 2010s. Furthermore, cultivars released after 1990 from five main wheat producing provinces in China, i.e., Anhui (9 cultivars), Hebei (13), Henan (44), Shandong (21) and Shaanxi (17), were used to investigate inter-province differences of SWSCC and TKW (Table S1).

All accessions were grown in four environments including Dezhou (37° 27′ N, 116° 18′ E; Shandong Province) and Gaoyi (37° 37′ N, 114° 34′ E; Hebei Province) during the 2016–2017 cropping season, and Luohe (33° 36′ N, 113° 58′ E; Henan Province) and Xinxiang (35° 18′ N, 113° 51′ E; Henan Province) during 2017–2018. These environments were designated as 17DZ, 17GY, 18LH and 18XX, respectively. The field trials at each location were carried out under well-watered conditions with two flood irrigations during jointing and flowering stages. All locations experienced warm temperatures during the later grain filling stage (Fig. S1). The cultivars were planted in randomized complete blocks with three replications. Each plot contained two 2-m rows spaced 20 cm apart, with about 50 seeds sown per row. Field managements followed local practices and fungicide applications were made to control diseases (powdery mildew, stripe rust and leaf rust). All accessions are available from the National Gene Bank of China, Chinese Academy of Agricultural Sciences.

Phenotypic evaluation

SWSCC was assayed at 10DPA, 20DPA and 30DPA using a near-infrared spectroscopy (NIRS) method following Wang et al. (2011), and the corresponding phenotypic data are referred to as WSC10, WSC20 and WSC30, respectively. About ten main culms were randomly taken from each plot, the leaf blades were removed, and the spikes were cut off at the spike collars. Fresh samples from each plot were put into a labeled paper bag and exposed to 105 °C for 30 min and then oven-dried at 80 °C for 24 h. The dried samples were cut into 3–5 mm lengths. Before NIRS assays, the cut samples were re-dried at 80 °C until a constant weight, and brought to room temperature in vacuum bags. The detailed procedure for NIRS assay was reported in Dong et al. (2016b), and SWSCC was reported as a percentage on a dry weight basis. Three technically independent assays were performed for each sample at 10DPA and 20DPA, and five assays were made for samples at 30DPA. The mean values for each sample with extreme outliers discarded were used in subsequent statistical analyses. TKW of wheat accessions were obtained using an automatic seed character analyzer with the SC-G V2.1.2.3 software (Wanshen Detection Technology Co., Ltd., Hangzhou, China, http://www.wseen.com/).

Statistical analysis

Analysis of variance (ANOVA) was performed using SAS 9.2 software (SAS Institute Inc., Cary, NC, USA). Mean squares of each source of variation were used to estimate the variance components for genotypes (σ_G^2), genotype × environment interaction (σ_{GE}^2) and residual error (σ_ε^2), respectively, and heritabilities were estimated using the formula $H^2 = \sigma_G^2/(\sigma_G^2 + \sigma_{GE}^2/e + \sigma_\varepsilon^2/(re))$, in which e and r were the numbers of environments and replicates per environment, respectively (Holland et al. 2003; Yin et al. 2015). Best linear unbiased estimations (BLUE) for phenotypic data across environments were extracted using the linear model described in Yin et al. (2015) which had been implemented in the ANOVA function in QTL IciMapping v4.1 software (Li et al. 2007). Correlation analyses and t-tests were performed using SAS 9.2.

Genotyping and physical map construction

Genomic DNA was extracted from young leaves using a modified CTAB method (Murray and Thompson 1980). All accessions were genotyped using the Affymetrix 660 K wheat SNP array (containing 630,517 SNPs, Cui et al. 2017) and the Illumina 90 K wheat SNP array (containing 81,587 SNPs, Wang et al. 2014) by CapitalBio Technology Co., Ltd. (http://www.capitalbiotech.com/). Minor allele frequency (MAF), polymorphism information content (PIC) and genetic diversity were computed by PowerMarker v3.25 (Liu and Muse 2005, http://statgen.ncsu.edu/powermarker/). The heterozygous genotypes were considered as missing data; markers with MAF < 5% and missing data > 20% were excluded to avoid spurious marker-trait associations (MTAs) in subsequent association mapping. Flanking sequences of SNP markers were used to blast against the Chinese Spring (CS) reference genome in IWGSC (RefSeq v1.0, http://www. wheatgenome.org/; IWGSC 2018), and corresponding physical positions were determined according to the best blast hit results. The positions of 1212 markers with multiple hit positions on different chromosomes were assigned according to the consensus 660 K-SNP (Cui et al. 2017) and 90 K-SNP (Wang et al. 2014) genetic linkage maps. High-quality markers from the two SNP arrays were integrated into a common physical map for association study.

Linkage disequilibrium and population structure

Linkage disequilibrium (LD) and population structure of the 166 accessions were analyzed in a previous study using the same population (Liu et al. 2017). Briefly, 12,324 SNPs evenly distributing on 21 wheat chromosomes were used to calculate LD using a full matrix and sliding window method implemented in Tassel v5.0 (Bradbury et al. 2007). It showed that the average LD decay distance for the whole genome was about 8 Mb; and LD decays were 6, 4 and 11 Mb for the A, B and D genomes, respectively (Fig. S2; Liu et al. 2017). A total of 2000 evenly distributed polymorphic SNPs were chosen to analyze the population structure and estimate the Q matrix by the software Structure v2.3.4 (Pritchard et al. 2000). A neighbor-joining tree was constructed and principal components analysis was performed by Tassel v5.0 to verify the population stratification. Obvious population stratification was observed and the entire panel comprised three subgroups (Fig. S3, Table S1; Liu et al. 2017).

Genome-wide association study

The mean values of three replicates in each environment and the BLUE values across environments for each trait were used for GWAS. To control background variation and eliminate spurious MTAs, associations between markers and traits were estimated using a $Q + K$ mixed linear model (MLM, Yu et al. 2006; Zhang et al. 2010) which was implemented in software TASSEL v5.0 (Bradbury et al. 2007). The Q matrix estimated by Structure v2.3.4 (Pritchard et al. 2000), defining the population structure was considered a fixed-effect factor. The kinship matrix (K matrix) computed by TASSEL v5.0, reflecting relationships among individuals, was incorporated as the variance–covariance structure of the random effect for individuals (Zhang et al. 2010). As the $Q + K$ MLM may also compromise true positives in some cases when it controls false positives, the fixed and random model circulating probability unification (FarmCPU) method, which was demonstrated to have a more improved statistical power than MLM (Liu et al. 2016), was also used to perform GWAS on SWSCC by R software (R Version 3.5.1, https ://www.r-project.org/; FarmCPU package at http://zzlab .net/FarmCPU/FarmCPU_functions.txt).

Different methods for multiple testing corrections were tried, including the Bonferroni-Holm correction method (Holm 1979) and the Benjamini and Hochberg's false discovery rate (FDR) procedure (Benjamini and Hochberg 1995), but few SNPs could be declared significant in some of the trait-by-environment conditions in the present study which may due to the higher extent of LD in wheat (Chao et al. 2010; Hao et al. 2011; Chen et al. 2012; Liu et al. 2017) and/or the complex underlying genetic architecture for WSC. Finally, a threshold of $P = 1.0 \times 10^{-3}$ ($-\log_{10}(P) = 3.0$) was adopted for calling significant MTAs. This threshold was also used in some other association studies on complex traits in hexaploid wheat (Liu et al. 2017; Muqaddasi et al. 2019; Rahimi et al. 2019).

The adjacent associated markers were grouped together as one locus if the inter-marker distance is smaller than the average LD decay for specific chromosome, which was reported in Liu et al. (2017). The most significant marker across environments for each locus was considered the representative, and the corresponding effect and R^2 (phenotypic variance explained) were estimated and outputted by TASSEL v5.0 (Bradbury et al. 2007). To further control the FDR of WSC-associated loci, those detected in at least two environments by either MLM or FarmCPU were considered to be stable. The genome-wide MTAs were visualized by Manhattan plots with $-\log_{10}(P)$ for each SNP displayed on the Y-axis and the corresponding genomic coordinates displayed along the X-axis. The quantile–quantile (Q–Q) plots (observed *versus* expected $-\log_{10}(P)$ values) were used to assess the association mapping models. Manhattan plots and Q–Q plots of the GWAS results were drawn using the CMplot code (https://githu b.com/YinLiLin/R-CMplot) in R software (Version 3.5.1).

Analyses of allele frequencies and effects of identified QTL in subgroups

Allele frequencies and effects of the identified QTL were analyzed in three subgroups (Table S1) of the population based on representative markers of each locus, and the averaged BLUE values for two genotypes at each locus were used to compare the effects among subgroups by *t*-test.

Comparison of two models for GWAS

To compare the GWAS performance of MLM and Farm-CPU, the linear model (LM) fitting analysis was performed using representative markers for the identified QTL, where markers were fitted as independent variables and observed phenotypes as dependent variables. The coefficients of determination from the LM were then calculated using the LM function in R 3.5.1 (https://stat.ethz.ch/R-manual/R-patched/library/stats/html/lm.html). Dot plots with the observed phenotypes on the *X*-axis and predicted values on the *Y*-axis were drawn to show the LM fitting results.

Comparison of identified WSC loci with reported QTL or genes

The WSC loci detected in this study were compared with WSC- and GY-related QTL or genes that were searched from the literature based on the physical positions (CS RefSeq v1.0; IWGSC 2018) of their flanking or associated markers. If the physical distances between two QTL were smaller than the average LD decay for a specific chromosome (Liu et al. 2017), they were considered to be at the same locus.

Putative candidate gene analyses

The genes located in the physical intervals of WSC-associated loci were screened based on the annotations in the wheat reference genome (CS RefSeq v1.0; IWGSC 2018), and those related to sugar metabolism or transportation were regarded as candidate genes. In addition, the sequences and corresponding physical positions of some known genes involved in WSC synthesis, degradation, and remobilization were obtained from the NCBI (https://www.ncbi.nlm.nih.gov/) and IWGSC (http://www.wheatgenome.org/). The positions of the known genes and their homoeologs were compared with the WSC loci, and some candidates were identified.

KASP marker development

High-throughput KASP markers for potentially important loci were developed based on corresponding representative SNPs. Flanking sequences of SNPs were used as queries to blast against the wheat reference genome in IWGSC (CS RefSeq v1.0; IWGSC 2018), and chromosome-specific KASP primers were developed based on alignment of homologous sequences. Allele-specific primers carrying FAM (5′ GAAGGTGACCAAGTTCATGCT 3′) and HEX (5′ GAAGGTCGGAGTCAACGGATT 3′) tails were designed with the targeted SNP at the 3′ end, and common reverse primer was designed for a chromosome-specific amplification with less than 200 bp of amplified sequence. The KASP assay mixture was prepared with 40 μL of common primer (100 μM), 16 μL of each tailed primer (100 μM), and 60 μL of ddH$_2$O. Each reaction mixture comprised 2.5 μL of 2×KASP master mixture (LGC Genomics, https://www.biosearchtech.com/), 0.056 μL of the assay mixture, and 2.5 μL of DNA (30–50 ng/μL). PCR were performed in a 384-well plate as follows: denaturation at 95 °C for 15 min, followed by 9 touchdown cycles (95 °C for 20 s; touchdown at 65 °C initially then decreasing by 1 °C per cycle for 1 min), and 32 additional cycles of denaturing, annealing and extension (95 °C for 10 s; 57 °C for 1 min). Consistency between KASP genotyping results and the original chip-based genotypes was investigated, and *t*-tests were conducted to confirm the effectiveness of the KASP markers.

Results

Phenotypic variation

Continuous variations among the 166 wheat accessions were observed for WSC10, WSC20, WSC30 and TKW (Table S1, Fig. S4). The resulting BLUE values for WSC10, WSC20, WSC30 and TKW across the four environments were 9.41–18.81% (mean, 14.71%), 8.49–17.87% (mean, 13.62%), 1.04–9.71% (mean, 4.14%) and 26.9–56.6 g (mean, 43.1 g), respectively (Tables S1, S2). Correlations among environments for WSC10, WSC20, WSC30 and TKW showed ranges of 0.56–0.75, 0.51–0.71, 0.50–0.60 and 0.66–0.87 with $P < 0.0001$, respectively (Table S3). ANOVA revealed that genotypes, environments and genotype×environment interactions had significant effects on SWSCC at all three time-points (Table 1). The H^2 of WSC10, WSC20, WSC30 and TKW across four environments were 0.90, 0.87, 0.85 and 0.93, respectively (Table 1), suggesting that most of the phenotypic variation was determined by genetic factors.

Correlation analyses showed that WSC10 was significantly correlated with WSC20 with correlation coefficients (*r*) ranging from 0.27 ($P < 0.001$) to 0.61 ($P < 0.0001$) in different environments and in BLUE; WSC20 was significantly correlated with WSC30, with $r = 0.43–0.61$ ($P < 0.0001$), whereas the correlations between WSC10 and WSC30 were much lower, ranging from 0.04 (not significant) to 0.31 ($P < 0.001$) (Table 2). Correlations between WSC10

Table 1 Analysis of variance for stem WSC contents and thousand-kernel weight in the 166 wheat accessions

Trait[a]	Analysis of variance[b]					Broad-sense heritability
	Genotype (G)	Environment (E)	G × E interaction	Replicate	Error	
WSC10	42.09***	341.64***	4.59***	6.78***	1.17	0.90
WSC20	41.35***	340.88***	6.02***	3.90***	1.51	0.87
WSC30	33.84***	408.22***	5.88***	6.29***	1.40	0.85
TKW	268.55***	6701.84***	18.48***	94.10***	3.64	0.93

[a]WSC10, WSC20 and WSC30 indicate stem WSC contents at 10 days post-anthesis (10DPA), 20DPA and 30DPA, respectively. TKW, thousand-kernel weight

[b]Mean square values from the analysis of variance are reported. ***$P < 0.0001$

Table 2 Correlation analyses for stem WSC contents and thousand-kernel weight of the 166 wheat accessions in different environments

Environment[a]	Trait[b]	WSC10	WSC20	WSC30
17DZ	WSC20	0.27**		
	WSC30	0.04	0.57***	
	TKW	0.45***	0.32***	0.15
17GY	WSC20	0.44***		
	WSC30	0.16*	0.61***	
	TKW	0.33***	0.53***	0.42***
18LH	WSC20	0.55***		
	WSC30	0.19*	0.43***	
	TKW	0.51***	0.50***	0.20*
18XX	WSC20	0.61***		
	WSC30	0.31***	0.52***	
	TKW	0.63***	0.59***	0.34***
BLUE	WSC20	0.57***		
	WSC30	0.20**	0.52***	
	TKW	0.56***	0.51***	0.34***

[a]17DZ and 17GY, Dezhou and Gaoyi locations, respectively, 2016–2017; 18LH and 18XX, Luohe and Xinxiang locations, respectively, 2017–2018; BLUE, best linear unbiased estimation across environments for each trait

[b]WSC10, WSC20 and WSC30 indicate stem WSC contents at 10 days post-anthesis (10DPA), 20DPA and 30DPA, respectively. TKW, thousand-kernel weight

***$P < 0.0001$; **$P < 0.001$; *$P < 0.01$

and TKW ($r = 0.33$–0.63, $P < 0.0001$) were similar to those between WSC20 and TKW ($r = 0.32$–0.59, $P < 0.0001$), whereas r between WSC30 and TKW ranged from 0.15 (not significant) to 0.42 ($P < 0.0001$) (Table 2).

Marker coverage and genetic diversity

A total of 373,106 high-quality SNPs with 359,760 (96.42%) and 13,346 (3.58%) from the 660 K and 90 K SNP arrays, respectively, were used in GWAS (Table S4). The A, B and D genomes were represented by 39.8, 49.3 and 10.8% of the markers, respectively. Chromosome 3B possessed the most markers (46,708), whereas chromosome 4D had

the least (2375). The markers covered the whole genome (14,061.15 Mb) with an average density of 0.038 Mb per marker, and for respective chromosomes, the marker densities ranged from 0.018 (3B) to 0.214 (4D). The B genome showed the highest marker density (0.028 Mb per marker), genetic diversity (0.367) and PIC (0.293) compared with the A (0.033, 0.357 and 0.286) and D (0.098, 0.334 and 0.270) genomes. The detailed information of marker number, marker density, genetic diversity and PIC is provided in Table S4, in addition to a density map showing marker distribution along chromosomes (Fig. S5).

Marker-trait associations and stable WSC-associated loci

Significant MTAs for WSC10, WSC20 and WSC30 analyzed using the MLM and FarmCPU methods are listed in Tables S5–S10. In total, 1095, 652 and 597 significant markers corresponding to 168, 93 and 14 loci were detected for WSC10, WSC20 and WSC30, respectively, using MLM in TASSEL; and 813, 1344 and 1415 significant markers corresponding to 134, 112 and 23 loci were identified for the three traits using FarmCPU. Manhattan plots for WSC contents analyzed by MLM and FarmCPU using BLUE values are shown in Fig. 1, and Manhattan and Q–Q plots for each trait in each environment analyzed by both methods are shown in Figs. S6 and S7.

The WSC loci detected in at least three out of the four environments by either MLM or FarmCPU are summarized in Table 3. Furthermore, all stable WSC loci that detected in at least two environments by either MLM or FarmCPU are reported in Table S11. The numbers of stable loci for WSC10, WSC20 and WSC30 were 36, 24 and 19, respectively (Table S11). Overall, 62 stable loci for SWSCC across the three developmental time-points were detected on all 21 chromosomes except 5D, and 16 of these loci were associated with SWSCC at two or more time-points (Table S11, Fig. 2). In terms of the stable loci, 6 (17%), 7 (29%) and 8 (42%) for WSC10, WSC20 and WSC30, respectively, were detected by both MLM and FarmCPU.
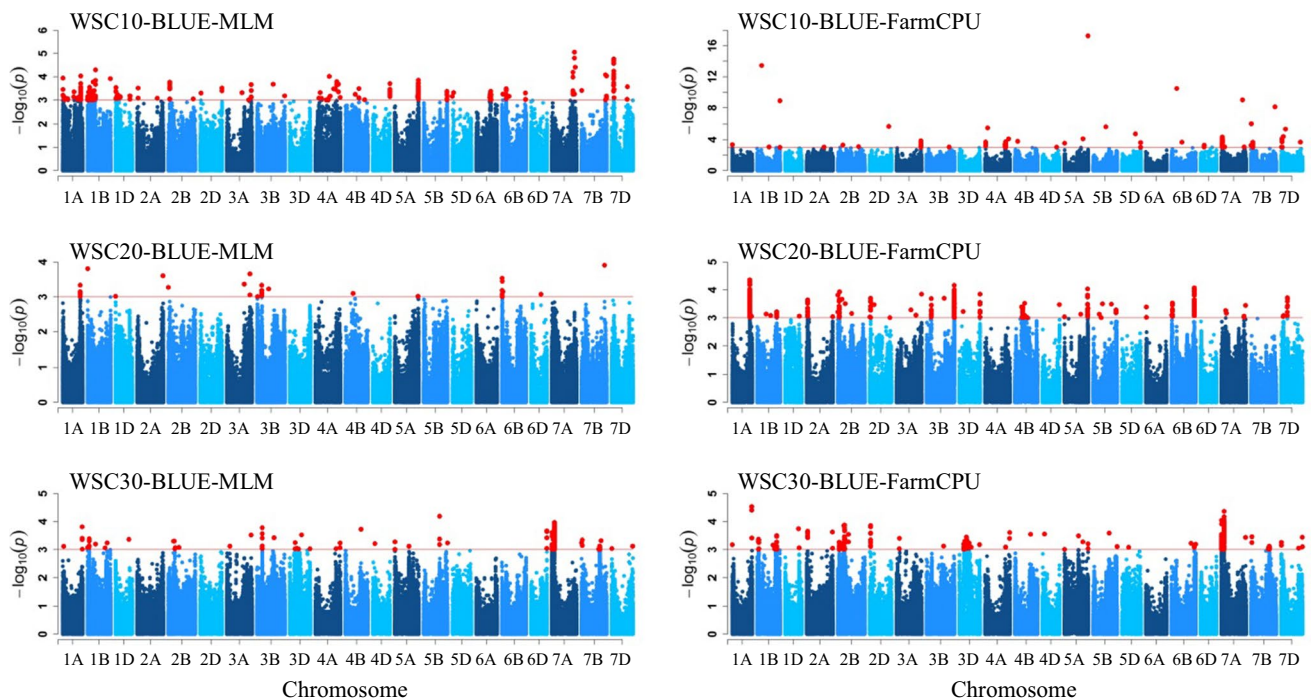
Fig. 1 Manhattan plots for stem WSC contents analyzed by $Q+K$ mixed linear model (MLM) and FarmCPU. The threshold of $P = 1.0 \times 10^{-3}$ ($-\log_{10}(P) = 3.0$) was used for calling significant marker-trait associations (MTAs). WSC10, WSC20 and WSC30 indicate stem WSC contents at 10 days post-anthesis (10DPA), 20DPA and 30DPA, respectively. BLUE indicates the best linear unbiased estimations for each trait across four environments in this study

Allele frequencies and QTL effects in subgroups

For each WSC QTL, the allele frequencies and effects in three subgroups of the population were investigated (Table S12). For most of the loci, similar allele frequencies and effects were observed among subgroups. But some large inter-subgroup variations of frequencies and effects were detected in a few loci. For instance, the allele frequency of a WSC10 locus on chromosome 5B (692.45 Mb, *AX-111100349*) were 0.37 for the entire panel, while the frequencies were 0.11, 0.57 and 0.49 in subgroups 1, 2 and 3, respectively (Table S12). The effect of a WSC10 locus on 1B (681.04 Mb, *AX-110000934*) was 0.57% for the entire population, whereas the effects were 2.48, 0.28 and 0.05% in subgroups 1, 2 and 3, respectively (Table S12). These indicated that the allele frequencies and effects of a few loci might be significantly affected by the population stratification, and we need to further confirm their effects in future studies.

Pleiotropic loci for SWSCC and TKW

By comparing the physical positions of significant markers, 13 stable loci associated with TKW were co-localized with the WSC loci on chromosomes 1A, 1B (2), 1D, 2A, 2B, 3A (2), 4A, 4B, 5B and 6B (2), explaining 7.69–16.38%

and 8.17–15.53% of the phenotypic variations for TKW and SWSCC, respectively (Tables 4, S11). Because high TKW is a main objective in wheat breeding, the alleles associated with increased TKW and SWSCC were considered favorable. The frequencies of favorable alleles at five loci on 1A, 1B, 2B and 3A (2) were 0.86–0.93 (Table 4), indicating that the favorable alleles at these loci had already been widely selected in traditional breeding programs, whereas the other eight loci exhibited favorable allele frequencies (FAF) of 0.45–0.70 (Table 4), indicating more potential possibilities for selection in future breeding.

Cumulative effect of increasing-effect alleles on SWSCC and TKW

To further investigate the effects of combined alleles on SWSCC, the number of increasing-effect alleles in each accession was investigated. The number of WSC-increasing alleles possessed by each accession had ranges of 6–32, 5–22 and 2–16 for WSC10, WSC20 and WSC30, respectively (Table S1, Fig. S8). Significant correlations were observed between SWSCC and number of WSC-increasing alleles with $r = 0.84$, 0.68 and 0.63 ($P < 0.001$) for WSC10, WSC20 and WSC30, respectively. Linear regressions using the BLUE values were determined to further investigate the relationships between SWSCC and the number

Table 3 WSC-associated loci detected in at least three environments by either MLM or FarmCPU

Trait[a]	Environment-MLM[b]	Environment-FarmCPU[b]	Chr[c]	Marker interval (Mb)[d]	Representative SNP[e]	Major/Minor allele[f]	MAF[f,g]	P value[f]	R² (%)[f,h]	Effect[f]
WSC10	2	1, 2, 4, B	1A	445.51–461.52	AX-109989656	C/T	0.22	2.72E–04	8.86	1.80
	1, 2, 4	1	1A	513.74–522.15	AX-109293426	T/C	0.12	2.29E–05	12.98	2.68
	2, 3	2, 3, 4, B	1B	39.60–41.64	AX-111648440	C/T	0.24	1.09E–05	13.68	1.89
	2, 3, 4, B	–	1D	4.25–8.62	AX-94811887	C/A	0.06	2.09E–04	10.38	2.99
	–	1, 3, 4	2A	701.48–709.03	BS00098423_51	C/A	0.14	3.38E–06		1.21
	1, 3, 4, B	–	4A	735.30–737.58	AX-110010963	C/T	0.08	2.52E–04	10.16	2.42
	2, 3, 4	2, 4	6B	0.49–14.26	AX-108959247	A/T	0.37	8.86E–05	12.33	1.58
	1, 2, B	1, 2, 4, B	7D	60.58–78.26	AX-110514480	T/G	0.24	5.50E–06	15.43	2.19
WSC20	1, 2, 4, B	1, 3, 4	1B	4.35–9.56	AX-111668935	G/A	0.27	1.21E–04	12.07	2.22
	4	1, 3, 4, B	1D	414.59–420.74	AX-111640220	C/T	0.16	4.07E–04	9.85	2.12
	3	1, 2, 3, B	2A	27.51–32.40	AX-111552571	T/C	0.16	2.06E–05	14.58	2.70
	1, 3, 4	1, 3	2B	747.81–752.49	AX-108729610	T/C	0.45	1.37E–04	11.24	1.67
	–	1, 2, 3, B	3A	737.44–744.29	AX-94829137	A/G	0.2	7.37E–04		1.61
	2, 3, 4, B	2, 3, 4, B	5A	688.14–690.41	AX-111031397	A/G	0.06	1.55E–05	13.8	3.56
	2	1, 2, 3, 4, B	6B	674.84–677.48	AX-111274903	A/G	0.42	8.97E–04	8.17	1.31
WSC30	2, 3, 4, B	2, 3, 4, B	1A	572.25–581.75	BobWhite_c39668_143	G/A	0.21	1.57E–04	11.77	1.89
	1, 2	1, 2, 3, B	2A	763.02–769.51	AX-110028549	C/G	0.49	2.52E–05	17.44	1.78
	–	1, 2, 3, B	7A	38.33–46.90	AX-108784564	G/A	0.07	9.48E–04		2.49
	2, B	1, 2, 4, B	7A	81.88–90.63	AX-108972618	T/C	0.36	1.32E–04	11.59	1.23

[a] WSC10, WSC20 and WSC30 indicate stem WSC contents at 10 days post-anthesis (10DPA), 20DPA and 30DPA, respectively

[b] Environments of the corresponding locus have been detected by mixed linear model (MLM) or FarmCPU; 1, 2, 3 and 4 indicate Dezhou, Gaoyi, Luohe and Xinxiang locations, and B indicates the best linear unbiased estimation across four environments; "–" indicates the locus has not been detected by the corresponding GWAS model

[c] The chromosome of the corresponding locus located on

[d] Physical positions of SNP markers were based on the Chinese Spring reference genome in IWGSC (RefSeq v1.0, http://www.wheatgenome.org/)

[e] The most significant SNP across environments for the corresponding locus was reported as a representative

[f] The information in corresponding columns are based on the representative SNP

[g] MAF indicates the minor allele frequency

[h] R^2 indicates the percentage of phenotypic variance explained by the SNP marker; the data were not provided if a locus was detected only by FarmCPU
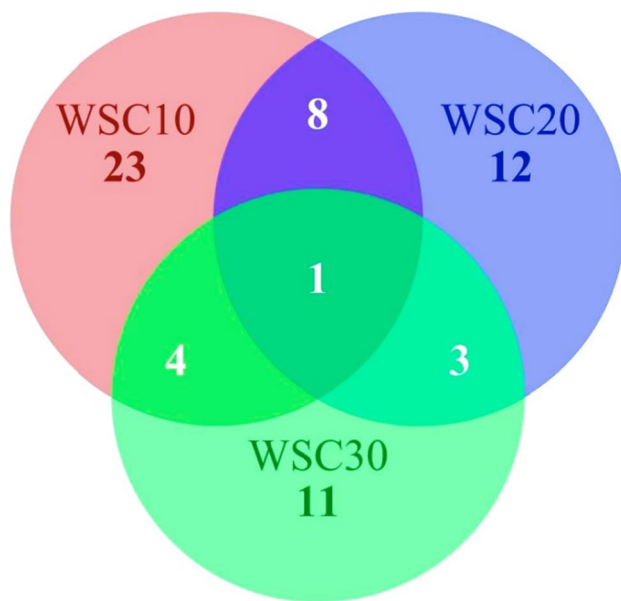
Fig. 2 The Venn diagram of stable loci associated with stem WSC contents. Loci detected in at least two environments by either mixed linear model (MLM) or FarmCPU were considered stable (Table S11). WSC10, WSC20 and WSC30 indicate the stem WSC contents at 10 days post-anthesis (10DPA), 20DPA and 30DPA, respectively

of WSC-increasing alleles, and it showed significant linear associations between SWSCC and number of alleles with regression slopes of 0.31 (coefficient of determination (r^2) = 0.95), 0.44 (r^2 = 0.95) and 0.40 (r^2 = 0.95) for WSC10, WSC20 and WSC30, respectively (Fig. 3).

The accessions contained 0–13 increasing-effect alleles at the 13 pleiotropic loci for both SWSCC and TKW (Table S1, Fig. S8). There were clear cumulative effects on WSC10, WSC20, WSC30 and TKW, with cumulative number of increasing-effect pleiotropic alleles (Fig. 4). Linear regression slopes for WSC10, WSC20, WSC30 and TKW *versus* corresponding numbers of increasing-effect alleles were 0.51 (r^2 = 0.88), 0.49 (r^2 = 0.91), 0.30 (r^2 = 0.82) and 1.56 (r^2 = 0.96), respectively.

Pyramiding WSC-increasing alleles improved SWSCC and TKW over past decades

To explore the roles of WSC-associated loci in improving GY, the genetic progresses of SWSCC and TKW have been investigated. The results showed that WSC10, WSC20 and WSC30 had increased on average from 13.02%, 12.55% and 3.45% before 1980 (9 cultivars) to 16.27%, 15.41% and 4.63% after 2010 (13 cultivars), respectively (Fig. 5a–c); accompanied by increased TKW from 41.1 to 46.4 g (Fig. 5d) and increased numbers of increasing-effect

Table 4 Thirteen TKW-associated loci simultaneously associated with stem WSC contents

Representative marker[a]	Chromosome	Physical interval (Mb)[b]	*P* value	R^2 (%)[c]	Effect (g)	Allele[d]	Frequency of favorable allele	Environment[e]
AX-110587308	1AL	587.83–587.84	3.22E−04	8.50	5.29	<u>G</u>/A	0.86	E3, E4
BS00023084_51	1BS	7.48	1.49E−04	9.44	3.41	<u>A</u>/G	0.68	E1, E2, BLUE
AX-109463291	1BL	580.63–581.21	1.72E−04	9.56	5.90	<u>T</u>/C	0.91	E1, E2, E3, E4, BLUE
AX-111496323	1DS	8.01–11.49	2.87E−04	8.51	3.86	<u>C</u>/G	0.70	E1, E3, BLUE
AX-109293110[†]	2AS	27.33–31.96	2.44E−04	8.89	3.89	<u>T</u>/C	0.51	E2, E4, BLUE
wsnp_Ex_c163_320858	2BS	41.96–46.88	3.92E−05	11.55	5.31	<u>G</u>/A	0.89	E1, E2, E3, E4, BLUE
Kukri_c7087_896	3AL	594.00–595.44	3.06E−04	8.45	5.81	<u>A</u>/G	0.93	E2, E4
AX-109483618	3AL	714.36–716.25	2.56E−04	8.76	5.72	<u>G</u>/C	0.93	E2, E4, BLUE
AX-109832317[†]	4AL	738.89–742.84	3.49E−04	9.55	3.63	<u>C</u>/T	0.45	E1, E2
AX-110974144[†]	4B	298.11–302.17	1.63E−04	11.37	4.60	<u>A</u>/G	0.50	E3, E4, BLUE
AX-111044647[†]	5BL	694.03–697.61	3.00E−04	8.33	3.56	<u>A</u>/G	0.67	E1, E3, BLUE
AX-94910312	6BS	30.01–32.33	5.57E−04	7.69	3.62	<u>G</u>/T	0.69	E1, E3
AX-111494281[†]	6BL	670.21–678.06	1.35E−06	16.38	4.51	<u>T</u>/C	0.50	E1, E2, E3, E4, BLUE

[a]SNP markers indicated with † were converted to KASP markers (Tables S12, S13; Figs. 7, S10)

[b]Physical positions of SNP markers were based on the Chinese Spring reference genome in IWGSC (RefSeq v1.0, http://www.wheatgenome.org/)

[c]R^2, percentage of the phenotypic variance explained by the QTL

[d]The underlined allele indicates the favorable allele that showed increasing-effects on thousand-kernel weight and stem WSC content

[e]E1, Dezhou 2016–2017; E2, Gaoyi 2016–2017; E3, Luohe 2017–2018; E4, Xinxiang 2017–2018; BLUE, best linear unbiased estimation
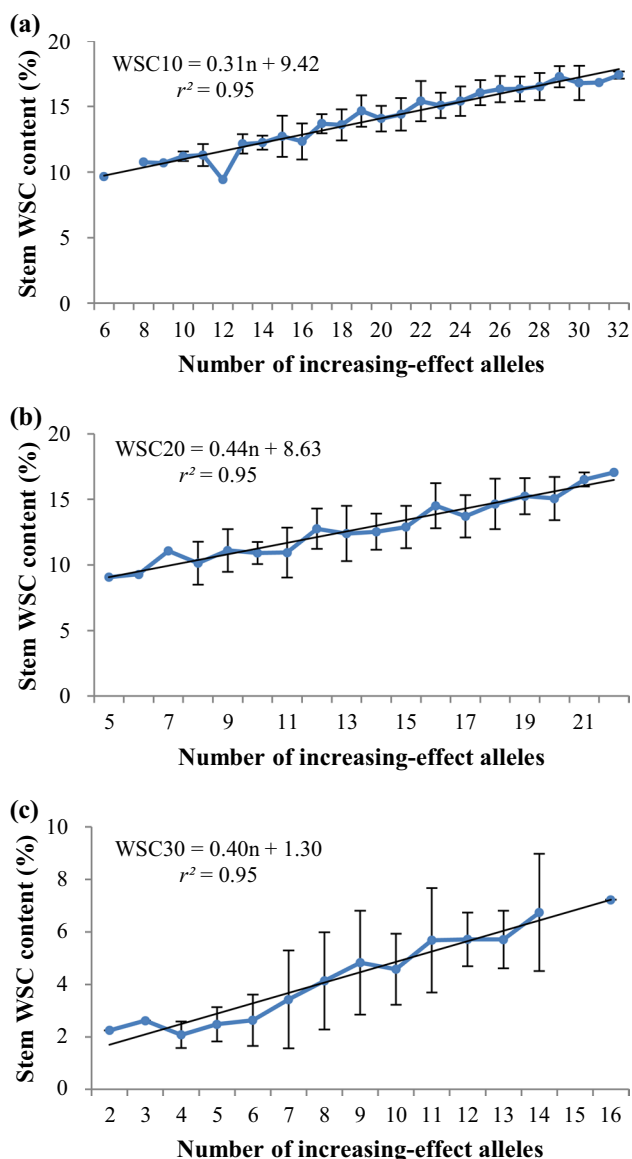
Fig. 3 Cumulative effects of increasing-effect alleles on stem WSC contents. WSC10, WSC20 and WSC30 indicate stem WSC contents at 10, 20 and 30 days post-anthesis, respectively. "n", number of increasing-effect alleles. Dots and bars represent the mean values and standard deviations, respectively. Linear regressions were performed to investigate the relationships between stem WSC contents and number of increasing-effect alleles in the 166 wheat accessions. Calculations were based on representative markers of the 36, 24 and 19 stable loci associated with WSC10, WSC20 and WSC30, respectively (Table S11). Best linear unbiased estimations (BLUE) for each trait across four environments were used

alleles from 15, 14, 7 and 7 to 26, 18, 10 and 11 for WSC10, WSC20, WSC30 and TKW, respectively (Fig. 5e–h).

In addition, we investigated the frequency-changes in favorable alleles for the 13 pleiotropic loci (Table 4). It showed that FAFs for all these loci increased in recent decades except the one on 4AL with a representative marker

AX-109832317 (Fig. 6). Furthermore, differences of SWSCC and TKW among five main wheat producing provinces in China have been investigated. The results showed that cultivars in Henan, which is the top wheat producing province in China, had the most WSC10 and TKW, as well as the most numbers of increasing-effect WSC alleles (Fig. S9).

KASP marker development

Among the 13 pleiotropic loci showing effects on both WSC and TKW, the alleles at loci on 1AL, 1BL, 2BS and 3AL (2) were almost fixed in this panel of wheat accessions with FAF of 0.86–0.93, whereas the FAF of 0.45–0.70 were found in the other eight loci leaving more values for MAS, and these have been focused to develop KASP markers (Table 4). Finally, five KASP markers were successfully developed for representative SNP markers of pleiotropic loci on chromosomes 2A (physical interval 27.51–34.67 Mb), 4A (728.51–739.58 Mb), 4B (298.15–299.53 Mb), 5B (689.91–696.19 Mb) and 6B (674.84–677.48 Mb) (Table 4; Fig. 7). The primers and information for the corresponding loci are listed in Table S13. Comparison between the KASP genotyping results and the chip-based genotypes showed consistencies of 0.96–0.98. *T*-tests in different environments confirmed the effectiveness of the KASP markers (Fig. S10).

Discussion

Comparison of GWAS results by MLM and FarmCPU

In the present study, 19 and 20 stable WSC loci identified by MLM and FarmCPU, respectively, were located at similar positions with reported WSC-related QTL or genes (Table S11); 6, 7 and 8 loci for WSC10, WSC20 and WSC30, respectively, were detected by both MLM and FarmCPU (Table S11). All these indicated the reliability of GWAS results from both models. Some different results were also observed in two models. The WSC20 loci on chromosomes 1D (414.59–420.74 Mb), 3A (737.44–744.29 Mb) and 6B (674.84–677.48 Mb) were detected by FarmCPU in three or more environments but not stably identified by MLM, whereas the WSC10 loci on 1B (236.96–237.38 Mb) and 4A (735.30–737.58 Mb) were specifically detected by MLM (Table S11).

Liu et al. (2016) indicated that the $Q + K$ MLM may lead to false negatives in some cases, although it generally performs well in controlling false positives, while FarmCPU could control false positives and simultaneously avoid model over-fitting by using fixed and random effect models iteratively. In the present study, the Q–Q plots showed that the population structure and kinship were over-corrected by MLM in some environments, especially for WSC20 and

Fig. 4 Cumulative effects of increasing-effect alleles on stem WSC contents and thousand-kernel weight based on the 13 pleiotropic loci. Calculations were based on representative markers for the 13 pleiotropic loci simultaneously affecting stem WSC contents and TKW (Table 4). Best linear unbiased estimations (BLUE) for phenotypic data across four environments were used. WSC10, WSC20 and WSC30 indicate stem WSC contents at 10, 20 and 30 days post-anthesis, respectively. TKW, thousand-kernel weight
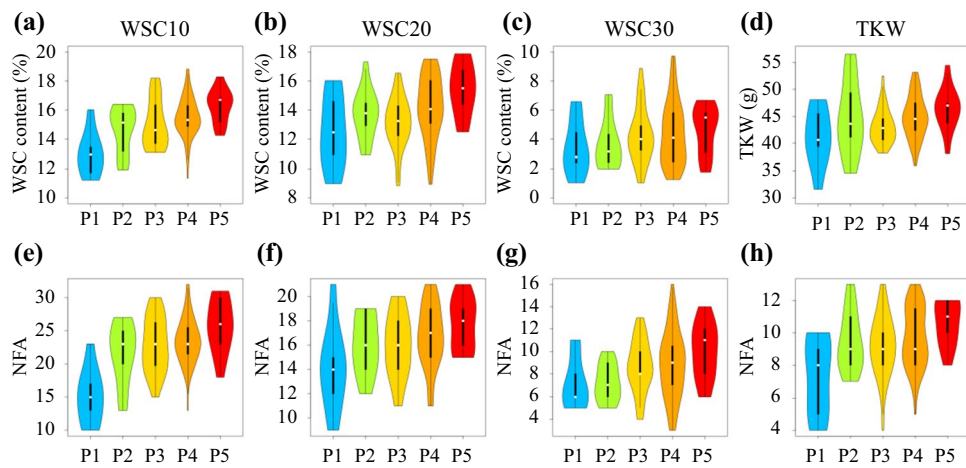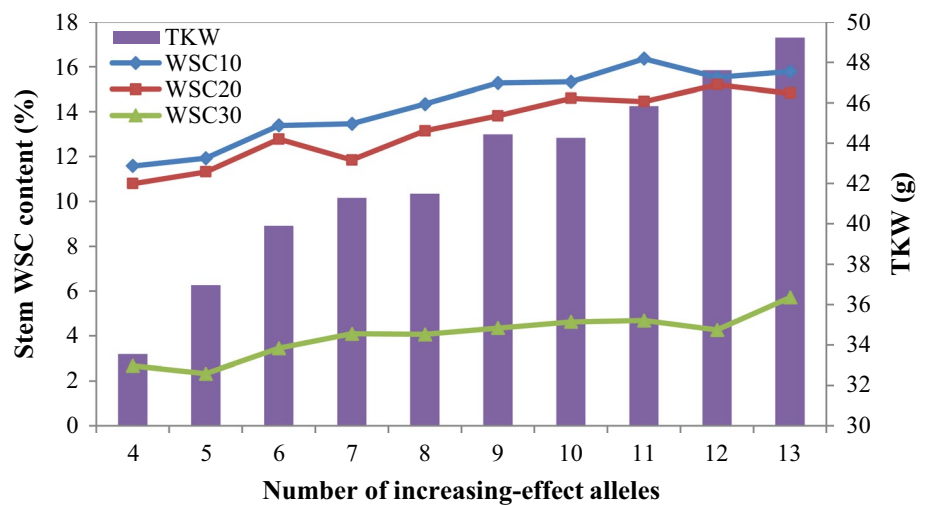
Fig. 5 Genetic progresses of stem WSC contents and thousand-kernel weight over past decades. Genetic progress was based on five groups from a panel of 130 Chinese wheat cultivars released from 1947 to 2016 (Table S1). WSC10, WSC20 and WSC30 indicate the stem WSC contents at 10 days post-anthesis (10DPA), 20DPA and 30DPA, respectively. TKW, thousand-kernel weight. P1, P2, P3, P4 and P5 indicate 1947–1979, the 1980s, the 1990s, the 2000s and 2010–2016, respectively. Violin plots **a** to **d**, phenotypic changes in WSC10, WSC20, WSC30 and TKW, respectively; **e** to **h**, changes in numbers of increasing-effect alleles for WSC10, WSC20, WSC30 and TKW, respectively. Each violin plot shows the median (indicated by the small, white dot), first through third interquartile range (black, thick, solid vertical band), and estimator of the density (thin vertical curves) of the corresponding observations. NFA indicates the number of increasing-effect alleles contained in a cultivar

WSC30 (Figs. S6, S7), indicating that there might be some false negatives for MLM. In addition, when we focused on the WSC loci detected in at least three environments, more loci were observed by FarmCPU (Table 3).

To further compare the results from MLM and Farm-CPU, LM fitting analysis was conducted based on representative markers identified by each model. For WSC10, the GWAS performance of MLM and FarmCPU was similar with coefficients of determination (r^2) of 0.76 and 0.72, respectively (Fig. S11). While for WSC20 and WSC30, FarmCPU performs better with $r^2 = 0.56$ and 0.58, respectively, than MLM with r^2 values of 0.43 and 0.34, respectively (Fig. S11). In addition, LM fitting analyses using all the identified representative markers from MLM and Farm-CPU (Table S11) gave the best results with $r^2 = 0.86$, 0.62 and 0.58 or WSC10, WSC20 and WSC30, respectively (Fig. S11). All these indicated that the detection powers of MLM and FarmCPU may vary according to different traits, and the two models are complementary in detecting QTL.

Additionally, although many loci were detected by both FarmCPU and MLM (Table S11), it seems that the Manhattan plots corresponding to the two models are quite different (Figs. S6, S7). Some single significant SNPs hang on the plots for FarmCPU results, while strings of SNPs clustered as peaks in the plots for MLM. This may be attributed to different methodologies for the two methods (Liu et al. 2016).
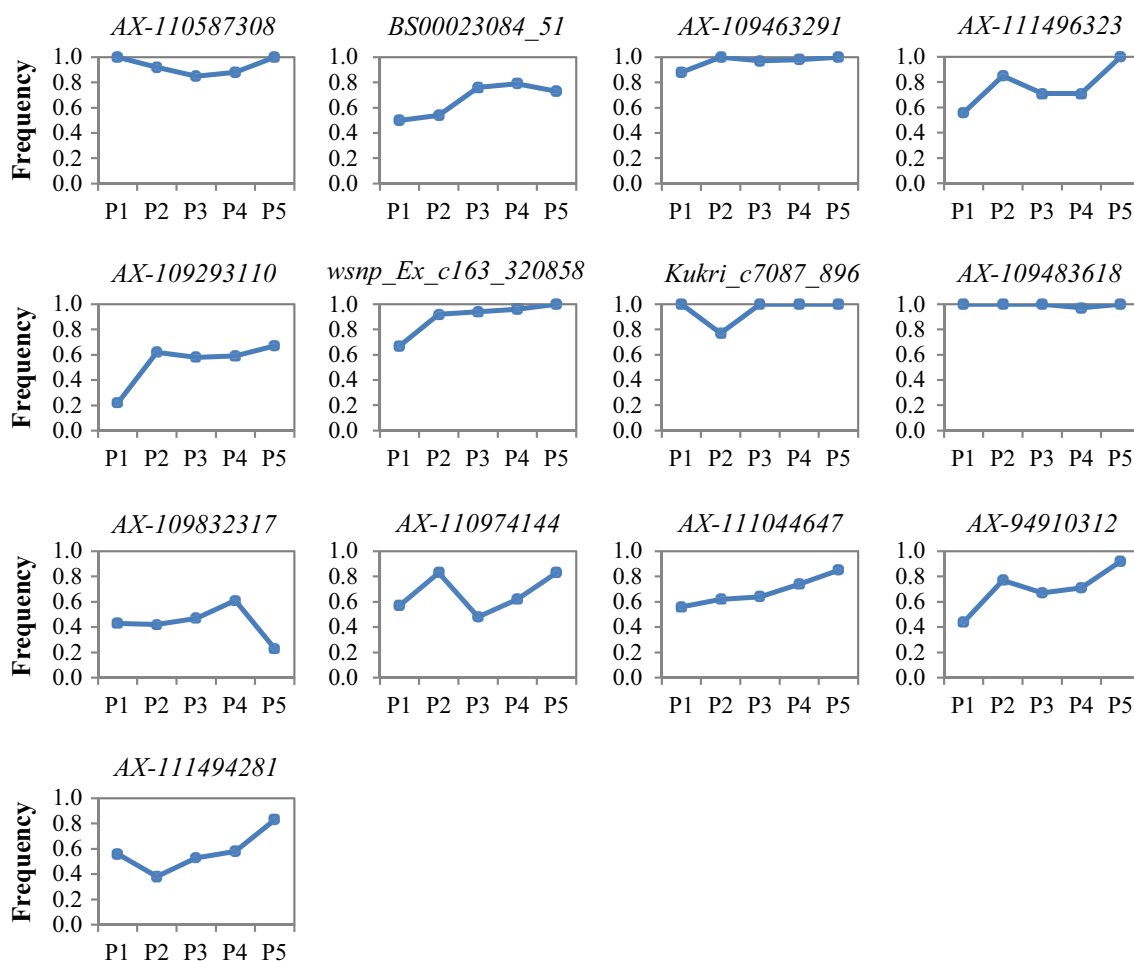
Fig. 6 Frequency changes in increasing-effect alleles in past decades based on representative markers of the 13 pleiotropic loci. The 13 pleiotropic loci were associated with both WSC and thousand-kernel weight (Table 4). The analysis was based on 130 wheat cultivars in five groups released from 1947 to 2016 (Table S1). P1, P2, P3, P4 and P5 indicate 1947–1979, the 1980s, the 1990s, the 2000s and 2010–2016, respectively. The name of each representative marker is indicated in the corresponding chart

## Comparison of identified WSC loci with reported QTL or genes

Based on the physical positions (CS RefSeq v1.0; IWGSC 2018) of markers or genes, the WSC loci identified in this study were compared with reported QTL or genes. Of the 62 WSC loci, 26 were located at similar positions to previously reported WSC-related QTL or genes (Table S11), including 15 (41.7%), 9 (37.5%) and 8 (42.1%) stable loci for WSC10, WSC20 and WSC30, respectively. This indicated the importance of these loci, reflecting the reliability of our findings. The remaining 36 new WSC loci, comprising 21, 15 and 11 loci for WSC10, WSC20 and WSC30, respectively (Table S11), provided us a basis to more comprehensively understand the complex genetic architecture underlying SWSCC.

It should be mentioned that some loci might be related to WSC content in different parts of stem, and some were associated with SWSCC under diverse conditions. For instance, a locus on 2DS (*Xcfd53*, 23.02 Mb) was associated with WSC in the uppermost internode at 14DPA under drought stress conditions (Zhang et al. 2014), whereas in the present study, a similar locus (29.17 Mb) was associated with WSC content of the whole stem (Table S11). A locus on 3AL (*Xbarc314*, 712.49 Mb) was associated with WSC in the lower internodes at the grain filling stage under simulated terminal drought stress conditions (Zhang et al. 2014), in agreement with the WSC10- and WSC20-associated locus (711.30–726.13 Mb) identified in the present study (Table S11).

Although diverse bi-parental and natural populations were used across the world, many WSC-associated loci identified in this study were located at similar positions to the previously reported ones based on the physical positions (CS RefSeq v1.0; IWGSC 2018) of QTL-flanking or associated markers (Table S11). This indicated that (a) linkage

Fig. 7 Genotype calling screenshots of five KASP markers. The KASP markers were developed based on representative SNP markers for five pleiotropic loci on both stem WSC contents (SWSCC) and thousand-kernel weight (TKW) (Tables 4, S11). Corresponding names for the KASP markers are presented on screenshots. Two homozygous genotypes are indicated beside the corresponding dots, and favorable genotypes showed increasing-effects on SWSCC and TKW (Fig. S10) are shown in red font (color figure online)
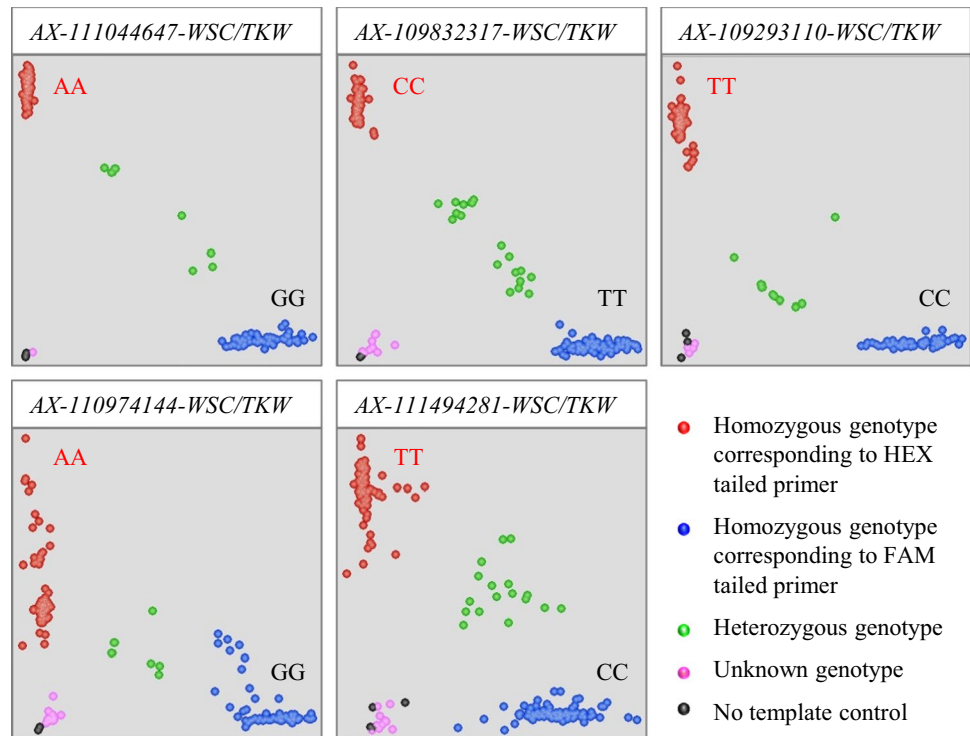
mapping and GWAS are complementary in identifying genes, and (b) many major genes controlling carbohydrate metabolism might be common in different wheat production regions. Considering the high marker density in the present study and consistency of many loci between the present and past researches, the loci and associated SNP markers detected here appear to be reliable and hence valuable for further genetic research on WSC metabolism and for MAS in breeding.

Putative candidate genes related to stem WSC content

Fructans are major components of WSC (Ruuska et al. 2006), and genes involved in fructan metabolism play important roles in controlling SWSCC. In regions adjacent to the WSC loci on chromosomes 2A (763.02–769.51 Mb), 4A (735.30–737.58 Mb) and 7D (2.23–17.34 Mb), some known genes involved in fructan synthesis and hydrolysis were identified, including *1-SST*, *6-SFT*, *1-FFT* and *6-FEH* (Table S11; McIntyre et al. 2011). SWSCC was significantly associated with *1-FFT-A1* (Yue et al. 2017) and *1-SST-D1* (Dong et al. 2016b), and the corresponding gene-specific markers had been reported. In addition to the genes directly involved in carbohydrate metabolism, transcription factors also play important roles in modulating metabolic pathways. *TaMYB13* was identified as a transcriptional activator of fructosyltransferase genes, and expression levels of *TaMYB13* were positively correlated with the mRNA levels

of *1-SST* and *6-SFT* in wheat stems (Xue et al. 2011; Kooiker et al. 2013). Based on the genomic DNA sequences, we found that *TaMYB13-1* (ID: TraesCS3A02G535100, 746.63 Mb) was located near the WSC20 locus on 3A (737.44–744.29 Mb).

Stem reserved WSC can be remobilized and transported to sink organs such as seeds and roots, and sugar transporters are critical in these processes (Halford et al. 2011). Sucrose transporters (SUT) and hexose and sucrose transporters are major sugar transporter families (Eom et al. 2015). *TaSUT1* genes on homoeologous group 4 chromosomes were the predominant sucrose transporter group in wheat stems and leaf sheaths, and significantly positive correlations were observed between TKW and expression level of *TaSUT1* genes (Ahmed et al. 2018). A *TaSUT* gene (TraesCS2A02G505000, 733.56 Mb) was identified in the vicinity of the WSC30 locus on chromosome 2A (734.16–734.26 Mb) (Table S11). In addition, two potentially new sugar transporter genes (TraesCS6B01G421500, 692.23 Mb; TraesCS7D01G521400, 618.68 Mb) belonging to the hexose and sucrose transporters family were identified in the intervals of WSC loci on 6B (692.03–702.39 Mb) and 7D (616.95–621.14 Mb) (Table S11), indicating that these genes might have important roles in WSC remobilization.

Sucrose synthase (SUS) controls carbon flow in starch biosynthesis, and dry matter accumulation of plants is positively correlated with SUS activity (Kato 1995), and *TaSus2-2B* was significantly associated with TKW (Jiang et al. 2011). Based on gene sequences, *TaSus2* (171.03 Mb)

was located in a WSC30 locus on 2B (168.78–184.77 Mb) (Table S11). A diverse family of glycosyltransferases participates in glycan and glycoside biosynthesis during plant development and growth (Lao et al. 2014), and glycoside hydrolases are involved in hydrolysis of complex sugars (Bourne and Henrissat 2001). In the present study, six potentially new glycosyltransferase candidate genes were located in the intervals of WSC loci on chromosomes 4B, 5B, 6D, 7A, 7B and 7D, and two glycoside hydrolase genes were in the intervals of WSC loci on 6B and 7A (Table S11). These candidate genes provide a basis for further exploration of the genetic mechanism of WSC metabolism, and might be useful in breeding programs targeting increased TKW.

Relationship between WSC and GY

In the present study, TKW was significantly and positively correlated with SWSCC, especially with WSC10 and WSC20 (Table 2). This was consistent with previous findings about relationship of WSC and TKW (Zhang et al. 2015; del Pozo et al. 2016). In addition, there were evidences that cultivars carrying favorable WSC alleles had higher TKW than those without favorable alleles in both well-watered and drought and heat stress conditions (Zhang et al. 2014). Similar findings were also observed in the present study that 13 pleiotropic loci simultaneously affecting SWSCC and TKW (Table 4) showed clear cumulative effects on both SWSCC and TKW (Fig. 4). Besides the 13 pleiotropic loci, other 20 WSC loci also located at similar positions to reported GY-related QTL, including 12, 5, 2 and 1 related to TKW, spike number, GY and grain filling rate, respectively (Table S11). All these indicated the positive contributions of WSC to TKW.

In many studies, significant and positive correlations ($r = 0.47–0.80$) were observed between SWSCC and GY under diverse conditions (Foulkes et al. 2007; Snape et al. 2007; Xue et al. 2008; Gao et al. 2017), whereas in some other researches, low or non-significant correlations between SWSCC and GY were observed (Ruuska et al. 2006; Zhang et al. 2015; del Pozo et al. 2016). Using 384 wheat cultivars, del Pozo et al. (2016) found that WSC content at anthesis was negatively correlated with spikes per square meter, but positively correlated with kernels per spike and TKW under water stress and full irrigation conditions; consequently, the correlation between WSC with GY were low or not significant.

By comparing the WSC QTL and reported GY-related QTL, we found some clues that might genetically explain the complicated relationship between WSC and GY. A pleiotropic locus simultaneously affecting WSC20, WSC30 and TKW was identified on chromosome 2AS (27.33–34.67 Mb) in this study (Table S11). At a similar position, a QTL for TKW and grain filling rate (*Xgwm359*, 28.20 Mb; Wang

et al. 2009) and a QTL for GY (30.3–31.9 Mb; Li et al. 2019) were reported. These indicate that the underlying gene (or genes) might control grain filling rate by affecting WSC content in stems, and finally affect the TKW and GY.

Plant density and flowering date could also affect SWSCC (Rebetzke et al. 2008; del Pozo et al. 2016). In the present study, five WSC loci were located at similar positions with reported plant density-related QTL (Table S11), indicating that SWSCC might be affected by plant density, and this also explains the complex relationship between SWSCC and GY. At the WSC10-associated locus on 5A (697.77–699.48 Mb), a vernalization gene *Vrn2-5A* (698.2 Mb) had been reported (Yan et al. 2004). To further investigate the potential impact of vernalization on stem WSC, we conducted GWAS by controlling the *Vrn2-5A* (using its genotypic data) as a fixed-effect in the $Q + K$ MLM. The results showed that 66.7% (24 out of 36) of the loci could still be detected (Table S14), including the WSC locus that collocated with *Vrn2-5A*. It seems that the effect of this WSC locus is independent from *Vrn2-5A*.

Pyramiding WSC-increasing alleles by MAS could improve GY

Significant and positive correlations between SWSCC and TKW (Table 2) were consistent with previous reports (Ehdaie et al. 2008; Li et al. 2015; Gao et al. 2017). Increased SWSCC has contributed to genetic gains in GY in the UK, Australia and China (Shearman et al. 2005; Sadras and Lawson 2011; Xiao et al. 2012; Gao et al. 2017). Similar results were evident in the present study that improvements in TKW from pre-1980 to post-2010 were accompanied by increased WSC10, WSC20 and WSC30 (Fig. 5a–d), demonstrated by increased numbers of favorable allele (Fig. 5e–h). Linear regressions also showed that pyramiding favorable alleles was effective for improving both SWSCC and TKW (Fig. 4). Furthermore, cultivars from Henan and Shandong, two high-yield wheat producing provinces in China, also had higher SWSCC and more favorable WSC alleles (Fig. S9). All these indicated that improvement in stem WSC content is a promising way to improve GY.

In addition, representative markers for pleiotropic loci on chromosomes 1BS (*BS00023084_51*), 2AS (*AX-109293110*), 4BS (*AX-110974144*), 5BL (*AX-111044647*), 6BS (*AX-94910312*) and 6BL (*AX-111494281*) exhibited FAF of 0.67–0.92 for cultivars in the 2010s (Fig. 6). These indicate that there are still opportunities for further improvement of SWSCC and TKW by pyramiding more increasing-effect alleles. High-throughput KASP markers for five pleiotropic loci were developed and validated in different environments (Fig. S10). These markers should be useful for MAS targeting improved SWSCC and TKW. Additionally, considering the important role of WSC in coping with

diverse stresses (Rebetzke et al. 2008; Livingston et al. 2009; Trouvelot et al. 2014; Ovenden et al. 2017), these markers may also have potential values in developing stress-resistant cultivars by MAS.

The present study was conducted in well-watered field conditions with the same field management as used in YHVWWZ of China, and the findings would be useful for wheat breeding in this area and similar environments. Cultivars Luyuan 502, Huaimai 20, Jinmai 61, Luomai 21, Yumai 13, Zhongmai 875, Zhongmai 895, Zhou 8425B, Zhoumai 22, Zhoumai 30, Zhoumai 31, Zhoumai 32 and Zimai 12 possess high SWSCC (WSC10 and WSC20) and TKW with high numbers of increasing-effect alleles (Tables 5, S11). These cultivars could be used as elite germplasms for improving WSC and TKW.

Advantages and disadvantages of this study

In the present study, GWAS was conducted to identify QTL associated with SWSCC based on a high density of SNP markers. The high marker density has significant impact on the number of haplotypes within trait-associated loci by change in LD pattern and would result in discovering marker-trait associations in low-recombination regions on chromosomes (Kim and Yoo 2016; Andrade et al. 2019). Previously, Dong et al. (2016a) performed GWAS on SWSCC with 18,207 markers from the 90 K SNP array using the same panel of wheat accessions as that used in the present study, but only 11 significant MTAs were identified

at two or more environments due to the presence of large gaps, particularly in the D genome (Liu et al. 2017). To resolve this problem, the present GWAS for WSC content was performed using 373,106 SNPs from the 660 K and 90 K SNP arrays. The integrated physical map showed a significantly improved marker density of 0.038 Mb per marker for the whole genome compared with 0.772 Mb per marker (14,061 Mb/18,207 markers) of the 90 K genetic map in Dong et al. (2016a). Consequently, 36, 24 and 19 stable loci were detected for WSC10, WSC20 and WSC30, respectively (Table S11). This indicated that the high-density physical map constructed here gave a significant advantage for GWAS on complex traits like SWSCC.

A relatively small population in this study might be the reason why a stringent threshold like Bonferroni correction (Holm 1979) could not be used for GWAS. Nevertheless, the 166 wheat accessions employed in this study were chosen from more than 400 cultivars, and they are good representatives of wheat germplasms from YHVWWZ of China. This population was previously used for GWAS on yield and quality-related traits, and many important loci were identified and validated (Liu et al. 2017; Zhai et al. 2018; Li et al. 2019). In the present study, 26 out of 62 WSC loci were identified at similar positions to previously reported QTL or genes (Table S11), also indicating the reliability of the results. Furthermore, based on the identified WSC loci, we observed a clear pyramiding effect of favorable WSC alleles in wheat accessions released in past decades, which was consistent with the improvement in SWSCC and TKW (Fig. 5). Therefore, the population is

Table 5 Elite germplasms for improving stem WSC contents and grain yield

Cultivar	Stem WSC content and TKW[a]			Number of favorable alleles[a,b]		
	WSC10 (%)	WSC20 (%)	TKW (g)	WSC10	WSC20	TKW
Huaimai 20	17.59	16.91	45.6	32	18	10
Jinmai 61	17.83	16.06	47.6	29	19	10
Luomai 21	16.70	15.18	49.9	29	19	12
Luyuan 502	16.73	15.83	50.5	26	18	11
Yumai 13	17.72	16.53	46.8	30	18	10
Zhongmai 875	16.69	16.74	54.5	30	21	12
Zhongmai 895	17.04	15.49	46.8	26	19	10
Zhoumai 22	17.20	16.31	51.3	32	20	12
Zhoumai 30	16.69	17.38	48.2	26	20	12
Zhoumai 31	16.47	15.91	47.2	26	20	10
Zhoumai 32	18.28	17.87	47.4	30	19	12
Zimai 12	17.92	17.03	45.6	27	21	10

[a]WSC10 and WSC20 indicate stem WSC contents at 10 and 20 days post-anthesis, respectively. TKW, thousand-kernel weight. The best linear unbiased estimations across four environments for WSC10, WSC20 and TKW are reported here. The ranges for WSC10, WSC20 and TKW were 9.41–18.81%, 8.49–17.87% and 26.9–56.6 g, respectively, in the 166 accessions (Table S1)

[b]Ranges for number of favorable alleles of WSC10, WSC20 and TKW were 6–32, 5–22 and 0–13, respectively, in the 166 accessions (Table S1, Figs. S5, S8). And the potential numbers of favorable alleles in an accession were 36 and 24 for WSC10 and WSC20, respectively (Table S11), and 13 for TKW based on the 13 pleiotropic loci in Table 4

suitable for GWAS of complex traits. In addition, the present study was conducted under well-watered conditions following the local practices in YHVWWZ, so the findings would be useful for breeding in this zone and similar environments.

Conclusion

TKW was significantly correlated with WSC10 and WSC20, indicating the important contribution of WSC to GY. We identified 62 stable loci for WSC at three grain filling time-points, with 36, 24 and 19 loci for WSC10, WSC20 and WSC30, respectively. Of these loci, 36 are potentially new, 16 affected SWSCC at two or more time-points, and 13 showed pleiotropic effects on both WSC and TKW. Linear regression analyses showed clear cumulative effects of increasing-effect alleles on SWSCC and TKW. In past decades, SWSCC and TKW of wheat cultivars in the YHVWWZ of China were significantly improved due to the pyramiding of WSC-increasing alleles. The present study showed a genome-wide genetic landscape of WSC, providing a perspective for understanding the relationship between WSC and GY. The identified WSC loci, especially the 13 pleiotropic loci for both WSC and TKW, are valuable targets for further dissection of the genetic basis underlying SWSCC and TKW. The five WSC- and TKW-associated KASP markers would be valuable tools for improving WSC and GY by MAS.

Authors' contribution statement LF performed the experiments, analyzed data, and wrote the manuscript; SY, YJ, MY and YZ participated in field trials; RJ and JW provided assistance for stem WSC content assays; JW assisted in marker development; JL and JW assisted in data analysis; YX, XX and ZH designed the experiment. YX, XX, AR and ZH assisted in writing the manuscript.

Compliance with ethical standards

References

Aggarwal PK, Sinha SK (1984) Effect of water stress on grain growth and assimilate partitioning in two cultivars of wheat contrasting in their yield stability in a drought-environment. Ann Bot 53:329–340

Ahmed SAS, Zhang J, Ma W, Dell B (2018) Contributions of TaS-UTs to grain weight in wheat under drought. Plant Mol Biol 98:333–347

Andrade ACB, Viana JMS, Pereira HD, Pinto VB, Fonseca e Silva F (2019) Linkage disequilibrium and haplotype block patterns in popcorn populations. PLoS ONE 14:e0219417

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300

Bian Y, Gu X, Sun D, Wang Y, Yin Z, Deng D, Wang Y, Li G (2015) Mapping dynamic QTL of stalk sugar content at different growth stages in maize. Euphytica 205:85–94

Bidinger F, Musgrave RB, Fischer RA (1977) Contribution of stored pre-anthesis assimilate to grain yield in wheat and barley. Nature 270:431–433

Blum A, Sinmena B, Mayer J, Golan G, Shpiler L (1994) Stem reserve mobilisation supports wheat-grain filling under heat stress. Aust J Plant Physiol 21:771–781

Bolouri-Moghaddam MR, van den Ende W (2013) Sweet immunity in the plant circadian regulatory network. J Exp Bot 64:1439–1449

Bourne Y, Henrissat B (2001) Glycoside hydrolases and glycosyltransferases: families and functional modules. Curr Opin Struct Biol 11:593–600

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Brenton ZW, Cooper EA, Myers MT, Boyles RE, Shakoor N, Zielinski KJ et al (2016) A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. Genetics 204:21–33

Chao S, Dubcovsky J, Dvorak J, Luo MC, Baenziger SP, Matnyazov R et al (2010) Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (*Triticum aestivum* L.). BMC Genom 11:727

Chen X, Min D, Yasir TA, Hu YG (2012) Genetic diversity, population structure and linkage disequilibrium in elite Chinese winter wheat investigated with SSR markers. PLoS ONE 7:e44510

Cui F, Zhang N, Fan XL, Zhang W, Zhao CH, Yang LJ et al (2017) Utilization of a Wheat660K SNP array-derived high-density genetic map for high-resolution mapping of a major QTL for kernel number. Sci Rep 7:3788

del Pozo A, Yáñez A, Matus IA, Tapia G, Castillo D, Sanchez-Jardón L, Araus JL (2016) Physiological traits associated with wheat yield potential and performance under water-stress in a Mediterranean environment. Front Plant Sci 7:987

Dong Y, Liu J, Zhang Y, Geng H, Rasheed A, Xiao Y et al (2016a) Genome-wide association of stem water soluble carbohydrates in bread wheat. PLoS ONE 11:e0164293

Dong Y, Zhang Y, Xiao Y, Yan J, Liu J, Wen W et al (2016b) Cloning of *TaSST* genes associated with water soluble carbohydrate content in bread wheat stems and development of a functional marker. Theor Appl Genet 129:1061–1070

Ehdaie B, Alloush GA, Waines JG (2008) Genotypic variation in linear rate of grain growth and contribution of stem reserves to grain yield in wheat. Field Crops Res 106:34–43

Eom JS, Chen LQ, Sosso D, Julius BT, Lin IW, Qu XQ et al (2015) SWEETs, transporters for intracellular and intercellular sugar translocation. Curr Opin Plant Biol 25:53–62

Foulkes MJ, Scott RK, Sylvester-Bradley R (2002) The ability of wheat cultivars to withstand drought in UK conditions: formation of grain yield. J Agr Sci 138:153–169

Foulkes MJ, Sylvester-Bradley R, Weightman R, Snape JW (2007) Identifying physiological traits associated with improved drought resistance in winter wheat. Field Crops Res 103:11–24

Foulkes MJ, Slafer GA, Davies WJ, Berry PM, Sylvester-Bradley R, Martre P et al (2010) Raising yield potential of wheat. III.

Optimizing partitioning to grain while maintaining lodging resistance. J Exp Bot 62:469–486

Gao F, Ma D, Yin G, Rasheed A, Dong Y, Xiao Y et al (2017) Genetic progress in grain yield and physiological traits in Chinese wheat cultivars of Southern Yellow and Huai Valley since 1950. Crop Sci 57:760–773

Goggin DE, Setter TL (2004) Fructosyltransferase activity and fructan accumulation during development in wheat exposed to terminal drought. Funct Plant Biol 31:11–21

Halford NG, Curtis TY, Muttucumaru N, Postles J, Mottram DS (2011) Sugars in crop plants. Ann Appl Biol 158:1–25

Hao C, Wang L, Ge H, Dong Y, Zhang X (2011) Genetic diversity and linkage disequilibrium in Chinese bread wheat (*Triticum aestivum* L.) revealed by SSR markers. PLoS ONE 6:e17279

Holland JB, Nyquist WE, Cervantes-Martínez CT (2003) Estimating and interpreting heritability for plant breeding: an update. Plant Breed Rev 22:11–112

Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70

Hou J, Huang X, Sun W, Du C, Wang C, Xie Y et al (2018) Accumulation of water-soluble carbohydrates and gene expression in wheat stems correlates with drought resistance. J Plant Physiol 231:182–191

IWGSC (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361:eaar7191

Jiang Q, Hou J, Hao C, Wang L, Ge H, Dong Y, Zhang X (2011) The wheat (*T. aestivum*) sucrose synthase 2 gene (*TaSus2*) active in endosperm development is associated with yield traits. Funct Integr Genomics 11:49–61

Kato T (1995) Change of sucrose synthase activity in developing endosperm of rice cultivars. Crop Sci 35:827–831

Kim SA, Yoo YJ (2016) Effects of single nucleotide polymorphism marker density on haplotype block partition. Genomics Inform 14:196–204

Kobata T, Palta JA, Turner NC (1992) Rate of development of postanthesis water deficits and grain filling of spring wheat. Crop Sci 32:1238–1242

Kooiker M, Drenth J, Glassop D, McIntyre CL, Xue GP (2013) TaMYB13-1, a R2R3 MYB transcription factor, regulates the fructan synthetic pathway and contributes to enhanced fructan accumulation in bread wheat. J Exp Bot 64:3681–3696

Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9:29

Lao J, Oikawa A, Bromley JR, McInerney P, Suttangkakul A, Smith-Moritz AM et al (2014) The plant glycosyltransferase clone collection for functional genomics. Plant J 79:517–529

Li H, Ye G, Wang J (2007) A modified algorithm for the improvement of composite interval mapping. Genetics 175:361–374

Li W, Zhang B, Li R, Chang X, Jing R (2015) Favorable alleles for stem water-soluble carbohydrates identified by association analysis contribute to grain weight under drought stress conditions in wheat. PLoS ONE 10:e0119438

Li F, Wen W, Liu J, Zhang Y, Cao S, He Z et al (2019) Genetic architecture of grain yield in bread wheat based on genome-wide association studies. BMC Plant Biol 19:168

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Liu X, Huang M, Fan B, Buckler ES, Zhang Z (2016) Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. PLoS Genet 12:e1005767

Liu J, He Z, Rasheed A, Wen W, Yan J, Zhang P et al (2017) Genome-wide association mapping of black point reaction in common wheat (*Triticum aestivum* L.). BMC Plant Biol 17:220

Livingston DP, Hincha DK, Heyer AG (2009) Fructan and its relationship to abiotic stress tolerance in plants. Cell Mol Life Sci 66:2007–2023

Lopes MS, Reynolds MP, Manes Y, Singh RP, Crossa J, Braun HJ (2012) Genetic yield gains and changes in associated traits of CIMMYT spring bread wheat in a "historic" set representing 30 years of breeding. Crop Sci 52:1123–1131

McIntyre CL, Casu RE, Rattey A, Dreccer MF, Kam JW, van Herwaarden AF et al (2011) Linked gene networks involved in nitrogen and carbon metabolism and levels of water-soluble carbohydrate accumulation in wheat stems. Funct Integr Genomics 11:585–597

Muqaddasi QH, Zhao Y, Rodemann B, Plieske J, Ganal MW, Röder MS (2019) Genome-wide association mapping and prediction of adult stage *Septoria tritici* blotch infection in European winter wheat via high-density marker arrays. Plant Genome 12:1

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucl Acids Res 8:4321–4326

Nagata K, Shimizu H, Terao T (2002) Quantitative trait loci for non-structural carbohydrate accumulation in leaf sheaths and culms of rice (*Oryza sativa* L.) and their effects on grain filling. Breed Sci 52:275–283

Ovenden B, Milgate A, Lisle C, Wade LJ, Rebetzke GJ, Holland JB (2017) Selection for water-soluble carbohydrate accumulation and investigation of genetic × environment interactions in an elite wheat breeding population. Theor Appl Genet 130:2445–2461

Pheloung PC, Siddique KHM (1991) Contribution of stem dry matter to grain yield in wheat cultivars. Funct Plant Biol 18:53–64

Phung HD, Sugiura D, Sunohara H, Makihara D, Kondo M, Nishiuchi S, Doi K (2019) QTL analysis for carbon assimilate translocation-related traits during maturity in rice (*Oryza sativa* L.). Breeding Sci 69:289–296

Platten JD, Cobb JN, Zantua RE (2019) Criteria for evaluating molecular markers: comprehensive quality metrics to improve marker-assisted selection. PLoS ONE 14:e0210529

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rahimi Y, Bihamta MR, Taleei A, Alipour H, Ingvarsson PK (2019) Genome-wide association study of agronomic traits in bread wheat reveals novel putative alleles for future breeding programs. BMC Plant Biol 19:541

Rebetzke GJ, van Herwaarden AF, Jenkins C, Weiss M, Lewis D, Ruuska S et al (2008) Quantitative trait loci for water-soluble carbohydrates and associations with agronomic traits in wheat. Aust J Agr Res 59:891–905

Rolland F, Baena-Gonzalez E, Sheen J (2006) Sugar sensing and signaling in plants: conserved and novel mechanisms. Annu Rev Plant Biol 57:675–709

Ruuska SA, Rebetzke GJ, van Herwaarden AF, Richards RA, Fettell NA, Tabe L, Jenkins CL (2006) Genotypic variation in water-soluble carbohydrate accumulation in wheat. Funct Plant Biol 33:799–809

Ruuska SA, Lewis DC, Kennedy G, Furbank RT, Jenkins CL, Tabe LM (2008) Large scale transcriptome analysis of the effects of nitrogen nutrition on accumulation of stem carbohydrate reserves in reproductive stage wheat. Plant Mol Biol 66:15–32

Sadras VO, Lawson C (2011) Genetic gain in yield and associated changes in phenotype, trait plasticity and competitive ability of South Australian wheat varieties released between 1958 and 2007. Crop Pasture Sci 62:533–549

Sharma RC, Crossa J, Velu G, Huerta-Espino J, Vargas M, Payne TS, Singh RP (2012) Genetic gains for grain yield in CIMMYT spring bread wheat across international environments. Crop Sci 52:1522–1533

Shearman VJ, Sylvester-Bradley R, Scott RK, Foulkes MJ (2005) Physiological processes associated with wheat yield progress in the UK. Crop Sci 45:175–185

Snape JW, Foulkes MJ, Simmonds J, Leverington M, Fish LJ, Wang Y, Ciavarrella M (2007) Dissecting gene × environmental effects

on wheat yields via QTL and physiological analysis. Euphytica 154:401–408

Teulat B, Borries C, This D (2001) New QTLs identified for plant water status, water-soluble carbohydrate and osmotic adjustment in a barley population grown in a growth-chamber under two water regimes. Theor Appl Genet 103:161–170

Thévenot C, Simond-Côte E, Reyss A, Manicacci D, Trouverie J, Le Guilloux M et al (2005) QTLs for enzyme activities and soluble carbohydrates involved in starch accumulation during grain filling in maize. J Exp Bot 56:945–958

Trouvelot S, Héloir MC, Poinssot B, Gauthier A, Paris F, Guillier C et al (2014) Carbohydrates in plant immunity and plant protection: roles and potential application as foliar sprays. Front Plant Sci 5:592

Turner LB, Cairns AJ, Armstead IP, Ashton J, Skøt K, Whittaker D, Humphreys MO (2006) Dissecting the regulation of fructan metabolism in perennial ryegrass (*Lolium perenne*) with quantitative trait locus mapping. New Phytol 169:45–58

Veenstra LD, Jannink JL, Sorrells ME (2017) Wheat fructans: a potential breeding target for nutritionally improved, climate-resilient varieties. Crop Sci 57:1624–1640

Wang RX, Hai L, Zhang XY, You GX, Yan CS, Xiao SH (2009) QTL mapping for grain filling rate and yield-related traits in RILs of the Chinese winter wheat population Heshangmai × Yu8679. Theor Appl Genet 118:313–325

Wang Z, Liu X, Li R, Chang X, Jing R (2011) Development of near-infrared reflectance spectroscopy models for quantitative determination of water-soluble carbohydrate content in wheat stem and glume. Anal Lett 44:2478–2490

Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE et al (2014) Characterization of polyploid wheat genomic diversity using a high-density 90000 single nucleotide polymorphism array. Plant Biotechnol J 12:787–796

Wang DR, Han R, Wolfrum EJ, McCouch SR (2017) The buffering capacity of stems: genetic architecture of nonstructural carbohydrates in cultivated Asian rice, *Oryza sativa*. New Phytol 215:658–671

Wardlaw IF, Willenbrink J (2000) Mobilization of fructan reserves and changes in enzyme activities in wheat stems correlate with water stress during kernel filling. New Phytol 148:413–422

Xiao YG, Qian ZG, Wu K, Liu JJ, Xia XC, Ji WQ, He ZH (2012) Genetic gains in grain yield and physiological traits of winter wheat in Shandong Province, China, from 1969 to 2006. Crop Sci 52:44–56

Xu S (2003) Theoretical basis of the Beavis effect. Genetics 165:2259–2268

Xue GP, McIntyre CL, Jenkins CLD, Glassop D, Herwaarden AF, Shorter R (2008) Molecular dissection of variation in carbohydrate metabolism related to water soluble carbohydrate accumulation in stems of wheat (*Triticum aestivum* L.). Plant Physiol 146:441–454

Xue GP, Kooiker M, Drenth J, McIntyre CL (2011) TaMYB13 is a transcriptional activator of fructosyltransferase genes involved in β-2,6-linked fructan synthesis in wheat. Plant J 68:857–870

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, San-Miguel P et al (2004) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. Science 303:1640–1644

Yáñez A, Tapia G, Guerra F, del Pozo A (2017) Stem carbohydrate dynamics and expression of genes involved in fructan accumulation and remobilization during grain growth in wheat (*Triticum aestivum* L.) genotypes with contrasting tolerance to water stress. PLoS ONE 12:e0177667

Yang DL, Jing RL, Chang XP, Li W (2007) Identification of quantitative trait loci and environmental interactions for accumulation and remobilization of water-soluble carbohydrates in wheat (*Triticum aestivum* L.) stems. Genetics 176:571–584

Yin C, Li H, Li S, Xu L, Zhao Z, Wang J (2015) Genetic dissection on rice grain shape by the two-dimensional image analysis in one *japonica × indica* population consisting of recombinant inbred lines. Theor Appl Genet 128:1969–1986

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

Yue AQ, Ang LI, Mao XG, Chang XP, Li RZ, Jing RL (2017) Single-nucleotide polymorphisms, mapping and association analysis of *1-FFT-A1* gene in wheat. J Integr Agr 16:789–799

Zhai S, Liu J, Xu D, Wen W, Yan J, Zhang P et al (2018) A Genome-wide association study reveals a rich genetic architecture of flour color-related traits in bread wheat. Front Plant Sci 9:1136

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA et al (2010) Mixed linear model approach adapted for genome-wide association studies. Nat Genet 42:355–360

Zhang B, Li W, Chang X, Li R, Jing R (2014) Effects of favorable alleles for water-soluble carbohydrates at grain filling on grain weight under drought and heat stresses in wheat. PLoS ONE 9:e102917

Zhang J, Xu Y, Chen W, Dell B, Vergauwen R, Biddulph B et al (2015) A wheat *1-FEH w3* variant underlies enzyme activity for stem WSC remobilization to grain under drought. New Phytol 205:293–305