Genome-wide association study for yield components in spring wheat collection harvested under two water regimes in Northern Kazakhstan (#59889)

First submission

Guidance from your Editor

Please submit by **31 May 2021** for the benefit of the authors (and your $200 publishing discount) .

✓ **Structure and Criteria**
Please read the 'Structure and Criteria' page for general guidance.

▦ **Raw data check**
Review the raw data.

🖼 **Image check**
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files Download and review all files from the materials page.	6 Figure file(s) 6 Table file(s) 2 Raw data file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

🗋 You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- ✓ Clear, unambiguous, professional English language used throughout.
- ✓ Intro & background to show context. Literature well referenced & relevant.
- ✓ Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- ✓ Figures are relevant, high quality, well labelled & described.
- ✓ Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- ✓ Original primary research within Scope of the journal.
- ✓ Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- ✓ Rigorous investigation performed to a high technical & ethical standard.
- ✓ Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- ℹ Impact and novelty not assessed. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- ✓ All underlying data have been provided; they are robust, statistically sound, & controlled.
- ✓ Speculation is welcome, but should be identified as such.
- ✓ Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Genome-wide association study for yield components in spring wheat collection harvested under two water regimes in Northern Kazakhstan

Akerke Amalova [1, 2] , **Saule Abugalieva** [1, 2] , **Adylkhan Babkenov** [3] , **Sandugash Babkenova** [3] , **Yerlan Turuspekov** [Corresp. 1, 4]

[1] Institute of Plant Biology and Biotechnology, Almaty, Казахстан
[2] Faculty of Biology and Biotechnology, al-Farabi Kazakh National University, Almaty, Казахстан
[3] A.I. Barayev Research and Production Centre of Grain Farming, Shortandy, Aqmola region, Казахстан
[4] Faculty of Agrobiology, Kazakh National Agrarian University, Almaty, Казахстан

Corresponding Author: Yerlan Turuspekov
Email address: yerlant@yahoo.com

Background. Bread wheat is the most important cereal in Kazakhstan, where it is grown on over 12 million hectares. One of the major constraints in the grain yield of wheat is a drought caused by a limited water supply. Hence, the development of drought-resistant cultivars will be critical for ensuring food security in this country. Therefore, identifying quantitative trait loci (QTLs) associated with drought tolerance is an essential step in modern breeding activities that rely on a marker-assisted selection approach. **Methods.** In this study, the collection of 179 spring wheat accessions was tested under irrigated and rainfed conditions in Northern Kazakhstan using nine traits; heading date (HD), seed maturity date (SMD), plant height (PH), peduncle length (PL), umber of productive spikes (NPS), spike length (SL), number of kernels per spike (NKS), thousand kernels weight (TKW), Kernels yield per m^2 (YM2), during three years (2018-2020). The collection was genotyped using a 20000 (20K) Illumina iSelect SNP array, and 8662 polymorphic SNP markers were selected for genome-wide association study (GWAS) to identify QTLs for targeted agronomic traits. **Results.** Fifty stable QTLs out of 237 total revealed QTLs were identified for irrigated and rainfed conditions in the Akmola region, Northern Kazakhstan, by studying eight traits: HD, SMD, PL, SL, NPS, NKS, TKW, and YM2. The study suggested that 9 QTLs for HD and 11 QTLs for SMD were presumably novel genetic factors identified in irrigated and rainfed conditions of Northern Kazakhstan. dentified SNP markers of QTLs for targeted traits can be deployed to develop new competitive spring wheat cultivars in dry arid zones using a marker-assisted selection approach.

Genome-wide association study for yield components in spring wheat collection harvested under two water regimes in Northern Kazakhstan

Akerke Amalova [1, 2], Saule Abugalieva [1, 2], Adylkhan Babkenov [3],
Sandugash Babkenova [3], Yerlan Turuspekov [1, 4]

1. Institute of Plant Biology and Biotechnology, Almaty, Kazakhstan
2. Faculty of Biology and Biotechnology, Al-Farabi Kazakh National University, Almaty, Kazakhstan
3. A.I. Barayev Research and Production Centre of Grain Farming, Shortandy, Kazakhstan
4. Faculty of Agrobiology, Kazakh National Agrarian University, Almaty, Kazakhstan

Corresponding Author:
Yerlan Turuspekov[1]
Timiryazev street 45, Almaty 050040, Kazakhstan
Email address: yerlant@yahoo.com

Abstract

Background. Bread wheat is the most important cereal in Kazakhstan, where it is grown on over 12 million hectares. One of the major constraints in the grain yield of wheat is a drought caused by a limited water supply. Hence, the development of drought-resistant cultivars will be critical for ensuring food security in this country. Therefore, identifying quantitative trait loci (QTLs) associated with drought tolerance is an essential step in modern breeding activities that rely on a marker-assisted selection approach.

Methods. In this study, the collection of 179 spring wheat accessions was tested under irrigated and rainfed conditions in Northern Kazakhstan using nine traits; heading date (HD), seed maturity date (SMD), plant height (PH), peduncle length (PL), umber of productive spikes (NPS), spike length (SL), number of kernels per spike (NKS), thousand kernels weight (TKW), kernels yield per m^2 (YM2), during three years (2018-2020). The collection was genotyped using a 20000 (20K) Illumina iSelect SNP array, and 8662 polymorphic SNP markers were selected for genome-wide association study (GWAS) to identify QTLs for targeted agronomic traits.

Results. Fifty stable QTLs out of 237 total revealed QTLs were identified for irrigated and rainfed conditions in the Akmola region, Northern Kazakhstan, by studying eight traits: HD, SMD, PL, SL, NPS, NKS, TKW, and YM2. The study suggested that 9 QTLs for HD and 11 QTLs for SMD were presumably novel genetic factors identified in irrigated and rainfed conditions of Northern Kazakhstan. Identified SNP markers of QTLs for targeted traits can be deployed to develop new competitive spring wheat cultivars in dry arid zones using a marker-assisted selection approach.

Keywords: common wheat; drought tolerance; quantitative trait loci; GWAS; marker-trait associations

Introduction

Bread wheat (*Triticum aestivum* L.) is the most important cereal crop in Kazakhstan, where it is grown on over 12 million hectares. Kazakhstan produces up to 20 million tons of common wheat per year annually and exports up to 5-7 million tons of the grain (*USDA, 2018*). The country is typically growing spring type of wheat with over 80% of the sowing area under this crop is producing in Northern territories (http://stat.gov.kz). The average grain yield is around 1.1 tons per hectare and is constrained by abiotic stresses, including drought (*Shiferaw et al., 2013*). Drought is prevalent stress in spring wheat production in Kazakhstan's northern territories, and frequencies of severe drought are ranging from 15% in Kostanai to 22% in Akmola regions in average data for 1971-2011 (*Patrick, 2017*). Climate change observed over the past 20 years has significantly increased the importance of this challenging factor for crop management (*Skirycz & Inzé, 2010*). Still, due to high grain quality, the bread wheat is the major agricultural commodity in the country and preferable choice of farmers over other crops (*Abugaliyeva & Savin, 2018*). Although additional supplying crops with water can solve the problem of drought, this can lead to a substantial increase in growth costs. Therefore, the development of drought-resistant cultivars that show good productivity and high grain quality under stress will be critical for ensuring food security in the future (*Foley et al., 2011*). Drought has a significant influence on wheat's physiological functions, such as the closure of stomata, a decrease of photosynthesis, development of oxidative stress, production of toxic metabolites, etc. (*Bray, 2002*). All of these changes in plant physiology lead to decreased plant height, reduced total shoot length, diminished growth rates, decreased number of tillers, reduced relative water content (*Nezhadahmadi et al., 2013*), a decline of various grain quality parameters (*Tsenov et al., 2015*), and, finally, substantial yield losses (*Zhang et al., 2018*).

Drought stress affects gene expression, and detection of quantitative trait loci (QTL for single and QTLs for plural) for crops growing in water stress conditions is crucial to observe their responses. Various genes involved in plant drought response were distinguished and described earlier (*Ingram & Bartels, 1996; Agarwal et al., 2006; Wei et all., 2009; Huseynova et al., 2013; Nezhadahmadi et al., 2013; Hassan et al., 2015; Kulkarni et al., 2017*).

In wheat, several known specific genetic factors are responsible for drought stress tolerance, including late embryogenesis abundant, responsive to abscisic acid, rubisco, proline, etc. (*Nezhadahmadi et al., 2013*). In addition, there are several transcription factor families, such as ERF (ethylene response factors), DREB (dehydration responsive element binding), ZFP (zinc finger proteins), that were found to be associated with drought in wheat (*Agarwal et al., 2006; Kulkarni et al., 2017*). At the same time, information available on drought-responsive genes is still limited as their roles have not been thoroughly determined (*Bray, 2002; Nezhadahmadi et al., 2013*). One of the critical aspects in capturing important genes associated with drought tolerance is a strong influence of the growth environment in which yield QTLs are identified with significant genotype x environment interaction (GEI). For instance, results obtained from three different genome-wide association studies (GWAS) related to the identification of QTLs for yield performance in Europe (*Guo et al., 2017*), India (*Jaiswal et al., 2016*), and Mexico (*Sukumaran et al., 2015*) showed different responses and QTLs for yield components revealed in different parts of the wheat genome. The sensitivity of plants to environmental factors at crucial growth phases, which determines the tolerance to stressful factors and potential yield, can explain this outcome (*Reynolds et al., 2002*). Therefore, regional projects' success may largely depend on local GWAS using adapted germplasm and lead to the discovery of new genetic factors that prove drought tolerance and high yield potential in a given environment.

The discovery of new genes for specific agronomic traits became feasible after recent breakthrough achievement in SNP (single nucleotide polymorphism) arrays development (*Cavanagh et al., 2013, Wang et al., 2014; Boeven et al., 2016; Sun et al., 2020*). The availability of high-throughput SNP arrays led to a massive genotyping of wheat germplasm collections (*Allen et al., 2011; Wang et al., 2014; Sun et al., 2020*), including accessions from

Kazakhstan (*Turuspekov et al., 2017*). Hence, new molecular tools provided rich opportunities for discovering marker-trait associations (MTA) for agronomic traits via GWAS of wheat in different parts of the world (*Rahimi et al., 2019; Tsai et al., 2020*), including Kazakhstan (*Turuspekov et al., 2017; Anuarbek et al., 2020; Genievskaya et al., 2020*). Previously, a number of QTLs associated with various drought resistance traits in wheat had been identified using linkage mapping (*Quarrie et al., 2005; Verma et al., 2004; Tura et al., 2020*) and association mapping via GWAS (*Sukumaran et al., 2015; Li et al., 2019; Lin et al., 2019; Mathew et al., 2019*). The current work is the first attempt to identify drought resistance-associated QTLs under irrigated and rainfed conditions in Northern Kazakhstan by using GWAS. As Northern Kazakhstan is the region where more than 80% of the wheat-growing area is concentrated in the country, the study is of great importance for breeding programs to improve wheat germplasm.

Materials and methods

Phenotyping of the collection under irrigated and rainfed conditions. A collection of 179 spring bread wheat accessions included 92 commercial and prospective cultivars of Kazakhstan and Russia, 86 breeding lines from A.I. Barayev Scientific-Production Center for Grain Farming (SPCGF, Shortandy, Akmola region), and a check cultivar for the Akmola region – Tselinnaya yubileinaya (TY, Table S1). Field experiments were conducted on the experimental plots of SPCGF (51˚36′09″N and 71˚02′24″E, 391 m above sea level) in 2018 – 2020, both under irrigated and non-irrigated (rainfed) conditions. Each accession was grown in 1 m^2 randomized plots composed of 7 rows with 50 seeds per row in two repetitions. Irrigation (45 mm) was applied at two critical stages: tillering and booting. The meteorological data registered for experiments were shown in Raw data.

The list of nine agronomic traits associated with drought resistance and grain productivity and used for the phenology and phenotyping included the heading date (HD, days), the seed maturity date (SMD, days), plant height (PH, cm), peduncle length (PL, cm), number of productive spikes (NPS, pcs), spike length (SL, cm), number of kernels per spike (NKS, pcs), thousand kernels weight (TKW, g), kernels yield per m^2 (YM2, g/m^2).

Genotyping of the collection. Genomic DNA samples were extracted from a single seedling of each individual accession using the cetyltrimethylammonium bromide (CTAB) method (*Doyle & Doyle, 1990*). The DNA concentration for each sample was adjusted to 30 ng/μL. All samples of 179 wheat accessions were genotyped using a 20000 (20K) Illumina iSelect SNP array at the TraitGenetics Company (TraitGenetics GmbH, Gatersleben, Germany). A total of 8662 polymorphic SNP markers have been selected for GWAS using the criteria (*Miyagawa et al., 2008*). According to these criteria, markers with a call rate ≥ 90%, Hardy–Weinberg equilibrium fit at $P \geq 0.001$, a confidence score of 0.5, and minor allele frequency (MAF) ≥5% were considered to meet the requirements. Accessions with greater than 15% missing data were also removed.

Analysis of linkage disequilibrium, kinship, population structure, and statistics. The linkage disequilibrium (LD) in the studied collection was calculated for each of hexaplopid common wheat genome separately (genome A, genome B, genome D), as well as an average LD for three genomes using the Java-based software TASSEL v.5.2.53 (*Bradbury et al., 2007*). R statistical platform was used to build a plot between pairwise R^2 and the genetic distance (LD decay plot) (*RStudio Team, 2015*). TASSEL was also applied to calculate a population kinship matrix (Kin) based on the scaled identity by state (IBS) method (*Stevens et al., 2011*).

A model-based clustering method (admixture models with correlated allele frequencies) in the software STRUCTURE v.2.3.4 (*Pritchard, Stephens & Donnelly, 2000*) was used to study the population structure of the entire collection. Five runs were conducted for each K ranged from 2 to 10 with a 100,000 burn-in length and 100,000 Markov chain Monte Carlo (MCMC) iterations. The optimal number of clusters (K) was chosen based on the ΔK as described by *Evanno, Regnaut & Goudet, 2005*. Obtained values were then transformed into a population structure (Q) matrix.

The correlation analysis was calculated using the Rstudio software (*RStudio Team, 2015*). The statistics for yield trials were assessed using GraphPad Prism Version 9.0 (*GraphPad Prism, 2021*). The GEI was analyzed by using the GGE (genotype and genotype-environment) biplot method and the Finlay and Wilkinson (FW) regression analysis in GenStat software Version 19.1 (*VSN International,2020*).

Marker-trait association analysis. The TASSEL and mixed linear model (MLM) method with the application of K and Q matrices was used for the identification of QTL for agronomic traits both under irrigated and rainfed conditions. The analysis was based on phenotypic data for nine traits obtained from field trials in 2018 - 2020 and their average values over two years. $P < 1E\text{-}03$ was used as a significant threshold for identified MTAs. Positions and sequences of SNP markers were obtained from the 90K Array Consensus map set of the common wheat genome (*Wang et al., 2014*). For confirmation of the correction due to K and Q matrices, the distribution lines in each quantile-quantile (QQ) plot were analyzed. In the case of several significant MTAs positioned closely to each other, the SNP with the lowest P-value was chosen. For the search of protein-coding genes that overlap with identified significant MTAs, each marker's sequence was inserted into the BLAST tool of Ensembl Plants (*Ensembl Plants, 2020*) and compared with the reference genome of *T. aestivum*. MapChart v.2.32 (*Voorrips, 2002*) was used to draw a genetic map.

Results

1. Phenotypic Variation and Correlation Analysis

The field performance of the 179 local spring wheat accessions was analyzed at the SPCGF (Northern Kazakhstan) under irrigated and rainfed conditions during three field seasons in 2018 - 2020 years. The two-tailed *t-test* suggested that average values in all nine studied traits were significantly different between tested irrigated and rainfed conditions. The average PH values showed the largest difference ($P<0.0001$) between the two tested conditions (Table.1), 73.6 cm in irrigated in contrast to 61.6 cm in rainfed conditions.

Table 1. The significance of differences between irrigated and rainfed trials using average data in nine traits based on a two-tailed t-test.

On average, YM2 declined by 5.7% under the rainfed conditions (332.3±5.68 g/m^2) compared to the irrigated (352.3±4.30 g/m^2) conditions. In total, 51 accessions exceeded the YM2 of the local standard cultivar "Tselinnaya yubileinaya" (TY, 374.5 g/m^2) under rainfed conditions, including nine accessions that outperformed the standard also under irrigated conditions (Fig. 1.A). The Finlay and Wilkinson (FW) regression analysis (Fig. 1.B) suggested that YM2 of four wheat accessions, particularly WS10, WS32, WS82, and WS85, was stable in all three tested years (2018-2020), out of nine accessions highlighted in the box in Figure 1.A, showing the YM2 with 400 g/m^2 and higher.

Figure 1. The average yield performance of best accessions under rainfed conditions. **A.** The list of 51 spring wheat accessions outperformed the local check cultivar “Tselinnaya yubileinaya” (TY) on average yield per square meter (YM2) under rainfed conditions. Accessions in red color also outperformed TY on average YM2 under irrigated conditions as well. The nine accessions selected in the box with the highest average YM2 were analyzed using Finlay and Wilkinson (FW) regression analysis. **B.** The FW test is suggesting a different level of stability in four out of nine selected spring wheat accessions in the box of Fig. 1.A.

The analysis of averaged YM2 using scattered GGE biplot indicated that 52.8% of the total variance was explained by Principal Coordinate 1 (PC1), and 47.2% by PC2 (Fig.2). PC1 has effectively separated accessions that showed the highest yield performance in irrigated and rainfed conditions, while PC2 has split the entire collection into groups with higher and lower YM2 for both conditions. The GGE biplot graph has essentially confirmed the results shown in Fig 1.A and exposed the accessions with high average YM2 under irrigated and rainfed trials as well as lines that showed high yield under both conditions (f.i., WS93 and WKZ19).

Figure 2. The scattered GGE biplot graph on averaged yield per squared meter (YM2) data in the collection of 179 common wheat accessions tested in irrigated and rainfed conditions (2018-2020). Green and blue colors indicated Genotype and Environment scores, respectively.

Pearson's correlation assessment in both conditions indicated that average YM2 was positively correlated with NPS and TKW (Fig. 3). Interestingly, the highest correlation value of the YM2 under irrigated condition was with NPS (0.39), while under the rainfed condition with NKS (0.36). It was shown that earlier HD is advantageous for higher TKW under rainfed conditions, while it was not a significant factor for the yield under irrigated conditions. At the rainfed conditions, the higher PH value was not a contributing factor to YM2 (Fig. 3, A). Interestingly, under the rainfed condition, the HD influenced both SL and NKS (Fig. 3). Expectedly, the PL was highly correlated with PH ($P<0.0001$), while the former one has negatively associated with NPS (Fig. 3. A).

Figure 3. Correlation analysis for nine agronomic traits analyzed in the collection of 179 spring wheat accessions tested in rainfed (A) and irrigated (B) conditions. Blue color is positive correlation, and red color is negative correlation.

2. Genetic map, population structure, and LD

The DNA genotyping data for studied 179 spring wheat accessions based on the use of 20K SNP Illumina array has allowed the identification of 8662 polymorphic SNP markers. The total map length for those 8662 SNPs was 3407.6 cM, with an average chromosome length of 162.2 cM. The density of markers in chromosomes varied from 0.1 SNP per cM (chromosome 4D) to 1.7 SNP per cM (chromosome 6B).

Based on the results of population analysis of STRUCTURE for 179 accessions wheat genotypes and STRUCTURE Harvester analyses showed that ΔK was optimal at K=3. LD decay was at 18.5 cM (genome A), 13.1 cM (genome B) and 53.8 cM (genome D) in different genomic regions with a genome-wide LD decay of 5.0 cM at 0.1 R^2 (Fig. S1).

3. Marker-trait associations under irrigated and rainfed conditions.

The phenotypic data for nine agronomic traits of 179 wheat accessions harvested under rainfed and irrigated conditions were subjected to GWAS using the 8662 polymorphic SNP markers. Fifty stable QTLs out of 237 total QTLs were identified for irrigated and rainfed conditions in the Akmola region, Northern Kazakhstan for HD, SMD, PH, PL, SL, NPS, NKS, and TKW, while no QTL were detected for the PTI (Tables 2, Table S2, S3). The highest number of stable QTLs was localized on chromosomes of genome B (26), followed by genomes A (16) and D (8), respectively. In general, 25 QTLs were identified in both conditions, rainfed and irrigated (Table 2, Fig. S2).

Table. 2. The list of quantitative trait loci (QTLs) for eight studied traits identified using 179 spring wheat accessions tested under irrigated and rainfed conditions of Northern Kazakhstan (2018- 2020)

In total, eleven stable QTLs were identified for HD (Table 2). Three of them were detected only in rainfed conditions, while eight QTLs were found in the irrigated site. One of the QTLs that was common for both tested conditions (*QHD.ta.ipbb-3B*) has also affected NKS (Table 2). The largest number of QTLs identified in the rainfed trials were associated with SMD (8 out of 12 total QTLs). PH and PL were one of the key traits to this study, as two tested sites highly significantly differed in these traits (P<0.0001). However, only two stable QTLs (QPLtaipbb-3B1 and QPLtaipbb-3B2) were identified for PL, and no stable QTL was identified for PH. The small number of QTL found under both conditions were YM2 (two QTL), NKS and SL (5 QTLs), and NPS (6 QTLs). However, while all seven QTLs for TKW were identified on both tested sites, 2 QTLs for detected only in rainfed (*QTKW.ta.ipbb-7B*) and only irrigated (*QTKW.ta.ipbb-7A*) conditions (Table 2).

Among identified 46 QTLs, seven pleiotropic MTAs were detected under both conditions. Those pleiotropic MTAs were mapped on the 1B, 1D, 2B, 3B, 5A, 6A, and 7A chromosomes and associated with HD (*wsnp_Ex_c8240_13914674, Excalibur_c20376_615*) and spike-related traits (NPS, NKS, and TKW) (Table 3, Fig. S2).

Discussion

1. Yield assessment in the spring wheat collection under irrigated and rainfed conditions in Northern Kazakhstan

The analysis of averaged YM2 revealed that under rainfed conditions, more than 51 common wheat accessions exceeded the YM2 of the local standard cultivar "TY" in Northern Kazakhstan (Fig. 1.A). Hence, cultivars and promising lines in the studied collection have high variability and potential for usage in breeding projects in water-limited environments. Particularly, nine accessions showed an outstanding yield performance in comparison to the local check cultivar "TY", and four of those lines (WS10, WS32, WS82, and WS85) displayed the grain yield stability across all three years (Fig. 1.B). The GGE biplot for average YM2 has firmly confirmed this result and also revealed the accessions with a high yield potential both in trials under irrigated and rainfed conditions (Fig. 2). Notably, the scatter method of the GGE biplot suggested that two main principal coordinates provided 52,8% (PC1) and 47,2% (PC2) of total variability, reinforcing the assumption in a high diversity of the tested collection.

The analysis of the collection in two tested conditions suggested substantial differences between irrigated and rainfed fields (Table 1), indicating that limitation in water supply has significantly affected all nine analyzed traits in the study. Particularly, the biggest difference between two sites was PH, which congruent with results reported earlier (*Tsenov et al., 2015, Lehar et al., 2019, Pour-Aboughadareh et al., 2020*). The correlation analysis for traits under rainfed conditions suggested a negatively significant correlation between HD and TKW, an important yield component, hinting that early HD might be favorable for better yield under stressed conditions (Fig. 3A).

2. Comparative QTL identification for agronomic traits in irrigated and rainfed conditions using GWAS

The GWAS analysis of the wheat collection grown under irrigated and rainfed conditions in the Akmola region of North Kazakhstan allowed the identification of 50 stable out of 237 total QTLs that were significant for eight studied traits (HD, SMD, PL, NPS, SL, NKS, TKW, and YM2) in both conditions (Table 2, Fig.S2). In total, 25 common QTLs for these traits were identified in both conditions (Fig. 4). Thirteen QTLs under rainfed conditions were revealed for HD (4), SMD (8), and TKW (1) that were not detected in trials at the irrigated conditions. Twelve QTLs have been identified only under irrigated conditions for the following traits: SMD (2), NPS (6), NKS (1), TKW (1), and YM2 (2).

Figure 4. Number of QTLs identified under irrigated, rainfed, and in both conditions in Northern Kazakhstan in 2018-2020.

The assessment of QTLs for phenological traits (HD and SMD) showed that only three out of twenty-three QTLs were having effects with longer days. The remaining twenty associations were with QTL effects with shorter flowering and seed maturation time (Table 2). Two QTLs for HD, *QHD.ta.ipbb-2B* and *QHD.ta.ipbb-3B,* have shown pleiotropic effects for NPS, NKS, and TKW, respectively. Also, an SNP marker in QTL for SMD, *QSMD.ta.ipbb-2A,* was also significant in QTL for TKW on chromosome 2A (Table 2). Interestingly, despite the sharp differences between the PH in irrigated and rainfed conditions, no QTL was detected in GWAS for this trait. Evidently, the collection's accessions were fixed for this trait, and low variation within studied conditions was not enough to identify any MTA.

Total eight QTLs were identified in irrigated and rainfed trials for spike-related traits SL and NKS. Interestingly, seven out of those eight QTLs were revealed both in irrigated and rainfed conditions, and three QTLs were detected only under rainfed trials (Table 2). The *QNKS.ta.ipbb-3B*, one of those QTLs identified only in irrigated conditions, should be particularly highlighted, as it was the factor with the highest QTL effect for the trait. Still, this QTL effect is rather the result of the pleiotropic effect of the QTL for HD, which was identified under both conditions (Table 2). TKW is known as one of the major yield components in wheat (*Quarrie et al., 2005*). In this study, eight out of a total of nine QTLs for TKW were detected using trials in rainfed conditions (Table 2). Particularly, *QTKW.ta.ipbb-1B* and *QTKW.ta.ipbb-2B* have shown equally high QTL effects for TKW.

A comparison of the mapped QTLs analyzed in this study with those from other previous studies indicated that ten QTLs matched known associations. For TKW, the two QTLs *QTKW.ta.ipbb-2B.1* and *QTKW.ta.ipbb-6A* were genetically mapped close to genomic regions to QTLs for this trait identified by Tura and his co-authors (*Tura et al., 2020*). Another two associations (*QNPS.ta.ipbb-1B*, *QYM2.ta.ipbb-7B*) were identical with the genetic positions of QTLs identified with the analyses of six traits using GWAS based on the assessment of spring wheat in Kazakhstan (*Turuspekov et al., 2017*). *QTKW.ta.ipbb-5A* was identified in the same genetic position as in the GWAS of 285 elite spring wheat lines of wheat association mapping initiative population grown in temperate irrigated environments (*Sukumaran et al., 2015*). The only QTL for TKW that seems to be putatively novel is *QTKW.ta.ipbb-7B*, which was identified under rainfed conditions and presumably was not reported earlier. In respect to other identified MTAs of yield components, it is worthy to note that *QNPS.ta.ipbb-2B, QNPS.ta.ipbb-5A, QSL.ta.ipbb-2D, QTKW.ta.ipbb-6A* were located in the close proximities of QTLs for the same traits in the study of UK reference mapping population Avalon x Cadenza in the Northern, Central, and Southern regions of Kazakhstan (*Amalova et al., 2021*).

The assessment of the identified QTLs only in rainfed conditions suggested that 12 out of 13 QTLs were associated with phenological traits HD and SMD (Fig. 4), which underlined the importance of plant growth-related traits in avoiding water deficiency stress. Particularly, *QHD.ta.ipbb-6A* and *QHD.ta.ipbb-6D* for HD, and seven QTLs for SMD (from chromosomes

3B to 6A), were essential for early flowering and seed maturation time under rainfed conditions (Table 2). The locations of QTLs for HD compared with known flowering genes showed that the position of the most important SNP for the *QHD.ta.ipbb-5A.1* has completely coincided with the physical position of the *Vrn1* (587,4 Mb; Table S4). Also, *QHD.ta.ipbb-5A.3* for HD and *QSMD.ta.ipbb-3B.3* for SMD were previously identified in Sukumaran et al. (2015), which is confirming the robustness of the identified QTL for HD in this study. To our best knowledge, the literature survey indicated that the remaining nine HD and eleven SMD associations seem to be putatively novel QTLs, as none of them were reported elsewhere (Table S4).

3. Localization of significant SNPs in identified QTLs for studied agronomic traits in the wheat physical map

The alignment of most significant SNPs in identified stable 50 MTAs with sequences in the Wheat Ensembl database (https://plants.ensembl.org/Triticum_aestivum/Info/Index) suggested that SNPs in 43 MTAs were in genic and 7 MTAs in intergenic positions (Table S4). Interestingly, two SNPs in QTLs for HD under rainfed conditions on homeological chromosomes 6A and 6B (*QHD.ta.ipbb-6A* and *QHD.ta.ipbb-6B*) were aligned with E3 ubiquitin-protein ligase (UPL) (Table S4). Alike the above two QTL locations for HD, the homeological region on the distal part of the short arm of the chromosome 6D has also carried MTA for HD (Table 2, Fig. 4A,B). However, the SNP in this MTA was aligned to a different gene (Table S4), most probably because of poor representation of polymorphic markers in the genome D. The UPL, along with E1 ubiquitin-activating and E2 ubiquitin-conjugating enzymes, is known to be participating in the ubiquitylation of proteins (*Liu et al., 2020*). Ubiquitylation is essential in the regulation of various biological processes, including growth and development, response to biotic and abiotic stress signaling, and regulation of chromatin structure (*Ramadan et al., 2015*). Additional confirmations for the relationship between ubiquitylation and HD in this study were the SNP alignment of *QHD.ta.ipbb-1B* and *QHD.ta.ipbb-3B* with the Wheat Ensembl database. Particularly, *QHD.ta.ipbb-1B* was aligned with a putative ubiquitin carrier protein, and *QHD.ta.ipbb-3B* with ubiquitin core domain-containing protein (Table S4). In other sequence alignments of identified QTLs, the SNP in the most significant QTL for SMD (*QSMD.ta.ipbb-3B.3*) was aligned with the unknown function protein. Also, the position of SNP for the *QTKW.ta.ipbb-1B* has significantly aligned with the position of a gene from *Tetratricopeptide-like helical domain superfamily*, which enables plants to cope up with adverse environmental stresses and allow them to rapidly acclimate to new conditions (*Sharma et al., 2017*).

Hence, the identified SNP markers for discovered 50 stable QTLs of eight agronomic traits, including eleven QTLs for HD and twelve QTLs for SMD, can be efficiently used in spring wheat projects to target the construction of new highly competitive cultivars in arid zones based on the utilization of marker-assisted selection approach.

Conclusion

The collection of spring wheat consisting of 179 local cultivars and promising lines showed a wide range of grain yield under two water regimes (irrigated and rainfed) in the

Akmola region of northern Kazakhstan during 2018-2020. In total, 51 accessions exceeded the YM2 of the local standard cultivar "Tselinnaya yubileinaya" under rainfed conditions, including four accessions, WS10, WS32, WS82, and WS85, that were stable in all three tested years. The GGE biplot method's application using two principal coordinates has confirmed the collection's high yield variability under both tested conditions. Pearson's correlation test suggested that earlier HD is advantageous for higher TKW, which one of the main yield components, under rainfed conditions. The SNP genotyping of the studied collection using the 20K Illumina SNP array has allowed the identification of 8662 polymorphic SNP markers. Field phenotypic data of nine agronomic traits and polymorphic SNP data were used to identify MTA based on GWAS. Fifty stable QTLs out of 237 total QTLs were identified in irrigated and rainfed conditions in the Akmola region, Northern Kazakhstan, by studying HD, SMD, PL, SL, NPS, NKS, TKW, and YM2. In general, 12 QTLs were identified only in irrigated, 13 QTLs only in rainfed, and 25 QTLs both in rainfed and irrigated conditions. Twelve out of 13 QTLs identified only under rainfed conditions were associated with flowering and seed maturation time, suggesting that early flowering time is essential for avoiding water deficiency stress. The literature survey indicated that 9 QTLs for HD and 11 QTLs for SMD were presumably novel genetic factors identified in irrigated and rainfed conditions.

Additional Information and Declarations

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

Akerke Amalova performed field work, analyzed the data, prepared figures and tables, prepared the draft, and approved the final draft.

Saule Abugalieva conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft

Adylkhan Babkenov, Sandugash Babkenova conceived and designed the experiments, performed and designed the experiments in experimental station, and approved the final draft

Yerlan Turuspekov raised funding for the research, conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft

Data Availability

The following information was supplied regarding data availability:

Raw data is available in the Supplementary file

Funding

This study was supported by the grant AP08855387 “Nested association mapping for gene discovery and deployment for improvement of yield, quality, and disease resistance in bread wheat” for 2020-2022 by the Ministry of Education and Science of the Republic of Kazakhstan.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. USDA. 2018. "Kazakhstan - Republic of Grain and Feed Update Kazakhstan Grain and Feed July Report" 2019: 11.
2. Shiferaw B., Smale M., Braun H. J., Duveiller E., Reynolds M., & Muricho G. 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security*, **5(3):** 291-317. DOI: 10.1007/s12571-013-0263-y
3. Patrick E. 2017. Drought characteristics and management in Central Asia and Turkey. *FAO Water Reports; Food and Agriculture Organization of the United Nations Rome: Rome, Italy*.
4. Skirycz A., & Inzé D. 2010. More from less: plant growth under limited water. *Current Opinion in Biotechnology*, **21(2):** 197-203. DOI: 10.1016/j.copbio.2010.03.002
5. Abugaliyeva A.I., Savin T.V. 2018. The wheat introgressive form evaluation by grain biochemical and technological properties. *Vavilovskii Zhurnal Genetiki i Selektsii=Vavilov Journal of Genetics and Breeding*. **22(3):** 353-362. DOI: 10.18699/VJ18.371
6. Foley J. A., Ramankutty N., Brauman K.A., Cassidy E.S., Gerber J.S., Johnston M., Mueller N.D., O'Connell C., Ray D.K., West P.C., Balzer C., Bennett E., Carpenter S., Hill J., Monfreda C., Polasky S., Rocksrtrom J., Sheehan J., Siebert S., & Zaks, D. P. 2011. Solutions for a cultivated planet. *Nature*, **478(7369):** 337-342. DOI: 10.1038/nature10452
7. Bray E. A. 2002. Classification of genes differentially expressed during water-deficit stress in Arabidopsis thaliana: an analysis using microarray and differential expression data. *Annals of botany*, **89(7):** 803-811. DOI: 10.1093/aob/mcf104
8. Nezhadahmadi A., Prodhan Z. H., & Faruq G. 2013. Drought tolerance in wheat. *The Scientific World Journal*, **2013** DOI:10.1155/2013/610721
9. Tsenov N., Atanasova D., Stoeva I., & Tsenova E. 2015. Effects of drought on grain productivity and quality in winter bread wheat. *Bulgarian Journal of Agricultural Science*, **21(3):**592-598.
10. Zhang J., Zhang S., Cheng M., Jiang H., Zhang X., Peng C., Lu X., Zhang M., & Jin J. 2018. Effect of Drought on Agronomic Traits of Rice and Wheat: A Meta-Analysis. *International journal of environmental research and public health*, **15(5):**839. DOI: 10.3390/ijerph15050839
11. Ingram J. & Bartels D. 1996. The Molecular Basis of Dehydration Tolerance in Plants. *Annual Review of Plant Physiology and Plant Molecular Biology,* **47:** 377-403. DOI:10.1146/annurev.arplant.47.1.377
12. Agarwal P.K., Agarwal P., Reddy M.K., & Sopory S.K. 2006. Role of *DREB* transcription factors in abiotic and biotic stress tolerance in plants. *Plant cell reports*, **25(12)**:1263-1274. DOI: 10.1007/s00299-006-0204-8
13. Wei B., Jing R., Wang C., Chen J., Mao X., Chang X., & Jia J. 2009. *Dreb1* genes in wheat (*Triticum aestivum* L.): development of functional markers and gene mapping based on SNPs. *Molecular Breeding*, **23(1):** 13-22. DOI: 10.1007/s11032-008-9209-z

14. Huseynova I. M., Rustamova S. M., & Mammadov A. C. 2013. Identification of *Dreb1* genes involved in drought tolerance in wheat (*Triticum* L.). *In Photosynthesis Research for Food, Fuel and the Future*, 552-555.
15. Hassan N.M., El-Bastawisy Z.M., El-Sayed A.K., Ebeed H.T., & Alla M.M.N. 2015. Roles of dehydrin genes in wheat tolerance to drought stress. *Journal of advanced research*, **6(2):** 179-188. DOI: 10.1016/j.jare.2013.11.004
16. Kulkarni M., Soolanayakanahally R., Ogawa S., Uga, Y., Selvaraj, M. G., & Kagale S. 2017. Drought response in wheat: key genes and regulatory mechanisms controlling root system architecture and transpiration efficiency. *Frontiers in chemistry*, **5**:106. DOI: 10.3389/fchem.2017.00106
17. Guo Z., Chen D., Alqudah A.M., Röder M.S., Ganal M.W., & Schnurbusch T. 2017. Genome-wide association analyses of 54 traits identified multiple loci for the determination of floret fertility in wheat. *New Phytologist*, ***214*(1),** 257-270. DOI: 10.1111/nph.14342.
18. Jaiswal V., Gahlaut V., Meher P.K., Mir R.R., Jaiswal J.P., Rao A.R., Balyan H.S., Gupta P.K. 2016. Genome Wide Single Locus Single Trait, Multi-Locus and Multi-Trait Association Mapping for Some Important Agronomic Traits in Common Wheat (*T. aestivum* L.) *PloS one*, **11(7),** e0159343. DOI:10.1371/journal.pone.0159343
19. Reynolds M.P., Trethowan R., Crossa J., Vargas M., & Sayre K.D. 2002. Physiological factors associated with genotype by environment interaction in wheat. *Field crops research*, **75(2-3):** 139-160. DOI:10.1016/S0378-4290(02)00023-0
20. Cavanagh C.R., Chao S., Wang S., Huang B.E., Stephen S., Kiani S., Forrest K., Saintenac C., Brown-Guedira G.L., Akhunova A. & See D., 2013. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proceedings of the national academy of sciences*, **110(20):** 8057-8062. DOI: 10.1073/pnas.1217133110
21. Wang S., Wong D., Forrest K., Allen A., Chao S., Huang E., Maccaferri M., Salvi S., Milner S., Cattivelli L., Mastrangelo A. M., Whan A., Stephen S., Barker G., Wieseke R., Plieske J., International Wheat Genome Sequencing Consortium, Lillemo M., Mather D., Appels R., Dolferus R., Brown-Guedira G., Korol A., Akhunova A. R., Feuillet C., Salse J., Morgante M., Pozniak C., Luo M., Dvorak J., Morell M., Dubcovsky J., Ganal M., Tuberosa R., Lawley C., Mikoulitch I., Cavanagh C., Edwards K. J., Hayden M., Akhunov E. 2014. Characterization of polyploid wheat genomic diversity using a high-density 90.000 SNP array. *Plant Biotechnology Journal* **12(6):** 787-796. doi: 10.1111/pbi.12183.
22. Boeven P.H., Longin C.F.H., Leiser W.L., Kollers S., Ebmeyer E. & Würschum, T. 2016. Genetic architecture of male floral traits required for hybrid wheat breeding. *Theoretical and Applied Genetics*. **129:** 2343-2357.
23. Sun C., Dong Z., Zhao L., Ren Y., Zhang N., & Chen F. 2020. The Wheat 660K SNP array demonstrates great potential for marker-assisted selection in polyploid wheat. *Plant biotechnology journal*, **18(6):**1354-1360. DOI: 10.1111/pbi.13361
24. Allen A. M., Barker G. L., Berry S.T., Coghill J.A., Gwilliam R., Kirby S., Robinson P., Brenchley R.C., D'Amore R., McKenzie N., Waite D., Hall A., Bevan M., N.Hall, Edwards K.J. 2011. Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant biotechnology journal* **9(9)**:1086-1099. DOI 10.1111/j.1467-7652.2011.00628.x

25. Turuspekov Y., Baibulatova A., Yermekbayev K., Tokhetova L., Chudinov V., Sereda G., Ganal M., Griffiths S. & Abugalieva S. 2017. GWAS for plant growth stages and yield components in spring wheat (*Triticum aestivum* L.) harvested in three regions of Kazakhstan. *BMC plant biology*, **17(1):**190. DOI: 10.1186/s12870-017-1131-2
26. Rahimi Y., Bihamta M.R., Taleei A., Alipour H., & Ingvarsson P. K. 2019. Genome-wide association study of agronomic traits in bread wheat reveals novel putative alleles for future breeding programs. *BMC plant biology*, **19(1):** 1-19. DOI: 10.1186/s12870-019-2165-4
27. Tsai H.Y., Janss L.L., Andersen J.R., Orabi J., Jensen J.D., Jahoor A., & Jensen J. 2020. Genomic prediction and GWAS of yield, quality and disease-related traits in spring barley and winter wheat. *Scientific reports*, **10(1):** 1-15. DOI:10.1038/s41598-020-60203-2
28. Anuarbek S., Abugalieva S., Pecchioni N., Laidò G., Maccaferri M., Tuberosa R., & Turuspekov Y. 2020. Quantitative trait loci for agronomic traits in tetraploid wheat for enhancing grain yield in Kazakhstan environments. *PloS one,* **15(6):** e0234863. DOI:10.1371/journal.pone.0234863.
29. Genievskaya Y., Turuspekov Y., Rsaliyev A., & Abugalieva S. 2020. Genome-wide association mapping for resistance to leaf, stem, and yellow rusts of common wheat under field conditions of South Kazakhstan. *PeerJ*, **8**: e9820.
30. Quarrie S. A., Steed A., Calestani C., Semikhodskii A., Lebreton C., Chinoy C., Steele N., Pljevljakusić D., Waterman E., Weyen J., Schondelmaier J., Habash D. Z., Farmer P., Saker L., Clarkson D. T., Abugalieva A., Yessimbekova M., Turuspekov Y., Abugalieva S., Tuberosa R., Sanguineti M-C., Hollington P. A., Aragués R., Royo A.& Dodig D. (2005). A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring× SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theoretical and Applied Genetics* **110(5):** 865-880. DOI 10.1007/s00122-004-1902-7.
31. Verma V., Foulkes M.J., Worland A.J., Sylvester-Bradley R., Caligari, PDS, & Snape J.W. 2004. Mapping quantitative trait loci for flag leaf senescence as a yield determinant in winter wheat under optimal and drought-stressed environments. *Euphytica*, **135(3):** 255-263. DOI: 10.1023/B:EUPH.0000013255.31618.14
32. Tura H., Edwards J., Gahlaut V., Garcia M., Sznajder B., Baumann U. Baumann, F. Shahinnia, Reynolds M., Langridge P., Balyan H.S., Gupta P.K., Schnurbusch T.& Fleury D. 2020. QTL analysis and fine mapping of a QTL for yield-related traits in wheat grown in dry and hot environments. *Theoretical and Applied Genetics*, **133(1):** 239-257. DOI: 10.1007/s00122-019-03454-6
33. Sukumaran S., Dreisigacker S., Lopes M., Chavez P., & Reynolds M. P. 2015. Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theoretical and applied genetics*, **128(2):** 353-363. DOI: 10.1007/s00122-014-2435-3
34. Lin Y., Yi X., Tang S., Chen W., Wu F., Yang X., Jiang X., Shi H., Ma J., Chen G., Chen G., Zheng Y., Wei Y., & Liu Y. 2019. Dissection of Phenotypic and Genetic Variation of Drought-Related Traits in Diverse Chinese Wheat Landraces. *The Plant Genome*, **12(3):** 190025. DOI: 10.3835/plantgenome2019.03.0025

35. Li L., Mao X., Wang J., Chang X., Reynolds M., & Jing R. 2019. Genetic dissection of drought and heat-responsive agronomic traits in wheat. *Plant, cell & environment.* **42(9):** 2540-2553. DOI: 10.1111/pce.13577
36. Mathew I., Shimelis H., Shayanowako A.I.T., Laing M., Chaplot V. 2019. Genome-wide association study of drought tolerance and biomass allocation in wheat. *PloS one,* **14(12):** e0225383. DOI:10.1371/journal.pone.0225383
37. Doyle J.J., Doyle J.L. 1990. Isolation of plant DNA from fresh tissue. *Focus* **12:**13-15.
38. Miyagawa T., Nishida N., Ohashi J., Kimura R., Fujimoto A., Kawashima M., Koike A., Sasaki T., Tanii H., Otowa T., Momose Y., Nakahara Y., Gotoh J., Okazaki Y., Tsuji S., Tokunaga K. 2008. Appropriate data cleaning methods for genome-wide association study. *Journal of Human Genetics* **53(10):**886-893 DOI: 10.1007/s10038-008-0322-y.
39. Bradbury P.J, Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler ES. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **23(19):**2633-2635 DOI: 10.1093/bioinformatics/btm308.
40. RStudio Team 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/ .
41. Stevens E.L., Heckenberg G., Roberson E.D.O., Baugher J.D., Downey T.J., Pevsner J. 2011. Inference of Relationships in Population Data Using Identity-by-Descent and Identity-by-State. *PLOS Genetics*, **7(9):** e1002287. DOI: 10.1371/journal.pgen.1002287
42. Pritchard J.K., Stephens M., & Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, **155(2):** 945-959.
43. Evanno G., Regnaut S., & Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, **14(8):** 2611-2620.
44. GraphPad Prism 2021. Software, Inc. *Available at* https://www.graphpad.com/scientific-software/prism/ (Accessed 22 March 2021).
45. VSN International (2020). Genstat for Windows 21st Edition. VSN International, Hemel Hempstead, UK. Web page: Genstat.co.uk
46. Voorrips R.E. 2002 MapChart: Software for the graphical presentation of linkage maps and QTLs. *The Journal of Heredity,* **93(1):**77-78. DOI: 10.1093/jhered/93.1.77.
47. Ensembl Plants 2020 https://plants.ensembl.org/Triticum_aestivum/Info/Index
48. Lehari K., Kumar M., Burman V., Aastha, Vaishali, Kumar V., Chand P., Singh R. Morphological, physiological and biochemical analysis of wheat genotypes under drought stress. 2019. *Journal of Pharmacognosy and Phytochemistry* **8(2S):** 1026-1030.
49. Pour-Aboughadareh A., Mohammadi R., Etminan A., Shooshtari L., Maleki-Tabrizi N., Poczai P. (2020). Effects of Drought Stress on Some Agronomic and Morpho-Physiological Traits in Durum Wheat Genotypes. *Sustainability*, **12(14):** 5610.
50. Amalova A., Abugalieva S., Chudinov V., Sereda G., Tokhetova L., Abdikhalyk A., Turuspekov Y. 2021. QTL mapping of agronomic traits in wheat using the UK Avalon×Cadenza reference mapping population grown in Kazakhstan. *PeerJ* **9**:e10733 https://doi.org/10.7717/peerj.10733
51. Liu W., Tang X., Qi X., Fu X., Ghimire S., Ma R., Li S., Zhang N., & Si H. 2020. The ubiquitin conjugating enzyme: an important ubiquitin transfer platform in ubiquitin-proteasome system. *International journal of molecular sciences* **21(8),** 2894. DOI: 10.3390/ijms21082894

52. Ramadan A., Nemoto K., Seki M., Shinozaki K., Takeda H., Takahashi H., Sawasaki T. .2015. Wheat germ-based protein libraries for the functional characterisation of the Arabidopsis E2 ubiquitin conjugating enzymes and the RING-type E3 ubiquitin ligase enzymes. *BMC Plant Biology,* **15**: 1–15 DOI: 10.1186/s12870-015-0660-9
53. Sharma M., & Pandey G.K. 2016. Expansion and function of repeat domain proteins during stress and development in plants. *Frontiers in plant science*, **6**, 1218. DOI: 10.3389/fpls.2015.01218

Figure 1

The average yield performance of best accessions under rainfed conditions.

(A) The list of 51 spring wheat accessions outperformed the local check cultivar Tselinnaya yubileinaya (TY) on average yield per square meter (YM2) under rainfed conditions. Accessions in red color also outperformed TY on average YM2 under irrigated conditions as well. The nine accessions selected in the box with the highest average YM2 were analyzed using Finlay and Wilkinson (FW) regression analysis. (**B)** The FW test is suggesting a different level of stability in four out of nine selected spring wheat accessions in the box of Fig. 1.A.

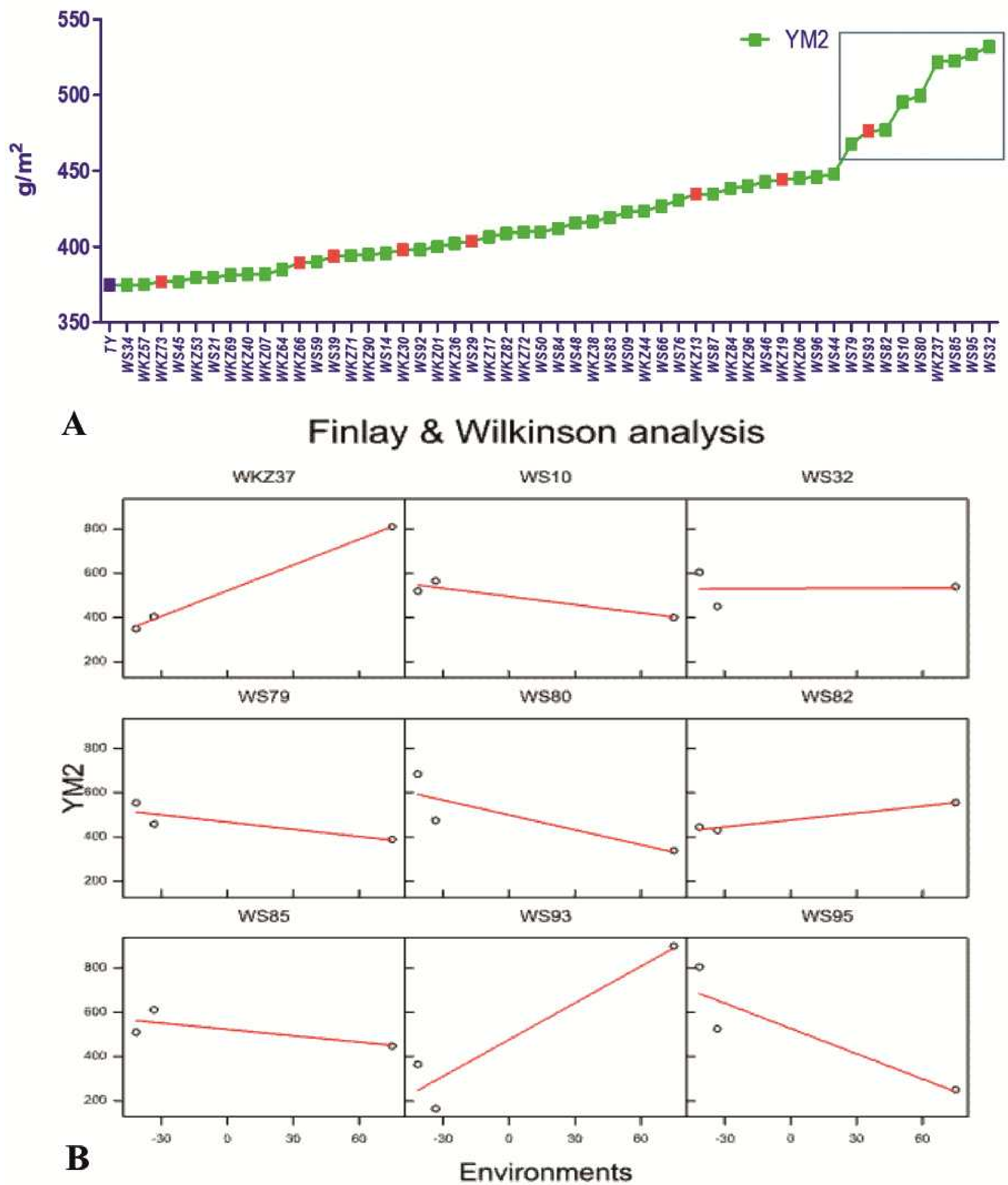

550
500
450
400
350
g/m^2
YM2
TY
WS34
WKZ57
WKZ73
WS45
WKZ53
WS21
WKZ69
WKZ40
WKZ07
WKZ64
WKZ66
WS59
WS39
WKZ71
WKZ90
WS14
WKZ30
WS92
WKZ01
WKZ36
WS29
WKZ17
WKZ82
WKZ72
WS50
WS84
WS48
WKZ38
WS83
WS09
WKZ44
WS66
WS76
WKZ13
WS87
WKZ84
WKZ96
WS46
WKZ19
WKZ06
WS96
WS44
WS79
WS93
WS82
WS10
WS80
WKZ37
WS85
WS95
WS32
A
Finlay & Wilkinson analysis
WKZ37
WS10
WS32
WS79
WS80
WS82
WS85
WS93
WS95
800
600
400
200
YM2
-30
0
30
60
Environments
B

Figure 2

The scattered GGE biplot graph on averaged yield per squared meter (YM2) data in the collection of 179 common wheat accessions tested in irrigated and rainfed conditions (2018-2020).

Green and blue colors indicated Genotype and Environment scores, respectively.

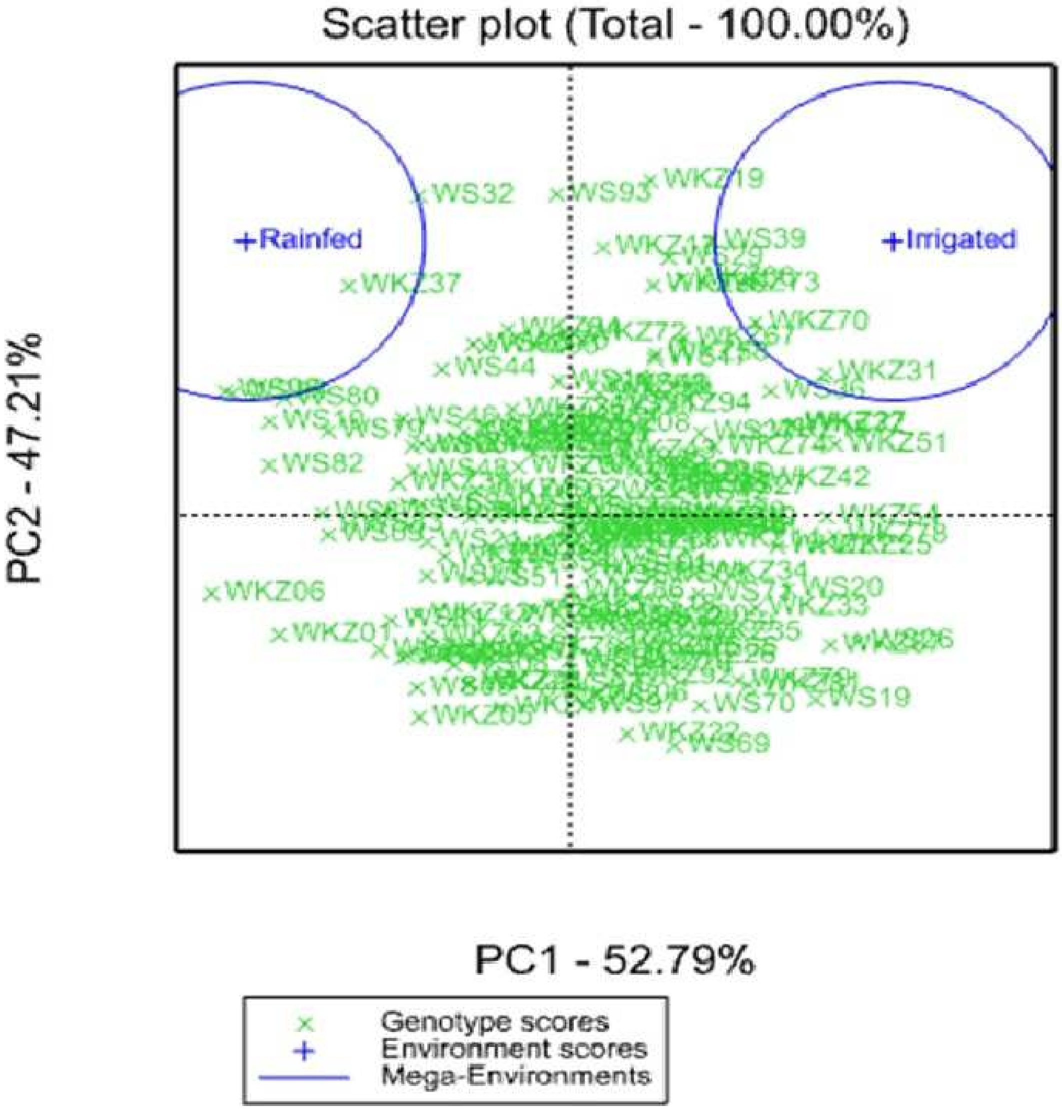
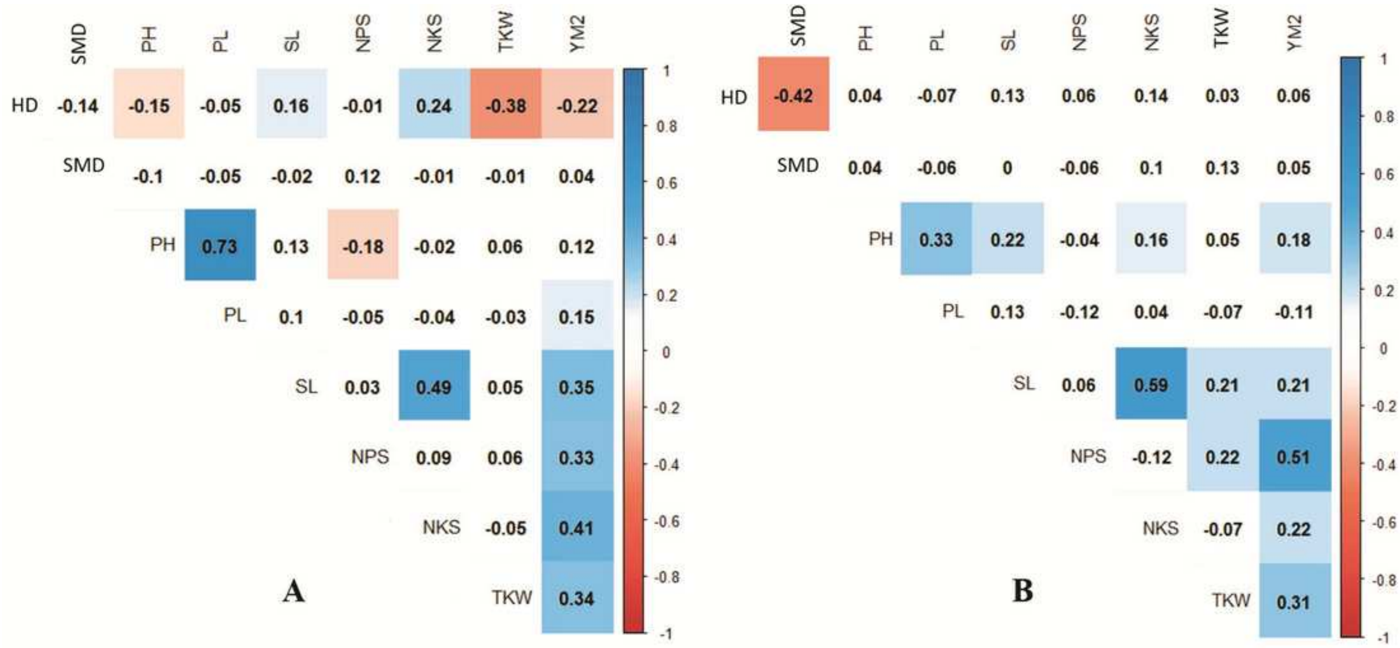

Scatter plot (Total - 100.00%)
+Rainfed
+Irrigated
WS32
WS93
WKZ19
WS39
WKZ37
WKZ70
WKZ72
WKZ67
WS44
WKZ31
WS80
WS46
WKZ94
WKZ27
WKZ51
WKZ74
WS82
WS48
WKZ42
WKZ54
WS89
WKZ78
WKZ25
WS51
WKZ34
WS20
WKZ06
WKZ33
WKZ01
WKZ35
WKZ05
WS70
WS19
WKZ22
WS69
PC2 - 47.21%
PC1 - 52.79%
Genotype scores
Environment scores
Mega-Environments

Figure 3

Correlation analysis for nine agronomic traits analyzed in the collection of 179 spring wheat accessions tested in rainfed (A) and irrigated (B) conditions.

Blue color is positive correlation, and red color is negative correlation.



A

	SMD	PH	PL	SL	NPS	NKS	TKW	YM2
HD	-0.14	-0.15	-0.05	0.16	-0.01	0.24	-0.38	-0.22
SMD		-0.1	-0.05	-0.02	0.12	-0.01	-0.01	0.04
PH			0.73	0.13	-0.18	-0.02	0.06	0.12
PL				0.1	-0.05	-0.04	-0.03	0.15
SL					0.03	0.49	0.05	0.35
NPS						0.09	0.06	0.33
NKS							-0.05	0.41
TKW								0.34

B

	SMD	PH	PL	SL	NPS	NKS	TKW	YM2
HD	-0.42	0.04	-0.07	0.13	0.06	0.14	0.03	0.06
SMD		0.04	-0.06	0	-0.06	0.1	0.13	0.05
PH			0.33	0.22	-0.04	0.16	0.05	0.18
PL				0.13	-0.12	0.04	-0.07	-0.11
SL					0.06	0.59	0.21	0.21
NPS						-0.12	0.22	0.51
NKS							-0.07	0.22
TKW								0.31

Figure 4

Number of QTLs identified under irrigated, rainfed, and in both conditions in Northern Kazakhstan in 2018-2020.

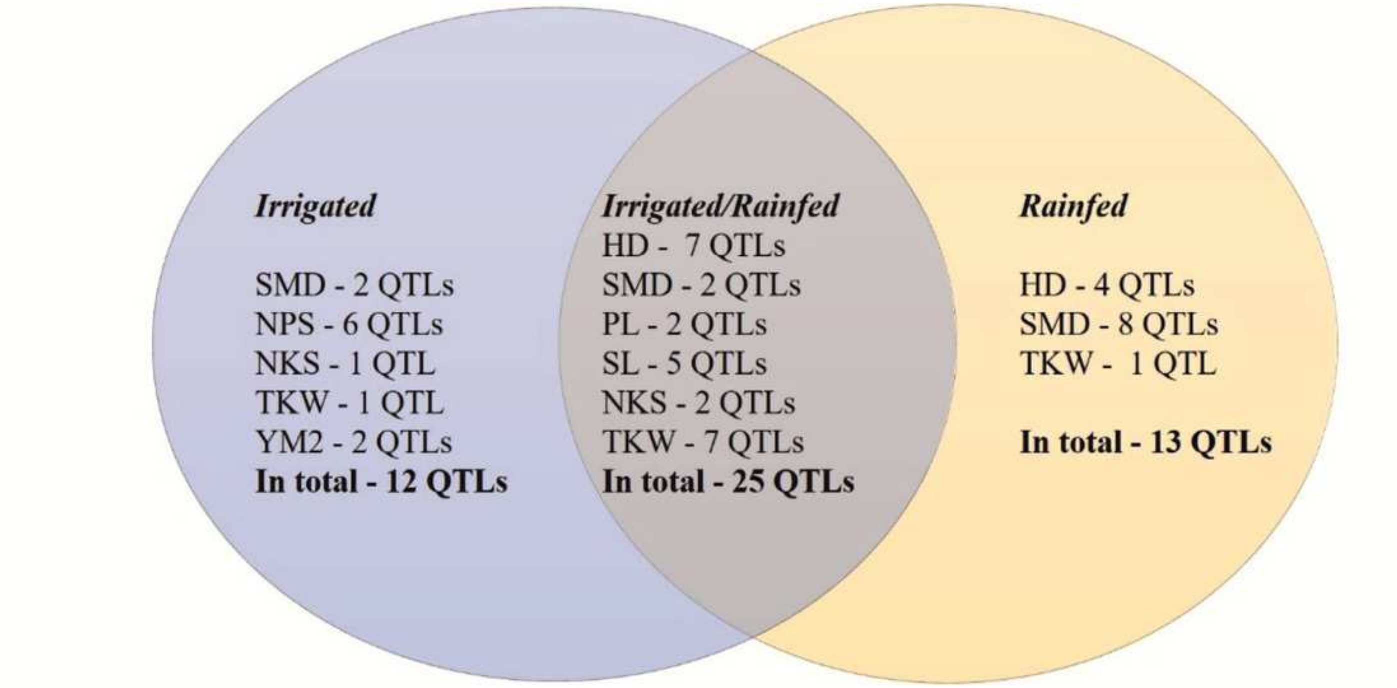


Table 1(on next page)

The significance of differences between irrigated and rainfed trials using average data in nine traits based on a two-tailed t-test.

Table 1. The significance of differences between irrigated and rainfed trials using average data in nine traits based on a two-tailed t-test

№	Traits	Rainfed (average)	Irrigated (average)	Significance (P-value)
1	Heading date (HD, days)	48.5±0.14	48.1±0.15	0.00260
2	Seed maturation date (SMD, days)	49.3±0.09	46.9±0.10	1.97E-45
3	Plant height (PH, cm)	61.6±0.39	73.6±0.44	4.4E-52
4	Peduncle length (PL, cm)	28.8±0.25	32.7±0.30	1.24E-19
5	Spike length (SL, cm)	9.05±0.05	8.59±0.05	8.18E-17
6	Number of productive spikes (NPS, pcs)	1.92±0.03	2.02±0.03	0.01322
7	Number of kernels per spike (NKS, pcs)	34.9±0.28	33.6±0.24	1.97E-06
8	Thousand kernel weight (TKW, g)	35.68±0.19	37.4±0.18	6.3E-14
9	Yield per square meter (YM2, g)	332.3±5.68	352.3±4.30	0.00592

Table 2(on next page)

The list of quantitative trait loci (QTLs) for eight studied traits identified using 179 spring wheat accessions tested under irrigated and rainfed conditions of Northern Kazakhstan (2018- 2020).

Table 2. The list of QTLs for eight studied traits identified using 179 spring wheat accessions tested under irrigated and rainfed conditions of Northern Kazakhstan (2018- 2020)

№	QTL name	SNP	Chr	Pos	P-value	Effect	Irrigated	Rainfed
1	QHD.ta.ipbb-1B	Kukri_c39223_871	1B	75.6	3.77E-04	4.84		+
2	QHD.ta.ipbb-2A	RAC875_c1706_1888	2A	151.2	3.33E-04	-1.61	+	+
3	QHD.ta.ipbb-2B	Excalibur_c20376_615	2B	76.8	6.54E-05	-1.42	+	+
4	QHD.ta.ipbb-3B	wsnp_Ex_c8240_13914674	3B	32.9	1.49E-06	2.06	+	+
5	QHD.ta.ipbb-5A.1	BobWhite_c10385_374	5A	0.00	1.57E-05	-5.68	+	+
6	QHD.ta.ipbb-5A.2	wsnp_BF293620A_Ta_2_1	5A	58.27	1.94E-05	-2.13	+	+
7	QHD.ta.ipbb-5A.3	BS00022071_51	5A	90.5	4.89E-05	-1.96	+	+
8	QHD.ta.ipbb-5B	RAC875_rep_c109634_90	5B	125.0	3.45E-04	-1.67		+
9	QHD.ta.ipbb-6A	Excalibur_c28854_1580	6A	0.88	1.94E-06	-6.89		+
10	QHD.ta.ipbb-6B	RAC875_c13610_1599	6B	0.37	3.76E-05	-5.39	+	+
11	QHD.ta.ipbb-6D	Excalibur_rep_c106566_371	6D	2.56	8.82E-06	-6.34		+
12	QSMD.ta.ipbb-2A	RAC875_c57998_165	2A	101.9	3.23E-04	-3.66	+	
13	QSMD.ta.ipbb-2B.1	Kukri_c9785_1472	2B	75.7	3.74E-04	-0.31	+	+
14	QSMD.ta.ipbb-2B.2	CAP8_c5161_541	2B	107.5	2.07E-04	0.47		+
15	QSMD.ta.ipbb-2D	Excalibur_c23239_961	2D	129.0	1.58E-04	1.75	+	
16	QSMD.ta.ipbb-3B.1	IMX3190	3B	56.6	5.05E-04	-1.40		+
17	QSMD.ta.ipbb-3B.2	BobWhite_c5095_634	3B	69.7	5.05E-04	-3.39		+
18	QSMD.ta.ipbb-3B.3	BS00078844_51	3B	85.0	5.00E-06	-6.32		+
19	QSMD.ta.ipbb-3D	GENE-1805_65	3D	71.9	6.63E-04	-3.39		+
20	QSMD.ta.ipbb-4A	RAC875_c40654_206	4A	120.1	1.76E-04	-1.28		+
21	QSMD.ta.ipbb-5D	Jagger_c8037_96	5D	167.0	6.62E-06	-5.35		+
22	QSMD.ta.ipbb-6A	BS00009985_51	6A	60.9	8.25E-05	-5.22		+
23	QSMD.ta.ipbb-6B	Excalibur_c15744_322	6B	0.37	8.66E-04	-3.36	+	+
24	QPL.ta.ipbb-3B.1	wsnp_Ra_c12935_20587578	3B	52.8	2.75E-04	-0.26	+	+
25	QPL.ta.ipbb-3B.2	BS00030534_51	3B	67.4	3.34E-04	4.76	+	+
26	QSL.ta.ipbb-1A	wsnp_Ku_c1818_3557408	1A	16.7	7.81E-04	-0.78	+	+
27	QSL.ta.ipbb-1B	wsnp_Ex_c26419_35667216	1B	65.4	5.99E-04	-0.95	+	+
28	QSL.ta.ipbb-2B	BS00093993_51	2B	108.3	5.64E-06	-1.09	+	+
29	QSL.ta.ipbb-2D	TA001453-0801	2D	96.1	3.11E-04	-0.62	+	+

30	QSL.ta.ipbb-5B	Excalibur_c9391_1016	5B	109.5	1.55E-04	-0.78	+	+
31	QNPS.ta.ipbb-1B	BS00078431_51	1B	70.8	7.91E-05	0.31	+	
32	QNPS.ta.ipbb-1D	BS00063511_51	1D	167.1	8.34E-05	0.29	+	
33	QNPS.ta.ipbb-2B	Excalibur_c20376_615	2B	76.8	1.12E-05	0.34	+	
34	QNPS.ta.ipbb-5A	RAC875_rep_c112818_307	5A	98.9	3.40E-05	-0.29	+	
35	QNPS.ta.ipbb-6A	TA003021-1057	6A	56.1	6.16E-04	0.02	+	
36	QNPS.ta.ipbb-7A	TA003458-0086	7A	133.9	3.54E-05	0.17	+	
37	QNKS.ta.ipbb-2B	Ku_c77612_301	2B	77.6	8.27E-05	-4.14	+	+
38	QNKS.ta.ipbb-3B	wsnp_Ex_c8240_13914674	3B	32.9	1.05E-06	5.13	+	
39	QNKS.ta.ipbb-4B	RAC875_c5087_310	4B	71.3	3.28E-04	-5.10	+	+
40	QTKW.ta.ipbb-1B	BS00078431_51	1B	70.8	2.75E-06	3.49	+	+
41	QTKW.ta.ipbb-1D	BS00063511_51	1D	167.1	3.64E-05	2.84	+	+
42	QTKW.ta.ipbb-2A	wsnp_Ex_rep_c101866_87158671	2A	101.9	4.40E-04	2.27	+	+
43	QTKW.ta.ipbb-2B	Excalibur_c20376_615	2B	76.8	1.98E-06	3.49	+	+
44	QTKW.ta.ipbb-4B	Excalibur_c27349_166	4B	77.9	2.67E-04	-2.65	+	+
45	QTKW.ta.ipbb-5A	RAC875_rep_c112818_307	5A	98.9	2.23E-05	-3.02	+	+
46	QTKW.ta.ipbb-6A	TA003021-1057	6A	56.1	1.67E-06	-3.34	+	+
47	QTKW.ta.ipbb-7A	TA003458-0086	7A	133.9	4.56E-05	2.92	+	
48	QTKW.ta.ipbb-7B	BS00063744_51	7B	99.2	2.83E-05	2.68		+
49	QYM2.ta.ipbb-3D	BS00061125_51	3D	149.8	3.10E-04	25.39	+	
50	QYM2.ta.ipbb-7B	wsnp_Ex_c11003_17857759	7B	77.2	5.26E-04	20.88	+	