

# Genomic regions of durum wheat involved in water productivity

Meryem Zaïm<sup>a,b</sup>, Miguel Sanchez-Garcia<sup>b</sup>, Bouchra Belkadi<sup>a</sup>, Abdelkarim Filali-Maltouf<sup>a</sup>, Ayed Al Abdallat<sup>c</sup>, Zakaria Kehel<sup>b</sup>, Filippo Maria Bassi<sup>b</sup>

<sup>a</sup> Laboratory of Microbiology and Molecular Biology, Faculty of Sciences, University Mohammed V in Rabat, Morocco.

<sup>b</sup> ICARDA, Biodiversity and Integrated Gene Management, P.O. Box 6299 Rabat Instituts, Rabat, Morocco.

<sup>c</sup> Faculty of Agriculture, The University of Jordan, Amman 11942, Jordan

Meryem Zaïm	m.zaim@cgiar.org
Miguel Sanchez-Garcia	m.sanchez-garcia@cgiar.org
Bouchra Belkadi	bbelkadi@gmail.com
Abdelkarim Filali-Maltouf	filalimaltouf@gmail.com
Ayed Al Abdallat	a.alabdallat@ju.edu.jo
Zakaria Kehel	z.kehel@cgiar.org
Filippo Maria Bassi	f.bassi@cgiar.org
(Corresponding author)	

## Highlight

Loci controlling higher water productivity under drought conditions were identified using a strategy involving 3 distinct germplasm panels.

© The Author(s) 2023. Published by Oxford University Press on behalf of the Society for Experimental Biology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

Durum wheat is a staple food of the Mediterranean Basin, mostly cultivated under rainfed conditions. As such, the crop is often exposed to moisture stress. Therefore, the identification of genetic factors controlling the capacity of genotypes to convert moisture into grain yield (i.e., water productivity) is quintessential to stabilize production despite climatic variations. A global panel of 384 accessions was tested across eighteen Mediterranean environments (Morocco, Lebanon, and Jordan) representing a vast range of moisture levels. The accessions were assigned to water responsiveness classes, with genotypes 'Responsive to Low Moisture' reaching an average + 1.5 kg ha<sup>-1</sup> mm<sup>-1</sup> yield advantage. Genome wide association studies (GWAS) revealed that six loci explained most of this variation. A second validation panel tested under moisture stress confirmed that carrying the positive allele at three loci on chromosomes 1B, 2A, and 7B generated an average water productivity gain of + 2.2 kg ha<sup>-1</sup> mm<sup>-1</sup>. These three loci were tagged by Kompetitive Allele Specific PCR (KASP) markers, and these were used to screen a third independent validation panel composed of elites tested across moisture stressed sites. The three KASP combined predicted up to 10% of the variation for grain yield at 60% accuracy. These loci are now ready for molecular pyramiding and transfer across cultivars to improve the moisture conversion of durum wheat.

## Keywords

GWAS, moisture stress, QTL, water productivity, wide adaptation, yield stability.

Accepted Manuscript

## Abbreviations

**AMMI**: additive main effect and multiplicative interaction, **AWAI**: AMMI wide adaptation index, **bi**: slope value, **BLUE**: Best linear unbiased estimates, **DTH**: Days to heading, **E**: environment, **G**: genotype, **GxE**: genotype by environment, **GpS**: Grain per spike, interaction, **GWAS**: genome wide association study, **GY**: grain yield, **IDON**: International Durum Observatory Nursery, **KASP**: Kompetitive Allele Specific PCR, **KFD**: Kfardan, **MTA**: marker trait association, **MCH**: Marchouch, **MKZ**: Melk Zhar, **MUS**: Musghar, **PCA**: principal component differentiation, **PCR**: Polymerase chain reaction, **PLH**: Plant height, **SPK**: Spike density per m<sup>2</sup>, **QTL**: quantitative trait loci, **RGA**: root growth angle, **SAD**: Sidi El Aidi, **TER**: Terbol, **TES**: Tessaout, **TKW**: 1,000 Kernel weight, **WP**: water productivity, **Z**: Zadok's scale.

Accepted Manuscript

## Introduction

Durum wheat ( $2n = 28$ , AABB, *Triticum turgidum* L. ssp. *durum*) is a staple and cash crop grown on over 17 million ha worldwide (Tidiane *et al.*, 2019; Xynias *et al.*, 2020). Approximately 2/3 of durum wheat is grown in the Mediterranean basin, but this area contributes to only half of the worldwide production (Kabbaj *et al.*, 2017; Li *et al.*, 2013). In fact, climate change has and will continue to affect this region, with annual precipitation projected to decrease by 20–40% by the second half of the 21<sup>st</sup> century (Zittis *et al.*, 2021). Rainfall and temperatures in the Mediterranean dryland areas are largely unpredictable within and between cropping seasons. In the past years, North African countries have witnessed a raise in the frequency of drought events, an extension in their length, and an anticipation in their time of occurrence, substantially shifting from late spring to the middle of winter (Belaid *et al.*, 2005; Qi *et al.*, 2022; Trambly *et al.*, 2020). Since drought stress has a devastating effect on yield and its related traits (Bilal *et al.*, 2015; Kiliç and Tacettin, 2010), the North African durum wheat farmers have experienced strong reductions in their productivities. Under such conditions, breeders have committed to the delivery of new varieties with enhanced adaptation mechanisms, by avoiding or tolerating these stresses. Genetic improvement programs have for a long time attempted to balance the needs of raising overall yield potential, while ensuring cultivars with stable yield performances across seasons. In fact, the final productivity of a variety results from the combined effects of genotype (G), environment (E), and their interaction (GxE) (Mohammadi *et al.*, 2015). Thus, the development of superior cultivars requires strategic approaches to combine good stress tolerance with strong yield stability (Bassi and Sanchez-Garcia, 2017; Mohammadi *et al.*, 2011).

Yield stability refers to the ability of certain genotype to ensure good yield performances despite the fluctuations of growing conditions occurring across environments, and it is normally associated with the GxE component. Several decades of studies have demonstrated that stability is controlled by genetic factors interacting with the environment. As such, it is possible to improve the stability of a genotype via pyramiding multiple positive alleles for this trait. Breeders approach this need by testing the genotypes under a vast range of environments and seasons, to then derive what are defined as stability scores (Malosetti *et al.*, 2013) and then use these to identify stable genotypes across environments. One such score widely used in durum wheat breeding is the AMMI wide adaptation index (AWAI) score that utilizes the AMMI capacity to partition the GxE into sub-factors, to then estimate a weighted value to be assigned to the genotype (Bassi and Sanchez-Garcia, 2017). However, a stable variety can also be obtained by pyramiding multiple positive alleles at loci controlling discrete interactions with the environment. For instance, a drought tolerant variety would be able to maintain its yield performance (i.e., stability) even when moisture stress occurs. The concept of water productivity is linked to yield stability and potential as it has been used in plant breeding to define genotypes capable of using moisture in a more efficient way, and hence achieve higher productivity at the same level of moisture input (Anyia *et al.*, 2008). While the application of this concept was originally proposed to define the response of genotypes to increasing irrigation rates, it has become even more important to assess the

response to moisture stress, when water availability is particularly scarce (Bhouri Khila *et al.*, 2021). In fact, wheat's most sensitive growth stages to water stress are mainly stem elongation and booting, followed by anthesis and grain filling (Blum and Pnuel, 1990; del Moral *et al.*, 2003; Shpiler and Blum, 1990). Water deficit around anthesis may lead to a loss in yield by reducing spike and spikelet number and the fertility of surviving spikelets, while water deficit during grain-filling period reduces grain weight (Karam *et al.*, 2009). Geerts and Raes (2009) indicate that scarce moisture can increase water productivity for various crops without causing severe yield reductions. Zhang *et al.* (2006) demonstrated that under rainfed conditions, wheat grain yield, harvest index and water productivity were greatly improved under regulated deficit irrigation when compared to the non-water stressed treatment. Maximizing WP may be economically more profitable for the farmer than maximizing yields or land productivity (English, 1990) in areas where water is the most limiting factor. Karrou and Oweis (2012) results showed that in general a reduced irrigation of 1/3 of full supplemental irrigation gave the highest rate of increase in grain yield and water productivity. Grain yield reductions due to the application of 2/3 supplemental irrigation were around 10% on average, while differences in total water productivity of crops grown under full irrigation compared to deficit irrigation were not significant.

Beyond the application of stability systems to adapt to all conditions, there are several discreet traits that have been proposed as favoring the adaptation of durum wheat to moisture stress. A simplified list of these would include early maturity to avoid terminal stress (Gupta *et al.*, 2020), good coverage of ground to favor shading and prevent moisture transpiration from the soil (Yadvinder *et al.*, 2014), access residual moisture in deeper soil layers (El Hassouni *et al.*, 2019; Lilley and Kirkegaard, 2016; Yadvinder *et al.*, 2014), and improvement of specific yield components, with a particular attention to grain size (Mohammadi *et al.*, 2019). In that sense, breeders seek to identify and pyramid these traits to achieve better stability when moisture stress occurs (Araus *et al.*, 2008; Reynolds *et al.*, 2009; Sukumaran *et al.*, 2018; Tuberosa, 2012). Therefore, the knowledge of genetics and gene action of these traits is essential for generating stable varieties (Habash *et al.*, 2009). Molecular markers technology offers the possibility to identify and track these positive alleles (Ceccarelli, 2015; Collard and Mackill, 2008). Genome wide association study (GWAS) is an approach that helps determine significant relationships between the allelic make up (i.e., haplotypes) of a genotype and its field response. Such approach was used for the identification of novel QTLs with potential implications for durum wheat breeding programs, such as loci associated with variation in kernel size (Fiedler *et al.*, 2017), grain yield and its components (Mangini *et al.*, 2018; Sukumaran *et al.*, 2018; Wang *et al.*, 2019), but also response to moisture changes and roots. Maccaferri *et al.* (2011) used an association mapping to dissect the genetic basis of drought-adaptive traits and grain yield (GY) in a collection of 189 elite durum wheat accessions evaluated in 15 environments highly different for water availability during the crop cycle (from 146 to 711 mm). For GY, significant associations were mostly detected in one environment only, while decreasing rapidly from two to five environments and with only one marker found significant in six environments. While in another study, Maccaferri *et al.* (2016) used linkage and association mapping for root system architecture in two recombinant inbred line populations and one association mapping panel of

183 elite durum wheat accessions evaluated as seedlings revealed 20 clusters of QTLs for root length and number, as well as 30 QTLs for root growth angle (RGA). Divergent RGA phenotypes observed by seminal root screening were validated by root phenotyping of field-grown adult plants.

In the present study, we aimed at broadening the understanding of the genetic factors involved in controlling water productivity and moisture stress adaptation in durum wheat. Therefore, we investigated the performances of a large ‘discovery panel’ of durum wheat accessions across 18 environments experiencing different degrees of in season moisture. Beyond the identification of stable and top performing entries, this investigation sought to define discrete clusters of water productivity types. GWAS on the ‘discovery panel’ was then used to identify haplotypes more frequently present in the most water responsive genotypes, which were then investigated in a second ‘confirmation panel’. Finally, to ensure these critical loci can be readily incorporated into novel cultivars, Kompetitive Allele Specific PCR (KASP) markers were developed to tag them and then confirmed for their ability to predict moisture stress adaptation on a third ‘validation panel’.

## Material and methods

### Plant material

This study evaluated three discrete germplasm panels and all associated phenotypic and genotypic data are made available as Table S1. The first panel is defined as the ‘discovery panel’ and it includes 384 durum wheat entries including landraces, elites and cultivars (Table S3). The kinship of this panel was previously presented by Kabbaj *et al.* (2017), and it has already been used to identify the genomic loci involved in resistant to a damaging insect pest (Bassi *et al.*, 2019), phenology (Gupta *et al.*, 2020), and its response to heat stress (El Hassouni *et al.*, 2019). This panel was tested in its entirety at some environments, while a subset was used in other environments as explained in more details below. The second set of entries is defined as the ‘confirmation panel’ and it includes 80 ICARDA’s elites that constituted the 2019 international nurseries 42<sup>nd</sup> International Durum Observatory Nursery (IDON; Table S3). The third set is defined as the ‘validation panel’ and it includes 80 ICARDA’s elites that constituted the 2020 international nurseries 43<sup>rd</sup> IDON (Table S4). The last two panels share some co-ancestry as it can be expected from germplasm developed by a breeding program. The ‘discovery’ panel also includes few entries that were later used as parents to derive the following two panels.

### Field trials and management

The ‘discovery panel’ was assessed during 2014-15, 2015-16, 2016-17 and 2017-18 growing seasons in eighteen contrasting environments as described in **Fig. 1**. Four were in Morocco: Marchouch (MCH), Sidi El Aidi (SAD), Melk Zhar (MKZ) and Tessaout (TES), two in Lebanon: Terbol (TER) and Kfardan (KFD) and one in Jordan: Musghar (MUS). The experimental design was an augmented design with four replicated checks in the 2014-15 (15) and 2015-16 (16) growing seasons in MCH15, MCH16, SAD16, MKZ15, MKZ16, TES16, TER15, TER16, KFD16, and MUS18 during 2017-18 season (18). During 2016-17



(17) and 2017-18 (18) seasons, a subset of 144 genotypes was selected and used to run an alpha-lattice design with two replications and twelve incomplete blocks of size twelve, at MCH17, MCH18, SAD17, TES17, KFD17 and KFD18. Each entry was planted in plots of 6 rows of 5 m in length, row spacing was 0.2 m, for a total sown surface of 6 m<sup>2</sup> at a seeding rate of 120 kg ha<sup>-1</sup>. Agronomic practices follow a timely sowing date between 15<sup>th</sup> of November to 15<sup>th</sup> of December with a base pre-sowing fertilizer application of 50 kg ha<sup>-1</sup> of N, P and K. Planting occurred after a legume crop season. During 2016-17 and 2017-18 seasons in MCH, two management conditions have been used: normal sowing (MCHN) following standard land preparations and tillage, and zero tillage (MCHZ) on a fully retained faba bean stubble. Both sowings were conducted using the same seeder, even though it was specifically developed for zero till practices. At stage 14 of the Zadok's scale (Z) herbicide was applied in a tank mixture to provide protection against both monocots and dicots. A week after herbicide application, ammonium nitrate was provided to add 36 kg ha<sup>-1</sup> of N. When in season moisture exceeded 350 mm a final application of urea was used at flowering to deliver additional 46 Kg ha<sup>-1</sup> of N. In KFD17 and KFD18, two kind of fertilizer applications were done: KFDA with only basal fertilization 50 kg ha<sup>-1</sup> of N, P and K and KFDB with additional 50 kg ha<sup>-1</sup> of Urea at Z15. In MKZ, the first basal fertilization was followed by 5 split applications each of 20 Kg ha<sup>-1</sup> of N via fertigation through drip pipes. Three sites were irrigated: TES, where four gravity irrigations of 35 mm each were provided after Z10, Z18, Z45, and Z65; MKZ, where 12 irrigations of 10 mm each were provided via drip irrigation at one week interval from two weeks after Z10 to Z89 and TER, where two sprinkle supplemental irrigation of 20 mm each were provided before Z10 and after Z65. The remaining experiments were conducted under rainfed conditions with total rainfall values and other details presented in **Fig. 1**.

The station of Sidi El Aidi (SAD) in Morocco was identified by the global initiative CRP WHEAT as an ideal site to test for drought tolerance of wheat, and for this reason it was also used to screen the two other panels. The 'confirmation panel' was tested at SAD and MCH, while the 'validation panel' was tested just at SAD. Both panels were planted as augmented design with four replicated checks, during seasons 2018-19 and 2019-2020, respectively. The total moisture recorded during 2018-19 was 296 and 154.6 mm at SAD and MCH respectively, while during 2019-2020 at SAD 286 mm, which constitute strong moisture stress for durum wheat.

## Phenotyping

Days to heading (DTH) was recorded as days elapsing between sowing and 50% of plants showing emerging heads. At maturity, the number of fertile spikes were counted in 0.25 m<sup>2</sup> and this value was multiplied by four to derive the number of spikes per m<sup>2</sup> (SPK). Grain yield (GY, kg ha<sup>-1</sup>) was recorded by harvesting the central four rows of each plot, weighting it on a precision scale and dividing this value by the plot surface. From the harvest of each plot, 1,000-kernel weight (TKW, g) was determined by counting five hundred randomly selected grains on a 'Chopin Numigral' counter followed by weighting on a precision scale. The number of grains per meter square (Gr.m<sup>-2</sup>) was calculated using the total weight of the plot, divided by the harvested surface and the estimated weight of one kernel, derived from the TKW value, as per (1):

$$Gr.m^{-2} = \frac{\text{Harvested weight of plot}}{\text{Plot area} \times \frac{TKW}{1000}} \quad (1)$$

The number of grains per spike (GpS) was then derived by dividing the number of grains per meter square (1) by the number of spikes recorded for the same area as follows (2):

$$GpS = \frac{Gr.m^{-2}}{Spk.m^{-2}} \quad (2)$$

As detailed in **Fig. 1**, GY was recorded in all environments, while the other traits were collected only in some.

## Field data analysis

Analysis of variance was performed using Genstat for the augmented designs, while alpha lattice and the combined analysis were run on GEA-R 4.1 in the R environment (Pacheco *et al.*, 2015). Combined ANOVA across mega-environments was obtained by linear model fitted considering genotypes as fixed term (Team R. Core, 2017). Best linear unbiased estimates (BLUEs) were calculated for each genotype at each environment defining genotypes as fixed effect using the R package ASReml-R (Butler *et al.*, 2009). The package ASReml-R was also used to estimate the narrow-sense heritability. Broad-sense heritability was calculated separately for each design by Genotype x Environment Analysis with R (GEA-R) version 4.1. The ratio of variance accounted by each source of variations (G, E, and GxE) was calculated dividing the sum of square of each source by the total sum of the square.

For grain yield, GxE was partitioned by additive main effects and multiplicative interaction (AMMI) model using R package Agricolae (De Mendiburu and Yaseen, 2020). The 'AMMI wide adaptation index' (AWAI) measures the distance of each genotype from each significant IPCs axis and it was calculated using the following formula, presented by Bassi and Sanchez-Garcia (2017):

$$AWAI = \sum_i s_i \times |PC_i|$$



Where  $i$  is the number of significant IPCs determined by classical F-test in R (Team R. Core, 2017),  $s_i$  is the percentage of total GxE variance explained by each IPC, and PC is the actual IPC value. AWAI values close to '0' are obtained for the most widely adapted and stable germplasm (Malosetti *et al.*, 2013). As indicated by Bassi and Sanchez-Garcia (2017), a biplot between the genetic (G) component of yield (i.e. yield potential) and the interaction (GxE) component (i.e. yield stability) was used to determine the best genotypes combining both G and GxE for grain yield. The AWAI index explaining GxE was presented as ratio to minimum value, and values close to '1' were obtained for the most widely adapted and stable genotypes. To define the genetic component of GY obtained across environments with vast differences in the average performances, the actual values were converted to a ratio of the top performing entry at each environment and then averaged across.

A climate matrix was developed for each environment, splitting the records into five growth stages: one month before sowing, sowing until the end of the vegetative stage, flowering stage, grain filling period, and physiological maturity period. Simple linear regression was conducted between the climatic matrix and the response of genotypes at each site for GY. The climatic factors having a significant effect ( $p < 0.05$ ) were used to perform hierarchical clustering among environments using the R package FactoMineR (Josse *et al.*, 2008) (**Fig. 2**).

### Assignment of genotypes to water productivity classes

The water productivity (WP) is calculated using the following formula:

$$WP = \frac{\text{Grain yield}}{\text{Total moisture}}$$

To define the average trend, the average GY performances of each environment was plotted against the total moisture of that environment, which corresponds to a graphical representation of the average WP trend. To increase the accuracy of this trend, the environments were split into two groups, one defined as 'stressed' including 11 environments experiencing moisture stress, and the second defined as 'non-stressed' including seven environments where moisture stress did not occur. The slope (b) of WP was calculated for each group, reaching 5.08 for moisture stress cluster and 4.85 for non-moisture stress. These values represent then the hypothetical average performance of a given genotype tested at that group of clusters. Hence, higher values (steeper response to water increase) would be obtained by genotypes with higher WP, while lower values (flatter curve) would be associated to less responsive genotypes. To assess this, the GY performance of each genotype at each environment was plotted against the moisture level of that environment and the actual slope value ( $b_i$ ) for the two trend lines (moisture stressed and non-stressed) were calculated (**Fig. 2**). Based on these  $b_i$  values, genotypes were assigned to four different water responsive classes representing more or less responsiveness compared to the average trend (Table 1). However, to ensure that the trendline explained the observed changes in moisture,

genotypes for which the regression value between GY and moisture levels was not significant ( $p < 0.01$ ) were assigned to a fifth class of water unresponsive behavior.

### **Genotyping and association mapping analysis**

The ‘discovery panel’ was genotyped with 35K Affymetrix Axiom wheat breeders array ([www.affymetrix.com](http://www.affymetrix.com)) to generate 7,652 polymorphic SNP with 98 to 100% identity when blast aligned to the Svevo genome (Maccaferri *et al.*, 2019), less than 1% missing data, minor allele frequency higher than 5% and heterozygosity less than 5% as detailed in Kabbaj *et al.* (2017). These authors also defined a kinship structure of 10 sub-clusters. Genome wide linkage disequilibrium (LD) decay analysis were performed by Bassi *et al.* (2019) and defined as 51.3 Mbp. GWAS was performed for the panel using as phenotypic input the BLUEs of each trait at each environment, and using the BLUEs estimated from the combined analysis of the two moisture groups (stressed and non-stressed). TASSEL 5 software (Bradbury *et al.*, 2007) was used for the analysis imposing DTH as covariate to avoid identifying flowering genes, since these were already described in Gupta *et al.* (2020). Two models were used and compared using two additional covariate parameters, Q (population structure) and K (Kinship). Q model was performed using a general linear model (GLM), and Q + K model using a mixed linear model (MLM). The best model for each trait was selected based on the quantile-quantile (Q-Q) plots (Sukumaran *et al.*, 2012). Significant marker-trait associations (MTA) were determined using a Bonferroni correction by LD as suggested by Duggal and Beaty (2008) for  $p < 0.05$  corresponding respectively to  $LOD = 2.69$  (Bassi *et al.*, 2019). In addition, Pearson’s critical values (Pearson, 1895) for correlations  $r$  was squared to obtain a critical  $r^2$  of 0.024 ( $p < 0.01$ ) and used to determine significant markers explaining sufficient ratio of the total phenotypic variation. Any marker-traits associations (MTAs) with  $LOD$  and  $r^2$  superior to these cut-offs were considered valid and presented here. MTAs falling at a distance inferior to twice the LD (102.6 Mbp) were deemed to be too physically close to be resolved by this panel into distinct loci and hence were assigned the same QTL identifier. QTL were defined as “consistent” when it included MTA significant in both the combined analysis across sites, and in more than one individual environment.

The ‘confirmation panel’ was genotyped using a 23K array chip developed by SGS - Institut Fresenius TraitGenetics Section (Germany) which incorporates 14.5K SNPs from the 90K Infinium Array, 8.5K SNPs from the Axiom array, and 265 SNPs reported as linked to genes in the literature (Vitale *et al.*, 2021). Marker curation was conducted as for the ‘discovery panel’, to result in 6,325 polymorphic SNPs. These were also aligned to the Svevo genome assembly (Maccaferri *et al.*, 2019). A kinship structure of 8 sub-clusters was identified (Table S3) and linkage analysis revealed that the LD was 21.2 Mbp, which resulted in a significant  $r^2 = 0.05$  ( $p < 0.01$ ). The two genotyping platforms were merged using the Svevo genome assembly as scaffold based on physical overlap. Similarly to the ‘discovery panel’, GWAS was conducted using flowering time as covariate. Significant MTA were determined using Bonferroni correction for  $p < 0.05$  corresponding to  $LOD = 4.1$ . ShinyCircos software

(Yu *et al.*, 2018) was used to graphically represent the MTA and QTL identified by both ‘discovery’ and ‘confirmation’ panels. Only those “consistent” QTLs identified by GWAS in the ‘discovery panel’ and then also identified by GWAS in the ‘confirmation panel’ were deemed “true” positive and studied further.

The most representative marker for each “true” QTL was selected based on its higher LOD and explaining a broader fraction of the phenotypic variation. Discrete classes of genotypes from the ‘confirmation panel’ were then defined based on their allelic combinations at these representative markers. These classes were defined as ‘haplotypes’. The phenotypic performances of the ‘confirmation panel’ genotypes belonging to each haplotype class were defined as a random effect, and a linear model was run to determine significance difference by LSD using LSD.test function of *agricolae* package (De Mendiburu and Yaseen, 2020; Team R. Core, 2017).

The 35K or 25K array probe sequences underlying the most interesting QTLs were submitted to LGC to run their proprietary software to assess their suitability to design KASP primers. For each QTL, four potential primer sets were synthesized and run on the ‘validation panel’. For each KASP marker that amplified and showed polymorphism, its allelic score was regressed against the GY value and a significance threshold was set at  $r^2 > 0.053$  ( $p < 0.01$ ). In addition, the top 20 yielding genotypes were defined as the ‘positive’ cases and the worst 20 genotypes as the ‘negative’ cases. The marker score was then evaluated among the positive and negative cases to define correct SNP call (true positive or true negative) and the wrong SNP calls (false positive and false negative). The marker accuracy was then calculated as the ratio of the correct allelic call among all, sensitivity as the ratio of the correct positive allelic calls among all, and specificity as the ratio of the correct negative allelic calls among all. The primer sequence of the markers is protected by commercial rights and cannot be disclosed here, but these can be purchased by all users as service via LGC indicating the marker names provided.

## Results

### Phenotypic variation under moisture stressed and non-stressed environments.

Analysis of variance revealed significant differences ( $p < 0.01$ ) for the genotypes (G), environments (E), and their interaction (G×E) for most of the traits (Table S5). The E effect explained most of the variation for GY (73%), GpS (69.3%), TKW (83%), SPK (88%) and DTH (84%), while the G effect explained the largest variation for GpS (16%). The G×E interaction showed a larger contribution to the total variability compared to the G effect for GY and SPK. Good heritability was obtained at all environments for all traits.

Vast phenotypic variation was recorded for all traits across the eighteen environments (Fig. S2). The highest average GY was recorded in TER16 (6,413.5 kg ha<sup>-1</sup>), while MUS18 had the lowest average GY (330.1 kg ha<sup>-1</sup>) (Fig. 1). Moisture data shows patterns of variation across environments, with some sites having a prevalence of

drought events (MUS, KFD, SAD, MCH, Fig. 1). GY performances were significantly ( $p < 0.05$ ) influenced by the total water input during vegetative, flowering and grain filling stages and the maximum temperature during the flowering stage (Table S6). Because of their significant effect on yield, these climatic factors were used to cluster the environments by PCA in two mega-environments: *i.* moisture stressed (MUS18, KFD17, SAD16, MCH15, MCH17, KFD16, MCH16 and KFD18) and *ii.* non-moisture stressed (TES16, TER15, MKZ15, TES17, TER16 and MCH18).

To avoid range effects, grain yield (BLUE) was converted to 'ratio to the max', to scale the variation based on the best performing entry. Under moisture stressed conditions the CIMMYT line GID: 800032262 (3,041 kg ha<sup>-1</sup>) was the top yielding. Among the top highly performing entries, the ICARDA wide crosses GID:800032191 (2,564 kg ha<sup>-1</sup>), GID:800043267 (2,375 kg ha<sup>-1</sup>), GID: 800032178 (2,242 kg ha<sup>-1</sup>) and the elite lines GID: 800032351 (2,204 kg ha<sup>-1</sup>) and GID: 800030179 (2,191 kg ha<sup>-1</sup>). The top yielding line under non moisture stressed conditions was the ICARDA elite GID: 800043103 (7,517 kg ha<sup>-1</sup>), the Moroccan line GID: 4984522 (6,407 kg ha<sup>-1</sup>) was also among the highly yielding.

Partitioning the GxE effect by AMMI defined seventeen significant components (PCs), of which the first three combined accounted for 67.9% of the variation. The definition of the AWAI score determined an average performance equal to 0.2, with the two most stable lines being GID: 800032258 a CIMMYT elite line and GID: 800032351 an ICARDA elite line. The bi-plot combining GY performances across sites and stability (AWAI) provides an ideal selection index to combine G and GxE effects (Fig. 3) (Bassi and Sanchez-Garcia, 2017). Combined analysis under moisture stress identified 24% of genotypes having higher than GY and AWAI average. Lines GID: 800032262, GID: 800032261 and GID: 800032280 were the top 3 stable and yielding. The ICARDA elite line GID: 800030179 and GID: 800032191 were also among the highly performing. While under non-stress conditions, 32% of the tested entries had higher yield and AWAI than the average. The highly yielding genotypes GID: 800043103 and GID: 4984522 were not stable. Contrary to the Australian elite line GID: 800032336, CIMMYT line GID: 800032267 and the ICARDA line GID: 800032342 were the top 3 stable and highly performing.

Beside stability per se, several traits contribute to the adaptation of genotypes to the environment. To determine which traits contributed to GY variation, correlation analysis was performed for each individual and mega-environment. This interaction (Table 2) revealed that GpS influenced ( $p < 0.001$ ) GY in all environments. TKW had an effect only in 5 out of 9 moisture stressed environments and in all the non-stressed ones. Overall, SPK was not significantly correlated with GY under non moisture stress, only two environments showed highly significant relationship, while it accounted for 60% of yield variation in moisture stressed environments. In most of the environments flowering time shows a highly significant correlation with GY.

## Water productivity performance of genotypes

Climatic regression against GY identified that moisture amount during the vegetative, flowering and grain filling stages are the most significant climatic factors, explaining more than 70% of its variation (Table S6). To better elucidate the relationship between GY and moisture, a water productivity (WP) value was calculated for each genotype. The average WP was estimated at 7.2 kg ha<sup>-1</sup> mm<sup>-1</sup> showing a significant linear relationship ( $r^2 = 0.327$ ) to the increase in moisture levels, resulting to an average WP of 5.1 kg ha<sup>-1</sup> mm<sup>-1</sup> across moisture stressed environments, while it reached 9.7 kg ha<sup>-1</sup> mm<sup>-1</sup> across non stressed environments (Fig. S3).

A subset of 120 genotypes that have been assessed at all environments, was assigned to WP classes based on their respective trend of yield variations plotted against the moisture levels across environments. One quarter of the tested entries were assigned to class 3 'Responsive to high moisture' and one quarter to class 4 'Highly water responsive', while class 2 'Responsive to low moisture' incorporated 20% of genotypes, 18% belong to class 5 'No water response' and 13% as class 1 'Stable water response'. From a breeding perspective, class 2, 3, and 4 are the most interesting because they identify genotypes capable of producing more yield per water input compared to the average. Interestingly, genotypes GID: 800030179, GID:800043267 and GID: 800032178 resulted among the highly WP performing elite lines under moisture stress, belonging to class 2 'Responsive to low moisture'. While under non moisture stress, the ICARDA elite line GID: 800028568 belonging to class 3 'Responsive to high moisture' was among the highest for WP. Instead, genotype GID: 800032054, a CIMMYT line with top yield under both moisture conditions, belongs to class 4 'Highly water responsive' (Table S7).

## QTLs controlling traits under moisture stressed and non-stressed conditions

Initially, GWAS was conducted on the 'discovery panel', involving individual and both combined mega-environments analyses for all traits, resulting in the identification of 280 significant MTAs. The MTAs explained from 3% to 22% of the phenotypic variation and LOD ranged from 2.7 to 7.2. MTAs were distributed across 47 discrete QTL (Fig. 4, Table S8).

Under non-moisture stress environments, four consistent QTLs (Q.ICD.04, Q.ICD.07, Q.ICD.37 and Q.ICD.39) associated with GY were identified on chromosomes 1A, 1B, 6A and 6B (Fig. 4). Q.ICD.37 was also associated with GpS and SPK, while Q.ICD.39 and Q.ICD.04 also controlled TKW and GpS.

Under moisture stressed conditions, GY was associated to 14 loci (Fig. 4). Among these, Q.ICD.08, Q.ICD.11, Q.ICD.17, Q.ICD.20, Q.ICD.28 and Q.ICD.44 on chromosomes 1B, 2A, 3A, 3B, 4B, 7B, respectively were identified in two or more stressed environments. Interestingly, locus Q.ICD.28 was also associated with TKW, SPK and GpS, while Q.ICD.44 controlled TKW, in addition to GY.



A comparison of significant loci for GY across stressed and non-stressed conditions identified a consistent locus on chromosome 7A (Q.ICD.41), also controlling GpS and SPK, on chr 5 A (Q.ICD.31) associated with GY, TKW, and SPK, and on chr 1A (Q.ICD.01) controlling GY, TKW, SPK and GpS. This last QTL was the most frequently identified region across all environments.

The GWAS conducted for yield stability (AWAI) revealed two QTLs (Q.ICD.02 and Q.ICD.13) on chromosomes 1A and 2B. Interestingly, both QTLs were linked to GpS and TKW. In addition, for TKW, two additional loci (Q.ICD.18 and Q.ICD.35) not associated with GY were identified on chromosomes 3A and 5B.

Conducting GWAS for the 'confirmation panel' tested at SAD and MCH, confirmed the importance of Q.ICD.08, Q.ICD.11, and Q.ICD.44, which were also identified as important QTL in the 'discovery panel' under moisture stress-conditions.

### **Effect of different allele combination on water productivity classes**

To determine the allelic effect on grain yield across environments, the main representative marker of each QTL was investigated as single marker regression at all locations (Fig. 5). The major allele of AX-94549122 is strongly correlated with grain yield under both moisture conditions, while for AX-95631864 is the minor allele that is associated with both conditions. Hence, these two loci contribute to yield overall. Instead, AX-94910470 major allele is mostly important for non-stressed environments, while AX-95191125 minor allele is linked only to stressed conditions. Hence, these two loci control yield performances under different moisture conditions. The combination of major allele at all QTLs explained variation at all non-stressed environment, and only in few stressed ones.

To better assess the interaction between QTL, the 'discovery panel' entries assigned to the five WP classes were investigated for their haplotype composition at these three QTLs (Q.ICD.08, Q.ICD.11 and Q.ICD.44). Four haplotypes groups could be identified (Fig. 5). Haplotype 1 with favorable alleles at all loci reached the highest average water productivity of  $6.5 \text{ kg ha}^{-1} \text{ mm}^{-1}$ , with 40% of the genotypes belonging to class 2 'Responsive to low moisture'. Interestingly, the same allelic combination is responsible for high grain yield under drought (Fig. S3). Haplotypes 2 and 3 reach an average water productivity of respectively  $5.3$  and  $5.4 \text{ kg ha}^{-1} \text{ mm}^{-1}$ , with only one positive allele. Haplotype 2 contributes equally to the four classes. While haplotype 3 is mainly expressed by class 4 'highly water responsive'. Haplotype 4 harboring three negative alleles reaches  $5 \text{ kg ha}^{-1} \text{ mm}^{-1}$ , with classes 3 and 4 'responsive to high moisture' and 'highly water responsive' corresponding to the highest portions of this haplotype.



### **Confirmation of haplotype effect**

The three main QTLs (Q.ICD.08, Q.ICD.11, and Q.ICD.44) were investigated for their additive effect within the ‘confirmation panel’. A total of five haplotypes were identified within the panel for these three loci (Fig. 6). The panel was tested under moisture stress at two environments (Sidi el Aydi and Marchouch) during season 2018-19. The linear model confirmed that the haplotype groups represented discrete classes with significant difference. Haplotype 1 (Hap1) carrying only favorable alleles at all QTLs showed a GY advantage of more 705 kg ha<sup>-1</sup> compared with haplotype 7 with no positive alleles at the three loci and a consequent gain in water productivity of 2.2 kg ha<sup>-1</sup> mm<sup>-1</sup>. Also, Hap3, with 2 positive alleles, except for Q.ICD.11 was significantly superior to Hap7 (no positive alleles), but it was not superior to Hap 4 (only 1 positive allele). This suggests that Q.ICD.11 has the strongest effect, followed in order by 8 and 44.

### **Conversion and validation to KASP**

Markers conversion and validation are quintessential steps to convert the discovery of QTLs into usable tools for breeders. Out of 36 array probes known to span the three major QTLs (Q.ICD.08, Q.ICD.11 and Q.ICD.44), KASP primers could be designed for 32 of them; of these 17 were purchased and used to screen the ‘validation panel’. Nine of these detected a polymorphism within this elite set (MAF>3%). Two explained a significant ( $p>0.05$ ) and three a highly significant ( $p<0.01$ ) portion of the phenotypic variation for grain yield (Fig. 7), when assessing the panel at the severely drought affected station of Sidi el Aydi during season 2019-20. KASP were validated for Q.ICD.08 located on chromosome 1B, and one each for Q.ICD.11 and Q.ICD.44 on chromosome 2A and 7B, respectively. All five markers are suitable for use in MAS, and their use in combination shall further increase their independent scores. In fact, AX-95631864 (Q.ICD.08) has overall the best average performance for all criteria, and the highest prediction of phenotypic variation ( $r^2 = 0.10$ ), while AX-94507963 (Q.ICD.08) is particularly suitable to identify the top yielders (true positive) with the highest overall sensitivity but it has low precision, instead AX-94549122 (Q.ICD.08) and AX-95191125 (Q.ICD.44) have perfect precision (true negative) in identifying the lines to be discarded, but low sensitivity. AX-94910470 tags the hypothetically strongest QTL (Q.ICD.11) but within the ‘validation panel’ its contribution was minor. However, the combined selection for carrying the positive allele at all five markers resulted in a drastic increase in precision, with only few top performing lines being selected.

### **Discussion**

#### **Water productivity classes explain the genotype response to moisture stress**

This study evaluated a global ‘discovery panel’ at 18 environments. Climatic regression (Table S6) confirmed that grain yield variation at these environments was mainly controlled by the moisture availability during the vegetative, flowering, and grain filling phases. Liliane and Charles (2020) and El Haddad *et al.* (2021) also

found that moisture availability during the vegetative stage is a critical climatic factor influencing the response of durum wheat genotypes. In fact, the major negative impact of drought stress on wheat is the reduction in fresh and dry biomass production (Farooq *et al.*, 2009), which affects grain number and grain size (Dickin and Wright, 2008). A common response to cope with drought stress is stomatal closure, since that also alters the photosynthetic rate, plants must constantly adjust stomatal conductance to maintain a balance between sufficient CO<sub>2</sub> uptake and water loss. Toulotte *et al.* (2022) hypothesized that reduced rates of stomatal conductance and subsequently decreased water loss due to reduced stomatal density, allowed the available plant resources to be allocated to seed propagation and aboveground Biomass. A previous study (Sokoto and Singh, 2013) revealed that water stress at vegetative stage significantly reduced spike length and grains per spike. While water stress at flowering and grain filling significantly reduced 1000 kernel weight, grain yield and harvest index. In fact, water stress induced accelerated senescence after anthesis shortens the duration of grain filling by causing premature desiccation of the endosperm and by limiting embryo volume has also been reported (Westgate, 1994).

Principal component clustering based on the most critical climatic factors allowed us to classify the sites into two mega environments (Fig. 1): moisture stressed and non-stressed. The effect of moisture was further partitioned by assigning genotypes to five classes of WP (Table 2). The same set of genotypes was tested for yield potential under the two mega-environments as well as yield stability overall. Interestingly, the highly yielding genotypes under moisture stress belong to class 2 'Responsive to low moisture'. While under non stress, the highly yielding belongs to class 3 'Responsive to high moisture'. And the class 4 'Highly water responsive' represent mainly the genotypes highly yielding under both conditions (Fig. 2). Similarly, to Siahpoosh and Dehghanian (2012) genotypes were significantly different for WP. The highly phenotypic variation was due to the environmental effect. The plant response to water stress varied. The decrease of production can be due to the plant defense by reducing stomatal conductance and CO<sub>2</sub> assimilation rate (Catola *et al.*, 2016). While, when the plant had high water productivity under moisture shortage, Stallmann *et al.* (2020) explained this reaction by the increase of the intrinsic plant water use efficiency caused by the stomatal closure, which restrict transpiration before it inhibits photosynthesis.

Correlation analysis (Table 2) was done to determine the main traits contributing to grain yield under drought. Interestingly, grain yield was positively correlated with yield components in moisture stressed environments. These findings were consistent with Kiliç and Tacettin (2010) and Al-Ghzawi *et al.* (2018), who reported that spike per m<sup>2</sup>, grains per spike and TKW were directly related to grain yield. Since, genetic research has shown that it is possible to increase grain size without a negative effect on grain number (Rivera-Amado *et al.*, 2019). The negative correlation between yield under stress and heading date has frequently been reported (Dodig *et al.*, 2010; Gonzalez-Ribot *et al.*, 2017), indicating that the most precocious genotypes would be

desirable in accordance with other reports for Mediterranean environments (Acevedo and Ceccarelli, 1989; Quarrie *et al.*, 1999; Richards *et al.*, 2001).

### Genetic dissection of drought tolerance in durum wheat

Breeding cultivars able to thrive under moisture stressed conditions is challenging, since wide adaptation is hindered by high genotype by environment interaction. Drought tolerance is a complex quantitative trait controlled by an army of loci interacting with the environment. Blanco *et al.* (2011a) reported that genomic regions linked with GY are present in all chromosomes, and that the magnitude of their effect varies based on the environment. Eleven loci were identified as responsible for the control of GY in more than one environment in our study (Fig. 4, Table S8). Four QTLs (Q.ICD.04, Q.ICD.07, Q.ICD.37, Q.ICD.39) were active under non moisture stressed conditions located on chromosomes 1A, 1B, 6A and 6B, six QTLs (Q.ICD.08, Q.ICD.11, Q.ICD.17, Q.ICD.20, Q.ICD.28, Q.ICD.44) on chromosomes 1B, 2A, 3A, 3B, 4B, 7B under moisture stress, and QTL.ICD.41 and Q.ICD.01 on chromosome 1A and 7A were common under both conditions. Among these QTLs, those located on 1A-1B (Q.ICD.04 and Q.ICD.07), 3A-3B (Q.ICD.17 and Q.ICD.20), and 6A-6B (Q.ICD.37 and Q.ICD.39) control similar functions, are located on homoeologous physical positions, and hence could represent homoeologous loci on different genomes. Sukumaran *et al.* (2018), Rahimi *et al.* (2019) and Muhu-Din Ahmed *et al.* (2020) all identified QTLs on chromosomes 1A, 6A and 6B related to GY under non moisture stressed conditions in durum wheat. Interestingly, most QTLs controlled also at least one of the yield components. In particular, Q.ICD.01 was linked to all measured four traits (TKW, SPK and GpS) and a similar region was already identified by other authors for its importance in wheat for GY under different water regimes (Ain *et al.*, 2015; Charmet *et al.*, 2001; Gupta *et al.*, 2017; Muhu-Din Ahmed *et al.*, 2020; Rahimi *et al.*, 2019; Zandipour *et al.*, 2020), yield stability (Sehgal *et al.*, 2020), TKW (Lozada *et al.*, 2017; Neumann *et al.*, 2010; Ogbonnaya *et al.*, 2017), GpS and SPK (El Hassouni *et al.*, 2019). Similarly, to previous studies (Shokat *et al.*, 2020; Xin *et al.*, 2020), TKW was positively correlated with grain yield under drought and irrigated conditions (Table 2), indicating that plant genotypes having higher TKW under irrigated conditions often have a chance to maintain higher TKW under drought conditions (Shokat *et al.*, 2020). Less reduction in TKW will ideally allow good yield under drought conditions. Previous studies have found QTLs for TKW on almost all chromosomes of the wheat genome (Cabral *et al.*, 2018; McCartney *et al.*, 2005; Okamoto *et al.*, 2013; Pradhan *et al.*, 2019; Williams *et al.*, 2012; Zhang *et al.*, 2015), we found the same except on chromosome 7A. The most consistent loci have been detected on chromosomes 3A and 4B; Q.ICD.18 and Q.ICD.28 can be compared with previous finding by Pinto *et al.* (2010) and Sun *et al.* (2017). Grain per spike is correlated with grain yield under both moisture and no moisture stressed conditions (Pradhan *et al.*, 2019), preserving high GpS during drought conditions is important in order to keep good yield. We found significant associations of GpS (Q.ICD.07 and Q.ICD.31) with chromosomes 1B and 5A under

both conditions. While significant regions for spike per m<sup>2</sup> (Q.ICD.25 and Q.ICD.46) were mainly located on chromosomes 4B and 7B. For AMMI wide adaptation index, 4 MTA spanned on 2 QTLs (Q.ICD.02 and Q.ICD.13) were detected on chromosomes 1A and 2B. Contrary to the finding of Sehgal *et al.* (2017), both QTLs were not associated to GY instead were linked to TKW and grain per spike. Recently, Sehgal *et al.* (2020) identified haplotypes blocks associated with stability index Pi on chromosome 1A, using advanced bread wheat lines under contrasting environments. While the role of chromosome 2B in controlling stability was previously reported in a large elite panel of wheat (Sehgal *et al.*, 2017) and a winter wheat population (Lozada and Carter, 2020), as most of the significant MTAs controlling yield trait stability were detected in this chromosomal region.

GY showed a significantly positive correlation with TKW, GpS and SPK (Table 2), indicating that the increased GY under moisture stress resulted from increased yield components. While under non moisture stress, the GY increase is due to the significant relationship with TKW and GpS. Consequently, it is feasible to improve GY by selecting these yield related traits in breeding programs because of the more accurate measurement across moisture and non-moisture stressed environments in comparison with yield.

The three QTLs Q.ICD.08, Q.ICD.11 and Q.ICD.44 on chromosomes 1B, 2B and 7B linked to grain yield under low moisture (Fig. 5), were used to investigate the allelic combination responsible for water productivity. Interestingly, class 2 genotypes had the positive alleles for all three loci providing a significant water productivity advantage of +1.5 kg ha<sup>-1</sup> mm<sup>-1</sup> under low moisture environments. While class 4 had one positive allele at AX-94910470 belonging to Q.ICD.11.

The three main QTLs were confirmed by a second independent ‘investigation panel’ grown under moisture stress (Fig. 6). The haplotype assessment confirmed that carrying the positive alleles at all loci increased grain yield by +704.6 kg ha<sup>-1</sup> and water productivity by 2.2 kg ha<sup>-1</sup> mm<sup>-1</sup>. Juliana *et al.* (2021) by using a large scale GWAS, reported the highest number of consistent GY GBS markers on chromosomes 2A, 6B, 6A, 5B, 1B and 7B. Similarly, several studies have also reported QTLs on chromosome 1B responsible for the control of GY in durum wheat (Rehman Arif *et al.*, 2020; Roncallo *et al.*, 2017; Xu *et al.*, 2017). The effect of Q.ICD.08 under moisture stress was also identified in Juliana *et al.* (2021) study, loci controlling GY under optimum and drought. That can be explained by Mathew *et al.* (2019) findings, who reported root and shoot biomass association region on chromosome 1B. Further, Pshenichnikova *et al.* (2021) found 53 QTLs associated with physiological and agronomic traits under contrasting water supply. Those findings may explain the importance of Q.ICD.11.

Q.ICD.44 was associated with TKW and grain yield in combined and four environments under drought represent important loci. Zaïm *et al.* (2020) found in a recent mapping populations study tested across dry environments, a consistent QTL

for GY in the same chromosome. Similarly, by using a diverse population of winter wheat, Lozada and Carter (2020) found a site controlling multiple yield trait and trait stability measures in the same chromosome.

Interestingly, within Q.ICD.08 a gene encoding hydroxyproline-rich glycoprotein (HRGP) from the late embryogenesis abundant (LEA) family has been pinpointed (Yates *et al.*, 2021). This gene plays a crucial role in enhancing the plant's ability to withstand drought stress (Ali *et al.*, 2020). The HRGP assists in maintaining optimal cellular hydration, protecting against water loss, stabilizing cell structures, and reducing oxidative damage. Its upregulation during drought stress indicates its involvement in stress-responsive pathways (Liu *et al.*, 2020; Ringli *et al.*, 2010). Further, Q.ICD.11 contains key genes related to water stress response, including Ethylene-responsive transcription factor, Dehydration-responsive element binding protein, and AP2-like ethylene-responsive transcription factor (Yates *et al.*, 2021). These genes play pivotal roles in enhancing the plant's ability to face drought conditions. The Ethylene-responsive transcription factor is involved in regulating stress-related gene expression (Djemal and Khoudi, 2015), while the Dehydration-responsive element binding protein contributes to water conservation mechanisms (Agarwal *et al.*, 2017; Buffagni *et al.*, 2020). The AP2-like ethylene-responsive transcription factor aids in orchestrating various stress responses (Yu *et al.*, 2022). While Q.ICD.44 contains genes encoding aquaporin-like proteins (Yates *et al.*, 2021), essential for regulating water movement within the plant. These proteins facilitate efficient water uptake and distribution, aiding the plant in coping with water scarcity during drought conditions (Ayadi *et al.*, 2019).

### **Markers validation for marker assisted selection**

Axiom to KASP marker conversion and validation was conducted for 17 MTA. Only five KASP markers generated polymorphic haplotypes in the independent set of ICARDA elites IDON43. All five demonstrated significant ( $p < 0.05$ ) correlation to grain yield assessed under the severely drought affected station of Sidi el Aydi (Fig. 7). AX-95631864, AX-94507963 and AX-94549122 tag Q.ICD.08 located on chromosome 1B, one of the three main loci identified in this study. The first two markers revealed good correlation, accuracy, precision and sensitivity, while the third one has medium sensitivity. Similarly, to AX-94910470 and AX-95191125 tag respectively Q.ICD.11 and Q.ICD.44 on chromosomes 4B and 7B. Those markers are protected against Type I errors, but they are prone to Type II errors, with several lines identified as not carrying the positive allele while instead being tolerant to drought. The five markers could be considered as validated and useful for wheat breeding.

### **Conclusion**

Drought tolerance is a complex quantitative trait that is influenced by genetic background and highly hindered by genotype by environment interactions. To understand the mechanism and the implied loci, a panel was tested under eighteen environments, clustered as moisture stressed and no-moisture stressed environments.



Our results confirmed that besides grain components, water productivity is the most critical trait to drive tolerance to moisture stress, and hence should be the primary targets of durum wheat breeders. A total of six QTLs were associated with GY under drought, some of them were linked with TKW, GpS and SPK. The haplotype diversity of three markers each from the three most promising QTLs Q.ICD.08, Q.ICD.11 and Q.ICD.44 size 19, 83 and 20 Mbp, causes water productivity of up to +1.5 kg ha<sup>-1</sup> mm<sup>-1</sup> across moisture stressed conditions. Five markers were validated into KASP markers and can further be utilized in marker assisted selection. Besides, the remaining QTLs might also prove useful after validation into easier assay to help improve the drought tolerance and yield stability in wheat. The genotypes GID: 800032178, GID: 800030179 and GID: 800043267 were confirmed to be drought tolerant and carrying the positive alleles of the three main QTLs. Those genotypes may serve as ideal crossing material in breeding programs.

Accepted Manuscript



## **Acknowledgments**

We wish to acknowledge Rached Abdelaziz and the other field technical staff in Morocco, Lebanon and Jordan for the intense support.

## **Author contributions**

Conceptualization and methodology, M.Z. and F.M.B.; Statistical analysis, M.Z., Z.K. and F.M.B.; validation, F.M.B.; writing original draft preparation, M.Z.; project administration, F.M.B. All authors have read and agreed to the published version of the manuscript.

## **Conflict of interest**

The authors declare no conflict of interest.

## **Funding**

This work was supported by the Australian Grains Research and Development Corporation (GRDC) project [ICA00012]: Focused improvement of ICARDA/Australia durum germplasm for abiotic tolerance.

## **Data availability**

All data are provided in the article as Table S1

Accepted Manuscript

## References

- Acevedo E, Ceccarelli S.** 1989. Role of the physiologist-breeder in a crop breeding program for drought resistance. *CIMMYT*.
- Agarwal PK, Gupta K, Lopato S, Agarwal P.** 2017. Dehydration responsive element binding transcription factors and their applications for the engineering of stress tolerance. *Journal of Experimental Botany* **68**, 2135-2148.
- Ain QU, Rasheed A, Anwar A, Mahmood T, Imtiaz M, Mahmood T, Xia X, He Z, Quraishi UM.** 2015. Genome-wide association for grain yield under rainfed conditions in historical wheat cultivars from Pakistan. *Frontiers in Plant Science* **6**, 743.
- Al-Ghzawi A, Khalaf Y, Al-Ajlouni Z, Al-Quraan N, Musallam I, Hani N.** 2018. The Effect of Supplemental Irrigation on Canopy Temperature Depression, Chlorophyll Content, and Water Use Efficiency in Three Wheat (*Triticum aestivum* L. and *T. durum* Desf.) Varieties Grown in Dry Regions of Jordan. *Agriculture* **8**.
- Ali M, Gul A, Hasan H, Alipour H, Abbasi AA, Zahra Khan Ft, Abbas S, Fatima T, Taimoor Z.** 2020. Chapter 12 - LEA proteins and drought stress in wheat. In: Ozturk M, Gul A, eds. *Climate Change and Food Security with Emphasis on Wheat*: Academic Press, 193-205.
- Anyia A, Slaski J, Capo-Chichi L, Chen J, Chang S.** 2008. Physiological traits contributing to water productivity and yield stability of barley on the Canadian Prairies. The 5th International Crop Science Congress, Jeju Island, South Korea. April, 13-18.
- Araus JL, Slafer GA, Royo C, Serret MD.** 2008. Breeding for Yield Potential and Stress Adaptation in Cereals. *Critical Reviews in Plant Sciences* **27**, 377-412.
- Ayadi M, Brini F, Masmoudi K.** 2019. Overexpression of a Wheat Aquaporin Gene, TdPIP2;1, Enhances Salt and Drought Tolerance in Transgenic Durum Wheat cv. Maali. LID - 10.3390/ijms20102389 [doi] LID - 2389. *International Journal of Molecular Sciences*.
- Bassi FM, Sanchez-Garcia M.** 2017. Adaptation and Stability Analysis of ICARDA Durum Wheat Elites across 18 Countries. *Crop Science* **57**, 2419-2430.
- Bassi FM, Brahma H, Sabraoui A, Amri A, Nsarellah N, Nachit MM, Al-Abdallat A, Chen MS, Lazraq A, El Bouhssini M.** 2019. Genetic identification of loci for Hessian fly resistance in durum wheat. *Molecular Breeding* **39**.
- Belaïd A, Nsarellah N, Laamari A, Nachit M, Amri A.** 2005. Assessing the economic impact of durum wheat research in Morocco. *International Centre for Agricultural Research in the Dry Areas (ICARDA)*, Aleppo, Syria, 51.
- Bhouri Khila S, Douh B, Mguidiche A, Boujelben A.** 2021. Assessing the Water Productivity of Durum Wheat in Tunisian Semi-Arid Conditions. In: Khebour Allouche F, Abu-hashim M, Negm AM, eds. *Agriculture Productivity in Tunisia Under Stressed Environment*. Cham: Springer International Publishing, 195-211.
- Bilal M, Iqbal I, Rana RM, Rehman SU, Haidery Q-a, Ahmad F, Ijaz A, Umar HMI.** 2015. A comprehensive review of effects of water stress and tolerance in wheat (*Triticuma estivum* L.). *Tropical Plant Research* **2**, 271-275.
- Blanco A, Mangini G, Giancaspro A, Giove S, Colasuonno P, Simeone R, Signorile A, De Vita P, Mastrangelo AM, Cattivelli L, Gadaleta A.** 2011a. Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars. *Molecular Breeding* **30**, 79-92.
- Blum A, Pnuel Y.** 1990. Physiological attributes associated with drought resistance of wheat cultivars in a Mediterranean environment. *Australian Journal of Agricultural Research - Australian Journal of Agricultural Research* **41**.
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES.** 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633-2635.

- Buffagni V, Vurro F, Janni M, Gullì M, Keller AA, Marmioli N.** 2020. Shaping Durum Wheat for the Future: Gene Expression Analyses and Metabolites Profiling Support the Contribution of BCAT Genes to Drought Stress Response. *Frontiers in plant science* **11**.
- Butler DG, Cullis BR, Gilmour AR, Gogel BJ, Thompson R.** 2009. ASReml-R reference manual version 4. The State of Queensland, Department of Primary Industries and Fisheries: Brisbane, Qld.
- Cabral AL, Jordan MC, Larson G, Somers DJ, Humphreys DG, McCartney CA.** 2018. Relationship between QTL for grain shape, grain weight, test weight, milling yield, and plant height in the spring wheat cross RL4452/'AC Domain'. *PLoS One* **13**, e0190681.
- Catola S, Marino G, Emiliani G, Huseynova T, Musayev M, Akparov Z, Maserti BE.** 2016. Physiological and metabolomic analysis of *Punica granatum* (L.) under drought stress. *Planta* **243**, 441-449.
- Ceccarelli S.** 2015. Efficiency of Plant Breeding. *Crop Science* **55**, 87-97.
- Charmet G, Robert N, Perretant MR, Gay G, Sourdille P, Groos C, Bernard S, Bernard M.** 2001. *Euphytica* **119**, 89-93.
- Collard BCY, Mackill DJ.** 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 557-572.
- De Mendiburu F, Yaseen M.** 2020. *Agricolae: Statistical Procedures for Agricultural Research*. R package version 1.4. 0.
- del Moral LFG, Rharrabti Y, Villegas D, Royo C.** 2003. Evaluation of Grain Yield and Its Components in Durum Wheat under Mediterranean Conditions. *Agronomy Journal* **95**, 266-274.
- Dickin E, Wright D.** 2008. The effects of winter waterlogging and summer drought on the growth and yield of winter wheat (*Triticum aestivum* L.). *European Journal of Agronomy* **28**, 234-244.
- Djemal R, Khoudi H.** 2015. Isolation and molecular characterization of a novel WIN1/SHN1 ethylene-responsive transcription factor TdSHN1 from durum wheat (*Triticum turgidum*. L. subsp. durum). *Protoplasma* **252**, 1461-1473.
- Dodig D, Zorić M, Kobiljski B, Šurlan-Momirović G, Quarrie SA.** 2010. Assessing drought tolerance and regional patterns of genetic diversity among spring and winter bread wheat using simple sequence repeats and phenotypic data. *Crop and Pasture Science* **61**.
- Duggal P, Beaty TH.** 2008. Genetic Epidemiology of Infectious Disease. *Genetic Susceptibility to Infectious Diseases*, 1.
- El Haddad N, Sanchez-Garcia M, Visioni A, Jilal A, El Amil R, Sall AT, Lagesse W, Kumar S, Bassi FM.** 2021. Crop Wild Relatives Crosses: Multi-Location Assessment in Durum Wheat, Barley, and Lentil. *Agronomy* **11**.
- El Hassouni K, Belkadi B, Filali-Maltouf A, Tidiane-Sall A, Al-Abdallat A, Nachit M, Bassi FM.** 2019. Loci Controlling Adaptation to Heat Stress Occurring at the Reproductive Stage in Durum Wheat. *Agronomy* **9**.
- English M.** 1990. Deficit Irrigation. I: Analytical Framework. **116**, 399-412.
- Farooq M, Wahid A, Kobayashi N, Fujita D, Basra SMA.** 2009. Plant drought stress: effects, mechanisms and management. *Agronomy for Sustainable Development* **29**, 185-212.
- Fiedler JD, Salsman E, Liu Y, Michalak de Jimenez M, Hegstad JB, Chen B, Manthey FA, Chao S, Xu S, Elias EM, Li X.** 2017. Genome-Wide Association and Prediction of Grain and Semolina Quality Traits in Durum Wheat Breeding Populations. *Plant Genome* **10**.
- Geerts S, Raes D.** 2009. Deficit irrigation as an on-farm strategy to maximize crop water productivity in dry areas. *Agricultural Water Management* **96**, 1275-1284.
- Gonzalez-Ribot G, Opazo M, Silva P, Acevedo E.** 2017. Traits Explaining Durum Wheat (*Triticum turgidum* L. spp. Durum) Yield in Dry Chilean Mediterranean Environments. *Frontiers in Plant Science* **8**, 1781.

- Gupta P, Balyan H, Gahlaut V.** 2017. QTL Analysis for Drought Tolerance in Wheat: Present Status and Future Possibilities. *Agronomy* **7**.
- Gupta P, Kabbaj H, El Hassouni K, Maccaferri M, Sanchez-Garcia M, Tuberosa R, Bassi FM.** 2020. Genomic regions associated with the control of flowering time in durum wheat. *Plants* **9**, 1628.
- Habash DZ, Kehel Z, Nachit M.** 2009. Genomic approaches for designing durum wheat ready for climate change with a focus on drought. *Journal of Experimental Botany* **60**, 2805-2815.
- Josse J, Husson F, Lê S.** 2008. FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software* **25**, 1-18.
- Juliana P, Singh RP, Poland J, Shrestha S, Huerta-Espino J, Govindan V, Mondal S, Crespo-Herrera LA, Kumar U, Joshi AK, Payne T, Bhati PK, Tomar V, Consolacion F, Campos Serna JA.** 2021. Elucidating the genetics of grain yield and stress-resilience in bread wheat using a large-scale genome-wide association mapping study with 55,568 lines. *Scientific Reports* **11**, 5254.
- Kabbaj H, Sall AT, Al-Abdallat A, Geleta M, Amri A, Filali-Maltouf A, Belkadi B, Ortiz R, Bassi FM.** 2017. Genetic Diversity within a Global Panel of Durum Wheat (*Triticum durum*) Landraces and Modern Germplasm Reveals the History of Alleles Exchange. *Frontiers in Plant Science* **8**, 1277.
- Karam F, Kaban R, Breidi J, Rouphael Y, Oweis T.** 2009. Yield and water-production functions of two durum wheat cultivars grown under different irrigation and nitrogen regimes. *Agricultural Water Management* **96**, 603-615.
- Karrou M, Oweis T.** 2012. Water and land productivities of wheat and food legumes with deficit supplemental irrigation in a Mediterranean environment. *Agricultural Water Management* **107**, 94-103.
- Kiliç H, Tacettin Y.** 2010. The Effect of Drought Stress on Grain Yield, Yield Components and Some Quality Traits of Durum Wheat (*Triticum turgidum* ssp. *durum*) Cultivars. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* **38**.
- Li Y-F, Wu Y, Hernandez-Espinosa N, Peña RJ.** 2013. Heat and drought stress on durum wheat: Responses of genotypes, yield, and quality parameters. *Journal of Cereal Science* **57**, 398-404.
- Liliane TN, Charles MS.** 2020. Factors affecting yield of crops. *Agronomy-Climate Change & Food Security*, 9.
- Lilley JM, Kirkegaard JA.** 2016. Farming system context drives the value of deep wheat roots in semi-arid environments. *Journal of Experimental Botany* **67**, 3665-3681.
- Liu X, McKenna S, Welch LR, Showalter AM.** 2020. Bioinformatic Identification of Plant Hydroxyproline-Rich Glycoproteins. In: Popper ZA, ed. *The Plant Cell Wall: Methods and Protocols*. New York, NY: Springer New York, 463-481.
- Lozada D, Carter A.** 2020. Insights into the Genetic Architecture of Phenotypic Stability Traits in Winter Wheat. *Agronomy* **10**.
- Lozada DN, Mason RE, Babar MA, Carver BF, Guedira G-B, Merrill K, Arguello MN, Acuna A, Vieira L, Holder A, Addison C, Moon DE, Miller RG, Dreisigacker S.** 2017. Association mapping reveals loci associated with multiple traits that affect grain yield and adaptation in soft winter wheat. *Euphytica* **213**.
- Maccaferri M, Sanguineti MC, Demontis A, El-Ahmed A, Garcia del Moral L, Maalouf F, Nachit M, Nserallah N, Ouabbou H, Rhouma S, Royo C, Villegas D, Tuberosa R.** 2011. Association mapping in durum wheat grown across a broad range of water regimes. *Journal of Experimental Botany* **62**, 409-438.
- Maccaferri M, El-Feki W, Nazemi G, Salvi S, Canè MA, Colalongo MC, Stefanelli S, Tuberosa R.** 2016. Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. *Journal of Experimental Botany* **67**, 1161-1178.

- Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, Ormanbekova D, Lux T, Prade VM, Milner SG, Himmelbach A, Mascher M, Bagnaresi P, Faccioli P, Cozzi P, Lauria M, Lazzari B, Stella A, Manconi A, Gnocchi M, Moscatelli M, Avni R, Deek J, Biyiklioglu S, Frascaroli E, Corneti S, Salvi S, Sonnante G, Desiderio F, Marè C, Crosatti C, Mica E, Özkan H, Kilian B, De Vita P, Marone D, Joukhadar R, Mazzucotelli E, Nigro D, Gadaleta A, Chao S, Faris JD, Melo ATO, Pumphrey M, Pecchioni N, Milanese L, Wiebe K, Ens J, MacLachlan RP, Clarke JM, Sharpe AG, Koh CS, Liang KYH, Taylor GJ, Knox R, Budak H, Mastrangelo AM, Xu SS, Stein N, Hale I, Distelfeld A, Hayden MJ, Tuberosa R, Walkowiak S, Mayer KFX, Ceriotti A, Pozniak CJ, Cattivelli L.** 2019. Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics* **51**, 885-895.
- Malosetti M, Ribaut JM, van Eeuwijk FA.** 2013. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Frontiers in Physiology* **4**, 44.
- Mangini G, Gadaleta A, Colasuonno P, Marcotuli I, Signorile AM, Simeone R, De Vita P, Mastrangelo AM, Laido G, Pecchioni N, Blanco A.** 2018. Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats. *PLoS One* **13**, e0190162.
- Mathew I, Shimelis H, Shayanowako AIT, Laing M, Chaplot V.** 2019. Genome-wide association study of drought tolerance and biomass allocation in wheat. *PLoS One* **14**, e0225383.
- McCartney CA, Somers DJ, Humphreys DG, Lukow O, Ames N, Noll J, Cloutier S, McCallum BD.** 2005. Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452x'AC Domain'. *Genome* **48**, 870-883.
- Mohammadi A, Rafiee S, Mohtasebi SS, Mousavi Avval SH, Rafiee H.** 2011. Energy efficiency improvement and input cost saving in kiwifruit production using Data Envelopment Analysis approach. *Renewable Energy* **36**, 2573-2579.
- Mohammadi M, Blake TK, Budde AD, Chao S, Hayes PM, Horsley RD, Obert DE, Ullrich SE, Smith KP.** 2015. A genome-wide association study of malting quality across eight U.S. barley breeding programs. *Theoretical and Applied Genetics* **128**, 705-721.
- Mohammadi R, Etminan A, Shoshtari LIA.** 2019. AGRO-PHYSIOLOGICAL CHARACTERIZATION OF DURUM WHEAT GENOTYPES UNDER DROUGHT CONDITIONS. *Experimental Agriculture* **55**, 484-499.
- Muhu-Din Ahmed HG, Sajjad M, Zeng Y, Iqbal M, Habibullah Khan S, Ullah A, Nadeem Akhtar M.** 2020. Genome-Wide Association Mapping through 90K SNP Array for Quality and Yield Attributes in Bread Wheat against Water-Deficit Conditions. *Agriculture* **10**.
- Neumann K, Kobiljski B, Denčić S, Varshney RK, Börner A.** 2010. Genome-wide association mapping: a case study in bread wheat (*Triticum aestivum* L.). *Molecular Breeding* **27**, 37-58.
- Ogbonnaya FC, Rasheed A, Okechukwu EC, Jighly A, Makdis F, Wuletaw T, Hagraas A, Uguru MI, Agbo CU.** 2017. Genome-wide association study for agronomic and physiological traits in spring wheat evaluated in a range of heat prone environments. *Theoretical and Applied Genetics* **130**, 1819-1835.
- Okamoto Y, Nguyen AT, Yoshioka M, Iehisa JC, Takumi S.** 2013. Identification of quantitative trait loci controlling grain size and shape in the D genome of synthetic hexaploid wheat lines. *Breeding Science* **63**, 423-429.
- Pacheco Á, Vargas M, Alvarado G, Rodríguez F, Crossa J, Burgueño J.** 2015. GEA-R (Genotype x Environment Analysis with R for Windows) Version 4.1. In: International M, Wheat Improvement C, eds: CIMMYT Research Data & Software Repository Network.
- Pearson K.** 1895. Correlation coefficient. Vol. 58, 214.



- Pinto RS, Reynolds MP, Mathews KL, McIntyre CL, Olivares-Villegas JJ, Chapman SC.** 2010. Heat and drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. *Theoretical and Applied Genetics* **121**, 1001-1021.
- Pradhan S, Babar MA, Robbins K, Bai G, Mason RE, Khan J, Shahi D, Avci M, Guo J, Maksud Hossain M, Bhatta M, Mergoum M, Asseng S, Amand PS, Gezan S, Baik BK, Blount A, Bernardo A.** 2019. Understanding the Genetic Basis of Spike Fertility to Improve Grain Number, Harvest Index, and Grain Yield in Wheat Under High Temperature Stress Environments. *Frontiers in Plant Science* **10**, 1481.
- Pshenichnikova TA, Osipova SV, Smirnova OG, Leonova IN, Permyakova MD, Permyakov AV, Rudikovskaya EG, Konstantinov DK, Verkhoturov VV, Lohwasser U, Börner A.** 2021. Regions of Chromosome 2A of Bread Wheat (*Triticum aestivum* L.) Associated with Variation in Physiological and Agronomical Traits under Contrasting Water Regimes. *Plants* **10**.
- Qi Y, Yu H, Fu Q, Chen Q, Ran J, Yang Z.** 2022. Future Changes in Drought Frequency Due To Changes in the Mean and Shape of the PDSI Probability Density Function Under RCP4.5 Scenario. *Frontiers in Earth Science* **10**.
- Quarrie SA, Stojanović J, Pekić S.** 1999. Improving drought resistance in small-grained cereals: A case study, progress and prospects. *Plant Growth Regulation* **29**, 1-21.
- Rahimi Y, Bihamta MR, Taleei A, Alipour H, Ingvarsson PK.** 2019. Genome-wide association study of agronomic traits in bread wheat reveals novel putative alleles for future breeding programs. *BMC Plant Biology* **19**, 541.
- Rehman Arif MA, Attaria F, Shokat S, Akram S, Waheed MQ, Arif A, Borner A.** 2020. Mapping of QTLs Associated with Yield and Yield Related Traits in Durum Wheat (*Triticum durum* Desf.) Under Irrigated and Drought Conditions. *International Journal of Molecular Sciences* **21**.
- Reynolds M, Manes Y, Izanloo A, Langridge P.** 2009. Phenotyping approaches for physiological breeding and gene discovery in wheat. *Annals of Applied Biology* **155**, 309-320.
- Richards RA, Condon AG, Rebetzke GJ.** 2001. Traits to improve yield in dry environments. *Application of Physiology in Wheat Breeding*.
- Ringli C.** 2010. The hydroxyproline-rich glycoprotein domain of the Arabidopsis LRX1 requires Tyr for function but not for insolubilization in the cell wall. *The Plant Journal* **63**, 662-669.
- Rivera-Amado C, Trujillo-Negrellos E, Molero G, Reynolds MP, Sylvester-Bradley R, Foulkes MJ.** 2019. Optimizing dry-matter partitioning for increased spike growth, grain number and harvest index in spring wheat. *Field Crops Research* **240**, 154-167.
- Roncallo PF, Akkiraju PC, Cervigni GL, Echenique VC.** 2017. QTL mapping and analysis of epistatic interactions for grain yield and yield-related traits in *Triticum turgidum* L. var. durum. *Euphytica* **213**.
- Sehgal D, Autrique E, Singh R, Ellis M, Singh S, Dreisigacker S.** 2017. Identification of genomic regions for grain yield and yield stability and their epistatic interactions. *Scientific Reports* **7**, 41578.
- Sehgal D, Rosyara U, Mondal S, Singh R, Poland J, Dreisigacker S.** 2020. Incorporating Genome-Wide Association Mapping Results Into Genomic Prediction Models for Grain Yield and Yield Stability in CIMMYT Spring Bread Wheat. *Frontiers in Plant Science* **11**, 197.
- Shokat S, Sehgal D, Vikram P, Liu F, Singh S.** 2020. Molecular Markers Associated with Agro-Physiological Traits under Terminal Drought Conditions in Bread Wheat. *International Journal of Molecular Sciences* **21**.
- Shpiler L, Blum A.** 1990. Heat tolerance for yield and its components in different wheat cultivars. *Euphytica* **51**, 257-263.
- Siahpoosh MR, Dehghanian E.** 2012. Water Use Efficiency, Transpiration Efficiency, and Uptake Efficiency of Wheat during Drought. *Agronomy Journal* **104**, 1238-1243.



- Sokoto M, Singh A.** 2013. Yield and Yield Components of Bread Wheat as Influenced by Water Stress, Sowing Date and Cultivar in Sokoto, Sudan Savannah, Nigeria. *American Journal of Plant Sciences* **04**, 122-130.
- Stallmann J, Schweiger R, Pons CAA, Müller C.** 2020. Wheat growth, applied water use efficiency and flag leaf metabolome under continuous and pulsed deficit irrigation. *Scientific Reports* **10**, 10112.
- Sukumaran S, Xiang W, Bean SR, Pedersen JF, Kresovich S, Tuinstra MR, Tesso TT, Hamblin MT, Yu J.** 2012. Association Mapping for Grain Quality in a Diverse Sorghum Collection. *The Plant Genome* **5**, 126-135.
- Sukumaran S, Reynolds MP, Sansaloni C.** 2018. Genome-Wide Association Analyses Identify QTL Hotspots for Yield and Component Traits in Durum Wheat Grown under Yield Potential, Drought, and Heat Stress Environments. *Frontiers in Plant Science* **9**, 81.
- Sun C, Zhang F, Yan X, Zhang X, Dong Z, Cui D, Chen F.** 2017. Genome-wide association study for 13 agronomic traits reveals distribution of superior alleles in bread wheat from the Yellow and Huai Valley of China. *Plant Biotechnology Journal* **15**, 953-969.
- Team R. Core.** 2017. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from ....
- Tidiane SA, Chiari T, Legesse W, Seid-Ahmed K, Ortiz R, van Ginkel M, Bassi FM.** 2019. Durum Wheat (*Triticum durum* Desf.): Origin, Cultivation and Potential Expansion in Sub-Saharan Africa. *Agronomy* **9**.
- Toulotte JM, Pantazopoulou CK, Sanclemente MA, Voeselek LACJ, Sasidharan R.** 2022. Water stress resilient cereal crops: Lessons from wild relatives. **64**, 412-430.
- Tramblay Y, Llasat MC, Randin C, Coppola E.** 2020. Climate change impacts on water resources in the Mediterranean. *Regional Environmental Change* **20**, 83.
- Tuberosa R.** 2012. Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology* **3**, 347.
- Vitale P, Fania F, Esposito S, Pecorella I, Pecchioni N, Palombieri S, Sestili F, Lafiandra D, Taranto F, De Vita P.** 2021. QTL Analysis of Five Morpho-Physiological Traits in Bread Wheat Using Two Mapping Populations Derived from Common Parents. *Genes* **12**, 604.
- Wang S, Xu S, Chao S, Sun Q, Liu S, Xia G.** 2019. A Genome-Wide Association Study of Highly Heritable Agronomic Traits in Durum Wheat. *Frontiers in plant science* **10**.
- Westgate ME.** 1994. Water Status and Development of the Maize Endosperm and Embryo during Drought. *Crop Science* **34**, 76-83.
- Williams K, Munkvold J, Sorrells M.** 2012. Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (*Triticum aestivum* L.). *Euphytica* **190**, 99-116.
- Xin F, Zhu T, Wei S, Han Y, Zhao Y, Zhang D, Ma L, Ding Q.** 2020. QTL Mapping of Kernel Traits and Validation of a Major QTL for Kernel Length-Width Ratio Using SNP and Bulk Segregant Analysis in Wheat. *Scientific Reports* **10**, 25.
- Xu Y-F, Li S-S, Li L-H, Ma F-F, Fu X-Y, Shi Z-L, Xu H-X, Ma P-T, An D-G.** 2017. QTL mapping for yield and photosynthetic related traits under different water regimes in wheat. *Molecular Breeding* **37**.
- Xynias IN, Mylonas I, Korpetis EG, Ninou E, Tsballa A, Avdikos ID, Mavromatis AG.** 2020. Durum Wheat Breeding in the Mediterranean Region: Current Status and Future Prospects. *Agronomy* **10**.
- Yadvinder S, Kukal SS, Jat ML, Sidhu HS.** 2014. Chapter Four - Improving Water Productivity of Wheat-Based Cropping Systems in South Asia for Sustained Productivity. In: Sparks D, ed. *Advances in Agronomy*, Vol. 127: Academic Press, 157-258.
- Yates Andrew D, Allen J, Amode RM, Azov AG, Barba M, Becerra A, Bhai J, Campbell Lahcen I, Carbajo Martinez M, Chakiachvili M, Chougule K, Christensen M, Contreras-Moreira B, Cuzick A, Da Rin Fioretto L, Davis P, De Silva Nishadi H, Diamantakis S, Dyer S,**

- Elser J, Filippi CV, Gall A, Grigoriadis D, Guijarro-Clarke C, Gupta P, Hammond-Kosack Kim E, Howe KL, Jaiswal P, Kaikala V, Kumar V, Kumari S, Langridge N, Le T, Luypaert M, Maslen GL, Maurel T, Moore B, Muffato M, Mushtaq A, Naamati G, Naithani S, Olson A, Parker A, Paulini M, Pedro H, Perry E, Preece J, Quinton-Tulloch M, Rodgers F, Rosello M, Ruffier M, Seager J, Sitnik V, Szpak M, Tate J, Tello-Ruiz Marcela K, Trevanion Stephen J, Urban M, Ware D, Wei S, Williams G, Winterbottom A, Zarowiecki M, Finn Robert D, Flicek P. 2021. Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. *Nucleic Acids Research* **50**, D996-D1003.
- Yu Y, Ouyang Y, Yao W. 2018. shinyCircos: an R/Shiny application for interactive creation of Circos plot. *Bioinformatics* **34**, 1229-1231.
- Yu YA-OX, Yu M, Zhang S, Song TA-O, Zhang M, Zhou H, Wang Y, Xiang J, Zhang X. 2022. Transcriptomic Identification of Wheat AP2/ERF Transcription Factors and Functional Characterization of TaERF-6-3A in Response to Drought and Salinity Stresses. LID - 10.3390/ijms23063272 [doi] LID - 3272. *International Journal of Molecular Sciences*.
- Zaïm M, Kabbaj H, Kehel Z, Gorjanc G, Filali-Maltouf A, Belkadi B, Nachit MM, Bassi FM. 2020. Combining QTL Analysis and Genomic Predictions for Four Durum Wheat Populations Under Drought Conditions. *Frontiers in Genetics* **11**, 316.
- Zandipour M, Majidi Hervan E, Azadi A, Khosroshahli M, Etminan A. 2020. A QTL hot spot region on chromosome 1B for nine important traits under terminal drought stress conditions in wheat. *Cereal Research Communications* **48**, 17-24.
- Zhang B, Li F-M, Huang G, Cheng Z-Y, Zhang Y. 2006. Yield performance of spring wheat improved by regulated deficit irrigation in an arid area. *Agricultural Water Management* **79**, 28-42.
- Zhang G, Wang Y, Guo Y, Zhao Y, Kong F, Li S. 2015. Characterization and mapping of QTLs on chromosome 2D for grain size and yield traits using a mutant line induced by EMS in wheat. *The Crop Journal* **3**, 135-144.
- Zittis G, Bruggeman A, Lelieveld J. 2021. Revisiting future extreme precipitation trends in the Mediterranean. *Weather and Climate Extremes* **34**, 100380.

Accepted Manuscript

## Tables

Table 1. Water productivity classes

Classes	Definition	$r^2$	$b_i$ stressed	$b_i$ non-stressed
1	Stable water response	**	$b_i < 5.08$	$b_i < 4.85$
2	Responsive to low moisture	**	$b_i > 5.08$	$b_i < 4.85$
3	Responsive to high moisture	**	$b_i < 5.08$	$b_i > 4.85$
4	Highly water responsive	**	$b_i > 5.08$	$b_i > 4.85$
5	No water response	Ns	.	.

\*\* ,  $p < 0.01$ ; ns, Not significant.

Accepted Manuscript

Table 2: Correlation analysis for all traits against grain yield across moisture and no-moisture stressed conditions.

Mega-env.	Environments	Moisture (mm)	DTH	TKW	SPK	GpS
Moisture stressed	SAD16	210.0	•	•	ns	•
	MUS18	247.0	•	-	-	-
	MCH17N	270.8	•	ns	ns	•
	MCH17Z	270.8	•	-	ns	-
	KFD16	282.3	•	•	-	-
	KFD18A	346.2	Ns	ns	-	-
	KFD18B	346.2	Ns	ns	-	-
	MCH15	363.7	•	•	•	•
	KFD17A	370.0	•	ns	•	•
	KFD17B	370.0	Ns	ns	•	•
	MCH16	444.0	•	•	•	•
	Average	320.1	•	•	•	•
	Non moisture stressed	TES16	455.4	Ns	•	•
TER16		457.6	•	•	ns	•
TES17		484.6	-	-	-	-
MCH18N		493.4	•	•	ns	•
MCH18Z		493.4	•	•	•	•
MKZ15		634.8	•	•	ns	•
TER15		641.4	•	•	-	-
Average		522.9	•	•	ns	•

Significant at the probability level 0.001 (•). ns: not significant. ‘-’ Not available.

## Figure legends

Figure 1. Description of the testing environments used for the ‘discovery panel’, and their principal component differentiation (PCA) hierarchical clustering based on climatic factors. GY: Grain yield (kg ha<sup>-1</sup>), TKW: 1,000 Kernel weight (g), SPK: Spike density per m<sup>2</sup>, GpS: Grain per spike, PLH: Plant height (cm), DTH: Days to heading.

Figure 2. Trendlines of linear regression representing water productivity as biplot of the total moisture at each environment against the grain yield at that environment. One representative genotype example for each water productivity class (Tab 1) is presented and compared to the average performance of all genotypes at each environment (dashed line). Trendline color cyan represent the ‘class 1’, the orange for ‘class 2’, green for ‘class 3’ and blue for ‘class 4’, Moisture stressed environments are presented at (A) and non-stressed at (B).

Figure 3: AMMI wide adaptation index (AWAI) against the ratio to the max of yield potential across 11 moistures stressed (up) and 7 no-stressed (down) environments. Dashed lines trace the average for each axis.

Figure 4. Circos representation of the QTL identification across germplasm panels. The outermost circle shows the Svevo durum wheat genome assembly (Macaferri et al. 2019), including its chromosomes, followed by the distribution of 35K Axiom polymorphic probes in the ‘discovery panel’ (35K) and the 23K SNP probes in the ‘confirmation panel’ (23K). The following tracks represent results of significant marker trait association (MTA) identified in the ‘discovery panel’ for AMMI wide adaptation index (AWAI) coded as green dots, spike density per m<sup>2</sup> (SPK) and grain per spike (GpS), 1,000 kernel weight (TKW) and grain yield (GYD) across individual and combined environments. The following tracks represent MTA for GY identified in the ‘confirmation panel’ (GYC). The MTA identified in moisture stressed environments are color coded as red dots, while those identified in non-moisture stressed conditions are coded as blue dots. The innermost circle provides the QTL labels for reference. \* represents the QTL confirmed by the GYC.

Figure 5. (A) . Correlation between allelic call of five major markers representing three QTLs and grain yield at eighteen environments, presented for the moisture stressed and non-stressed separately. Each environment is named as the total of its moisture, and color coded from darker to brighter orange for severity of drought, and darker to lighter blue for decreasing moisture content. The average performance is presented as a black triangle. The Pearson’s significant cut off are presented for both major and minor alleles. (B). Allelic haplotypes effects of the significant loci on water productivity of 120 genotypes grown under 18 environments. Left: The accessions were divided into four groups based on their haplotype for three major QTLs: ‘+’ mark the positive and ‘-’ the negative alleles. Right: The haplotype frequencies of each water response class.

Figure 6. Allelic effect for the combination of 3 loci associated with GY on the ‘confirmation panel’ under moisture stress. The black line inside the boxes indicates the median of each haplotype across each cluster. The ‘+’ for the positive and ‘-’ for the negative alleles. Letters (a, b, c) indicate the LSD test. The green dots and dashed line represent the average WP.



Figure 7. KASP markers validation on an independent set of 94 elite lines of ICARDA tested under severe drought. Correlation was measured between the BLUE for grain yield recorded at Sidi el Aydi and the haplotype score. Accuracy, sensitivity, and specificity were determined using only the top 20 and worst 20 yielding lines. AX-95631864, AX-94507963 and AX-94549122 tag Q.ICD.08, AX-94910470 tags Q.ICD.11 and AX-95191125 tags Q.ICD.44.

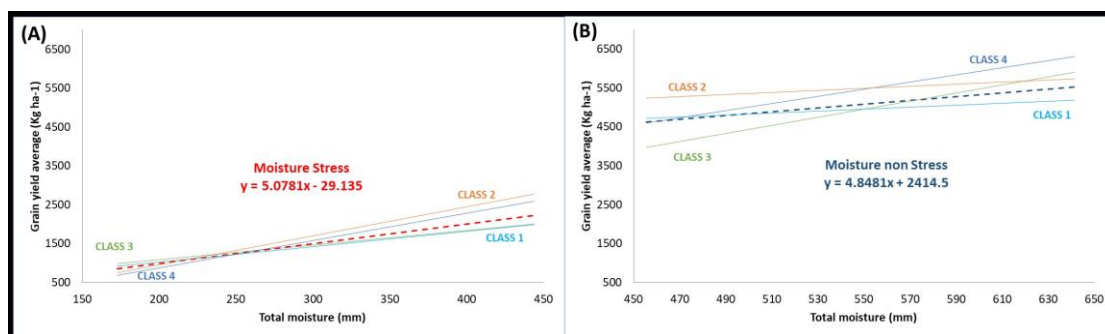
Accepted Manuscript

Figure 1

Code	Site	Country	Season	Coordinates	Altitude (m)	Soil type	Climate	Moisture	Fertilization	Total moisture (mm)	Average yield (kg ha <sup>-1</sup> )	Traits recorded			
												CV	TKM	SPK	GS
TES16	Tessaout	Morocco	2015/2016	29° 49' 48" N, 8° 34' 48" W	472	Calcic xerosols	Hot steppe	Gravity	30+30+20 U of N; 50 U of P and K	455.4	3,674.1				
TER15	Terbol	Lebanon	2014/2015	33° 48' 29" N, 35° 59' 22" W	897	Chromic vertisols	Mediterranean / temperate	sprinkle	50+50+50 U of N; 50 U of P and K	641.4	5,843.3				
TER16	Terbol	Lebanon	2015/2016	33° 48' 29" N, 35° 59' 22" W	897	Chromic vertisols	Mediterranean / temperate	sprinkle	50+50+50 U of N; 50 U of P and K	457.6	6,413.5				
TES17	Tessaout	Morocco	2016/2017	29° 49' 48" N, 8° 34' 48" W	472	Calcic xerosols	Hot steppe	gravity	30+30+20 U of N; 50 U of P and K	484.6	3,937.9				
MKZ15	Melk Zhar	Morocco	2014/2015	30° 2' 33" N, 9° 33' 4" W	137	Sandy limestone	Mediterranean / hot and temperate	drip (fertigation)	50+25+25+50 U of N; 25+25 U of P and K	634.8	5,302.5				
MCH18N & MCH18Z	Marchouch	Morocco	2017/2018	33° 34' 3.1" N, 6° 38' 0.1" W	398	Clay vertisol	Mediterranean / warm temperate	rainfed	50+50+50 U of N; 50 U of P and K	493.4	N: 4,424.2 Z: 4,670.9				
SAD16	Sidi el Aidi	Morocco	2015/2016	33° 9' 36" N, 7° 24' 0" W	226	Vertisol	Mediterranean / hot and temperate	rainfed	30+30 U of N; 25 U of P and K	210.0	1,532.5				
KFD18A & KFD18B	Kfardan	Lebanon	2017/2018	34° 02' 08" N, 36° 08' 18" E	1,000	Sandy clay	Mediterranean / temperate	rainfed	A: 25+25 U of N; 25 U P and K B: 50 U of N, P and K	346.2	A: 957.6 B: 858.7				
MCH15	Marchouch	Morocco	2014/2015	33° 34' 3.1" N, 6° 38' 0.1" W	398	Clay vertisol	Mediterranean / warm temperate	rainfed	50+50+50 U of N; 50 U of P and K	363.7	2,753.8				
MCH17N & MCH17Z	Marchouch	Morocco	2016/2017	33° 34' 3.1" N, 6° 38' 0.1" W	398	Clay vertisol	Mediterranean / warm temperate	rainfed	50+50+50 U of N; 50 U of P and K	270.8	N: 1,663.2 Z: 1,761.4				
MUS18	Mushaqar	Jordan	2017/2018	1.77°N, 35.77°E	790	clay loam	Mediterranean / warm temperate	rainfed	20 U of N	274.1	330.1				
KFD16	Kfardan	Lebanon	2015/2016	34° 02' 08" N, 36° 08' 18" E	1,000	Sandy clay	Mediterranean / temperate	rainfed	50+50 U of N; 50 U of P and K	282.3	2,274.9				
KFD17B & KFD17A	Kfardan	Lebanon	2016/2017	34° 02' 08" N, 36° 08' 18" E	1,000	Sandy clay	Mediterranean / temperate	rainfed	B: 50 U of N, P and K A: 25+25 U of N; 25 U P and K	370.0	B: 1,792.4 A: 2,581.1				
MCH16	Marchouch	Morocco	2015/2016	33° 34' 3.1" N, 6° 38' 0.1" W	398	Clay vertisol	Mediterranean / warm temperate	rainfed	50+50+50 U of N; 50 U of P and K	444.0	1,907.9				

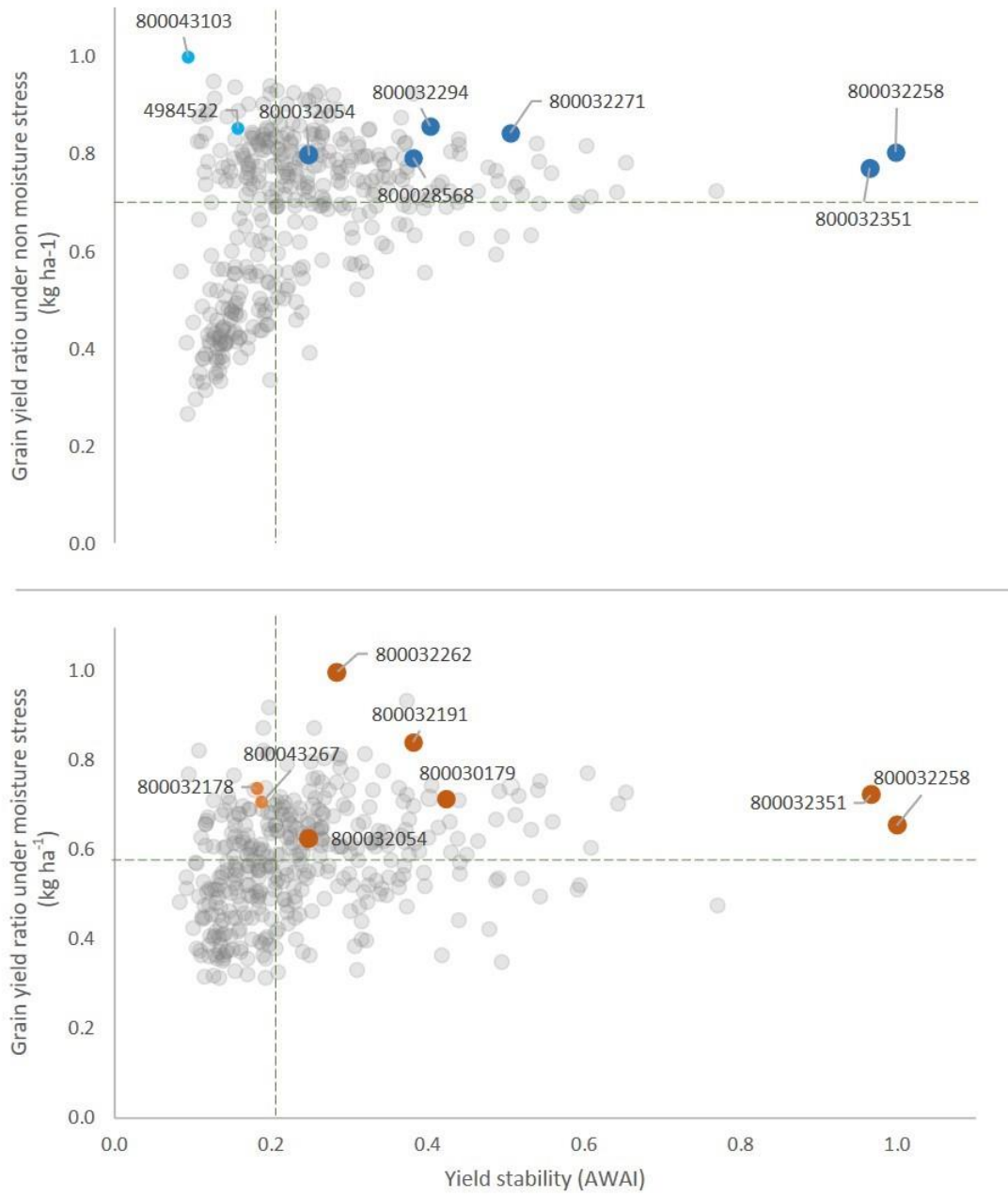
Accepted Manuscript

Figure 2



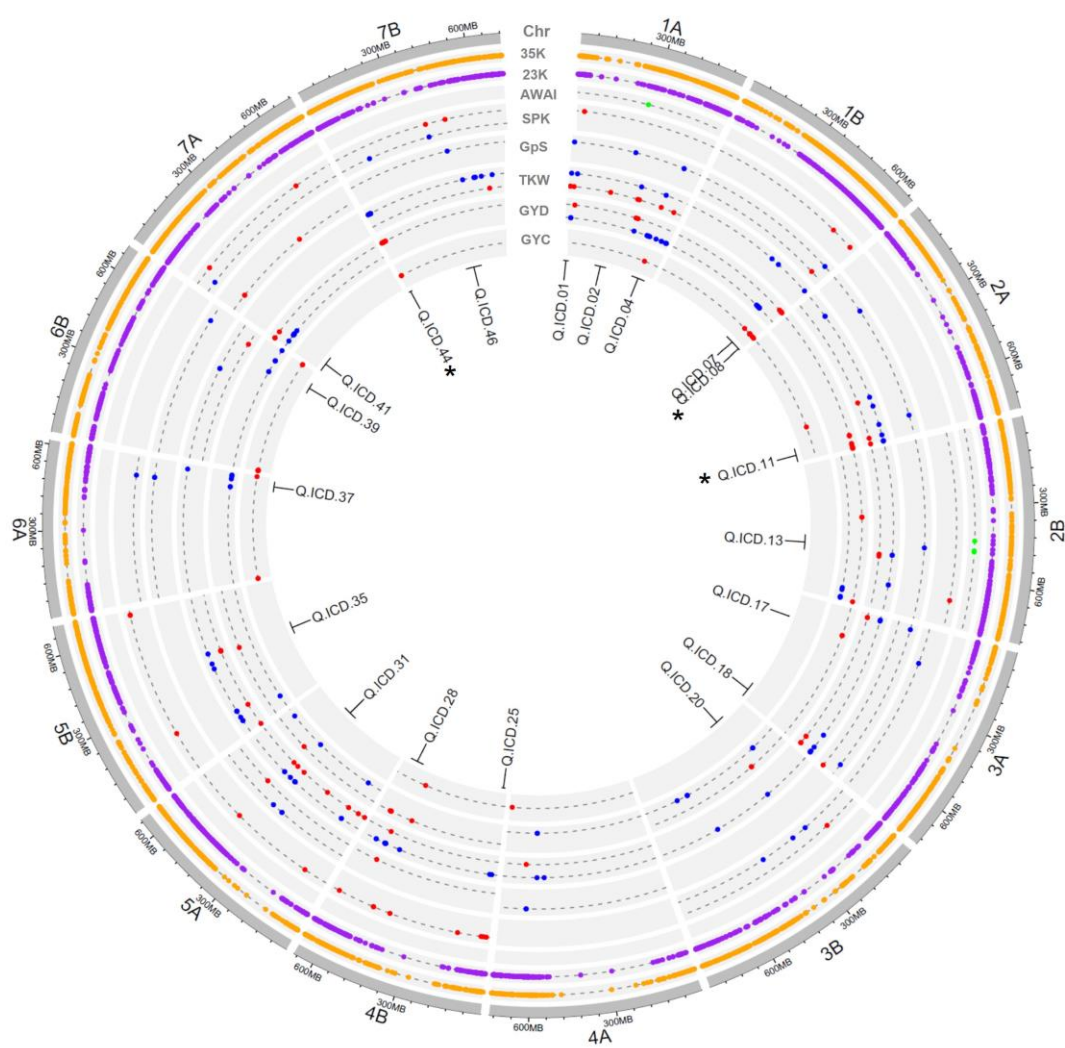
Accepted Manuscript

Figure 3



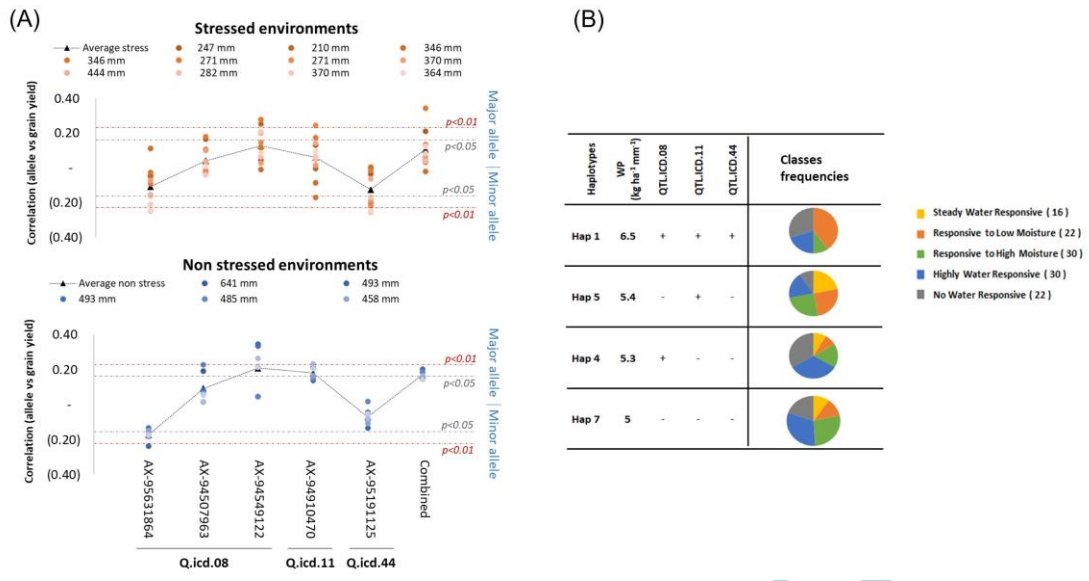
AC

Figure 4



Accept

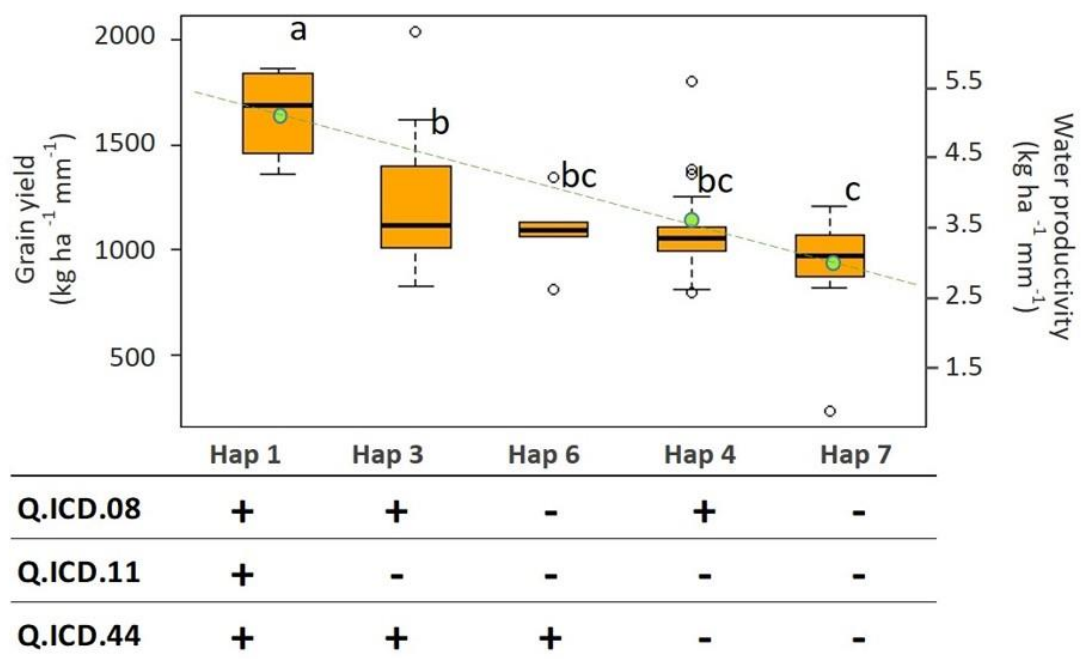
Figure 5



Accepted Manuscript

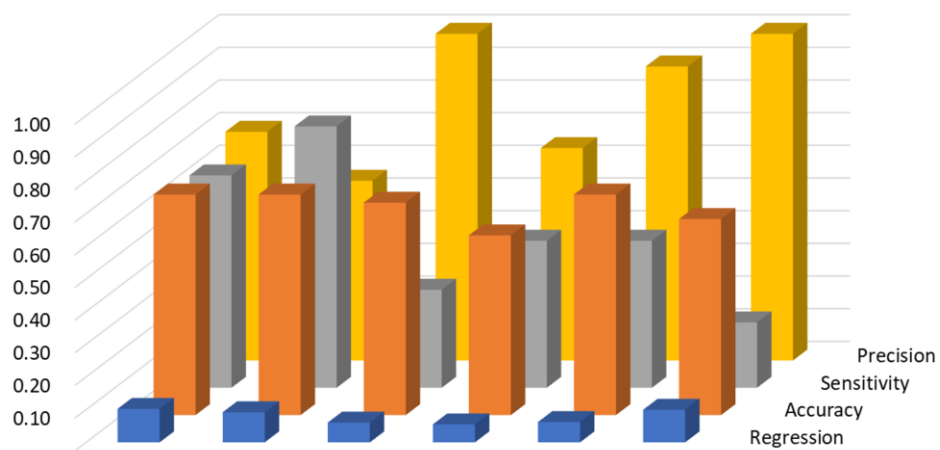


Figure 6



Accepted Manuscript

Figure 7



ID	AX-95631864	AX-94507963	AX-94549122	AX-94910470	AX-95191125	Combined
Chr	1B	1B	1B	2A	7B	1B + 2A + 7B
QTL	Q.ICD-08	Q.ICD-08	Q.ICD-08	Q.ICD.11	Q.ICD.44	3 QTLs
Allele 2	G:G	G:G	A:A	C:C	G:G	-
Allele 1	A:A	A:A	G:G	G:G	A:A	-
MAF	48%	36%	91%	43%	33%	5%

Accepted Manuscript