Association mapping of quantitative trait loci for agronomic traits in spring wheat collection tested under two water regimes in Northern Kazakhstan

Amalova A. [1,2], Abugalieva S.[1,2], Babkenov A.[3], Babkenova S.[3], Turuspekov Y.[1,2]

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

Introduction

The **purpose of this work** was to identify drought resistance-associated QTLs under irrigated and rainfed conditions in Northern Kazakhstan by using a genome-wide association study. 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 (Fig.1).
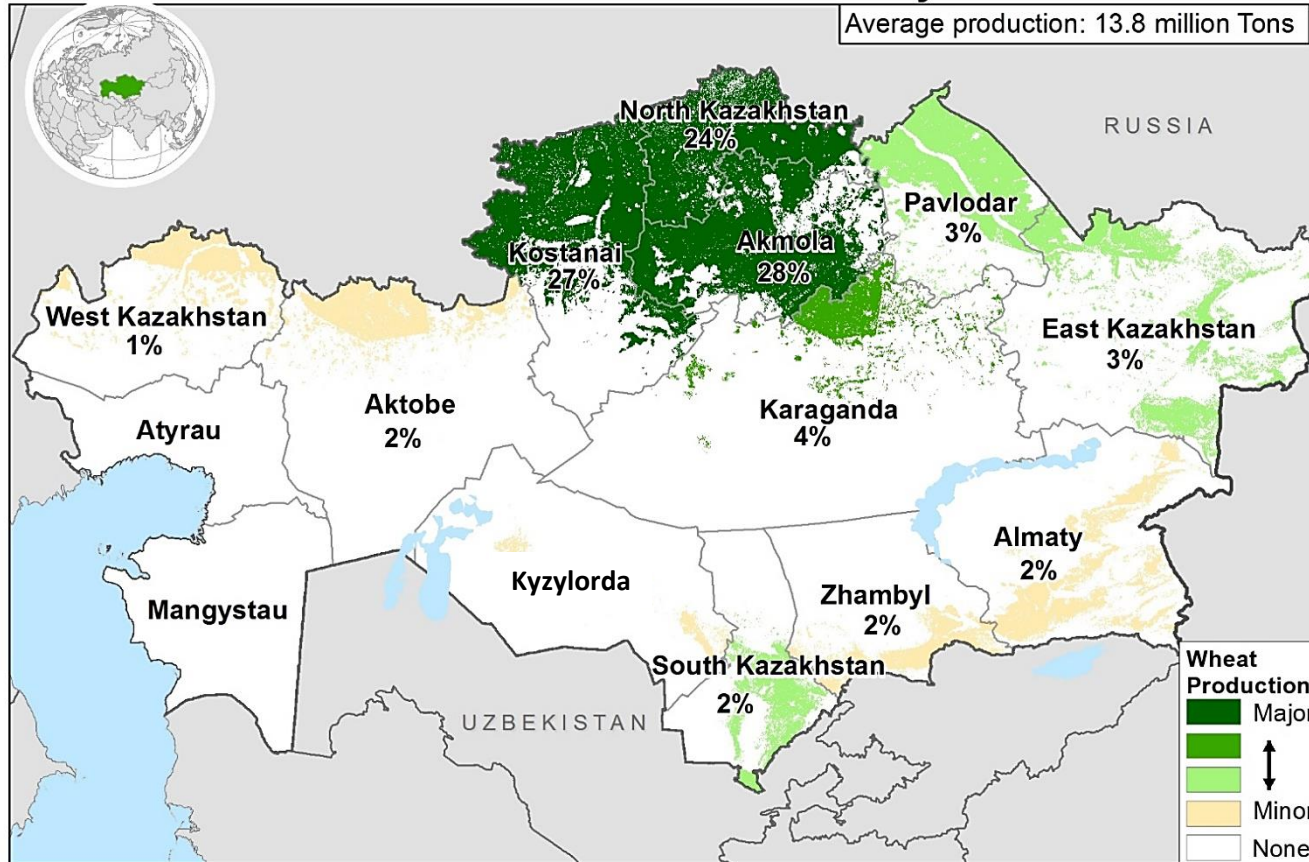
KAZAKHSTAN: Wheat Production by Oblast

Figure 1. Wheat production in Kazakhstan (USDA, 2015-2019)

Materials and methods

Materials: The 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 (Shortandy, Akmola region), and a check cultivar for the Akmola region – Tselinnaya yubileinaya.

Methods:

- The collection 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, 2019, 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.

- 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*).

Phenotypic Variation and Correlation Analysis

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

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.

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. 2.A). The Finlay and Wilkinson (FW) regression analysis (Fig. 2.B) suggested that YM2 of four wheat accessions, particularly WS10, WS32, WS82, and WS85, was stable in all three tested years (2018, 2019, 2020), out of nine accessions highlighted in the box in Figure 2.A, showing the YM2 with 400 g/m^2 and higher.
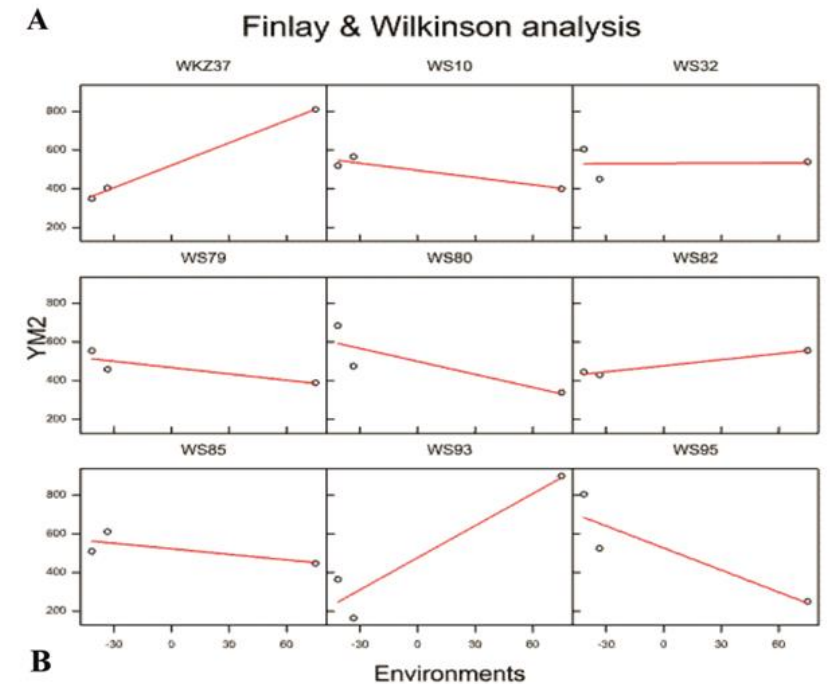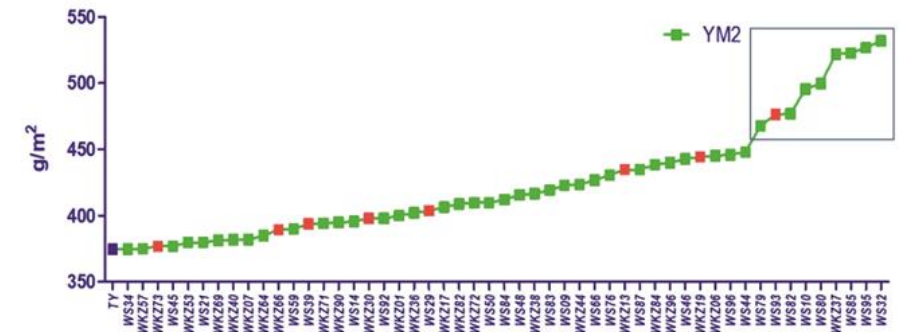

Figure 2. The average yield performance of best accessions under rainfed conditions

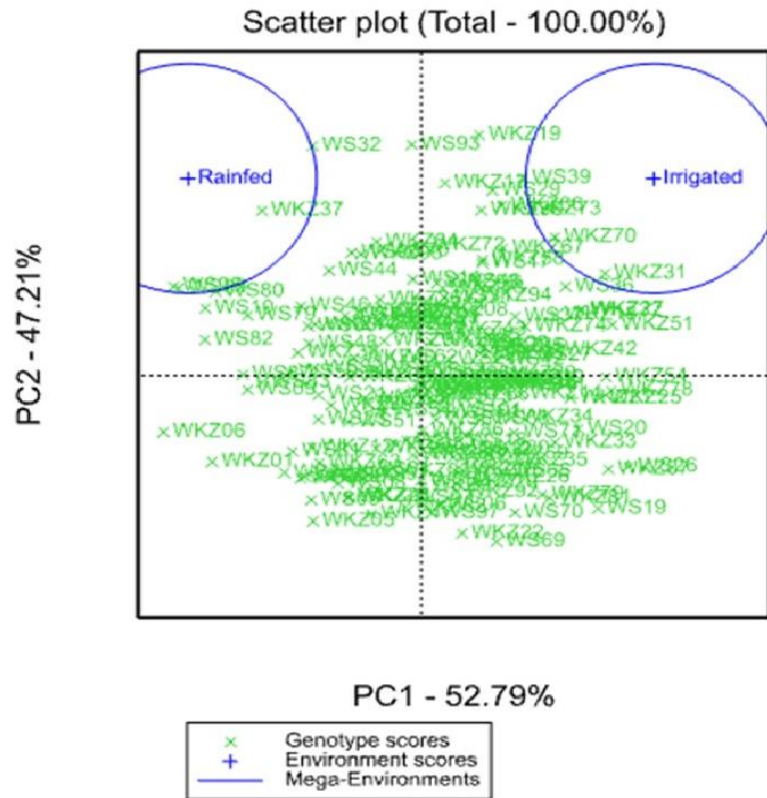# Phenotypic Variation and Correlation Analysis

Figure 3. The scattered GGE biplot graph on averaged yield per squared meter (YM2)

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.3). 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.

Figure 4. Correlation analysis for nine agronomic traits analyzed in the collection of 179 spring wheat accessions tested in rainfed (A) and irrigated (B) conditions (blue highlight – positive correlation, red – negative correlation)

Pearson's correlation assessment in both conditions indicated that average YM2 was positively correlated with NPS and TKW (Fig. 4). 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. 4, A). Interestingly, under the rainfed condition, the HD influenced both SL and NKS (Fig. 4). Expectedly, the PL was highly correlated with PH (P<0.0001), while the former one has negatively associated with NPS (Fig. 4. A).

4

Marker-trait associations under irrigated and rainfed conditions.
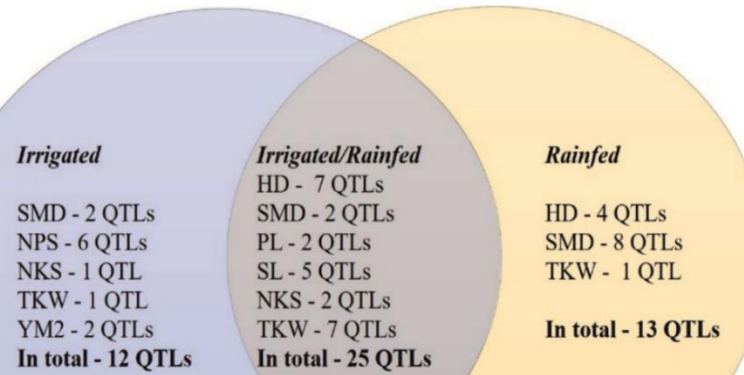

Figure 5. Number of QTLs identified under irrigated, rainfed, and both conditions in Northern Kazakhstan in 2018, 2019, 2020

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 PH.

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. Among identified 46 QTLs, seven pleiotropic MTAs were detected under both conditions (Fig.5).

Also, pleiotropic QTLs 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) (Fig. 6).
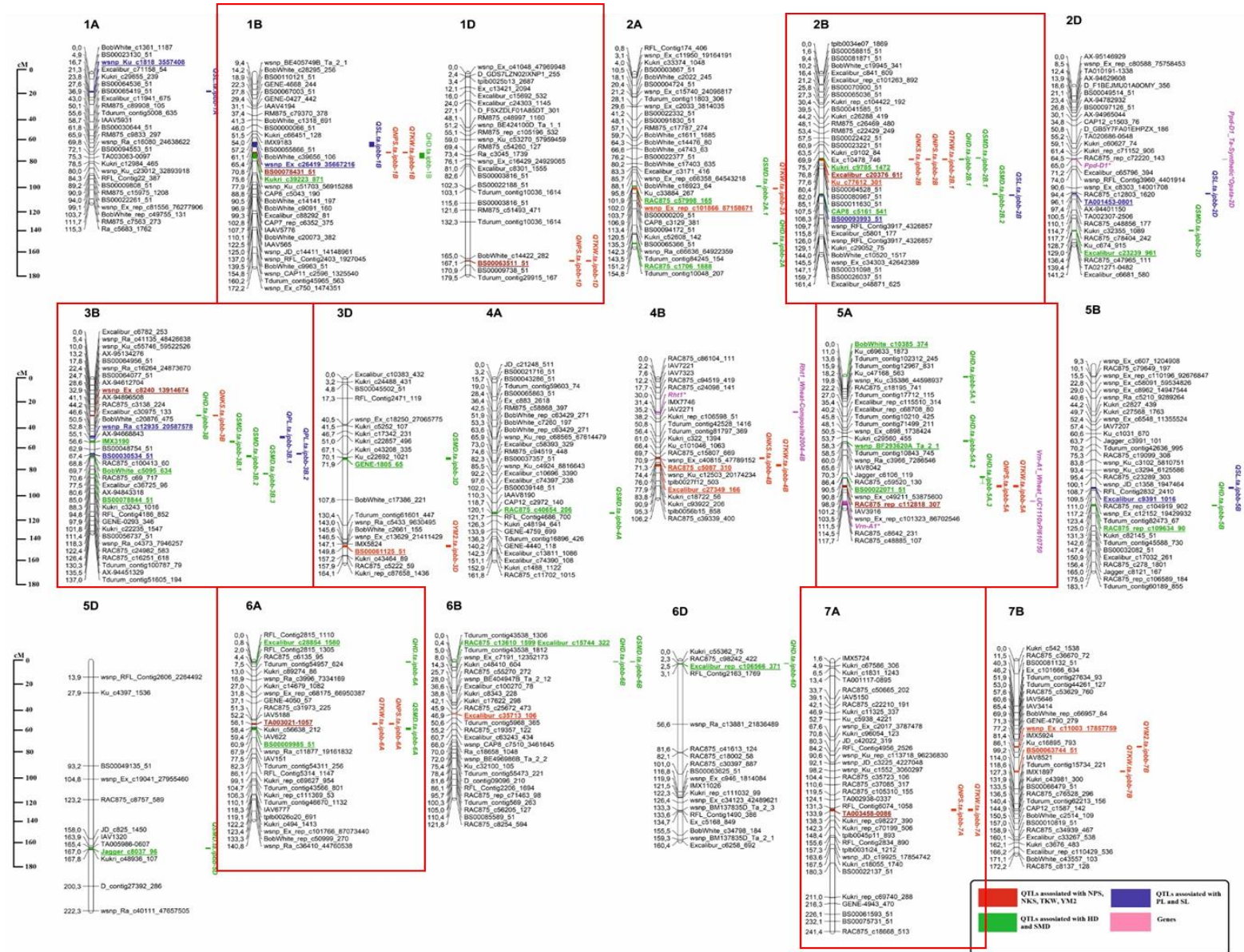

Figure 6. Localization of significant SNPs in identified QTLs for studied agronomic traits in the wheat genetic map

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, 2019, 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.