

Genome-wide association studies on heat stress tolerance during grain development in wheat (*Triticum aestivum* L.)

Jie Guo

Shanxi Agricultural University

Weiping Shi

Shanxi Agricultural University

Jiahui Guo

Shanxi Agricultural University

Linqi Yue

Shanxi Agricultural University

Lei Zhuang

Shanxi Agricultural University

Wenying Zhang

Yangtze University

Xiaohai Tian

Yangtze University

Yanhao Xu (✉ xyh09@yangtzeu.edu.cn)

Yangtze University

Meixue Zhou

University of Tasmania

Research article

Keywords: Common wheat, genome-wide association analysis, selective sweeps, heat stress; agronomic and quality traits

Posted Date: January 9th, 2020

DOI: <https://doi.org/10.21203/rs.2.20460/v1>

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

Abstract

Background

Heat stress at the late reproductive stages is a common problem encountered in autumn-sown wheat production regions in China with the affected area covering as much as two-thirds of the crop. In order to develop wheat cultivars with heat-tolerance, it is crucial to explore favorable alleles for use in breeding programs.

Results

In this study, we performed a 90K iSelect SNP genotyping assay on a collection of 207 wheat cultivars subjected to heat stress during grain-fill growth stage in three years (2015-2017). Genotypic analyses of 19 phenotypic traits revealed that heat stress had major impacts on grain weight, size, and quality. Correlation analyses indicated that thousand kernel weight (TKW) was significantly correlated with grain width (GW) and grain perimeter (GP), whereas grain protein content (GPC) was negatively correlated with total starch content (TSC) ($P < 0.01$). We applied heat susceptibility indices (HSI) for different traits to assess heat tolerance. Genome-wide association studies (GWAS) revealed a total of 125 marker-trait associations (MTAs) at 63 SNP loci on 16 chromosomes each accounting for phenotypic variation (R^2) of 3.0-21.4%. 17 loci showed significant associations in three environments. The analysis of selective sweeps indicated that *RAC875_c19042_2102* (2B), *wsnp_Ex_c257_491667* (3B), *wsnp_Ex_rep_c101323_86702413* (5A) and *BS00061911_51* (7A) were selected between two subpopulations (top 5%).

Conclusions

These four key MTAs detected in the present study are candidates for further genetic dissection and development of molecular markers.

Background

Climate change is leading not only to rising global mean temperatures but also more frequent occurrences of extreme weather events. The average temperature of the earth-surface increased by approximately 0.85°C from 1880 to 2012 and is predicted to continue to increase [1]. Wheat, the third largest grain crop, provides food for 60% of the world population [2]. High temperatures during the late reproductive growth stage has become a major factor affecting wheat production and quality [3]. It is predicted that each °C of further increase in global mean temperature will cause approximate 6% decrease in global wheat production [4]. Heat stress at the late reproductive stages is a common problem encountered in autumn-sown wheat production regions in China with the affected area covering as much as two-thirds of the crop [5, 6]. In order to develop wheat cultivars with heat-tolerance, it is crucial to explore favorable alleles for use in breeding programs.

Heat stress has different impacts on wheat during different developmental stages. During the grain filling stage, with average temperatures ranging from 15.8 to 27.7°C, each °C increase, shortens the grain filling period by approximately 3.1 d, and reduces thousand-grain weight by 2.8g [7-9]. Accumulation of starch and protein during grain filling is affected when the day temperature exceeds 30°C. Starch synthesis is hypersensitive to heat, and the relative protein content is significantly higher in heat-sensitive cultivars (0.22% / d) than in heat-tolerant cultivars (0.13% / d) [10]. Under heat stress, the supply of substrate (sucrose) in starch synthesis is not affected; the reduced rate of grain filling rate is mainly due to suppressed activity of starch biosynthetic enzymes, such as soluble starch synthase, ADP-glucose-pyrophosphate transferase, granule-bound starch synthase, and starch branch enzyme, that reduces the rate of transformation from sucrose to starch [11].

In recent years various studies on heat stress tolerance have addressed the effects of crown temperature [12, 13], grain-fill duration [14], chlorophyll content and chlorophyll fluorescence parameters [15-17], and heat sensitivity indices of yield-related traits [18-20]. Many QTL have been identified for heat resistance and these QTL are distributed on almost all chromosomes. Among them, QTL on homoeologous group 3 chromosomes were identified in multiple populations [13, 15, 18, 21, 22]. Using a DH population from a cross between RAC875 and Kukri, Bennett et al. [13] identified two major QTL on chromosome 3B with the one on the long arm explaining 22% of the phenotypic variation of the yield under heat stress. Using chlorophyll content and chlorophyll fluorescence parameters as indices Vijayalakshmi et al. [15] identified heat stress-related QTL on chromosomes 3A and 3B in a RIL population derived from a cross between heat-tolerant winter wheat cultivar Ventnor and heat-sensitive cultivar Karl92.

Artificial selections contributed to the major differences in population structure and traits between cultivated and wild species and between modern varieties and landrace varieties [23]. Relatively few alleles are conserved in genomic regions controlling important selected traits in a range of important crops resulting in significantly lower genetic diversity at some loci than the average genome-wide level; examples include the *fw2.2* in tomato [24], *tga1* in maize [25] and *sh4* in rice [26]. Selection of specific alleles at these loci will further reduce the genetic diversity in closely flanking regions, known as the selective sweep or linkage drag [27]. Lukens and Doebley [28] analyzed genes controlling branching (*tb1*) and yellow endosperm color (*y1*) in maize and found that selection at these two loci during domestication significantly reduced the genetic diversity in flanking regions. Genomic regions with low genetic diversity may be adjacent to genes or principal QTLs controlling important agronomical traits [29]. During evolution, vital traits generally have strong linkage effects, whereas other traits have small effects. Several important genes controlling such traits have been cloned using canonical map-based cloning approaches, with examples including *PROG1* in rice [30], *GIF1* in rice [31], and *sh1* in different cereals [32].

In this study, we used heat susceptibility indices of a range of traits to 1) screen a set of wheat accessions for heat tolerance; and 2) perform GWAS (genome-wide association studies) to identify QTL for heat tolerance. We also combined the nucleotide diversity (π) and F_{ST} parameters to assess

selective sweeps that helped to confirm marker trait associations.

Results

Heritability for 19 traits under natural condition

The broad sense heritabilities of all traits grown under natural condition were calculated that varied from 0.57 for KNPS to 0.78 for DTH (Table 1). There were significant differences ($P<0.01$ or $P<0.001$) among genotypes (G), years (E) and G×E interactions for all 19 traits in three years (Table 1).

Effects of heat stress on grain development

Heat stress caused significant decreases in TKW, GW, GP, GR, GAR, TSC, PV, TV, BD, FV and SB, and significant increases in RLW, GPC, PT and PTP; KNPS and GL were not affected (Fig. 1). GW was significantly correlated with grain morphology traits GL, RLW, GP, GR, and GAR ($P<0.01$). TKW showed a significant positive correlation with GW and GP under both control ($r=0.64^{**}$, $r=0.46^{**}$ for GW and GP, respectively) and heat stress conditions ($r=0.25^{**}$, $r=0.36^{**}$ for GW and GP, respectively). GPC and TSC showed a significant negative correlation under both control ($r=-0.81^{**}$) and heat stressed ($r=-0.77^{**}$) conditions (Fig. 2a, b). Most parameters for pasting properties were significantly correlated (Fig. 2c, d).

Allelic diversity and genetic structure analysis

Of the 81,587 SNP, 25,115 (30.78%) were polymorphic. The highest number of polymorphic markers was in the B genome (12,712, 50.62%), followed by the A genome (9,728, 38.73%) and D genome (2,675, 10.65%), respectively.

Genetic structure (Q value) and the kinship coefficient between individuals (K value) in the population were assessed (Fig. 3b). As shown in Fig. 3a ΔK reached its maximum at $K=2$. NJ clustering analysis showed that the accessions clustered into two groups (Q1 and Q2) (Fig. 3c). With accessions from the same ecological region generally clustering in the same group, among which Q1 group mainly consisted of accessions from the Yellow and Huai River Valleys' winter wheat zone including provinces of Hebei, Henan, Shandong and Shaanxi, Q2 from Hubei, Hunan and Jiangsu provinces belong to the Middle and Lower Yangtze Valleys' winter wheat zone. Principal coordinate analysis (PCA) of the SNPs revealed that the variation explained by the first three principal coordinates were 7.27%, 4.80% and 3.70%, respectively (Fig. 3d).

Detection of MTAs

Genome-wide association studies was performed on the 19 phenotypic traits using 25,115 markers; 125 MTAs (marker-trait associations) were detected at 63 SNP loci ($P<5.00\times10^{-5}$). The associated loci were distributed on 16 chromosomes including 1B (4.8%), 1D (5.6%), 2A (7.2%), 2B (28.8%), 3B (8.8%), 4A (6.4%), 4D (3.2%), 5A (7.2%), 5B (2.4%), 5D (1.6%), 6A (2.4%), 6B (3.2%), 6D (5.6%), 7A (7.2%), 7B (3.2%), and 7D (2.4%), with levels of explained variation (R^2) ranging from 3.00 to 21.38% (Fig. 4, Additional file 1: Table S1). Fourteen of these loci were detected in all three years.

RAC875_rep_c69546_73 (2B, TKW_HSI), *wsnp_Ex_rep_c101323_86702413* (5A, TKW_HSI), *GENE-1350_36* (2A, BD_HSI), *Excalibur_c224_1383* (2B, BD_HSI), *RAC875_c24205_110* (2B, BD_HSI), *RAC875_c19042_2102* (2B, BD_HSI) and *Kukri_c64744_1087* (4D, SB_HSI) showed significant associations in each year and in BLUP of all three years (Additional file 1: Table S1).

Selection signals between Q1 and Q2

Selective sweep analysis on subpopulations Q1 and Q2 showed significant genotype-level differences in genetic structure and cluster analyses. F_{ST} and π were used to identify selective sweeps; 165 loci reached the top 5% according to $\pi_{Q1/Q2}$ and F_{ST} analyses (Fig. 5a, b). Significantly associated loci *RAC875_c19042_2102* (2B) and *wsnp_Ex_c257_491667* (3B) detected in multiple environments were selected in the $\pi_{Q1/Q2}$ analysis (top 5%) (Fig. 5a, c), and *wsnp_Ex_rep_c101323_86702413* (5A) and *BS00061911_51* (7A) were selected in the F_{ST} analysis (top 5%) (Fig. 5b, c).

Functional annotations of candidate regions

Within the regions of most significant markers we compared the local r^2 linkage disequilibrium (LD) pattern and four haplotype blocks were identified on each of chromosome 2B, 3B, 5A and 7A (Fig. 6-I). The haplotype blocks for each chromosome were 10.2 Mb (2B, 7.9 cM), 13.4 Mb (3B, 0.6 cM), 0.2 Mb (5A, 0.5 cM), and 7.0 Mb (7A, 0.9 cM), (Fig. 6-II, 7-III). The functions of genes in the haplotype blocks were identified as shown in Fig. 6-IV and Additional file 2: Table S2.

Stable MTAs and their genetic effects

Favorable alleles of loci showing the most significant association BLUP were selected to check their actual effects on the traits. *RAC875_c19042_2102*-2B_{CC} had a major effect on BD (HSI = 1.36, $P<0.05$), *wsnp_Ex_c257_491667*-3B_{TT} had a major effect on GR (HSI = 0.28, $P<0.0001$), *wsnp_Ex_rep_c101323_86702413*-5A_{TT} had a major effect on TKW (HSI = 0.93, $P<0.0001$), and *BS00061911_51*-7A_{CC} had a major effect on TSC (HSI=0.65, $P<0.0001$) (Fig. 7).

Discussion

Impacts of heat stress on wheat grain traits

Mature grains following subjection to heat stress have a dry, shrunken shape, with reduced GW but unchanged GL [33, 34]. TKW was significantly correlated with GW ($r=0.64^{**}$, $r=0.25^{**}$) ($P<0.01$), confirming that the heat stress caused insufficient grain filling. The significant negative correlation between TKW and GPC ($r=-0.13^*$) under heat stress (Fig. 2b) confirmed that the increased grain protein content under heat stress was caused by reduced grain weight [35].

Protein and starch contents are important indices of wheat grain quality. The relative protein content and protein: starch ratio determine development time, strength, ductility, and food processing quality of flour and dough [36]. Our study detected an increase in GPC and a decrease in TSC, as well as a significant negative correlation between these traits ($r=-0.77^{**}$) under heat stress, reflecting a shortened grain-fill period and reduced starch content (Fig. 1 and 2). In addition, the gelatinizing character of starch largely affects noodle quality and is an important criterion for starch quality assessment. Of the gelatinizing features, peak viscosity (PV) is the most important measure of starch quality and noodle production. Generally, higher PV indicates better starch and noodles qualities [37]. Our study detected a reduced PV under heat stress (Fig. 1). Although high temperatures increased protein content overall grain quality was compromised. Nonetheless, the influence of heat stress at the early-mid grain development stages is very complicated, and further investigation is still needed.

Indices used to assess heat stress tolerance in wheat

Different indices have been used in assessing heat stress tolerance in wheat. These include the physiological traits crown temperature, grain fill duration, plasma membrane stability and chlorophyll content, as well as ear number and thousand-kernel weight heat sensitivity index [12, 14, 15, 16, 18, 19]. It has been evident in literature that low HSI could be a more useful index in discriminating better genotypes under drought stress condition [38, 39]. Moreover, previous studies evaluated HSI<1 of yield and its different components of wheat genotypes for heat stress tolerance and grouped them into highly tolerant, tolerant and susceptible genotypes based on their HSI values [40, 41]. In our study, we used HSI for different traits to determine the impact of heat stress on each trait and larger HSI values represent relatively more sensitivity to stress, thus smaller HSI values are favored.

MTAs for heat tolerance

Heat tolerance in wheat is a quantitative trait controlled by polygenes [12, 13, 15-17]. Many QTL for heat stress tolerance were reported and these QTL were distributed on most chromosomes [15, 19-22, 52]. Several major MTAs were identified for heat stress tolerance. A major MTAs for GAR_HSI on 1D with the closest marker *BobWhite_c359_1026* was identified in all three years, whereas a QTL for heat stress tolerance of TKW was previously detected in this region [42]. Two major MTAs for TKW_HSI on 2B and 5A were identified in all three years. The 2B MTAs for TKW_HSI also showed a significant association with TSC_HSI and pasting properties (BD_HSI), suggesting that heat stress affected starch synthesis/translocation leading to reduced TKW and BD [13, 43]. Qaseem et al. [20] located QTL for GR_HSI on chromosomes 3B, 5A and 7A that were associated with peduncle length and awn length under heat stress conditions (Additional file 1: Table S1). A QTL for GP_HSI identified on 3B contained several candidate genes related to grain yield under heat stress [20].

Discovery of important loci for heat stress tolerance based on selection

Selection for agronomic traits of interest (grain size and tiller number/angle) by breeders is likely to cause strong selective sweeps. So far, a large number of QTLs related to such traits have been characterized in rice including *GS3* [44], *GW2* [45], *GW5* [46], *DEP1* [47] and *GS5* [48] for grain size; *Ghd7* [49], and *IPA1* [50] for plant architecture. Our study combined GWAS and selective sweep analysis, and revealed four loci significantly associated with heat tolerance in multiple years, namely, *RAC875_c19042_2102* (2B), *wsnp_Ex_c257_491667* (3B), *wsnp_Ex_rep_c101323_86702413* (5A), and *BS00061911_51* (7A) (top 5%), in multiple environments were all selected (Fig. 5a, 6b). Furthermore, these loci had major effects on BD, GR, TKW and TSC, respectively (Fig. 7). All except *BS00061911_51* (7A) of these loci were protein coding and their gene functions were unknown, which need to be further studied (Fig. 6, Additional file 2: Table S2).

Conclusion

Heat stress during the grain fill period showed significant effects on grain filling. For the first time, using heat stress indices, several major QTL were identified for grain filling, starch accumulation/transportation, and flour pasting properties. The QTL for starch accumulation/transportation on 2B also showed significant effects on grain weight and flour pasting properties.

Materials And Methods

Plant materials

A total of 207 wheat accessions used in the study were collected from the Yellow and Huai River Valleys' winter wheat zone (117, among which 11 from Hebei, 45 from Henan, 25 from Shandong and 36 from Shaanxi) and the Middle and Lower Yangtze Valleys' winter wheat zone (90, including 7 from Hubei, 14 from Hunan and 69 from Jiangsu) (Additional file 3: Table S3).

Phenotyping

The 207 wheat accessions were grown in pots under natural conditions at Yangtze University in the 2015, 2016 and 2017 growing seasons. Six seeds of each accession were sown in each pot filled with 7 kg soil and 8 g mixed fertilizer (26N: 10P: 15K). Two treatments (control and heat stress) were performed with six replicates for each treatment. The split plot in completely randomized design was used in this study. Fertilization, irrigation, and disease and pest control were performed on a regular basis. For each genotype, we recorded DTH (days to heading when 50% of the individual (main stem) heading) and DTF (days to flowering, when 50% of the individual (main stem) flowering), and tagged individuals that flowered on the same day. One week after flowering, the pots for heat stress were moved to a natural light phytotron (AGC-MR, Zhejiang Qiushi Environment Co., Ltd, Zhejiang, China) with a simulated heat environment, and control pots remained outside. To keep the photoperiod and light intensity close to that of the natural environment, no artificial light source was provided in the natural light phytotron. The temperature of the artificial chamber was set at 8°C higher than the outside temperature (Additional file 4: Fig. S1). Day- and night-time humidity was set to 75% and 65% in the natural light phytotron, respectively. To avoid drought stress, an automatic drip irrigation system (arrow dripper flow rate 1L/h) was used for both the heat stress treatments and the control treatments, running for 0.5 hour every 5.5 hours. Ten days after the initiation of heat stress, all six plots were moved out of natural light phytotron for further growth and maturation conditions were similar to the control pots.

After maturation five individuals from each pot were phenotyped for KNPS (kernel number per spike) and TKW (thousand kernel weight, based on 150 seeds). A Wseen SC-G Automatic Seed Analyzer (Hangzhou Wanshen Detection Technology Co., Ltd.) was used to assess grain size parameters (150 seeds for each pot) that included GL (grain length), GW (grain width), RLW (ratio of length to width), GP (grain perimeter), GR (grain roundness) and GAR (grain area ratio).

Pasting properties of the wheat flour were determined using a Perten RVA 4500 instrument [51]. Five grams of grains were ground into wholemeal flour in a Perten LABORATORY MILL 3100; 3.5 g of the wholemeal was added to 25 ml 0.1% silver nitrate solution. Standard tests and analysis with TCW3 software were performed to estimate a range of flour parameters including PV (peak viscosity), TV (trough viscosity), BD (breakdown viscosity), FV (final viscosity), SB (setback viscosity), PT (peak time), and PTP (pasting temperature). The TSC (total starch content) and GPC (grain protein content) were obtained by near-infrared reflectance (NIR) spectroscopy (DA-7200, Perten Instruments, Huddinge, Sweden). The NIR predicted TSC and GPC were calibrated using 20 calibration samples from each treatment at each year with the chemical methods (TSC: AACC Method 76-13.01, Megazyme amyloglucosidase/alpha-amylase method. GOC: AACC Method 46-11.02, improved Kjeldahl method).

Heat tolerance was assessed as a heat susceptibility index (HSI) for each accession and parameter using the following equation: $HSI = (1-Y_h/Y)/(1-X_h/X)$, where Y_h and Y are the phenotypic means were calculated using least square method for each genotype under heat stress and control conditions, respectively, and X_h and X are the phenotypic means for all lines under heat stress and control conditions, respectively [52].

Genotyping and data analysis

Genomic DNA was extracted by the CTAB method [53]. Descriptive statistical analyses and estimated standard deviations were carried out in SPSS 21.0 (<http://www.brothersoft.com/ibm-spss-statistics-469577.html>). Best linear unbiased estimates (BLUEs) were used to represent phenotypic means of each accession. The means were used for the association analysis while residual maximum likelihood (REML) was used to analyze phenotypic data [54]. Analysis of variance of the phenotypic data was carried out using SAS 9.4 (https://www.sas.com/en_us/software/sas9). Broad sense heritabilities (h^2) were calculated as $h^2 = \sigma_g^2 / [\sigma_g^2 + (\sigma_{ge}^2/n) + (\sigma_e^2/rn)]$, where σ_g^2 was the genetic variance; σ_{ge}^2 the genotype \times year (environment) variance; σ_e^2 the error variance; r the number of replications/year, and n was the number of years [55].

SNP markers were assayed at the Seed Biotechnology Center, Department of Plant Biology, University of California at Davis, with the Illumina SNP Genotyping Platform using BeadArray [56]. Data correction, reading and quality checking were performed in Genomestudio v2011.1. SNP markers with minor allele frequency (MAF) <0.05 and missing data >10% were excluded from further analyses. Genetic mapping information for SNP markers was obtained from Wang et al. [57].

Genetic structure of the natural population was estimated based on the 25,115 SNP markers distributed on all 21 chromosomes using Structure V 2.3.2 [58], and the number of subpopulations was determined using the ΔK model [59]. Based on the Q + K model [60, 61], neighbor-joining (NJ) cluster analysis was performed using MEGA 5.0 [62].

MTAs were calculated for different traits from four phenotypic data sets (2015, 2016, 2017 and BLUP) using TASSEL 5.0 software MLM (mixed linear model) program [63] (<http://www.maizegenetics.net/>), with deletion of SNP markers with frequencies <0.05. A standard way of checking GWAS analysis is based on the assumption that only a very small proportion of the variants affect the phenotype and therefore the test statistics of the majority of the variants should follow the null distribution [64]. The association signal threshold $P(P < 5.00 \times 10^{-5}, -\text{Log } P > 4.30)$ was evaluated using TASSEL with the permutation test. To estimate the genetic effect of the stable MTAs, trait values of different genotypes were compared and significance analysis were assessed by t-test.

The methods of the nucleotide diversity (π) and population differentiation statistic (F_{ST}) were combined to detect evidence of selection [65]. According to the π method, which based on multilocus allele frequency differentiation between two populations, the genetic diversity in the Q1 group (accessions mainly collected from the Yellow and Huai River Valleys' winter wheat zone) was compared with that in the Q2 (accessions mainly from the Middle and Lower Yangtze Valleys' winter wheat zone) group named Π_{Q1}/Π_{Q2} [66]. We also calculated the population fixation statistic F_{ST} to identify selective

sweeps, population-level F_{ST} was calculated for each variant site using VCFtools' `weir-fst-pop` function [67, 68]. In addition, both methods used a 10-SNP sliding window with a step size of 1-SNP for the top 5% of ratios set as the significant threshold for further analysis [65].

A threshold of $r^2 > 0.8$ for the pairwise LD surrounding the most significant markers was used to analyze the haplotype. Linkage disequilibrium and haplotype analyses of the significant SNPs were performed with HAPLOVIEW v.4.2 (<http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview>). Sequences of SNPs flanking MTAs were subject to a BLASTN (Basic search) on the IWGSC database (<http://wheat-urgi.versailles.inra.fr/Seq-Repository>) to obtain function annotations of candidate genes.

Declarations

Authors' contributions

JG, YX and MZ conceived and conducted experiments. JG, WS, JG, LY and LZ analysed phenotypic and genomic data. WZ and XT conducted heat treatment. JG, YX and MZ wrote the manuscript.

Funding

This work was supported by grants from the National Key R&D Program of China (2017YFD0101000), National Natural Science Foundation of China (31901541), National Key R&D Program of Shanxi Province (201703D211007, 201803D221019-5), and Science & Technology Innovation Foundation of Shanxi Agricultural University (2016YJ05).

Availability of data and materials

All datasets for this study are included in the manuscript and the Supplementary information.

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] IPPC. AR5 Synthesis Report: Climate Change 2014. 2014; Available: <https://www.ipcc.ch/>
- [2] FAO. Food and Agriculture Organisation FAOSTAT database. 2012; Available: <http://faostat.fao.org/>
- [3] Prasad PVV, Djanaguiraman M. Response of floret fertility and individual grain weight of wheat to high temperature stress: sensitive stages and thresholds for temperature and duration. *Funct Plant Biol.* 2014; 41: 1261-1269.
- [4] Asseng S, Ewert F, Martre P, Rötter RP, Lobell DB, Cammarano D, Kimball BA, Ottman MJ, Wall GW, White JW, et al. Rising temperatures reduce global wheat production. *Nat Clim Change.* 2015; 5: 143-147.
- [5] Jin SB. Wheat Cultivar and Pedigree in China. Agriculture Press. 1983; Beijing (In Chinese)
- [6] Tao F, Zhang Z, Zhang S, Rötter RP. Heat stress impacts on wheat growth and yield were reduced in the Huang-Huai-Hai Plain of China in the past three decades. *Eur J Agron.* 2015; 71: 44-52.
- [7] Wardlaw IF, Sofield I, Cartwright PM. Factors limiting the rate of dry matter accumulation in the grain of wheat grown at high temperature. *Funct Plant Biol.* 1980; 7: 387-400.
- [8] Tack J, Barkley A, Nalley LL. Effect of warming temperatures on US wheat yields. *Proc Natl Acad Sci USA.* 2015; 112: 6931-6936.
- [9] Lesk C, Rowhani P, Ramankutty N. Influence of extreme weather disasters on global crop production. *Nature.* 2016; 529: 84-87.
- [10] Sofield I, Evans LT, Cook MG, Wardlaw IF. Factors influencing the rate and duration of grain filling in wheat. *Aust J Plant Physiol.* 1977; 4: 785-797.
- [11] Hurkman WJ, Mccue KF, Altenbach SB, Korn A, Tanaka CK, Kothari KM, Johnson EL, Bechtel DB, Wilson JD, Anderson OD, et al. Effect of temperature on expression of genes encoding enzymes for starch biosynthesis in developing wheat endosperm. *Plant Sci.* 2003; 164: 1-9.

- [12] Pinto RS, Reynolds MP, Matthews KL, McIntyre CL, Olivares JJ, Chapman SC. Heat and drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. *Theor Applied Genet.* 2010; 121: 1001-1021.
- [13] Bennett D, Reynolds M, Mullan D, Izanloo A, Kuchel H, Langridge P, Schurbusch T. Detection of two major grain yield QTL in bread wheat (*Triticum aestivum* L.) under heat, drought and high yield potential environments. *Theor Applied Genet.* 2012; 125: 1473-1485.
- [14] Yang J, Sears RG, Gill BS, Paulsen GM. Quantitative and molecular characterization of heat tolerance in hexaploid wheat. *Euphytica.* 2002; 26: 275-282.
- [15] Vijayalakshmi K, Fritz AK, Paulsen GM, Bai GH, Pandravada S, Gill BS. Modeling and mapping QTL for senescence-related traits in winter wheat under high temperature. *Mol Breeding.* 2010; 26: 163-175.
- [16] Talukader SK, Babar MA, Vijayalakshmi K, Poland J, Prasad PVV, Bowden R, Fritz A. Mapping QTL for the traits associated with heat tolerance in wheat (*Triticum aestivum* L.). *BMC Genet.* 2014; 15: 97.
- [17] Sukumaran S, Reynolds MP, Sansaloni C. Genome-wide association analyses identify QTL hotspots for yield and component traits in durum wheat grown under yield potential, drought, and heat stress environments. *Front Plant Sci.* 2018; 9: 81.
- [18] Mason RE, Mondal S, Beecher FW, Hays DB. Genetic loci linking improved heat tolerance in wheat (*Triticum aestivum* L.) to lower leaf and spike temperatures under controlled conditions. *Euphytica.* 2011; 180: 181-194.
- [19] Paliwal R, Roder MS, Kumar U, Srivastava JP, Joshi AK. QTL mapping of terminal heat tolerance in hexaploid wheat (*T. aestivum* L.). *Theor Applied Genet.* 2012; 125: 561-575.
- [20] Qaseem MF, Qureshi R, Muqaddasi QH, Shaheen H, Kousar R, Röder MS. Genome-wide association mapping in bread wheat subjected to independent and combined high temperature and drought stress. *PLoS One.* 2018; 13: e0199121.
- [21] Acuña-Galindo MA, Mason RE, Subramanian NK, Hays DB. Meta-analysis of wheat QTL regions associated with adaptation to drought and heat stress. *Crop Sci.* 2015; 55: 477-492.
- [22] Ogbonnaya FC, Rasheed A, Okechukwu EC, Jighly A, Makdis F, Wuletaw T, Hargas A, Uguru MI, Agbo CU. Genome-wide association study for agronomic and physiological traits in spring wheat evaluated in a range of heat prone environments. *Theor Applied Genet.* 2017; 130: 1819-1835.
- [23] Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, et al. A map of rice genome variation reveals the origin of cultivated rice. *Nature.* 2012; 490: 497-501.
- [24] Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD. fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. *Science.* 2000; 289: 85-88.
- [25] Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, Bomblies K, Lukens L, Doebley JF. The origin of the naked grains of maize. *Nature.* 2005; 436: 714-719.
- [26] Li C, Zhou A, Sang T. Rice domestication by reducing shattering. *Science.* 2006; 311: 1936-1939.
- [27] Palaisa K, Morgante M, Tingey S, Rafalski A. Long-range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep. *Proc Natl Acad Sci USA.* 2004; 101: 9885-9890.
- [28] Lukens LN, Doebley J. Molecular evolution of the teosinte branched gene among maize and related grasses. *Mol Biol Evol.* 2001; 18: 627-638.
- [29] Andolfatto P. Adaptive hitchhiking effects on genome variability. *Curr Opin Genet Dev.* 2001; 11: 635-641.
- [30] Jin J, Huang W, Gao J, Yang J, Shi M, Zhu M, Luo D, Lin H. Genetic control of rice plant architecture under domestication. *Nat Genet.* 2008; 40: 1365-1369.
- [31] Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, Lu B, Lin H, et al. Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat Genet.* 2008; 40: 1370-1374.
- [32] Lin Z, Li X, Shannon LM, Yeh CT, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE, et al. Parallel domestication of the Shattering1 genes in cereals. *Nat Genet.* 2012; 44: 720-724.
- [33] Dholakia BB, Ammiraju JSS, Singh H, Lagu MD, Roder MS, Rao VS, Dhaliwal HS, Ranjekar PK, Gupta VS, Weber WE. Molecular marker analysis of kernel size and shape in bread wheat. *Plant Breed.* 2003; 122: 392-395.
- [34] Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW. A genetic framework for grain size and shape variation in wheat. *Plant Cell.* 2010; 22: 1046-1056.

- [35] Spiertz JHJ, Hamer RJ, Xu H, Primo-Martin C, Don C, van der Putten PEL. Heat stress in wheat (*Triticum aestivum* L.): Effects on grain growth and quality traits. Eur J Agron. 2006; 25: 89-95.
- [36] Mastilović J, Živančev D, Lončar E, Malbaša R, Hristov N, Kevrešan Ž. Effects of high temperatures and drought during anthesis and grain filling period on wheat processing quality and underlying gluten structural changes. J Sci Food Agric. 2018; 98: 2898-2907.
- [37] Awolu OO, Sudha LM, Manohar B. Influence of defatted mango kernel seed flour addition on the rheological characteristics and cookie making quality of wheat flour. Food Sci Nutr. 2018; 6: 2363-2373.
- [38] Khanna-Chopra R, Viswanathan C. Evaluation of heat stress tolerance in irrigated environment of *T. aestivum* and related species. I. Stability in yield and yield components. Euphytica. 1999; 106: 169- 180.
- [39] Rahaman M, Mamidi S, Rahman M. Genome-wide association study of heat stress-tolerance traits in spring-type *Brassica napus* L. under controlled conditions. Crop J. 2018; 6: 115-125.
- [40] Fischer RA, Maurer R. Drought resistance in spring wheat cultivars. I. Grain yield response. Aust J Agric Res. 1978; 29: 897-907.
- [41] Nouri A, Etminan A, Teixeira da Silva JA, Mohammadi R. Assessment of yield, yield-related traits and drought tolerance of durum wheat genotypes (*Triticum turgidum* var. *durum* Desf.). Aust J Crop Sci. 2011; 5: 8-16.
- [42] Turuspekov Y, Baibulatova A, Yermekbayev K, Tokhetova L, Chudinov V, Sereda G, Ganal M, Griffiths S, Abugalieva S. GWAS for plant growth stages and yield components in spring wheat (*Triticum aestivum* L.) harvested in three regions of Kazakhstan. BMC Plant Biol. 2017; 17: 190.
- [43] Valluru R, Reynolds MP, Davies WJ, Sukumaran S. Phenotypic and genome-wide association analysis of spike ethylene in diverse wheat genotypes under heat stress. New Phytol. 2017; 214: 271-283.
- [44] Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q. GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Applied Genet. 2006; 112 : 1164-1171.
- [45] Song X, Huang W, Shi M, Zhu M, Lin H. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet. 2007; 39: 623-630.
- [46] Weng J, Gu S, Wan X, Gao H, Guo T, Su N, Lei C, Zhang X, Cheng Z, Guo X, et al. Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. Cell Res. 2008; 18: 1199-1209.
- [47] Huang X, Qian Q, Liu Z, Sun H, He S, Luo D, Xia G, Chu C, Li J, Fu X. Natural variation at the DEP1 locus enhances grain yield in rice. Nat Genet. 2009; 41: 494-497.
- [48] Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J, et al. Natural variation in GS5 plays an important role in regulating grain size and yield in rice. Nat Genet. 2011; 43: 1266-1269.
- [49] Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X, et al. Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. Nat Genet. 2008; 40: 761-767.
- [50] Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X, et al. Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. Nat Genet. 2010; 42: 541-544.
- [51] Hansdah R, Prabhakar PK, Srivastav PP, Mishra HN. Physico-chemical characterization of lesser known palo (*Curcuma leucorrhiza*) starch. Int Food Res J. 2015; 22: 1368-1373.
- [52] Mason RE, Mondal S, Beecher FW, Pacheco A, Jampala B, Ibrahim AM, Hays DB. QTL associated with heat susceptibility index in wheat (*Triticum aestivum* L.) under short-term reproductive stage heat stress. Euphytica. 2010; 174: 423-436.
- [53] Sharp PJ, Chao S, Desai S, Gale MD. The isolation, characterization and application in the Triticeae of a set of wheat RFLP probes identifying each homoeologous chromosome arm. Theor Applied Genet. 1989; 78: 342-348.
- [54] VSN International. Genstat for Windows 16th Edition. VSN International, Hemel Hempstead, UK. 2015; Available: <https://www.vsni.co.uk/>
- [55] Nyquist WE. Estimation of heritability and prediction of selection response in plant populations. CRC Crit Rev Plant Sci. 1991; 10: 235-322.
- [56] Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A, et al. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc Natl Acad Sci USA. 2013; 110: 8057-8062.

- [57] Wang SC, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L, et al. Characterization of polyploid wheat genomic diversity using a high-density 90000 single nucleotide polymorphism array. *Plant Biotechnol J.* 2014; 12: 787-796.
- [58] Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. *Am J Hum Genet.* 2000; 67: 170-181.
- [59] 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.
- [60] Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet.* 2005; 38: 203-208.
- [61] Zhang ZW, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu JM, Arnett DK, Ordovas JM, et al. Mixed linear model approach adapted for genome-wide association studies. *Nat Genet.* 2010; 42: 355-360.
- [62] Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Mol Biol Evol.* 2016; 33: 1870.
- [63] Bradbury PJ, Zhang ZW, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics.* 2007; 23: 2633-2635.
- [64] Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol.* 2001; 60: 155-166.
- [65] Wang M, Tu L, Lin M, Lin Z, Wang P, Yang Q, Ye Z, Shen C, Li J, Zhang L, et al. Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication. *Nat Genet.* 2017; 49: 579-587.
- [66] Cruickshank TE, Hahn MW. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol.* 2014; 23: 3133-3157.
- [67] Bhatia G, Patterson N, Sankararaman S, Price AL. Estimating and interpreting FST: the impact of rare variants. *Genome Res.* 2013; 23: 1514-1521.
- [68] McGirr JA, Martin CH. Novel Candidate Genes Underlying Extreme Trophic Specialization in Caribbean Pupfishes. *Mol Biol Evol.* 2017; 34: 873-888.

Table

Table 1 Analysis of variance of 19 traits in three years of 207 wheat cultivars under natural condition.

Source of variation	DTH	DTF	KNPS	TKW	GL	GW	RLW	GP	GR	GAR	
	df	SS	SS	SS	SS	SS	SS	SS	SS	SS	
Genotypes	206	37225.3***	12880.7***	293721.2***	1538.0***	2652.8***	6125.6***	17012.1***	489096.7***	65528.1***	12404.3***
Years	2	5970.0***	4707.5***	3350.2***	6.4***	1358.0***	4.5***	1.1**	1276.8***	1.7**	3315.9***
GxE interaction	412	9416.3***	3931.9***	662073.0***	1798.9***	4056.0***	9014.8***	29581.8***	662927.9***	113112.2***	18771.9***
Error	1236	47.9	52.3	2579.0	14.4	22.9	46.8	133.0	3540.0	419.5	97.4
h^2 a		0.78	0.59	0.57	0.72	0.66	0.67	0.63	0.69	0.63	0.66
Source of variation	GPC	TSC	PV	TV	BD	FV	SB	PT	PTP		
	df	SS	SS	SS	SS	SS	SS	SS	SS		
Genotypes	206	799.2***	821.2***	1211.0***	34529.5***	949.5***	26658.6***	33089.6***	2047.2***	768.5***	
Years	2	0.8***	0.2***	3.5***	528.2***	10.4***	71.3***	195.9***	39.2***	28.7***	
GxE interaction	412	984.1***	1138.7***	1090.2***	41979.6***	1122.9***	28014.2***	31338.4***	2716.2***	800.3***	
Error	1236	10.4	11.2	12.2	240.1	13.2	157.7	149.3	16.1	10.1	
h^2		0.71	0.68	0.77	0.71	0.72	0.74	0.76	0.69	0.74	

a h^2 broad sense heritability; ** and ***, significant at $P < 0.01$ and $P < 0.001$, respectively.

Supplementary Information

Additional file 1: Table S1. One hundred and twenty-nine significant MTAs involving 59 SNP loci and fifteen phenotypic traits.

Additional file 2: Table S2. Best blast hits sequence flanking SNP against the Chinese Spring reference genomic sequence (IWGSC).

Additional file 3: Table S3. Information regarding the 207 wheat accessions used in this study.

Additional file 4: Fig. S1. The daily maximum, minimum and average temperature during the heat treatment period. (a) 2015, (b) 2016 and (c) 2017.

Figures

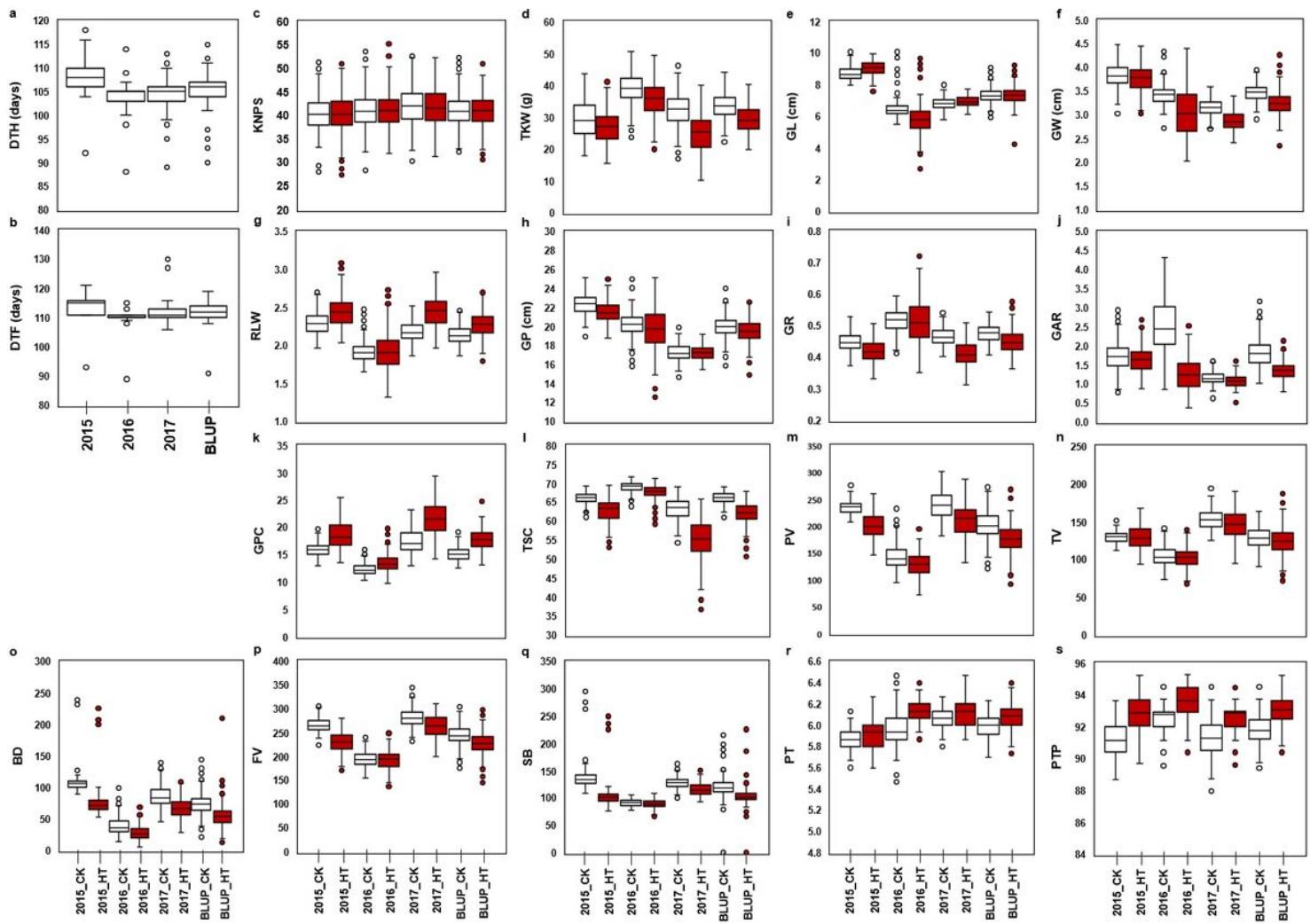


Figure 1

Boxplots of nineteen traits across three environments and BLUPs under heat stress (HT) in 2015-16, 2016-17 and 2017-18. (a) DTH, (b) DTF, (c) KNPS, (d) TKW, (e) GL, (f) GW, (g) RLW, (h) GP, (i) GR, (j) GAR, (k) GPC, (l) TSC, (m) PV, (n) TV, (o) BD, (p) FV, (q) SB, (r) PT and (s) PTP.

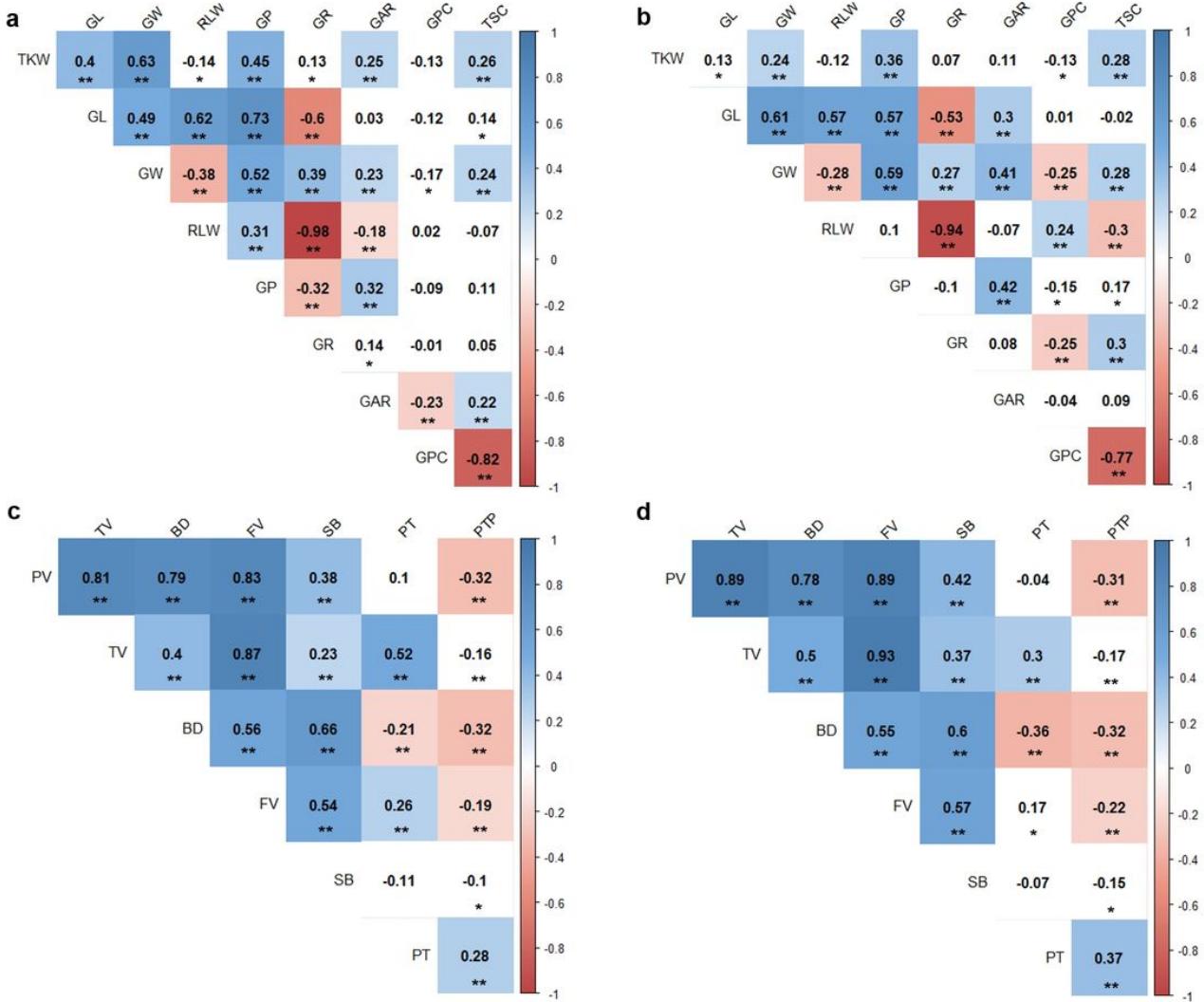


Figure 2

Correlations between traits (TKW, GL, GW, RLW, GP, GR, GAR, GPC, TSC, PV, TV, BD, FV, SB, PT and PTP) in two environments. (a) and (c) yield potential, (b) and (d) heat stress. *, P <0.05, **, P <0.01.

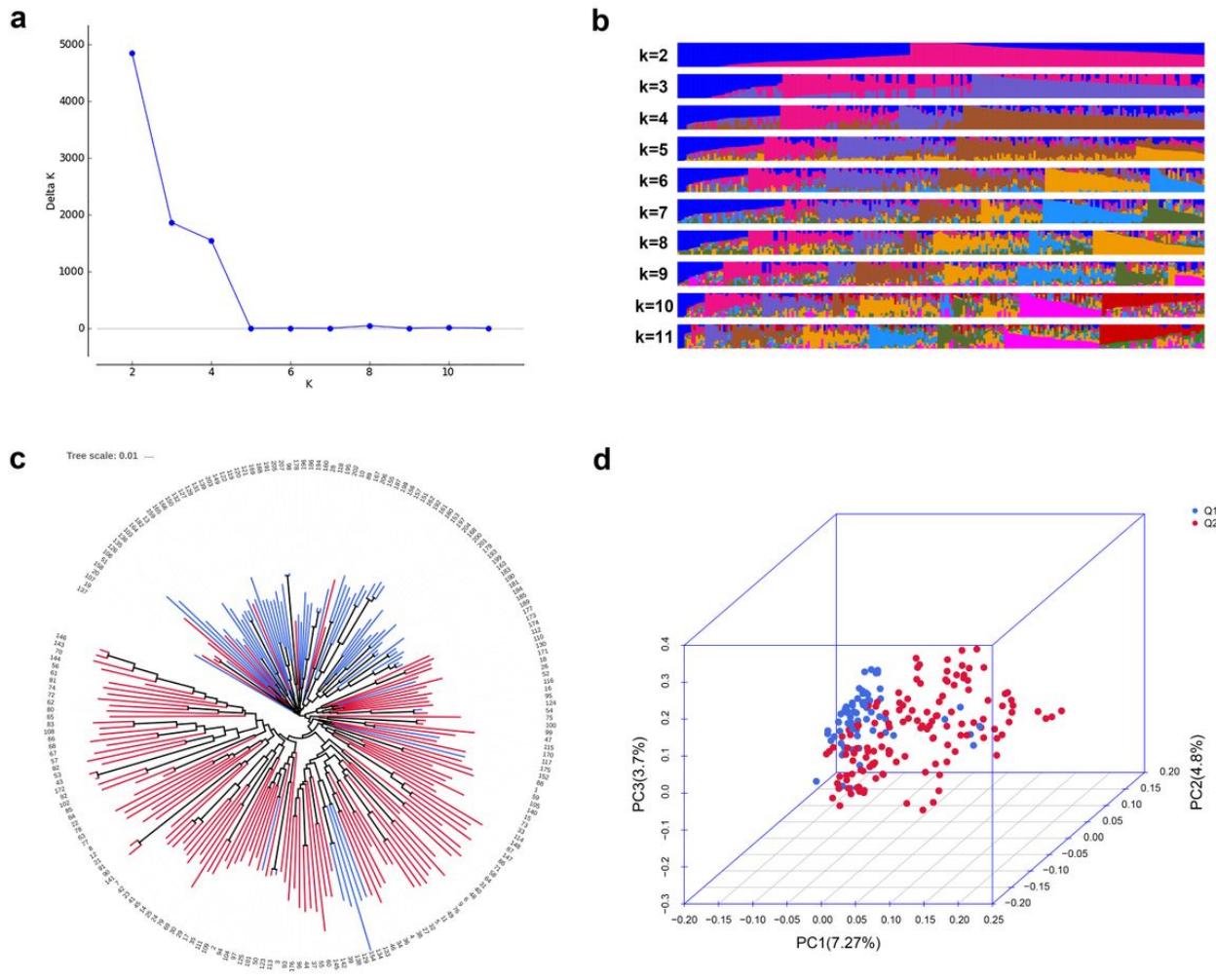


Figure 3

Population structure analysis of 207 wheat accessions. (a) Estimated ΔK over five repeats; (b) probability of population group based on STRUCTURE with subgroups ranging from 1 to 10; (c) neighbor-joining (NJ) tree; (d) principal components analysis (PCA) plots.

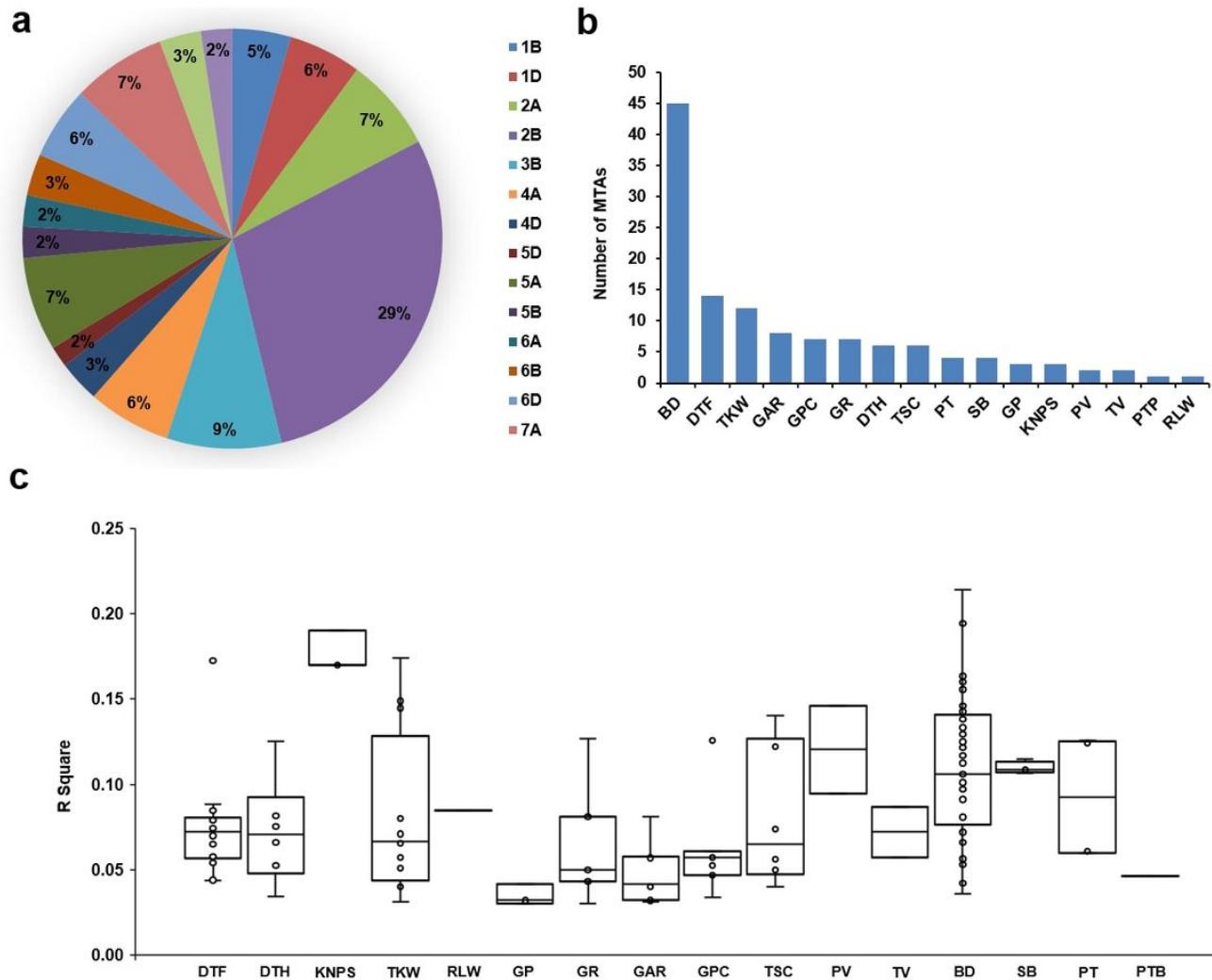


Figure 4

Summary of MTAs. (a) Pie chart showing the percentage of MTAs in different chromosomes; (b) number of MTA for each trait; (c) range of percentage variation explained (R square) for each trait.

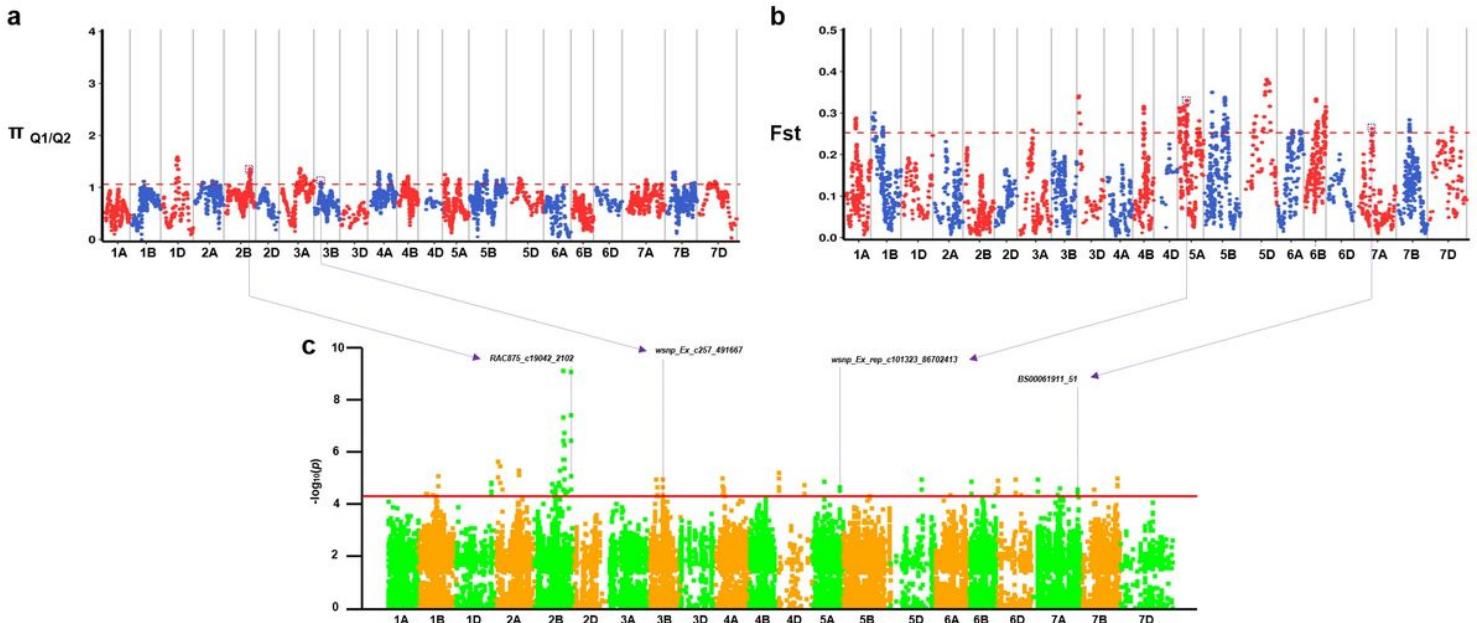


Figure 5

Genome-wide screening of group-specific selection sweeps and GWAS on grain-related traits under heat stress. Π (a) and FST (b) values for all SNP sites between Q1 and Q2. Red dots above the horizontal dashed line indicate highly divergent SNP (top 5%). (c) Manhattan plot of the GWAS for heat stress related traits using the compressed mixed linear model (MLM). Red horizontal line corresponds to the threshold value for significant association ($P < 5 \times 10^{-5}$).

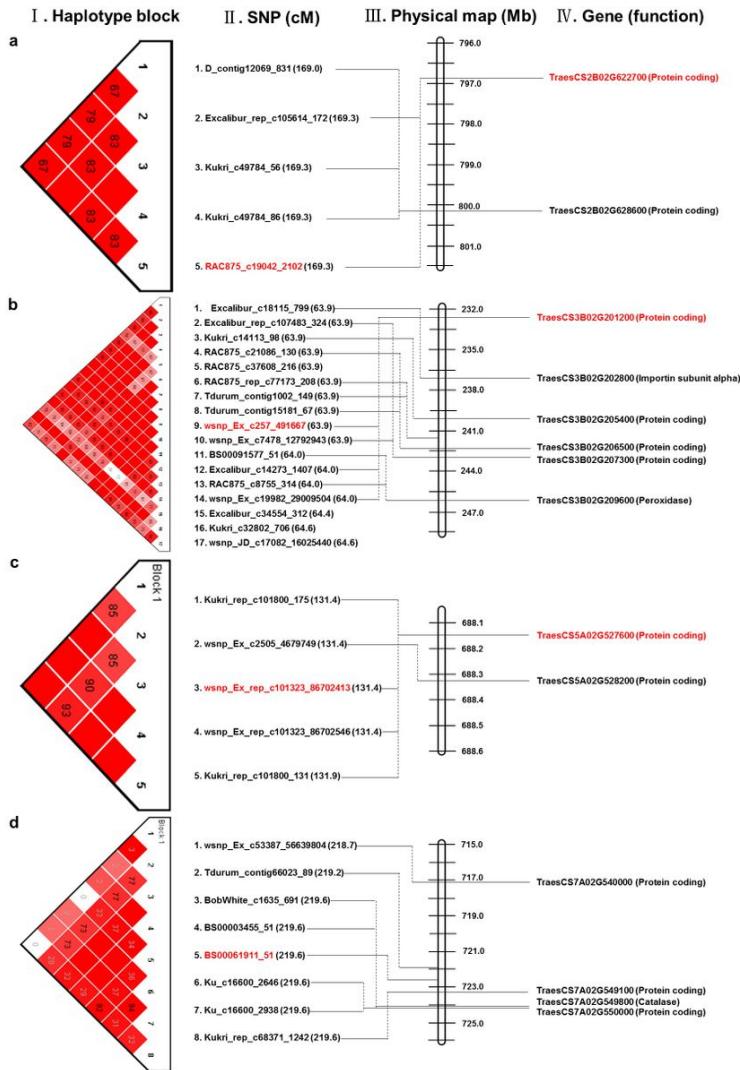


Figure 6

Gene annotation of MTAs identified on chromosome 2B (a), 3B (b), 5A (c) and 7A (d) for heat stress tolerance. Linkage disequilibrium (LD) plots of SNP markers falling significant haplotype blocks (I). Names of the markers in the region of interest along with cM (II), and physical (III) positions. Image at the far right is a physical map of candidate genes on chromosome segments (IV). The cM position is based on Wang et al. (2014) and the physical position is based on IWGSC 2018.

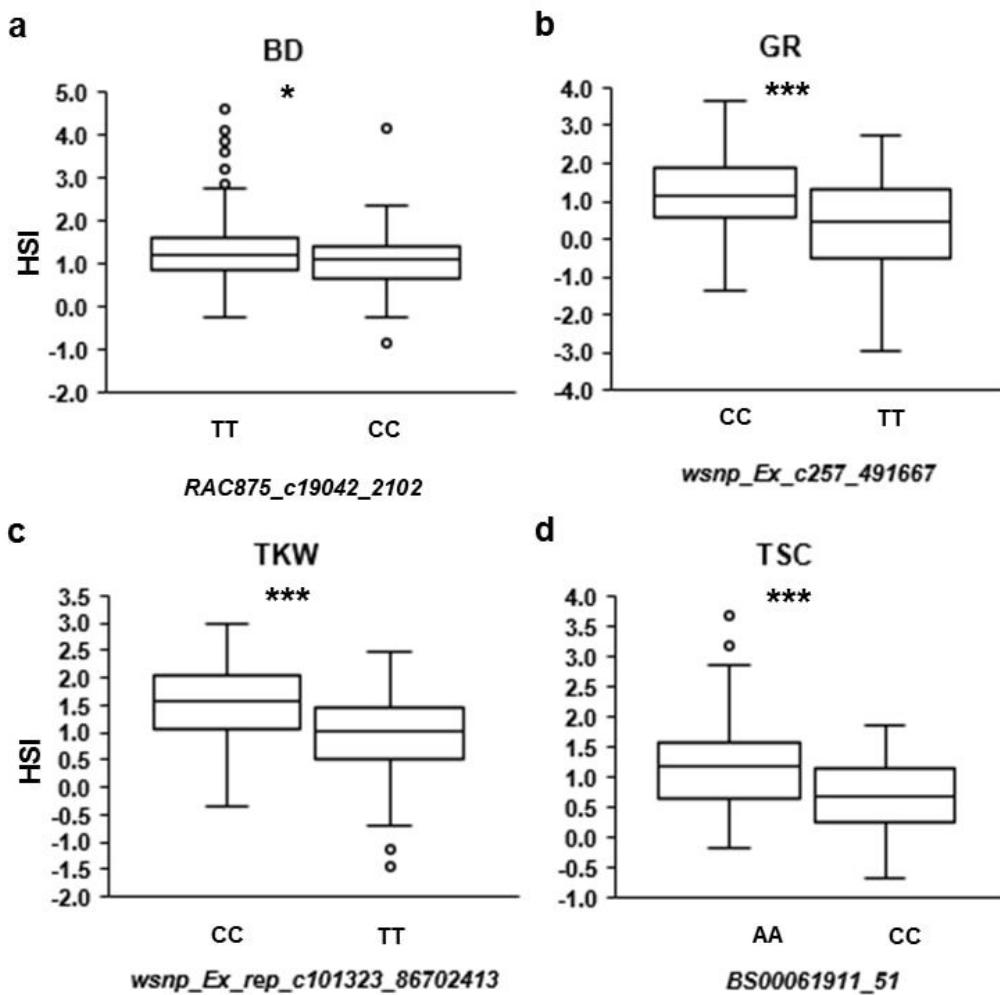


Figure 7

Boxplots of the relationship of HSI for BD (a), GR (b), TKW (c) and TSC (d), using marker polymorphisms at RAC875_c19042_2102, wsnp_Ex_c257_491667, wsnp_Ex_rep_c101323_86702413 and BS00061911_51, respectively. *, P < 0.05, ***, P < 0.0001.

Supplementary Files

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

- [TableS2.xlsx](#)
- [TableS1.xlsx](#)
- [Fig.S1.jpeg](#)
- [TableS3.xlsx](#)