Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 February 2020 doi:10.20944/preprints202002.0272.v1

GWAS analysis of wheat pre-breeding germplasm for terminal drought stress using next generation sequencing technology

Sajid Shokat[1,2], Deepmala Sehgal[3], Fulai Liu[1], Sukhwinder Singh[3,4]

[1]Department of Plant and Environmental Sciences, University of Copenhagen, Højbakkegård Allé 13, 2630, Taastrup, Denmark

[2]Nuclear Institute for Agriculture and Biology, 38000, Faisalabad, Pakistan

[3]International Maize and Wheat Improvement Centre (CIMMYT) km, 45, Carretera Mex-Veracruz, El-Batan, CP 56237, Texcoco, Mexico

[4]Geneshifters, 222 Mary Jena Lane, Pullman WA, USA 99163

Abstract

Bread wheat (*Triticum aestivum* L.) is one of the most important cereal crops for food security. Of all the stresses that curtail wheat productivity, drought has the most detrimental effects. Especially terminal drought stress i.e. at the time of flowering imposes a big challenge to sustain grain production. In the current study, 339 pre-breeding lines derived from three-way crosses of exotics x elite lines were evaluated in the irrigated and drought stress environments at Obregon, Mexico for the year 2016 and 2018. Drought significantly reduced yield (Y), spike length (SL), number of grains per spikes (NGS) and thousand kernel weight (TKW) by 46.4, 19.2, 23.5 and 25.9%, respectively as compared to irrigated conditions. Kernel abortion (KA), highly correlated with Y, increased significantly (11.6%) under drought stress environment. Population structure analysis in this panel revealed three sub-populations and a genome wide linkage disequilibrium (LD) decay at 2.5 cM. Single marker and haplotypes-based genome wide association study (GWAS) revealed significant associations on three chromosomes; 4A (HB10.7), 2D (HB6.10) and 3B (HB8.12) with Y, spike length (SL) and TKW, respectively. Likewise, associations on chromosomes 6B (HB17.1) and 3A (HB7.11, HB7.12) were identified for NGS and KA. Five traits *i.e.* normalized difference vegetation index (NDVI), canopy temperature depression (CTD) days to heading (DTH), NGS, KA were associated at chromosome 3A both under irrigated and drought conditions however, different haplotypes were estimated. Twenty-six SNPs were part of 10 haplotype blocks associated with Y, SL, TKW, NGS, KA. In *silico* analysis of the associated SNPs/haplotypes showed hits with candidate genes known to confer abiotic stress resistance in model species and crops. Potential candidate genes include those coding for sulfite exporter TauE/SafE family in *Arabidopsis thaliana*, TBC domain containing protein in *Oryza sativa subsp. Japonica* and heat shock proteins in *Aegilops tauschii subsp. tauschii* were revealed. The SNPs linked to the promising genes identified in the study can be used for marker-assisted selection.

Key words: Bread wheat, terminal drought, GWAS, grain yield, kernel abortion

1. Introduction

Uneven pattern of rainfalls and more frequent spells of drought are among fatal threats for crop production. Equally, global freshwater resources are shrinking and further creating harsh conditions for global agriculture. Therefore, there is a pressing need to develop crops that produce higher yield with less water. Wheat (*Triticum aestivum* L.) is grown in more than 85 countries of the globe with about 2.1 million km^2 total harvested area and contributes to about the 20% of the total dietary calories and proteins worldwide [1]. Due to its cultivation in wide range of environments ranging from very favorable ones in Western Europe to severely stressed ones in parts of Asia, Africa and Australia, it is highly prone to various biotic and abiotic stresses. Of all the stresses curtailing wheat productivity, drought has the most detrimental effects [2]. Especially terminal drought stress i.e. drought stress at the time of flowering reduces grain yield through ceased growth of spike and reduction in grain number and grain weight[3][4][5][5]. Improvement of drought stress tolerance has therefore become the prime target for wheat breeders.

Traditional breeding methods have achieved tremendous success in the last century [6] in pushing the annual genetic gain in wheat by using desirable genetic variations within primary gene pools. However, to feed the burgeoning world population and to develop climate smart varieties more efforts are required. Genetic resources such as landraces and wild relative collections, maintained in gene banks, are a reservoir of unexplored beneficial alleles. Worldwide wheat breeders have access to up to 800,000 accessions many of which show adaptation to different abiotic stresses [7], for example, Creole wheat landraces (the landraces introduced to Mexico from Europe). However, breeders are reluctant to introgress these rare unexplored alleles into wheat cultivars because of the challenges involved in identifying useful, novel diversity and introgressing it into well-adapted elite cultivars with minimum linkage drag.

Pre-breeding plays a key role in minimizing the barrier in mobilizing novel genetic variation into breeding programs introducing less undesired, linked genes. It refers to the development of bridging, semi-finished or intermediate germplasm having introgressions from exotic or unimproved germplasm [8]. The Seeds of Discovery (SeeD) project was launched at CIMMYT to create new pre-breeding germplasm in order to enhance the use of genetic resources in breeding programs (SeeD 2011). Within the framework of SeeD, a panel of pre-breeding lines (PBLs) has been generated by three-way (exotic/elite1//elite2) crosses of hundreds of exotics (landraces and synthetics) with 25 elites [9]. These PBLs have been distributed to different international partners for multi environment trials for the release of varieties and for novel gene discoveries for complex traits. We investigated a subset of this germplasm under terminal drought and irrigated environments to identify genomic regions associated with drought tolerance traits using genome wide association (GWA) approach. GWA studies have been reported for various traits in wheat, including grain yield and yield components and physiological traits under multiple environments [10][11][12][13][14][15][16]. GWA aided by haplotype based mapping has proven even better for complex traits[17][18] [19]). We conducted both single marker and haplotypes-based GWA in the present study.

2. Materials and Methods

2.1. *Plant material and crossings*

Three hundred and thirty-nine pre-breeding lines derived from the three-ways top-crosses and their performance was evaluated under drought conditions at Obregon, Mexico. In the first crossing scheme, the landraces/accessions from the gene bank collected from different areas of the world were crossed with best elite line of wheat and then F_1 hybrids were top crossed with second best elite line (landraces/elite1//elite2). Likewise, in the second scheme synthetics wheat derived from the crosses of *Triticum turgidum subsp durum*, *Triticum dicoccum* and *Aegilops tauschii* were crossed with best elite line and then F_1 top crossed with second best elite (synthetic/elite1//elite2). In total 25 elite parents developed at CIMMYT and other places, displaying higher yield potential and rust (leaf and stripe) resistance were used in different cross combinations. These crossing schemes resulted into 1,200 TC_1F_1 crosses. However, 244 of these TC_1F_2 populations were advanced, based on their performance in the field for different yield contributing traits. A selected bulk method was used to advance these generations up to TC_1F_5 and 339 best lines were selected for further evaluation.

2.2.*Experimental setup*

The plant material was sown at the experimental station "campo experimental Norman E Borlaug (CENEB)" of CIMMYT Obregon, Sanora during the years 2015-16 and 2017-18. This diverse germplasm was planted in two replicates following the alpha lattice design under well-watered and drought conditions. Each genotype was planted in 2 rows each with 2-meter length and 40 cm distance between each row. Five irrigations, each containing approximately 125 mm were given to normal irrigated treatment. In contrast, only two irrigations (approximately (55+125=180 mm)) were applied until tillering stage in drought treatment.

2.3.*Phenotypic characterization*

Different agro-physiological parameters were recorded during and after maturity and are as follow:

a) Canopy temperature depression (CTD) was measured with infrared thermometer sixth sense (LT300). The measurement was adjusted 0.5 m distance above the canopy and viewing angle maintained around 45° middle area about 2m of each plot was targeted.

b) Normalized difference in vegetative index (NDVI) was measured with the help of Green seeker variable rate application and mapping systems with portable sensor (Mod:505) for 5 seconds for both well-watered and drought genotypes. The average data was used with values 0-1 (no-green to maximum green).

c) Days to heading (DTH) were recorded by counting the number of days taken from sowing to 50% flowering.

d) Days to maturity (DTM) were counted by number of days taken from sowing to until all spikes were turn to their specific color of maturity.

e) Plant height (PH) was measured from 5 random plants of each replication from ground level to the end spike excluding awns

f) Spike length (SL) was calculated by taking 5 random spikes from each replication and length was measured from the start of spike to the end of spike excluding awns

g) Number of grains spike^{-1} (NGS) were counted by taking five random spikes from each replication and their averages were recorded

h) Thousand kernel weight (TKW) were recorded by counting thousand grains from each replication and their weight was recorded in grams

i) Kernel abortion (KA) was recorded using following formula:

(NGS for drought plants/NGS for well-watered) x 100

j) Grain yield (Y) of each genotype was recorded after harvesting the both the rows and yield was expressed in kilograms per hectare.

2.4. *Molecular characterization*

Leaf samples were taken at booting stage from TC_1F_5 plants, snap frozen in liquid nitrogen and stored at -80°C until further analysis. Genomic DNA was extracted using modified cetyltrimethylammonium bromide (CTAB) protocol [9], followed by quantification by Nano-Drop 8000 spectrophotometer V 2.1.0. Likewise, for genotypic characterization of these samples DArTseq™ technology (http://www.diversityarrays.com/dart-application-dartseq) of Genetic Analysis Service for Agriculture (SAGA) service unit at CIMMYT headquarters (Texcoco, Mexico) was used. High quality 58,378 SNP markers were generated and reported by Singh et al. 2018. The proportion of samples with genotypic score not recorded as missing data (call rate) and the proportion of technical replicate assay pairs for which the marker score was consistent (average reproducibility) were used to select the markers. The result was only 7,180 markers used for the final analysis out of 12,071 SNPs. 100K-marker DArT-seq consensus map released by Diversity Arrays Technology Pty Ltd. (DArT) (http://www.diversityarrays.com/sequence-maps) was used to define the chromosome location, marker order and genetic distances.

2.5. *Haplotype Characterization*

Haplotype map was generated in R using the algorithm-based script described by (Gabriel et al. (2002). In brief, D prime (D′) was generated through 95% confidence and three categories were made for the comparison called, "strong linkage disequilibrium (LD)", "inconclusive" or "strong recombination". Haplotype block was created using the R code as used in previous study Singh et al. [9]. Briefly, we calculated linkage disequilibrium parameter D' and D' 95% confidence intervals between SNPs and each comparison was categorized as “strong LD,” “inconclusive,” or "strong recombination." A haplotype block was created if 95% of the comparisons in one block were in "strong LD." Cut off p-value for Hardy Weinberg was established at 0.001 while minimum value for marker allele frequency was set to 0.05. Haplotypes were not constructed for the individuals having more than 75% missing data. If the multiple SNPs were indicating the same genetic position, only the first marker was considered to construct haplotype map. The haplotypes were displayed as blocks of marker numbers and alleles and "HB" prefix was used for each haplotype block followed by the chromosome number followed by dot and then the increment number from 1 to N where, N is the total number of the haplotype blocks along the chromosome.

2.6. *Genome Wide Association results*

Population structure and genome wide linkage disequilibrium (LD) decay has been reported in this panel before by Ledesma-Ramírez et al. [18]. Principal component analysis (PCA) file from Ledesma-Ramírez et al. [18] was used while kinship matrix was generated using VanRaden algorithm in GAPIT v 3.0. among the variables. A mixed linear model (MLM) was used with PCA as fixed effect and kinship as random effect. Significance of a marker was decided by two factors; (a) p-value (0.001 and 0.0001 corresponded to the bottom 0.1 percentile of yield and yield related traits) and (b) deviation of its p-value from the normal distribution curve. The following equations were used to determine GLM and MLM, respectively:

$$Y_{GLM} = A\alpha + B\beta + e;$$

$$Y_{MLM} = A\alpha + B\beta + C\mu + e;$$

where vector of phenotype is represented by Y, vector of marker genotypes is represented by A; principal components matrix is represented by B; C is representing the relative kinship matrix; α,

β, and μ are the corresponding effects; and residual effects are represented by e in the matrix. A and B matrix were fitted as fixed effects, and C and e matrix were fitted as random effects.

2.7. *Candidate gene analysis*

Significant markers from the above-mentioned analysis had their 69 bp sequences used as input query at EnsemblPlants using the BLAST tool at default settings. The search was performed primarily targeting the 5 *Triticeae* species available in this database. The results were selected based on minimum length of 69, score > 80 and sequence identity > 90 %. The full sequence of the hits that fulfilled the established criterion were retrieved and queried at NCBI BLASTN tool using default settings to investigate the underlying function, if any, and creating a list of potential gene candidates for future validation.

3. **Results**

3.1. *Analysis of variance and heritability*

Significant differences for grain yield (Y), canopy temperature depression (CTD) days to heading (DTH) and days to maturity (DTM) were observed for the years 2016 and 2018 (P<0.001) and the lowest grain yield was recorded in 2016 (Figure 1a and Figure S2a). However, DTH and DTM were significantly reduced during 2018 (Figure S4a and Figure S5a). Drought significantly reduced Y, spike length (SL), number of grains spike^{-1}, (NGS), thousand kernel weight (TKW) (Figure 1a, 2a, 3a and 5a), normalized difference in vegetative index (NDVI), CTD (Figure S1a and S2a), plant height (PH), DTH and DTM (Figure S3a, Figure S4a and S5a) in comparison to irrigated conditions (P<0.001). In contrast kernel abortion (KA) was significantly reduced (Figure 4a). All genotypes differed significantly for Y, SL, TKW, PH and DTM. Y was varying from 1653 kg ha^{-1} under drought condition to 5788 kg ha^{-1} under irrigated conditions. Maximum reduction

(46.43%) in Y was observed in pre-breeding line (GID-7641264) from 4437 kg/ha to 2556 kg/ha. The interaction of genotype*year and year*treatment were significant for Y (Table 1). Maximum reduction, 19.2% in SL was recorded in genotype GID-7645435 however, highest 12.1cm SL under drought was recorded in GID-7640992 (Table 1). Highest reduction 23.5% for NGS was recorded in GID-7644025 (Table 1) and GID-7643397 was most sensitive with 11.6% reduction in KA. (Table 1) Maximum reduction 25.9 % for TKW was recorded in GID-7644595 and highest average two years TKW (58.77g) under irrigated conditions was recorded in GID-7645143 (Table 1). Interaction between genotype*treatment was also significant ($P<0.001$). Interaction for NDVI between genotype*treatment was significant while interactions between genotypes*year and treatment*year were also significant for PH (Table 1). Interaction of genotype*year and treatment*year were significant for DTM ($P<0.01$).

Broad sense heritability (h_{bs}) was 0.61 and 0.65 for Y, 0.80 and 0.72 for SL, 0.72 and 0.6 for NGS, 0.67 and 0.84 for TKW, 0.25 and 0.68 for NDVI, 0.13 and 0.22 for CTD, 0.81 and 0.71 for PH, 0.77 to 0.08 for DTH and 0.85 and 0.82 for DTM for irrigated and drought conditions, respectively. Very low (0.16) h_{bs} was found for KA under drought conditions (Table 1)

3.2. *GWAS analysis in diverse pre-breeding germplasm*

Population structure in this panel has been investigated in a previous study by Ledesma-Ramírez et al. [18]. Briefly, three subpopulations were detected in the panel; subpopulation 1 was sharing the pedigrees of Baj#1, Reedling#1, Villa Juarez F2009, and Tacupeto F2001, subpopulation 2 was sharing the pedigrees of Seri.1b//Kauz/Hevo/3/Amad*2/4/Kiritati; and subpopulation 3 was sharing the pedigrees of Fret2*2/4/Sni/Trap#1/3/Kauz*2/Trap//Kauz/5/Kachu. Genome wide linkage disequilibrium (LD) decay was observed at 2.5 cM and, in subgenomes LD decay values were 0.25, 0.50 and 0.25 cM for A, B and D genome respectively.

i. Physiological traits

GLM and MLM analysis revealed significant ($P<0.001$) associations of NDVI on chromosome 3A for year 2018 (Figure S1.b-c and Table S1) For CTD, association was obtained on chromosome 7B for both the treatments (Figure S2.b-c and Table S1).

ii. Plant height and earliness

GLM and MLM of GWAS indicated a significant association ($P<0.05$) of PH at chromosome 19 (7A) for both years and treatments (Figure S3.b-e and Table S1). Significant associations of DTH were recorded (Figure S4.b-e and Table S1) on chromosomes 3A and 6B for both the years under irrigated conditions. On chromosome 3A, favorable allele TA in HB717 showed earliness in DTH in comparison to CC allele under irrigated conditions (Figure S4.f). A strong association for DTM on chromosome 1B during both years and treatments (Figure S5.b-e and Table S1) was identified. Haplotype HB2.5 with favorable allele AT showed earliness in comparison to other alleles (Figure S5.f).

iii. Grain yield (Y) and yield related traits

For Y, significant associations ($P<0.001$) were identified on chromosomes 4A and 6B in both conditions and years (Figure 1.b, 1c, 1d & 1e and Table 2). Haplotype based analysis also identified HB10.1 on chromosome 4A to be associated with Y with haplotype TCG as favorable allele as compared to CTC for the year 2018 in both conditions (Figure 1.f). Likewise, for another associations on chromosome 4A, HB10.6 and HB10.7, haplotype AC was favorable in comparison to CT for 2016 (Table S2) in both conditions (Figure 1.g, h). For SL, GWAS analysis indicated significant (P<0.001) associations on chromosomes 2D and 3D under drought and irrigated treatments in 2018 (Figure 2.b-c and Table 2). Chromosome 2D was further analyzed and HB6.10 with favorable allele TC showed higher spike length compared to other alleles (TG or CC) under drought and irrigated conditions (Figure 2.d). GLM and MLM identified significant associations of NGS on chromosomes 6B and 3A both under irrigated and drought conditions for the year 2018 (Figure 3.b-c and; Table 2). Haplotype CC of HB17.1 on chromosome 6B showed significantly higher NGS for 2018 under both treatments (Figure 3.d). Similarly, haplotype CT on chromosome 3A (HB7.11) is responsible for higher NGS under drought in comparison to allele TT or CG (Figure 3.f). Strong associations for KA were identified on chromosomes 3A and 3B. Haplotype CGGTC from HB7.12 showed lower KA compared to TTCGT under drought conditions (Figure 4.b-c and Table 2). For TKW, association was identified on chromosome 3B (Figure 5.b-c and; Table 2). The haplotype-based analysis indicates that haplotype TAGGCT (HB8.12) is favorable in comparison to TAGGCC or CGATC under irrigated condition of both the years (Figure 5.d).

3.3.*Multiple quantitative traits loci (QTL's) at same chromosome*

Single marker trait analysis revealed significant associations of NDVI, CTD, DTH, NGS and KA at chromosome 3A under both irrigated and drought conditions. TKW and KA were also significantly associated at chromosome 3B under drought conditions (Table 2 and S1). Moreover, maker-trait association of NGS and Y was significantly associated on chromosome 6B.

3.4.*In silico results*

In silico analysis of important associations revealed many interesting candidate gene hits. SNP 979934 of HB10.7 (chromosome 4A) showed hits with a gene called *TauE*/*SafE* responsible for taurine metabolism and anion export across cell membrane in *Ae. thaliana*. SNP 3029487 of HB6.10 (chromosome 2D) associated with SL and revealed a gene called "LOC109764454" coding for heat shock proteins in *Ae. tauschii*. SNPs 3222182 and 1050615 on chromosome 7B associated with NGS showed hits with candidate genes coding for omega gliadin-D1 and osa-like protein in bread wheat. BLASTN analysis of SNP 1038112 of HB7.12 associated with KA showed hits with a NRT1/PTR family (NPF) like 8.2 gene in *Ae. tauschii*. This gene has a role in transport of phytohormones in *A. thaliana* (Table S3).

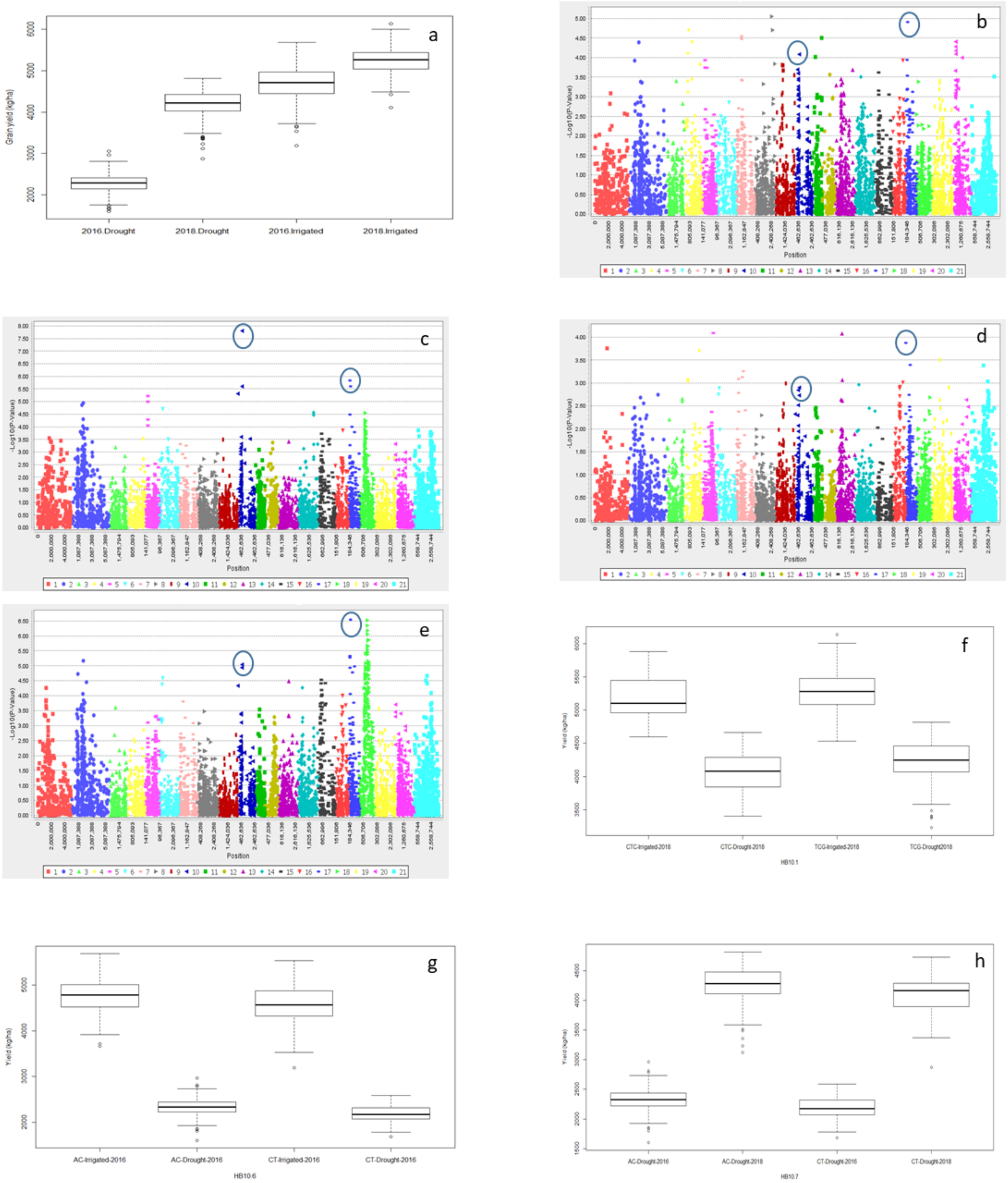

Figure 1. a) Boxplot indicating variation in grain yield under irrigated and drought condition for the year 2016 and 2018, b) marker trait association through GLM under irrigated conditions 2016, c) marker trait association through GLM under drought 2016, d) marker trait association through

GLM under irrigated 2018 e), marker trait association through GLM under drought 2018 f) haplotype block (HB) 10.1 showing differences in grain yield under irrigated and drought condition for 2018, g) haplotype block (HB) 10.6 showing differences in grain yield for irrigated and drought conditions in 2016 and f) haplotype block (HB) 10.7 showing differences in grain yield for drought conditions for 2016 and 2018

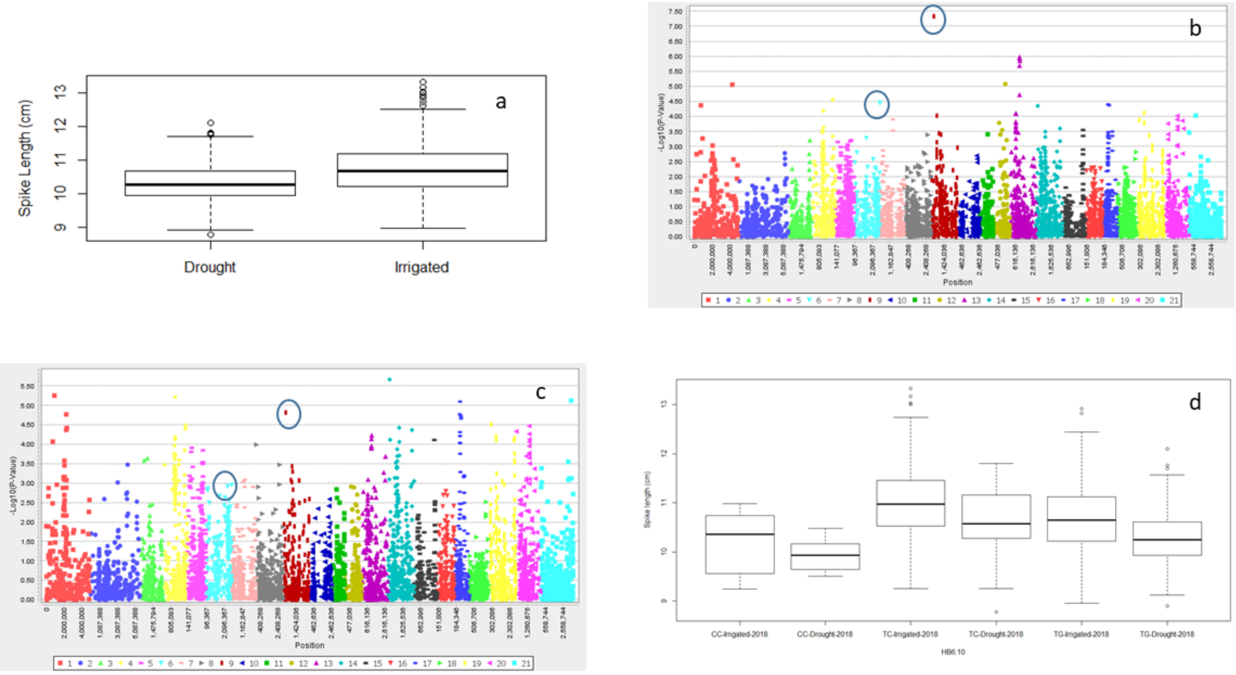

Figure 2. a) Boxplot showing variation for spike length under irrigated and drought condition for the year 2018 b) marker trait association through GLM under irrigated 2018 c) marker trait association through GLM under drought 2018 d) haplotype block (HB) 6.10 showing differences in spike length

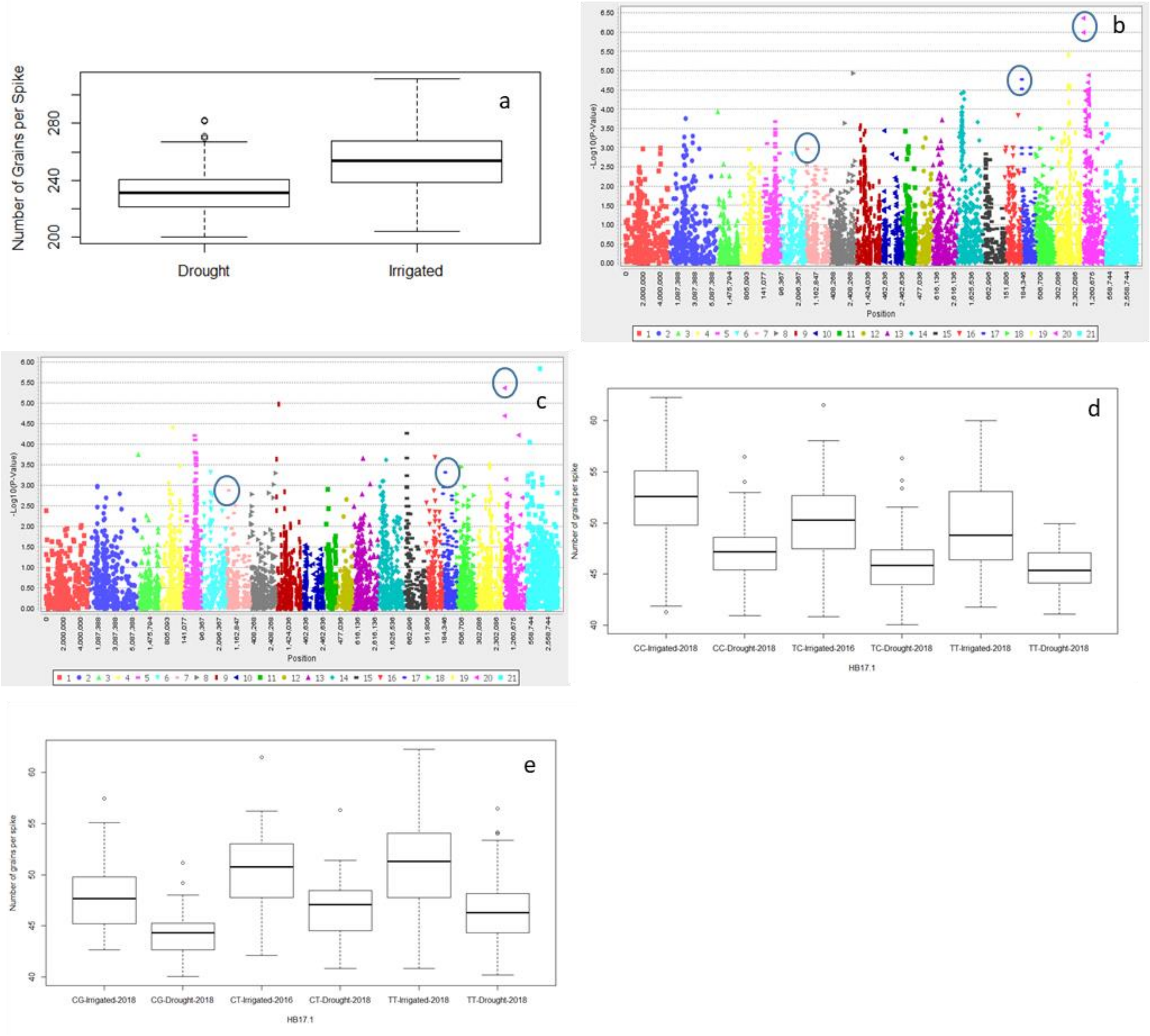

Figure 3. a) Boxplot showing variation in number of grains spike-1 under irrigated and drought conditions for 2018 b) marker trait association through GLM under irrigated 2018 c) marker trait association through GLM under drought 2018 d) haplotype block (HB) 17.1 showing differences in number of grains spike-1 e) haplotype block (HB) 7.11 showing differences in number of grains spike^{-1}.

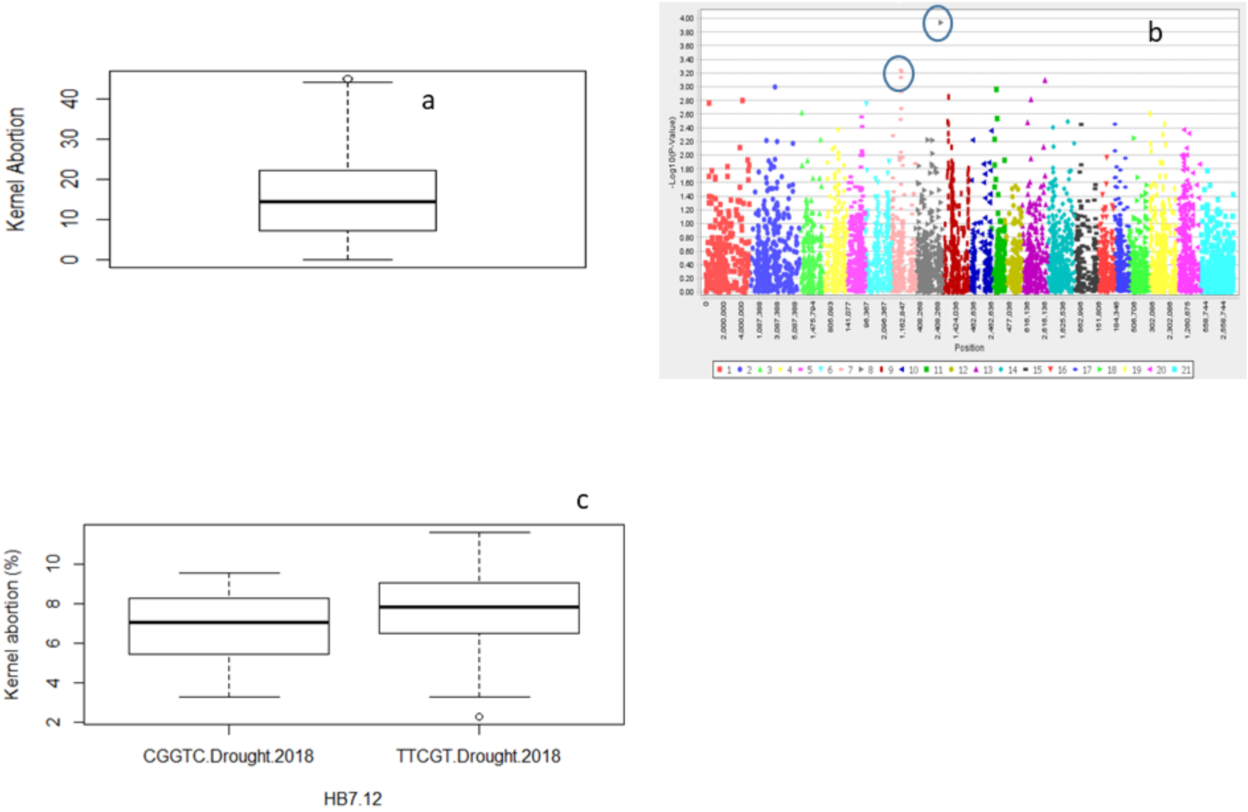

Figure 4. a) Boxplot showing variation in kernel abortion b) marker trait association through GLM under drought 2018 c) haplotype block (HB) 7.12 showing differences in kernel abortion

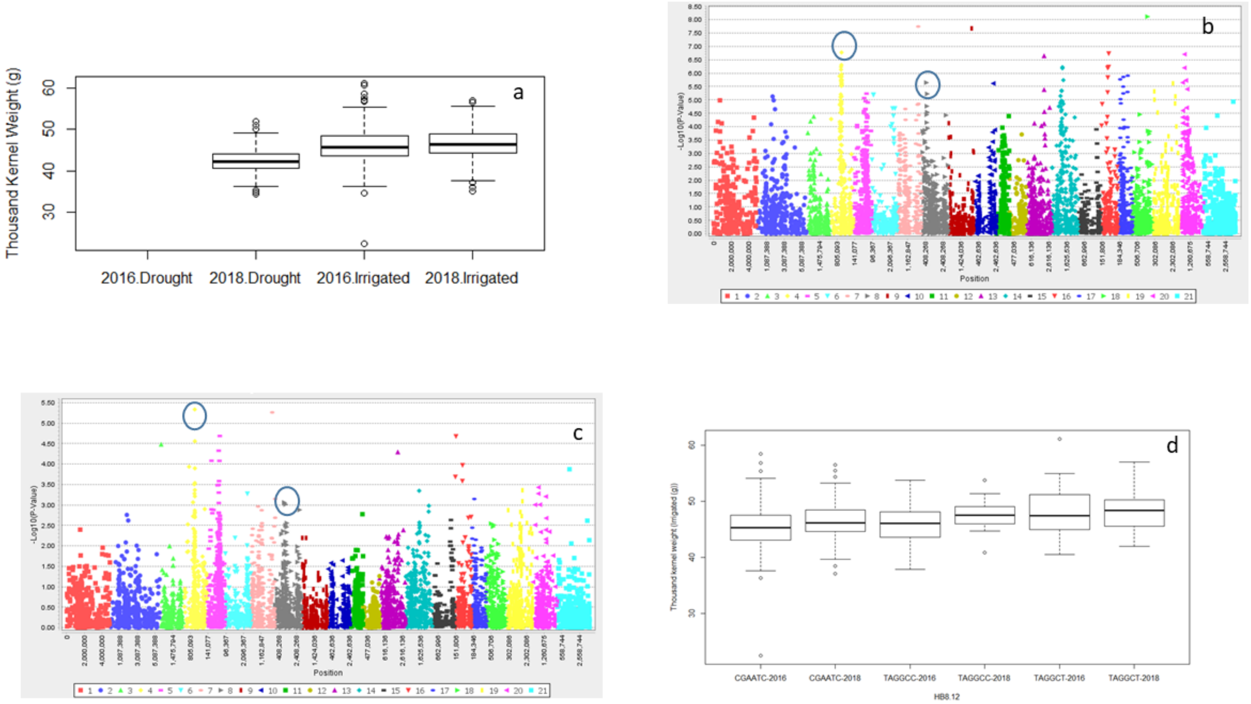

Figure 5. a) Boxplot showing variation for thousand kernel weight under irrigated conditions 2016 and irrigated and drought condition for 2018 b) marker trait association through GLM under irrigated 2016 c) marker trait association through GLM under irrigated 2018 d) haplotype block (HB) 8.12 showing differences in thousand kernel weight under irrigated condition for 2016 and 2018

4. Discussion

Use of haplotypes maximizes linkage disequilibrium between markers and QTLs as compared to SNPs. For traits with complex genetic architecture, use of multi-allelic haplotypes has significantly improved the power and robustness of GWAS studies in crops ([21][22][23]. In wheat too, recent studies have shown better resolving power of haplotypes in GWAS than SNPs [17] [18] [19][24]. We conducted both single marker and haplotypes-based GWAS to dissect QTLs for drought tolerance in a panel of pre-breeding lines derived by three-way crosses of exotics with CIMMYT's best 25 elites.

Panel investigated in the present study has been analyzed for population structure and LD decay which showed a moderate population structure in the panel and an LD decay of 2.5 cM genome wide [18]. The genome wide LD decay of 2.5 cM indicates its higher genetic diversity as compared to various other wheat panels used in previous studies reporting from 5 to 10 cM or even slower LD decay [14][25][26][27][10][28][29][30]. These results are expected since the panel has been drawn from a large set of pre-breeding lines developed by crossings among exotics and elites [9]. Each pre-breeding line acquired approximately 25% of the exotic genome and 75% of the elite genome at an early stage thus allowing recombination between exotic and elite genomes. Further, many of the agronomical and physiological traits investigated on the panel including DTM, SL, NGS, TKW and Y showed high heritability both under irrigated and drought conditions. Shokat et al.[31] reported that plant traits exhibiting higher heritability and genetic advance can directly be selected for crop improvement.

Both single marker and haplotype-based GWAS revealed the association of Y on chromosome 4A in both the years under irrigated and drought conditions (Figure 1.b-e, Table 2). Previous studies have also reported drought tolerance QTL on 4A in bread wheat [9] [14][27] [32][33] with percentage variation explained as large as 41% in some studies [34]. The present and previous studies therefore suggest an important role of 4A chromosome in drought adaptation in wheat. Three haplotypes blocks, i.e. HB10.1 (TCG), HB10.6 (AC) and HB10.7 (AC), within 9.3 to 60.4 cM resulted in significantly higher yield across years (Figure 1.f-h). BLASTN analysis indicates that SNPs in these genomic regions are part of a gene coding for sulfite exporter TauE/SafE family protein in *Arabidopsis thaliana*. This protein family is known to be involved in taurine metabolism [35] and transport of anion across the cell membrane [36]. However, role of TauE/SafE protein family under drought stress is not known yet.

SL is an important yield contributing trait and longer compact spikes are one of the attractive plant traits to obtain higher Y in wheat [37]. Mwadzingeni et al. [38] reported a positive correlation of SL with grain yield under drought stress indicating that the maintenance of higher SL under drought conditions can contribute towards the optimum yield under drought. In the present study, SL had high heritability and it showed significant associations on chromosome 2D with HB6.10 positioning between 277.5 cM to 287 cM under irrigated and drought conditions in both type of GWAS (Figure 2.b-c), where allele TC was linked with higher SL. Previously QTLs for SL were reported by Chai et al. [39] on short arm of chromosome 2D (2DS). Likewise, Zhai et al. [40] reported that QTL related to spike morphology is located on 2DS. BLASTN analysis showed hits with candidate gene “LOC109764454” in *Aegilops tauschii subsp. tauschii* indicating that this genomic region may have been contributed by synthetic parents of current germplasm. The locus LOC109764454 confers a heat shock protein in *Aegilops tauschii* [41].

Limited availability of water at the time of anthesis reduces pollen viability [42] and seed setting which increases kernel abortion (KA). Maintenance of high number of grains spike^{-1} (NGS) under drought is one of the important challenges in *T. aestivum* to obtain ideal yield under drought condition. This trait is positively correlated with grain yield under drought and irrigated conditions [43][44]. We found higher heritability and significant association of NGS with chromosome 6B and 3A both under irrigated and drought conditions and trait KA was also associated with chromosome 3A (Figure 3.b-c & Figure 4.b). Likewise, NDVI, CTD and DTH, were also associated with chromosome 3A. In line of these, Li et al. [19] reported the significant association of NGS and Y at chromosome 3A. Likewise, Pradhan et al. [45] also reported the marker traits association of NGS on chromosome 3A explaining 12-16% of variability. Other studies indicate that chromosome 6B is the third major chromosome of wheat covering more than 5% of wheat diversity with more 2000 genes [46]. Nadolska-Orczyk et al. [47] reported the presence of genes controlling supernumerary spikelets trait on this chromosome. In present study, HB 17.1 (TC) and HB7.11 (CT) were found to be associated with higher NGS under drought positioning at 5.7 to 6.4 cM and 106.6 to 106.9 cM, respectively (Figure 3.e-f). HB7.12 with allele CGGTC responsible for lower KA was present at 111.6 to 115.9 cM (Figure 4.c). HB7.11 and HB7.12 are closely placed indicating that NGS and KA may be linked together. BLASTN analysis indicated that SNP 3222182 (part of HB17.1) is the part of gene "ATY51569" coding for omega gliadin-D1, gluten protein in *T. aestivum* [48]. Additionally, another SNP 1050615 from the same haplotype block is the part of gene "LOC109767680" coding for osa-like proteins. The role of osa-MiRNAs has recently been demonstrated in rice in degrading and synthesizing nucleic acids and thus manipulating their abundance under drought stress [49].

The SNP 1038112 from HB7.12 associated with KA showed hits with gene "LOC109781107" of *Aegilops tauschii* coding for NPF 8.2-like protein [50]. In *A. thaliana*, NPF 8.2-like protein is reported to be responsible for the transport of phytohormones [51][52]. NPF family members display sequence and structural homologies with peptide transporter proteins involved in the uptake of di- and tri-peptides in organisms. In plants, they were initially characterized as nitrate or peptide transporters [53]. In recent years, several other substrates have been identified including abscisic acid (ABA). ABA has been well known to have a significant role under drought stress tolerance by controlling ABA-regulated mechanism [54] .

Thousand kernel weight (TKW) is one of the main contributors of grain yield [55] and less reduction in TKW enables plants to maintain ideal yield under drought conditions [56]. Liu et al.

[57] reported positive association of TKW with grain yield under irrigated and drought conditions indicating that plant genotypes having higher TKW under irrigated conditions often have a chance to maintain higher TKW under drought conditions. We found significant association of TKW at chromosome 3B (Figure 5.b-c) and haplotype HB8.12 with allele TAGGCT was giving higher TKW under irrigated conditions in both years in contrast to other alleles and positioning between 96.8 to 102.7 cM (Figure 5.d). Lehnert et al. [58] also reported the association of TKW at chromosome 3B under irrigated conditions. LujánBasile et al. [59] reported presence one haplotype block for TKW at chromosome 3B. Rustenholz et al.[60] mentioned 3000 genes are located on this chromosome indicating the richness and diversity of information on this chromosome. Likewise, among six of the corresponding markers, 1082914 and 1010250 (Table S3) were found to have their presence in bread wheat. The sequence of 1010250 is also reported in Chinese Spring of bread wheat by Ogihara et al. [50] however, function is not mentioned in wheat while in *Oryza sativa subsp. japonica* this sequence is a part of gene "Os01g0100100" coding for TBC domain containing protein [61].

Our experimental results indicate the transgressed genetic diversity from landraces and synthetic derivatives can improve Y under terminal drought through increase in TKW, SL and NGS. GWAS analysis coupled with haplotype-based association indicates stable association of certain markers for higher Y at chromosome 4A over both the years under drought conditions. These marker sequences have been studied in wheat before, but their function is not known. Our BLASTN results indicate this sequence of marker is coding for sulfite exporter TauE/SafE family protein in *A. thaliana* which is having role in taurine metabolism and anion export across cell membrane. Likewise, SL is found to be associated at chromosome 2D both under irrigated and drought condition where HB6.10 was associated with higher SL. Sequence of markers for this haplotype is coding for gene “LOC109764454”, responsible for heat shock protein in *A. thaliana*. Likewise, HB17.1 of chromosome 6B is having role in to obtain higher NGS in *T. aestivum* and gene “ATY51569” coding for omega gliadin-D1, gluten protein may have role to obtain higher NGS. Additionally, another gene “LOC109767680” coding for osa-like proteins can also play a role to sustain NGS under drought. GWAS and haplotype-based association indicates HB7.12 is showing lower KA and sequence for this haplotype block was coding for NPF proteins in *A. thaliana* involved in transport of phytohormones. Stable associations of certain markers were estimated at chromosome 3B to obtain higher TKW under irrigated conditions for both the years. These sequences have been reported in Chinese spring cultivar earlier, but their function was again unknow. This sequence is the part of gene "Os01g0100100", working in *O. sativa subsp. japonica* as a part of TBC domain containing protein. All these genes can further be studied to understand their role in bread wheat grown under drought stress andto find their association with yield and yield contributing traits. Likewise, validated genes can further be utilized in marker assisted selection to improve the drought tolerance in wheat. Moreover, reported germplasm can be helpful to improve the tolerance of bread wheat commercial cultivars.

Funding

Authors are thankful to Islamic Development Bank, Saudi Arabia and Wheat Seed of Discovery project at international maize and wheat improvement center (CIMMYT), Mexico for providing funds to conduct this research.

Authors’ contribution

Sajid Shokat, Sukhwinder Singh and Fulai Liu conceive the idea to plan this study. Sajid Shokat conducted an experiment with the help of Sukhwinder. Data analysis was done by Sajid

Shokat and Deepmala Sehgal. Manuscript was written by Sajid Shokat while Deepmala Sehgal and Sukhwinder Singh gave their valuable inputs to improve the manuscript.

Acknowledgement

Authors are grateful to Prashant Vikram for technical guidance to conduct an experiment and highly obliged to Bruno Trevenzoli Favero for technical help to use BLASTN.

References:

1. Food and Agriculture Organisation of the United Nations (FAO) *World Food and Agriculture - Statistical Pocketbook 2018. Rome. 254 pp. Licence: CC BY-NC-SA 3.0 IGO.*; 2018; ISBN 9789251310120.
2. Lobell, D.; Burke, M.; Tebaldi, C.; Mastrandrea, M.; Falcon, W.; Naylor, R. Needs for Food Security in 2030 Region. *Science (80-.).* **2008**, *319*, 607–610.
3. Amiri, R.; Bahraminejad, S.; Jalali-honarmand, S. Effect of terminal drought stress on grain yield and some morphological traits in 80 bread wheat genotypes. **2013**.
4. Giunta, F.; Motzo, R.; Deidda, M. Effect of drought on yield and yield components of durum wheat and triticale in a Mediterranean environment. **1993**, *33*, 399–409.
5. Saeidi, M.; Abdoli, M. Effect of Drought Stress during Grain Filling on Yield and Its Components , Gas Exchange Variables , and Some Physiological Traits of Wheat Cultivars. **2015**, *17*, 885–898.
6. Pingali, P.L. Green Revolution: Impacts, limits, and the path ahead. *Proc. Natl. Acad. Sci.* **2012**, *109*, 12302–12308.
7. Vikram, P.; Franco, J.; Burgueño-Ferreira, J.; Li, H.; Sehgal, D.; Saint Pierre, C.; Ortiz, C.; Sneller, C.; Tattaris, M.; Guzman, C.; et al. Unlocking the genetic diversity of Creole wheats. *Sci. Rep.* **2016**, *6*, 23092.

8. Moore, G. Strategic pre-breeding for wheat improvement. *Nat. Plants* **2015**, *1*, 10–12.

9. Singh, S.; Vikram, P.; Sehgal, D.; Burgueño, J.; Sharma, A.; Singh, S.K.; Sansaloni, C.P.; Joynson, R.; Brabbs, T.; Ortiz, C.; et al. Harnessing genetic potential of wheat germplasm banks through impact-oriented-prebreeding for future food and nutritional security. *Sci. Rep.* **2018**, *8*.

10. Crossa, J.; Burgueño, J.; Dreisigacker, S.; Vargas, M.; Herrera-Foessel, S.A.; Lillemo, M.; Singh, R.P.; Trethowan, R.; Warburton, M.; Franco, J.; et al. Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics* **2007**, *177*, 1889–1913.

11. Neumann, K.; Kobiljski, B.; Denčić, S.; Varshney, R.K.; Börner, A. Genome-wide association mapping: A case study in bread wheat (Triticum aestivum L.). *Mol. Breed.* **2011**, *27*, 37–58.

12. Bordes, J.; Ravel, C.; Jaubertie, J.P.; Duperrier, B.; Gardet, O.; Heumez, E.; Pissavy, A.L.; Charmet, G.; Le Gouis, J.; Balfourier, F. Genomic regions associated with the nitrogen limitation response revealed in a global wheat core collection. *Theor. Appl. Genet.* **2013**, *126*, 805–822.

13. Sukumaran, S.; Dreisigacker, S.; Lopes, M.; Chavez, P.; Reynolds, M.P. Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theor. Appl. Genet.* **2015**, *128*, 353–363.

14. Edae, E.A.; Byrne, P.F.; Haley, S.D.; Lopes, M.S.; Reynolds, M.P. Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. *Theor. Appl. Genet.* **2014**, *127*, 791–807.

15. Lopes, M.S.; Dreisigacker, S.; Peña, R.J.; Sukumaran, S.; Reynolds, M.P. Genetic characterization of the wheat association mapping initiative (WAMI) panel for dissection of complex traits in spring wheat. *Theor. Appl. Genet.* **2015**, *128*, 453–464.

16. Mora, F.; Castillo, D.; Lado, B.; Matus, I.; Poland, J.; Belzile, F.; von Zitzewitz, J.; del Pozo, A.

Genome-wide association mapping of agronomic traits and carbon isotope discrimination in a worldwide germplasm collection of spring wheat using SNP markers. *Mol. Breed.* **2015**, *35*, 1–12.

17. Sehgal, D.; Mondal, S.; Guzman, C.; Garcia Barrios, G.; Franco, C.; Singh, R.; Dreisigacker, S. Validation of Candidate Gene-Based Markers and Identification of Novel Loci for Thousand-Grain Weight in Spring Bread Wheat. *Front. Plant Sci.* **2019**, *10*, 1–12.
18. Ledesma-Ramírez, L.; Solís-Moya, E.; Iturriaga, G.; Sehgal, D.; Reyes-Valdes, M.H.; Montero-Tavera, V.; Sansaloni, C.P.; Burgueño, J.; Ortiz, C.; Aguirre-Mancilla, C.L.; et al. GWAS to Identify Genetic Loci for Resistance to Yellow Rust in Wheat Pre-Breeding Lines Derived From Diverse Exotic Crosses. *Front. Plant Sci.* **2019**, *10*, 1–14.
19. Li, F.; Wen, W.; Liu, J.; Zhang, Y.; Cao, S.; He, Z.; Rasheed, A.; Jin, H. Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biol.* **2019**, *19*, 1–19.
20. Gabriel, S..; Schaffner, S.F.; Nguyen, H.; Moore, J..; Roy, J.; Blumenstiel, B.; Higgins, J.; DeFelice, M.; Lochner, A.; Faggart, M.; et al. The structure of haplotype blocks in the human genome. *Science (80-.).* **2002**, *296*, 2225–2229.
21. Qian, L.; Hickey, L.T.; Stahl, A.; Werner, C.R.; Hayes, B.; Snowdon, R.J.; Voss-Fels, K.P. Exploring and harnessing haplotype diversity to improve yield stability in crops. *Front. Plant Sci.* **2017**, *8*, 1534.
22. Zhang, H.; Li, H.; Yuan, L.; Wang, Z.; Yang, J.; Zhang, J. Post-anthesis alternate wetting and moderate soil drying enhances activities of key enzymes in sucrose-to-starch conversion in inferior spikelets of rice. *J. Exp. Bot.* **2012**, *63*, 215–227.
23. Lorenz, A.J.; Hamblin, M.T.; Jannink, J.L. Performance of single nucleotide polymorphisms versus haplotypes for genome-wide association analysis in barley. *PLoS One* **2010**, *5*, e14079.
24. N'Diaye, A.; Haile, J.K.; Cory, A.T.; Clarke, F.R.; Clarke, J.M.; Knox, R.E.; Pozniak, C.J. Single marker

and haplotype-based association analysis of semolina and pasta colour in elite durum wheat breeding lines using a high-density consensus map. *PLoS One* **2017**, *12*, 1–24.

25. Würschum, T.; Langer, S.M.; Longin, C.F.H.; Korzun, V.; Akhunov, E.; Ebmeyer, E.; Schachschneider, R.; Schacht, J.; Kazman, E.; Reif, J.C. Population structure , genetic diversity and linkage disequilibrium in elite winter wheat assessed with SNP and SSR markers. *Theor. Appl. Genet.* **2013**, *126*, 1477–1486.
26. Zegeye, H.; Rasheed, A.; Makdis, F.; Badebo, A.; Ogbonnaya, F.C. Genome-Wide Association Mapping for Seedling and Adult Plant Resistance to Stripe Rust in Synthetic Hexaploid Wheat. *PLoS One* **2014**, *9*, e105593.
27. Sehgal, D.; Autrique, E.; Singh, R.; Ellis, M.; Singh, S.; Dreisigacker, S. Identification of genomic regions for grain yield and yield stability and their epistatic interactions. *Sci. Rep.* **2017**, *7*, 1–12.
28. Somers, D.J.; Banks, T.; Depauw, R.; Fox, S.; Clarke, J.; Pozniak, C.; Mccartney, C. Genome-wide linkage disequilibrium analysis in bread wheat and durum wheat. *Genome* **2007**, *50*, 557–567.
29. Benson, J.; Brown-guedira, G.; Murphy, J.P.; Sneller, C. Population Structure , Linkage Disequilibrium , and Genetic Diversity in Soft Winter Wheat Enriched for Fusarium Head Blight Resistance. *Plant Genome* **2012**, *5*, 71–80.
30. Chen, X.; Min, D.; Yasir, T.A.; Hu, Y. Genetic Diversity , Population Structure and Linkage Disequilibrium in Elite Chinese Winter Wheat Investigated with SSR Markers. *PLoS One* **2012**, *7*, e44510.
31. Shokat, S.; Azhar, F.M.; Nabi, G.; Iqbal, Q. Estimation of heritability and genetic advance for some characters related to earliness in tomato (Solanum lycopersicum L.). *J. Agric. Res.* **2015**, *53*, 351–356.
32. Pinto, R.S.; Reynolds, M.P.; Mathews, K.L.; McIntyre, C.L.; Chapman, J.-J.O.-V.; Scott, C. Heat and

drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. *Theor. Appl. Genet.* **2010**, *121*, 1001–1021.

33. Alexander, L.M.; Kirigwi, F.M.; Fritz, A.K.; Fellers, J.P. Mapping and Quantitative Trait Loci Analysis of Drought Tolerance in a Spring Wheat Population Using Amplifi ed Fragment Length Polymorphism and Diversity Array Technology Markers. *Crop Sci.* **2012**, *52*, 253–261.
34. Ballesta, P.; Mora, F.; Pozo, A. Del Association mapping of drought tolerance indices in wheat: QTL-rich regions on chromosome 4A. *Sci. Agric.* **2018**, *77*.
35. Weinitschke, S.; Denger, K.; Cook, A.M.; Smits, T.H.M. The DUF81 protein TauE in Cupriavidus necator H16 , a sulfite exporter in the metabolism of C-2 sulfonates The DUF81 protein TauE in Cupriavidus necator H16 , a sulfite exporter in the metabolism of C 2 sulfonates. **2007**, 3055–3060.
36. Krejcík, Z.; Denger, K.; Weinitschke, S.; Hollemeyer, K.; Paces, V.; Cook, A.; Smits, T. Sulfoacetate released during the assimilation of taurine-nitrogen by Neptuniibacter caesariensis : purification of sulfoacetaldehyde dehydrogenase. *Arch. Microbiol.* **2008**, *190*, 159–168.
37. Guo, Z.; Zhao, Y.; Röder, M.S.; Reif, J.C.; Ganal, M.W.; Chen, D.; Schnurbusch, T. Manipulation and prediction of spike morphology traits for the improvement of grain yield in wheat. *Sci. Rep.* **2018**, *8*, 1–10.
38. Mwadzingeni, L.; Shimelis, H.; Tesfay, S.; Tsilo, T.J. Screening of Bread Wheat Genotypes for Drought Tolerance Using Phenotypic and Proline Analyses. **2016**, *7*.
39. Chai, L.; Chen, Z.; Bian, R.; Zhai, H.; Cheng, X.; Peng, H.; Yao, Y. Dissection of two quantitative trait loci with pleiotropic effects on plant height and spike length linked in coupling phase on the short arm of chromosome 2D of common wheat (Triticum aestivum L .). *Theor. Appl. Genet.* **2019**, *132*, 1815–1831.

40. Zhai, H.; Feng, Z.; Li, J.; Liu, X.; Xiao, S.; Ni, Z. QTL Analysis of Spike Morphological Traits and Plant Height in Winter Wheat (Triticum aestivum L .) Using a High-Density SNP and SSR-Based Linkage Map. *PLoS One* **2016**, *7*, 1–13.

41. Zheng, Z.; Scott, S.; Lukas, W.; Webb, M. PREDICTED: Aegilops tauschii subsp. tauschii heat shock 70 kDa protein BIP5-like (LOC109764454), mRNA. *A greedy algorithm aligning DNA Seq.* 2000.

42. Ji, X.; Shiran, B.; Wan, J.; Lewis, D.C.; Jenkins, C.L.D.; Condon, A.G.; Richards, R.A.; Dolferus, R. Importance of pre-anthesis anther sink strength for maintenance of grain number during reproductive stage water stress in wheat. **2010**, 926–942.

43. Rajala, A.; Hakala, K.; Mäkelä, P.; Peltonen-Sainio, P. Drought Effect on Grain Number and Grain Weight at Spike and Spikelet Level in Six-Row Spring Barley. *J. Agron. Crop Sci.* **2011**, *197*, 103–112.

44. Shokat, S.; Großkinsky, D.K.; Roitsch, T.; Liu, F. Activity of leaf and spike carbohydrate-metabolic and antioxidant enzymes reveals yield performance in three spring wheat genotypes grown under well-watered and drought conditions. *Front. Plant Sci.* **2020**, *(under rev*.

45. Pradhan, S.; Robbins, K.; Bai, G.; Mason, R.E.; Khan, J.; Shahi, D.; Avci, M.; Guo, J.; Hossain, M.M.; Bhatta, M.; et al. Understanding the Genetic Basis of Spike Fertility to Improve Grain Number , Harvest Index , and Grain Yield in Wheat Under High Temperature Stress Environments. *Front. Plant Sci.* **2019**, *10*, 1–13.

46. Tanaka, T.; Kobayashi, F.; Joshi, G.P.; Onuki, R.; Sakai, H.; Kanamori, H.; Wu, J.; Šimková, H.; Nasuda, S.; Endo, T.R.; et al. Next-Generation Survey Sequencing and the Molecular Organization of Wheat Chromosome 6B. *DNA Res.* **2014**, *21*, 103–114.

47. Nadolska-Orczyk, A.; Izabela, O.; Wacław, K.R. Major genes determining yield-related traits in wheat and barley. *Theor. Appl. Genet.* **2017**, *130*, 1081–1098.

48. Huo, N.; Zhang, S.; Zhu, T.; Dong, L.; Wang, Y.; Mohr, T.; Hu, T.; Liu, Z.; Dvorak, J.; Luo, M.C.; et al. Gene duplication and evolution dynamics in the homeologous regions harboring multiple prolamin and resistance gene families in hexaploid wheat. *Front. Plant Sci.* **2018**, *9*, 1–13.

49. Awasthi, J.P.; Chandra, T.; Mishra, S.; Parmar, S.; Shaw, B.P.; Nilawe, P.D.; Chauhan, N.K.; Sahoo, S.; Panda, S.K. Plant Physiology and Biochemistry Identification and characterization of drought responsive miRNAs in a drought tolerant upland rice cultivar KMJ 1-12-3. *Plant Physiol. Biochem.* **2019**, *137*, 62–74.

50. Ogihara, Y.; Kawaura, K.; Mishina, K.; Tarui, H.; Kawai, J. *Functional partitioning of the expressed homoeoloci based on the fully collected cDNAs of allohexaploid wheat*; Contact:Kohei Mishina Yokohama City University, Kihara Institute for Biological Research; Maioka-cho 641-12, Yokohama, Kanagawa 244-0813, Japan, 2016;

51. Chiba, Y.; Shimizu, T.; Miyakawa, S.; Kanno, Y.; Koshiba, T.; Kamiya, Y.; Seo, M. Identification of Arabidopsis thaliana NRT1/PTR FAMILY (NPF) proteins capable of transporting plant hormones. *J. Plant Res.* **2015**, *128*, 679–686.

52. Tal, I.; Zhang, Y.; Jørgensen, M.E.; Pisanty, O.; Barbosa, I.C.R.; Zourelidou, M.; Regnault, T.; Crocoll, C.; Erik Olsen, C.; Weinstain, R.; et al. The Arabidopsis NPF3 protein is a GA transporter. *Nat. Commun.* **2016**, *7*.

53. Corratgé-faillie, C.; Lacombe, B. Substrate (un) specificity of Arabidopsis NRT1 / PTR FAMILY (NPF) proteins. *J. Exp. Bot.* **2017**, *68*, 3107–3113.

54. Li, S.; Li, X.; Wei, Z.; Liu, F. ABA-mediated modulation of elevated CO2 on stomatal response to drought. *Curr. Opin. Plant Biol.* **2019**, *13*, 1–7.

55. Wu, W.; Zhou, L.; Chen, J.; Qiu, Z.; He, Y. GainTKW: A Measurement System of Thousand Kernel Weight Based on the Android Platform. *Agronomy* **2018**, *8*, 178.

56. Duggan, B.L.; Domitruk, D.R.; Fowler, D.B. Yield component variation in winter wheat grown under drought stress. *Can. J. Plant Sci.* **2000**, *80*, 739–745.

57. Liu, Y.; Bowman, B.; Hu, Y.-G.; Liang, X.; Zhao, W.; Wheeler, J.; Klassen, N.; Bockelman, H.; Bonman, J.; Chen, J. Evaluation of Agronomic Traits and Drought Tolerance of Winter Wheat Accessions from the USDA-ARS National Small Grains Collection. *Agronomy* **2017**, *7*, 51.

58. Lehnert, H.; Serfling, A.; Friedt, W.; Ordon, F. Genome-Wide Association Studies Reveal Genomic Regions Associated With the Response of Wheat (Triticum aestivum L .) to Mycorrhizae Under Drought Stress Conditions. *Front. Plant Sci.* **2018**, *9*.

59. LujánBasile, S.M.; Ramírez, I.A.; Crescente, J.M.; Conde, M.B.; Demichelis, M.; Abbate, P.; Rogers, W.J.; Pontaroli, A.C.; Helguera, M.; Leonardo Sebastián Vanzetti Haplotype block analysis of an Argentinean hexaploid wheat collection and GWAS for yield components and adaptation. *BMC Plant Biol.* **2019**, *19*, 1–16.

60. Rustenholz, C.; Choulet, F.; Laugier, C.; Šafář, J.; Šimková, H.; Doležel, J.; Magni, F.; Scalabrin, S.; Cattonaro, F.; Vautrin, S.; et al. A 3,000-Loci transcription map of chromosome 3B unravels the structural and functional features of gene islands in hexaploid wheat. *Plant Physiol.* **2011**, *157*, 1596–1608.

61. Sakai, H.; Lee, S.S.; Tanaka, T.; Numa, H.; Kim, J.; Kawahara, Y.; Wakimoto, H.; Yang, C.; Iwamoto, M.; Abe, T.; et al. Rice Annotation Project Database (RAP-DB): An Integrative and Interactive Database for Rice Genomics Special Focus Issue – Databases. **2013**, *54*, 1–11.

Table-1 Significance level for the mean squares of different plant traits

Factors	NDVI	CTD	PH	DTH	DTM	SL	NGS	KA	TKW	Y
Genotypes	ns	ns	<0.001	ns	<0.001	<0.001	ns		<0.001	<0.001
Year			<0.001	<0.001	<0.001					<0.001
Treatment	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Genotypes:Year			<0.001	ns	<0.001					<0.05
Genotypes:Treatment	<0.05	ns	ns	ns	ns	ns	ns		<0.001	ns
Year:Treatment			<0.001	<0.001	<0.001					<0.001
Genotypes:Year:Treatment			ns	ns	ns					ns
Heritability broad sense (h_{bs}) Irrigated	0.25	0.13	0.81	0.77	0.85	0.80	0.72		0.84	0.61
Heritability broad sense (h_{bs}) Drought	0.68	0.22	0.71	0.08	0.82	0.72	0.6	0.17	0.67	0.65

ns = non-significant

NDVI = Normalized difference vegetation index, CTD = Canopy temperature depression, PH = Plant height, DTH = Days to heading, DTM = Days to maturity, SL = Spike length, NGS = Number of grains, KA = Kernel abortion, TKW = Thousand kernel weight and Y = Grain yield

Table-2 Mix linear model (MLM) describing the location of different traits on chromosomes

Year	Treatment	Trait	Marker	Chromosome	Position	Log10	P-value	R^2 (%)

2018	Irrigated	SL	1216938	3D	309140	5.46	3.43e-06	8.2
	Drought		1216938	3D	309140	2.39	0.0041	2.8
2018	Irrigated	SL	1113306	2D	2925497	3.32	4.76e-04	4.8
	Drought		1046144	2D	2178410	1.98	0.01047	2.1
2018	Irrigated	NGS	3064504	6B	45980	2.46	0.00346	2.9
	Drought		2256902	6B	1091090	2.54	0.00287	2.9
2018	Irrigated	NGS	1079850	3A	932130	2.12	0.00761	2.3
	Drought		1190017	3A	627301	2.27	0.00532	2.6
2018	Drought	KA	4910194	3B	2786440	2.78	0.00167	3.4
	Drought	KA	1076524	3A	1046210	2.71	0.0019	3.1
2016	Irrigated	TKW	2260800	3B	1334838	2.04	0.0096	2.5
2018	Irrigated	TKW	3064863	3B	1041390	3.02	9.61e-04	3.7
	Drought		988433	3B	1176780	3.08	8.36e-04	3.9
2016	Irrigated	Y	993866	6B	323301	1.99	0.01024	2.1
	Drought		1219677	6B	909568	2.47	0.0034	2.9
2018	Irrigated		3064504	6B	45980	3.28	5.22e-04	4.1

	Drought		1219677	6B	909568	4.57	2.70e-05	6.1
2016	Irrigated	Y	1101137	4A	457217	2.66	0.0022	3.6
	Drought		1041102	4A	0	3.40	3.94e-04	5.0
2018	Irrigated		1091070	4A	465850	2.64	0.0023	3.2
	Drought		1041102	4A	0	2.30	0.00502	2.4

SL = Spike length, NGS = Number of grains, KA = Kernel abortion, TKW = Thousand Kernel Weight and Y = grain yield