Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genetic Diversity and Genome-Wide Association Study for some agronomic traits in durum wheat (Triticum turgidum L.) Using whole genome DArTseq Markers

Peyman Ebrahimi
Islamic Azad University
Ezzat Karami (ezzatut81@yahoo.com)
Islamic Azad University
Alireza Etminan
Islamic Azad University
Reza Talebi
Islamic Azad University
Reza Mohammadi
Islamic Azad University

Research Article

Keywords: Association mapping, DArTseq technology, Durum wheat, Polymorphism information content

Posted Date: July 6th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3125419/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

This research was conducted to study genetic diversity, population structure, linkage disequilibrium, and marker-trait relationship in durum wheat using DArTseq technology. Therefore, 94 durum wheat genotypes were evaluated in the form of augmented design (augment) with six repeated controls during the two cropping years, 2016-2017 and 2017-2018. Some Agronomic traits including the number of days to spike emergence, number of days to physiological maturity, plant height, thousand kernel weight, and grain yield were measured and recorded. The DNA samples were processed for DArTseq genotyping, using the Genotyping by Sequencing (GBS) platform at Diversity Array Technology Pty, Ltd, Australia. The values of polymorphism information content (PIC) of SilicoDArT markers varied from 0.023 to 0.499 with an average of 0.38. The average repeatability and readability of the sequences in all linkage groups were above 0.98 and 0.92, respectively. The number of mapped SilicoDArT markers varied from 300 markers in the linkage group (Chr1A) to 853 markers in the linkage group (Chr7B). The chromosome size covered by SilicoDArT markers ranged from 829200.1 kbp in the linkage group (Chr3B) to 589293.79 kbp in the linkage group (Chr1A). The studied durum wheat genotypes were grouped into four separate groups based on the Bayesian model. The four subpopulations showed a relatively high genetic diversity, which was variable, ranging from 0.02 in the subpopulation (POP4) to 0.34 in the subpopulation (POP3). LD analysis between A and B genomes showed that there is a high number of significant marker pairs (82052) in the B genome compared to the A genome (68885). In general, during two cropping years, 29 markers related to the studied traits were found, and out of these 29 related markers, 19 markers were identified in the first year and 10 markers were identified in the second year. The results revealed that the DArTseq markers are a very powerful tool for evaluating the genetic diversity and population structure of durum wheat genotypes and can be used in genotype screening as breeding parents and marker-assisted selection in breeding programs.

Introduction

Among cereal crops, wheat is one of the most grown crops that is essential for the human diet due to providing carbohydrates, proteins, zinc, calcium, fiber, and energy (Mehrabi et al. 2020). Durum wheat (*Triticum turgidum* L. Var. *Durum*) is a tetraploid species with an AABB genome that has been created from hybridization between two wild species *Aegilops speltoides* Tausch (B genome) and *Triticum urartu* Ghandilyan (A genome) (Giunta et al. 2019). This crop is well-adapted to various climatic conditions, and it's mainly cultivated in the Mediterranean region, North Africa, Northern plains, Southwestern USA, and Southern Europe. However, it has an important role in supplying food to local people in the Mediterranean region, which produced more than 50% of the world's durum wheat production (Mérida-García et al. 2020). Hence, detailed genetic information on important agronomic traits and grain yield of durum wheat is required as a basis in breeding programs (Mantovani et al. 2008).

Investigation of genetic diversity among plant populations is an important component of genetics and plant breeding programs. Indeed, the study of phenotypic and genotypic variation in plant genetic materials provides various insights for better utilizing undiscovered features that can be used to improve crop productivity and its adaptation to a wide range of environmental conditions (Mehrabi et al. 2020). Population genetic structure analysis and evaluation of the level of genetic diversity in durum wheat have a long history and so far numerous researches have been done using different markers from agro-morphological to molecular markers (Etminan et al. 2018). Characterization of genetic diversity using phenotypic markers is often not completely successful due to environmental influences on them, whereas molecular markers disclose genetic similarities in a better context without interference from environmental factors. Hence, molecular markers can provide complete information regarding genetic diversity and population structure in the targeted plant populations (Pour-Aboughadareh et al. 2017).

Several types of genetic markers, such as microsatellites, start codon targeted polymorphism (Scot), sequence-tagged sites (STS), CAAT-box derived polymorphism (CBDP), diverse array technology (DArT), and single-nucleotide polymorphisms (SNP) have been used for evaluating the genetic diversity and population structure analyses in durum wheat e.g. (Mantovani et al. 2008; Pozniak et al. 2012; Hu et al. 2015; Etminan et al. 2016, 2018; Baloch et al. 2017; Mehrabi et al. 2020; Mérida-García et al. 2020; Sansaloni et al. 2020). Of these, the SNP and DArT are identified as higher throughput molecular markers. These techniques are the most common polymorphism among plant materials (Deschamps and Campbell 2010). SNPs show the most frequent type of DNA polymorphism markers that therefore can provide a high density of genetic markers near a target locus (Batley and Edwards 2007). Hence, they are useful DNA-based markers to use in genetic studies. Besides, DArT is another high-throughput genotyping platform based on hybridizing DNA to microarrays (Li et al. 2002). The main advantage of this technique is that it does not require prior sequence information of test individuals. The provision of high-quality dominant markers with a cost- and the time-competitive trade-off is another important advantage usage of the DArT technique in genetic studies (Li et al. 2002). In general, the accessibility of DArT and SNP genotyping techniques would facilitate the genetic analysis and the application of marker-assisted selection in breeding pragmas (Trebbi et al. 2011; Mehrabi et al. 2020).

One of the most efficient approaches currently used for the analysis of quantitative agronomic traits and quality features is association mapping or AM analysis (Flint-Garcia et al. 2003). This method is based on the conception that each quantitative trait that has entered a population will still be linked to the genetic background of the evolutionary ancestor (Gibson and Muse 2009). Hence, AM analysis can discover specific functional genetic alleles or loci associated with phenotypic variation in a trait to simplify the correlation between phenotypic and genome sequence polymorphisms (Ranieri 2015). In other words, this analysis is a powerful approach to discovering the relationship between phenotypic variations and genome polymorphisms in natural germplasm collections (Mehrabi et al. 2020). As a promising approach, several advantages such as much finer mapping resolution, providing broader genomic region coverage, and minimum confidence intervals of the detected loci have distinguished this approach compare to classical linkage analysis (Zhu et al. 2008). Albeit, many studies have represented the genetic diversity and AM analysis in durum wheat germplasm e.g. (Mantovani et al. 2008; Hu et al. 2015; Mwadzingeni et al. 2017; Fayaz et al. 2019; Mehrabi et al. 2020; Sansaloni et al. 2020), there are few studies on genetic diversity and AM analysis using SNP and DArT markers together in durum wheat e.g.(Baloch et al. 2017; Mérida-García et al. 2020). Hence, the main objectives of this study were (i) to the analysis of genetic diversity, genetic population structure, and linkage disequilibrium (LD) in a set of durum wheat genotypes, and (ii) to dissection of associations between the several agronomic characters with the SNP and DArT markers.

Materials and Methods

Genetic materials

Ninety-six durum wheat genotypes including 92 breeding lines along with 4 local cultivars (as the check genotypes) were investigated. All plant genetic materials were provided by the Dryland Agricultural Research Institute (DARI), Kermanshah, Iran. Additional information about the investigated genotypes pedigree is presented in Table 1.

Table 1. Names and code of the studied durum wheat (Triticum turgidum L.) genotypes in this research.

No	Genotype/pedigree	No	Genotype/pedigree
1	Saji	48	Ossl1/Stj5/5……. Waha/3/Stj/Mrb3/6/Stk/Hau//Heca1
2	Zahab	49	Omrabi5 (Jori c69/Hau)
3	SRN-1/KILL//2*FOLTA-1	50	Azeghar1/6/……./Cr/3/F9.3/7/Azeghar1//Msbl1/Quarmal
4	Imren	51	Aghrass1/3/HFN94N8/Mrb5//Zna1/4/IcamorTA0458
5	Baniswaf4/Miki2	52	Ouasloukos1/5/Azn1/……./Waha/3/Gdr2
6	Bcr/Lks4//Mrf1/Stj2/3/Mrf2/NormalHamari//Bcr/Lks4	53	Ter1/3/Stj3//Bcr/Lks4/4/Aghrass1/3/Mrf1//Mrb16/Ru
7	Bicrederaa1//Ossl1/Stj5/3/Ammar8	54	Ter1/3/Stj3//Bcr/Lks4/4/Aghrass1/3/Mrf1//Mrb16/Ru
8	CandocrossH25/3/.....Bcr/Lks4/4/Berghouata1	55	Korifla/AegSpeltoidesSyr//Amedakul
9	Korifla (DS15/Geier)	56	Korifla/AegSpeltoidesSyr//Mrb5
10	CandocrossH25/BEZAIZAHF//Adnan2	57	Korifla/AegSpeltoidesSyr//Loukos
11	CandocrossH25/BEZAIZAHF/3/Sebatel2//Wdz6/Gil4	58	MINIMUS/COMB DUCK_2//…… /11/ALTAR 84/S
12	CandocrossH25...../4/Bcrch1//Ossl1/Stj5/5/Ysf1/Otb6	59	Younes/TdicoAlpCol//Korifla
13	CandocrossH25/Ouasbar2//Berghouata1	60	Waha (Plc/Ruff//Gta/Rtte)
14	CandocrossH25/Ouasbar2//Miki2	61	Korifla/AegSpeltoidesSyr//Lahn
15	CM829/CandocrossH25//Massine	62	T.dic.Crosses61/4/…….NormalHamari//Bcr/Lks4
16	EMN093/Icarasha1	63	Younes1/6/Ouasloukos1/……. SD19539/Waha/3/Gd2
17	Geruftel2/Ammar8	64	Icarasha1/T.Aeg.Crosses
18	IcaJihan39//Marsyr3/Murlagost2	65	Terbol975/Geruftel2/7/…… /3/Mtl5/4/Lahn
19	Icamilmus1/Waha/4/Icasyr1/3/Bcr/Sbl5//Turartu	66	Icamilmus1/3/Marsyr3//……..Mgnl3/Ainzen1
20	Gidara2 (Stojocri/Omrabi3)	67	Icasyr1/3/Gcn//Stj…… .3/Bcr/Gro1//Mgnl1
21	Icamilmus3/Younes1/4/Marsyr3/3/Gcn//Stj/Mrb3	68	Marsyr3/Murlagost2//T.dicocconIG91683/Ammar9
22	Icamilmus4/3/Bcrch1//Ossl1/Stj5/4/Adnan2	69	Waha//Azeghar2/RedGrainTer112
23	Icamilmus4/Adnan2//Adnan2	70	Icasyr1/3/Bcr/Sbl5//T.urartu/4/T.Aeg.Crosses
24	Icamilmus5/Miki2//Geromtel1/Icasyr1	71	Icarasha2 (Stj3//Bcr/Lks4/3/Ter3)
25	Icarasha1/IcaKader2	72	Marsyr3/3/Gcn//Stj/Mrb3/4/T.Aeg.Crosses
26	Icasyr1/3/Bcr/Sbl5//Turartu/4/13376/Bcrch1//Ossl1/Stj5	73	Marsyr3//Saadi1989/Chan/3/T.Aeg.Crosses
27	Icasyr1/3/Bcr/Sbl5//Turartu/4/Icarasha1	74	ICAMORTA0469/……. Geromtel1/Icasyr1
28	Maamouri3/Sebatel2	75	Jk/Ch1613//TpolonicumIG97746
29	Mrb5/Tdicds600545.....Gil4/4/Saadi1989/Chan/5/Geruftel2	76	PLATA_7/ILBOR_1//……..//GUAYACAN INIA
30	Ossl1/Stj5/5/….SD19539/Waha/3/Stj/Mrb3/6/Icajihan1	77	Jk/Ch1523//Atlas2/3/Mrb5/Lahaucan
31	Quarmal/Gbch2/3/Mrf2/NormalHamari//Bcr/Lks4/4/IcaKader2	78	Jk/Ch1604//Ysf1/Otb6/3/Adnan2/Berghouata1
32	Sbh/4/D68193A1A.....5/Marsyr3//Saadi1989/Chan	79	Icamilmus1/3/Marsyr3//……Mgnl3/Ainzen1
33	SwAlg/Gd112/4/……/Mrf2/NormalHamari//Bcr/Lks4	80	Mgnl3/Aghrass2/4/……Bicrederaa1/Azeghar2
34	CM829/CandocrossH25//Icajihan7	81	Icasyr1//Mrf2/T……..Waha/3/Gd2
35	Mrf1/Stj2//Gdr2/Mgnl1/3/Bcrch1	82	Icasyr1//Mrf2/T.dids20123……Waha/3/Gd2
36	Ter1//Mrf1/Stj2/3/Icasyr1	83	Stj3//Bcr/Lks4/3/……Sebatel2//Wdz6/Gil4
37	Icarasha2 (Stj3//Bcr/Lks4/3/Ter3)	84	Icarasha1//Quabrach3/RedGrainTer109
38	Adnan1//Mgnl3/Ainzen1	85	Miki3 (Stj3//Bcr/Lks4)
39	Geromtel1/IRANYT053//Mgnl3/Ainzen1	86	Mrb3/Tboeoticum500648……..//Mrb16/Ru
40	IcamorTA0462/4/Gdr243.....TA0463/5/Ter1//Mrf1/Stj2	87	Mrb3/Tboeoticum500648……/3/Mrf1//Mrb16/Ru
41	IcamorTA0471//……../3/Mgnl3/Ainzen1	88	SwAlgia/Gd181…./4/EMN0917/Adnan2
42	IcamorTA041/4/……Bcr/Gro1//Mgnl1/5/MIKI2	89	ALTAR 84
43	Waha (Plc/Ruff//Gta/Rtte)	90	SORA/2*PLATA_12/3/SORA/2*PLATA_12//SOMAT_3/4/AJAIA_13/YAZI//DI

44	Mrb3/Mna1//Ter1/3/……../Ammar7/4/Beltagy2	91	ALTAR84/STINT//……..RASCON_21
45	Maamouri1/5/…….Bcr/Lks4/3/Icamor"s"/6/Mgnl3/Ainzen1	92	CAMAYO//HYDRANASSA30……2*RASCON_21
46	Azeghar2/5/……/Lks4/3/Icamor"s"/6/Stj3//Bcr/Lks4/3/Ter3	93	ALTAR 84/ ……DUCK_2//CHAM_3/3/GREEN_19
47	Mgnl3/Ainzen1/3/Bcr/Gro1//Mgnl1	94	PLATA_7/ILBOR_1…….POD_9/11/CIRNO C 2008

Phenotypic assay

The field experiment was performed at the Sara-rood Dryland Agricultural Research Institute, Kermanshah, Iran (Latitude 34°20 N, Longitude 47°17 E, and Altitude 1351 m above sea level) during the 2017–2018 and 2018–2019 cropping seasons. The weather condition of the genotype evaluation site is shown in Fig.1. In each cropping season, all genotypes were planted in an augmented design with 5 blocks and 4 repeated check genotypes. Sowing and crop management were done based on experts' advice. The plot size was six rows, 3 m in length, spaced 20 cm apart. Grain sowing was done manually. During plant growth and development, five agro-morphological traits were recorded based on the standard system of the National Gene Bank of Tunisia (NGBT). The measured traits were: (1) the number of days-to-heading, (2) the number of days-to-physiological maturities, (3) plant height, (4) thousand kernel weight, and (5) grain yield. After adjusting means for measured traits, Variance analysis on morphological data was done based on the Best Linear Unbiased Estimation model(BLUE) using the MIXED procedure of the SAS-STAT statistical package.

DNA Extraction

Ten grains from each genotype were sown in plastic pots under controlled conditions at the research greenhouse. The DNA of 94 durum wheat genotypes was extracted by the CTAB method with a few changes from fresh leaves(Doyle and Doyle 1990). To check the absence of breakage in the extracted DNA molecules, a standard 0.8% agarose gel was used. In this method, 5 microliters of the extracted DNA solution were stained and loaded with a dye, and gel electrophoresis was performed with a current of 75 volts for 30 minutes. To determine the quality of DNA, it is possible to compare the pattern and color intensity of the used lambda phage DNA with the position and intensity of the band obtained from the electrophoresed DNA. After determining the concentration of the extracted DNA by the spectrophotometric method in comparison with the control DNA, each of the extracted DNA tubes was diluted to a concentration of 50 ng/µl.

Genotypingassay

For genetic evaluation in 94 wheat genotypes, the extracted DNA was prepared at a concentration of 50 ng/µl. From each genotype, at least 20 microliters of each sample were loaded into 96 DNA plates. The plates containing the samples were sent to the DArt Diversity Array Technology Center located at the University of Canberra in Australia (https://www.diversityarrays.com) for genotypic evaluation by Sililco DArT and SNP markers. DArT markers were firstly filtered in terms of unsuitable parameters such as minor allele frequency > 2%, repeatability > 90%, and unread rate > 90%. Finally, 7882 Silico DArT and 5455 SNP markers were yielded for further analysis. The obtained data are of two categories: Sililco DArT markers based on the presence and absence of a gene array sequence in each sample. Another category of SNP markers is based on the presence or absence of a mutation in a gene region. Sililco DArT markers are dominant markers and SNP markers are common markers(Alam et al. 2018).

Genetic analysis

The amount of polymorphism information content (PIC) for SilicoDArT markers was calculated by Power Marker $_{V.3.25}$ software(Powell et al. 1996). Analysis of population genetic structure (Population Structure) was read using STRUCTURE HARVESTER software. The number of sub-populations resulting from the population analysis by SilicoDArT markers data was determined by the number of K, and based on this, the genetic information between the sub-populations and their graphic plot were extracted from the population software(Evanno et al. 2005). The amount of linkage disequilibrium(LD) of molecular data (SilicoDArT and SNP) was calculated using TASSEL software and its graphical analysis was done in the R environment using the GAPIT package (Lipka et al., 2012). Finally, Association mapping was calculated separately by TASSEL and GAPIT software and compared to ensure their correctness. In both software, CMLM (Compressed Mixed Linear Model) method was calculated using PCA, KINSHIP MATRIX, and Q matrix data. Markers that had a significant level ($P<0.001$) were considered correlated markers with each of the phenotypic data. Due to the high number of markers used, therefore, to determine markers with higher efficiency, Q-Q-Plot analysis was also performed simultaneously with CMLM, and markers that had Log10 (P) greater than 3 were selected as markers with very high correlation. To identify the candidate genes, the physical location of each marker was searched in the IWGSC Refseq v1.0 genome, and if a gene is found 5-10 kb above the desired location, it will be considered correlated with the desired marker.

Results And Discussion

Genetic diversity with pheno-morphological markers

One of the primary goals of any breeding program is to identify and select the desired genotypes in terms of various traits(Mohammadi et al. 2014). To achieve this goal, the traits used should be clearly defined in terms of quantity and quality, so that in combination with important economic traits, they can be used to maximize the Grain yield of the selected genotypes. In this regard, several statistical methods, including correlation analysis, multiple regression analysis, and path coefficient analysis, have been proposed and widely used by experts and breeders. In the meantime, one of the most important statistical methods introduced to achieve this goal is genotype-trait graphical analysis (GT), which is obtained through analysis into principal components(Yan 2001). In this method, the bi-plot used is obtained through the first two components extracted through the analysis of eigenvalues on the desired attribute data in

several environments (year, place, or a combination of year and place). Although this method was introduced to analyze multi-environment experiments, it can be used for any type of data that has a two-way structure(Yan 2001). In general, according to Yan and Rajcan (2002), the desired genotypes identified through this analysis can be used as parents in cross-breeding programs or directly for the production of commercial cultivars. In the present research, the GT charts related to 94 studied durum wheat genotypes based on the studied traits, separated by the two years of the experiment, are shown in Fig. 2. Based on the first year's data, 52.92% of the total genetic diversity among genotypes was justified by the first two components. By justifying the polygon diagram drawn in this bi-plot, genotypes No. 3, 4, 38, 77, 24, 43, and 86 are located at the vertices of the polygon. Therefore, it can be said that genotype No. 43 had the highest value in terms of thousand kernel weight, and genotypes No. 3 and 4 had the highest value in terms of plant height. Also, the highest amount of grain yield belonged to genotype No. 86 (Fig. 2A). The bi-plot drawn based on the obtained data from the second year of the experiment showed that the first two components accounted for 33.22 and 26.53 % of the total variation related to studied traits among genotypes, respectively (Fig. 2B). Based on the second year's data, genotypes No. 18, 14, 91, 44, 72, 41, 45, 47, and 65 were placed in the vertices of polygons. As shown in Fig.2B can be seen, genotype 72 was superior to other genotypes in terms of thousand kernel weight. On the other hand, the best genotypes in terms of plant height and the number of days to physiological maturity were genotypes No. 47 and 65. Genotype No. 41 was also identified as the latest genotype. In terms of grain yield, genotypes No. 9, 11, and 86 had better potential than other genotypes. Considering that the bi-plots showed all the available changes did not justify, these predictions may not accurately reflect the observed numbers. However, the predictions are very close to reality and it is possible to select genotypes with superior values compared to previously. Yan and Kang (2002) in soybean, Yan and Frégeau-Reid (2008) in oats, Dehghani et al. (2008) in winter rapeseed, and Rahmati et al. (2020) in cantaloupe applied this method to evaluate, compare and select superior cultivars and reported that this method in addition to showing the relationship between different traits graphically, can make the visual comparison between genotypes very possible.

Molecular diversity by SilicoDArT markers

The SilicoDArT markers with very high genome coverage were used to study genome diversity in a panel consisting of 92 durum wheat genotypes, which is a valuable and powerful tool for GWAS analysis in the future. Also, these germplasms will be very valuable resources for extracting alleles and selecting genotypes that give resistance genes to various biotic and abiotic stresses through linkage mapping in the whole genome. Candidate SilicoDArT markers can be used to establish marker-trait association (MTA) concerning specific genes or large QTLs controlling valuable traits in durum wheat. The information (data) of DArTseq markers was estimated by polymorphism information content (PIC). Polymorphism information content (PIC) is used to provide information about the variation of a gene or segment of DNA in a population to indicate evolutionary pressure on an allele and mutations at a gene locus that may have occurred over some time. When the scoring of a marker is (0) and (1) with ratios of 50% to 50%, the maximum PIC value is calculated as 0.5. In this study, a large number (7882) of high-quality SilicoDArT polymorphic markers were filtered out of 62269 markers and used for the analysis of genetic diversity and population structure in a set of 94 durum wheat genotypes. The average PIC values for all SilicoDArT markers were equal to 0.38, which indicates the high efficiency of these markers for measuring the genomic diversity of durum wheat (Table 2). In general, the distribution of PIC values among the markers was asymmetric and tended towards values higher than the average, so about 60% of the markers showed PIC above the average (Fig.3). Our ability to compare PIC values in this study was limited because there are few reports on genetic diversity using DArT, and we could not track documented reports using DArTseq markers in durum wheat genetic diversity studies. The average PIC values in the recent study were higher than the PIC value reported by Ren et al. (2013) using SNP markers in the global collection of 150 durum wheat genotypes (it was equal to 0.18888). Moragues et al.(2007) investigated the genetic diversity of 63 indigenous populations of durum wheat from Mediterranean countries using AFLP and SSR markers, the average PIC values for AFLP and SSR markers were 0.24 and 0.70, respectively. The distribution of PIC values for SilicoDArT markers is shown in Fig.3. The distribution of PIC values among markers was almost non-uniform and more than 3000 SilicoDArT markers had PIC values close to 0.5. About 75% of SilicoDArT markers had PIC values greater than 0.3. In addition to the PIC value, some other quality parameters, such as the call rate and the reproducibility of each marker within the panel of examined durum wheat genotypes, were also estimated. Sequence reading rate (Call Rate) is the percentage of valid scores among all possible scores for a marker, where reproducibility is measured as the percentage of scoring repeatability for repeated samples. The average call rate and reproducibility of SilicoDArT markers in this study were more than 0.98 and 0.92, respectively (Table 2). Although the composition of restriction enzymes as well as the amount of diversity in the presented panel is different from other plant species, the amount of technical measurement of repeating pairs for the score of markers and success in reading the sequence of markers in the investigated sample is high and comparable to the previous applications of SilicoDArT in other plant species(Grzebelus et al. 2014; Ndjiondjop et al. 2017; O'Connor et al. 2019). The genomic coverage of SilicoDArT markers in recent research was uniform for most chromosomes and the highest (853) and lowest (300) number of markers were observed in Chr7B and Chr1A, respectively. Chromosome size covered by SilicoDArT markers ranged from **829200** kbp in the linkage group (Chr3B) to **589293.786** kbp in the linkage group (Chr1A) which is similar to previous studies in durum wheat using high-performance GBS (Baloch et al. 2017).

Table 2. SilicoDArT Markers Distribution on Durum Wheat Genome, Polymorphism Information Content (PIC), Call Rate and Average of Reproducibility

Linkage Groups	No. Markers	Chromosome Size(kbp)	Call Rate (mean)	One Ratio(mean)	PIC (mean)	Average of Reproducibility
Chr1A	**300**	**589293.786**	0.926811	0.53253	0.392783	0.987783
Chr2A	574	779665.301	0.928376	0.531615	0.378022	0.987819
Chr3A	454	750610.4	0.926852	0.528354	0.371944	0.987748
Chr4A	489	742808.8	0.929105	0.507459	0.382376	0.986216
Chr5A	436	708977	0.928146	0.520776	0.375594	0.98842
Chr6A	456	616320.1	0.928312	0.510363	0.392759	0.987314
Chr7A	789	735262.1	0.929175	0.510735	0.384203	0.986959
Ch1B	681	688589.5	0.92758	0.533268	0.385003	0.987277
Chr2B	759	800769.3	0.927394	0.531767	0.393895	0.987843
Chr3B	582	**829200.1**	0.922085	0.564573	0.369268	0.9861
Chr4B	333	671754.7	0.924849	0.538444	0.37495	0.987985
Chr5B	586	708929.4	0.923233	0.566275	0.382289	0.986366
Chr6B	590	720840	0.926843	0.534025	0.382743	0.986644
Chr7B	**853**	749946.8	0.92991	0.515027	0.392338	0.987082
Mean	563	720926.2	0.927048	0.530372	0.382726	0.987254
Group A	499.72	703276.78	0.928111	0.520261	0.382525	0.987465
Group B	626.3	738575.7	0.925985	0.540483	0.382927	0.987042

Kinship coefficients between pairs of genotypes

To the acceptable evaluation of genetic diversity in durum wheat genotypes, relative kinship coefficients between each pair of genotypes were estimated based on DArTseq markers. The value of the matrix of kinship coefficients can vary from 3 to -3, as much as its average tends to be zero or less, it indicates a very high diversity in the population. Based on the kinship coefficients matrix, the investigated genotypes were divided into 3 groups for gene data. Also, the results showed that about 60% of the kinship coefficients between pairs of genotypes were between zero (except zero) and 0.1 and about 39% of the values were between 0.1 (except 0.1) and 0.5. Only about one percent of the coefficients had values greater than 0.5. Therefore, the results of this analysis indicated that there is very little kinship in the studied genotypes(Fig.4). Genotypes that had a high kinship coefficient are very likely to have a common ancestry(Salgotra et al. 2015).

Bayesian clustering

The existence of groups and subgroups in large populations can be due to reasons such as different geographical origins of genotypes, natural or human selection, genetic drift, etc. (Flint-Garcia et al. 2003). In this analysis, the possibility of subpopulations in the investigated germplasm was evaluated using STRUCTURE software. According to the results, K and Delta K statistics were extracted and two-dimensional diagrams were drawn, which clearly showed the curve at the maximum of K=4 (Fig. 5). Therefore, according to the obtained results, durum wheat genotypes were grouped into four separate subpopulations based on the Bayesian model (Fig. 6). The four subpopulations showed relatively high genetic diversity, which ranged from 0.02 in a subpopulation (POP4) to 0.34 in a subpopulation (POP3). Net Nucleotide Distance is a parameter to measure genetic diversity among populations that revealed the highest (0.270) and lowest (0.07) genetic distance between POP2 and POP4 subpopulations and POP1 and POP2 subpopulations (Table 3).

Table 3. Genetic divergence among (Net Nucleotide Distance) and within (expected heterozygosity) population, proportion of membership and mean value of Fst observed from the study of population structure of 92 durum wheat genotypes using SilicoDArT markers.

Population	Net Nucleotide Distance			Expected Heterozygosity	% 0f membership	Mean fixation index (Fst)
	POP2	POP3	POP4			
POP1	0.070	0.084	0.247	0.20	46	0.41
POP2		0.085	0.270	0.20	26	0.46
POP3			0.249	0.34	19.5	0.21
POP4				0.02	8.5	0.95

Analysis of linkage disequilibrium(LD)

Genome-wide Association study is the method that determines the QTLs location based on linkage disequilibrium(Gupta and Varshney 2005). Therefore, for association mapping in addition to the composition of the population structure, the extent of linkage disequilibrium in the genome has fundamental importance(Al-Maskri et al. 2012). The linkage disequilibrium index was evaluated using 605212 pairs of DArTseq markers with specific map locations along 14 durum wheat chromosomes. Pairwise linkage disequilibrium values were estimated using the square of allelic frequency correlations (r^2) between markers. To identify differences in intra-chromosomal linkage disequilibrium, the average values of r^2 between pairs of markers are divided into five groups based on the genetic distance between them: markers with very high linkage (Distance less than 5 cM), Contiguous markers (with a distance of 5-10 cM), markers with medium continuity (with a distance of 10-12 cM), weakly connected markers (with a distance of 20-50 cM) and independent markers (with a distance of more than 50 cM). The evaluation of intra-chromosomal linkage disequilibrium showed that the linkage disequilibrium decreases with increasing genetic distance. Also, the results revealed that linkage disequilibrium varies from chromosome to chromosome. The greatest imbalance was related to markers with a distance of less than 5 cM. Pairs of markers that are less than 5 cM away and have a lot of continuity showed the highest imbalance among all chromosomes. Previous studies have demonstrated that linkage disequilibrium varies along the chromosome and from chromosome to chromosome(Chao et al. 2010; Hao et al. 2011). The results also showed that increasing the distance between the markers does not completely remove the imbalance. This is due to the presence of other effective factors such as population structure, genetic drift, migration, selection, and mutation on linkage disequilibrium(Reich and Lander 2001). Analysis of linkage disequilibrium was also investigated at the genome level. To show the extent of linkage disequilibrium at the level of the genome, the average reduction in disequilibrium was obtained by plotting intra-chromosomal r^2 values against the genetic distance between markers (Fig.7). The amount of LD within the chromosomes based on both SilicoDArT and SNP markers for both durum wheat genomes (A and B) is shown in Table 4. The amount of LD in both types of markers was very wide, in such a way that 605,212 pairs of significant markers were observed in the entire population. LD analysis between A and B genomes revealed that there is a high number of significant marker pairs (82052) in the B genome compared to the A genome (68885). The average r^2 values for the whole population were equal to 0.14 and for genomes, A and B were equal to 0.12 and 0.11, respectively (Table 4). A graphic analysis of the distribution of LD markers is shown in Fig.7. In general, 15 kb of the genome was covered in LD. The largest amount of linkage disequilibrium was observed in chromosome 4B, some of the markers are shown on this chromosome (Fig.8).

Table 4 Characteristics of linkage disequilibrium in durum wheat genome based on DArTseq markers

Genome	Total pairs	Significant pairs (P≤0.001)	% of Significant pairs	Mean r^2	Pairs in complete LD	Critical r^2 value
A-genome	269891	68885	24.7	0.13	1337	0.12
B-genome	335322	82052	24.4	0.12	1146	0.11
Whole-genome	605212	148937	24.6	0.13	2438	0.14

Genome-Wide Association Study

The correlation analysis was performed using markers with a frequency of more than ten percent and the P-value statistic was considered with 1000 permutations. Also, the lowest P-value was used as the basis for selecting the correlated markers. The distribution of markers was studied based on the coefficient of determination of the marker in the regression model. The coefficient of determination (r^2) is the proportion of the phenotypic variance accounted for by the QTL for each locus. To association mapping of traits including day-to-flowering, day-to-physiological maturity, plant height, thousand seed weight, and grain yield of durum wheat, both molecular and phenotypic data were used. Association mapping analysis for each trait was done based on CMLM (Compressed Mixed Linear Model). To determine the significant markers associated with the studied traits, QQPlot analysis was also performed and finally, the markers with -log 10(p) were selected. The Manhattan plot of two cropping years based on each of the measured traits is shown in Fig.9 and Fig.10, respectively. In Tables 5 and 6, the full report of significant associations of markers related to all traits for both two cropping years is presented. The results of a genome-wide association study by the CMLM method in 2016 showed that 19 markers had a significant association with the studied traits (Table 5). The results of association mapping for each of the traits in the first cropping year are summarized as follows:

- Two significant relationships were identified for days to flowering by markers 1703829 and 1087984, which have located on chromosomes 6B and 7A, respectively.
- The correlation of 3940462 markers on chromosome 4A with days to maturity was significant.
- Ten significant relationships were identified between plant height and DArT markers, these markers had chromosomal locations 3A, 3B, 4B, 5A, 5B, 6A, 6B, and 7B.
- For the trait of thousand kernel weight were identified four significant associated markers were including 3570140, 992437, 3534094, and 3533907 on chromosomes 7A, 3A, 5B, and 3B, respectively.

- About grain yield as an important trait identified three significant relationships with 1108172, 4989009, and 1233550 markers that have located on chromosomal position 3A, 2B, and 1B, respectively.

The results of genome-wide association analysis by the CMLM method in 2018 indicated that 10 markers had a significant association with the studied traits (Table 6). The summary of the results was as follows:

- Association mapping for day-to-flowering revealed two significant markers (3935863 and 1211191) which have chromosomal locus 1A and 5B, respectively.
- The number of three markers 3064874, 1034732, and 1025860 with gene loci 794315843, 673038094, and 666991274 on chromosomes chr3B, chr4A, and chr4A respectively, had a significant relationship with the number of days to maturity.
- The relationship of plant height with marker 1057654 with gene locus 691163974 on chromosome chr7A was significant. This is even though 10 significant relationships were identified for plant height in the previous year.
- The weight of a thousand kernel had a significant relationship with 3025786 and 981221 markers. These markers with gene loci 609888794 and 684637596 are located on chr6A and chr3B chromosomes, respectively.
- Two significant associations were identified between grain yield and DArT markers. The locus of these markers is located on chromosomes 6A and 7A, respectively.

In general, according to the results, the same marker (1057654) was identified for grain yield and plant height in the second cropping year. This marker with genetic locus 691163974 has located on chromosome 7A. By comparing the significant marker-trait relationships as well as the durum wheat consensus map, the position of the marker can be identified and checked by comparing them with physical maps. If the marker distance is more than +10 or -10 cM that position can be identified as a new QTL. Based on this, 29 quantitative positions were identified and confirmed in this study. Among the identified loci, 9 loci were not previously recorded in Durum's Consensus map (Table 7).

Table 5 Associated markers with studied traits in 94 durum wheat genotypes using linkage analysis in 2016 crop year.

Traits	Marker	Linkage Group	Locus	value*P*-	Maf
Days to Flowering	1703829	Chr6B	4863112	0.000075	309783
	1087984	Chr7A	42448091	0.000083	0.141304
Days to Maturity	3940462	Chr4A	606757033	0.000087	0.943478
Plant Height	1228105	Chr5A	4994308	0.000000	0.288043
	3946051	Chr3A	697061451	0.000000	0.13587
	4004958	Chr5B	648289673	0.000001	0.130435
	4405812	Chr5A	109937709	0.000002	0.369565
	1068144	Chr6A	613084911	0.000013	0.70652
	1093423	Chr6B	33427529	0.000026	0.375
	2275869	Chr3B	19232431	0.000034	0.483696
	5410721	Chr7B	625701808	0.000041	0.065217
	1671913	Chr3A	497297304	0.000043	0.233696
	1090315	Chr4B	629706516	0.000067	0.413043
Thousand grain weight	3570140	Chr7A	71591095	0.000031	0.211957
	992437	Chr3A	541246518	0.000041	0.391304
	3534094	Chr5B	27137735	0.000044	0.255435
	3533907	Chr3B	739839965	0.000056	0.461957
Grain Yield	1108172	Chr3A	652905402	0.000071	0.059783
	4989009	Chr2B	106775933	0.000021	0.179348
	1233550	Chr1B	558575664	0.000033	0.190217

Table 6 Correlated markers with studied traits in 94 durum wheat genotypes using linkage analysis in 2017 crop year.

Table 7 Comparison of gene locations identified by Consensus map of durum wheat

Traits	Marker	Linkage Group	Locus	value*P*-	Maf
Days to Flowering	3935863	Chr1A	30133162	0.000032	0.086957
	1211191	Chr5B	700372980	0.000037	0.065217
Days to Maturity	3064874	Chr3B	794315843	0.000006	0.097826
	1034732	Chr4A	673038094	0.0000028	0.25
	1025860	Chr4A	666991274	0.000078	0.304348
Plant Height	1057654	Chr7A	691163974	0.000051	0.097826
Thousand Grain weight	3025786	Chr6A	609888794	0.000044	0.070652
	981221	Chr3B	684637594	0.000090	0.0157609
Grain Yield	1057654	Chr7A	691163974	0.000016	0.097826
	7918245	Chr6A	611855760	0.000040	0.141304

Traits	Marker	Linkage Group	cM
Days to Flowering	1703829	Chr6B	3.34
	1087984	Chr7A	3.34
	3935863	Chr1A	35.82
	1211191	Chr5B	137.12
Days to Maturity	3940462	Chr4A	54.26
	3064874	Chr3B	133.21
	1034732	Chr4A	96.08
	1025860	Chr4A	96.16
Plant Height	1228105	Chr5A	5.94
	3946051	Chr3A	-
	4004958	Chr5B	-
	4405812	Chr5A	-
	1068144	Chr6A	98.22
	1093423	Chr6B	13.14
	2275869	Chr3B	17.44
	5410721	Chr7B	81.31
	1671913	Chr3A	-
	1090315	Chr4B	-
	1057654	Chr7A	129.67
Thousand grain weight	3570140	Chr7A	149.67
	992437	Chr3A	68.59
	3534094	Chr5B	161.82
	3533907	Chr3B	-
	3025786	Chr6A	98.83
	981221	Chr3B	81.17
Grain Yield	1108172	Chr3A	65.11
	4989009	Chr2B	-
	1233550	Chr1B	-
	1057654	Chr7A	129.67
	7918245	Chr6A	-

Conclusions

Durum wheat in Iran and other main producing countries in the world such as Canada, Mexico, the United States of America, Australia, Italy, Turkey, etc. is mainly grown under rainfed conditions and due to environmental stresses, especially drought stress, mostly have a low grain yield(Beres et al. 2020). Wheat and its products can account for 20% of the protein and calorie consumption per capita of the world's 9.7 billion people in 2050(CRP-WHEAT 2016). Durum wheat is primarily produced for pasta but is also an important ingredient for couscous and bulgur, especially in North Africa and the Middle East(Beres et al. 2020). Therefore, the introduction of new germplasms in breeding programs for sustainable agriculture and food security for the growing world population is necessary and inevitable(Zeven 2000). Shahid Masood et al. (2005) showed that wild emmer wheat was domesticated in the Fertile Crescent. Therefore, as a consequence of this event, kinds of wheat (bread and durum) were domesticated, which have caused been narrowed their genetic base and diversity, now. The initial loss of durum wheat genetic diversity during the domestication process is estimated at 84%. Throughout their evolutionary history, wheat crops have been formed through the major contributions of farmers to meet consumer needs, cultural practices and respond to growing conditions and socio-economic change(Zeven 2000). It is expected that the comprehensive knowledge about the genetic diversity of durum wheat from various eco-geographic regions will have a significant impact on the protection and use of desirable germplasms. This can help breeders in designing new approaches to achieve profitable diversity in breeding programs. Genetic diversity is the basis of each breeding program. The study of genetic diversity subordination from geographical and climatic diversity indicates the possible adaptation of plant germplasms with different habitats and environments. Also, evaluating the genetic diversity and structure of the population in the germplasm sources enables the identification of genetic capacities associated with important plant breeding

goals(Mohammadi and Prasanna 2003). Due to limited genetic diversity in modified agronomy species for adaptation to climate change, and as a result, the poor chance of achieving new alleles in these plant materials, screening of germplasm resources, including genetic reserves and breeding lines, can provide a rich and varied gene source of new and ideal alleles for plant breeders (Pour-Aboughadareh et al. 2017). In this regard, in the current study, 94 durum wheat genotypes were evaluated in terms of some agronomic traits in field conditions during two consecutive cropping years. The obtained results from the analysis of variance (The table has not been brought.) using the BLUE model revealed the existence of a high level of genetic diversity among the set of studied durum wheat genotypes in terms of the number of days to flowering, the number of days to maturity, plant height, thousand kernel weight, and grain yield. These results show that among the evaluated genotypes, finding individuals with favorable agronomic characteristics is not far from expected, and they can be used in the breeding program to improve the existing cultivars (Fig. 2). Previously, in some studies conducted on the durum wheat has been reported high agronomic and morphological diversity (Mehrabi et al. 2020; Ouaja et al. 2021; Dastfal et al. 2022). Based on the graphical analysis of Genotype-Trait, the highest amount of grain yield in the first cropping year belonged to genotype number 86 (Fig. 2A) and in the second cropping year, genotypes number 9, 11, and 86 had the better grain yield potential than other genotypes (Fig. 2B). Corresponding to the fact that the bi-plots did not justify all the existing variations, these predictions may not accurately reflect the observed data. However, the predictions are very close to reality and it is possible to identify genotypes with superior values for each trait. Yan and Rajcan (2002) stated that the desired genotypes identified through genotype-trait analysis can be used as parents in breeding programs or directly for the production of commercial cultivars. Diversity Arrays (DArTseq) technology has proven to be a powerful molecular marker, so it can be reliably used to analysis of genetic diversity and gene mapping in many plant species(Akbari et al. 2006; Briñez et al. 2012; Grzebelus et al. 2014). In this study, in addition to evaluating the genetic diversity and population structure in the set of durum wheat genotypes, DArTseq markers, and FarmCPU analysis were used to identify gene loci related to the investigated traits. The results showed a high level of molecular genetic diversity among the investigated germplasm collection (Table 2). Although this result was not unexpected, it was consistent with other obtained results regarding the presence of a significant level of genetic diversity in durum wheat germplasm(Sukumaran et al. 2018; Wang et al. 2019; Mehrabi et al. 2020; Ávila et al. 2021; Alemu et al. 2021). The PIC values were estimated to check the efficiency of the markers for separating and differentiating the characteristics of genotypes (Table 2). Our ability to compare PIC values in this study was limited because there are few reports on genetic diversity using DArT and we could not track down documented reports using DArTseq markers in durum wheat genetic diversity studies. The average values of PIC in this study were higher than the PIC values in the study of Ren et al. (2013) which using SNP markers in the global collection of 150 durum wheat genotypes have reported equal to 0.18888. In the present study, the values of the kinship coefficient in the studied genotypes varied from 3 to -3. Based on the Kinship matrix information, the studied genotypes were divided into 3 groups. Relative kinship among individuals in the population can lead to the identification of false associations between markers and traits and therefore it should be taken into consideration in relational mapping studies. Estimation of coefficients of the kinship between pairs of genotypes showed that only about one percent of the coefficients had values above 0.5, which indicates that there is a weak genetic kinship between the studied genotypes. One of the reasons why some genotypes have a high kinship coefficient is their common pedigree, in other words, common parents played a role in their creation(Saura et al. 2013). In genetic research, it is possible to get a perspective of the evolutionary relationships in populations by the structure analysis, which is used to investigate the kinship of individuals within and between the populations. In genome-wide association study, due to the achievement of useful and valid results, the studied population should not be divided into subgroups, and using the structure and relationships of the populations in association analysis can be reduced false positive results to a great extent(Breseghello and Sorrells 2006). The results of population structure analysis based on the *Bayesian* model put the studied genotypes into four groups. Based on the obtained bar plot, 43 genotypes (46%), 25 genotypes (26%), 18 genotypes (19.5%), and 8 genotypes (8.5%) were grouped into subpopulations 1, 2, 3, and 4, respectively. Since some clusters showed a mixture of durum wheat genotypes, it can be concluded that most probably the same parental lines were used in the breeding programs of these genotypes. This result corresponded with previous studies conducted by SNP and SilicoDArT markers generated from DArTseq, on similar sets (in terms of population size) of durum wheat genotypes(Baloch et al. 2016; Fayaz et al. 2019). Laidò et al. (2013) investigated the genetic diversity of 230 durum wheat populations using 6 morphological traits, 9 biochemical traits related to grain protein, 26 SSR markers, and 970 DArT markers. In their research, both ISSR and DArT markers provided a correct insight of the genetic structure within the tetraploid wheat populations. In our research, the highest amount of linkage disequilibrium was observed for markers with a distance of less than 5 cM on chromosome 4B, which indicates that linkage is one of the most effective forces that cause disequilibrium in the population. In addition, high values of linkage disequilibrium can be the result of selection that increases inter-allelic correlations in certain loci(Schlötterer 2003). The average coefficient of r^2 in the whole genome was equal to 0.14, which corresponded with the reported values by Hao et al. (2010) and Horvath et al. (2009). The linkage disequilibrium analysis revealed that there is a high number of significant marker pairs in genome B compared to genome A which doesn't correspond with the results of Mehrabi et al. (2020). Hao et al. (2010) investigated genetic diversity and linkage disequilibrium in 376 Asian and European wheat populations on chromosome B3, which showed the average linkage disequilibrium of European populations (0.23) and Asian populations (0.18), which indicates the strong linkage disequilibrium in these genotypes. Concerning the use of DArT markers and their high efficiency in the investigation of the association between marker-trait, extensive studies have been reported in a variety of agronomic plants(Sodedji et al. 2022; Bukhari et al. 2022; Nantongo et al. 2022; Fufa et al. 2022). In the present study, the relationship between the marker-trait was investigated using the CMLM model. The CMLM model is significantly superior to other models, including GML, due to the control of false positive errors, and many studies have pointed out the high efficiency of this model in genome-wide association analysis (Hu et al. 2015; Maulana et al. 2018). In general, 29 significant "marker-trait" relationships were found during two cropping years, and out of these 29 related markers, 19 markers were identified in the first cropping year and 10 markers were identified in the second cropping year (Fig.9 &10). Among the identified markers, there were 11 markers related to plant height. One of the most promising approaches to improving traits is the introduction of new alleles. As a first step, allele extraction approaches can be performed in different ways and with different sets of plant germplasm. Therefore, performing multiple trials in different environments and the phenotypic analysis of the studied germplasm can be identified associated markers for a genetic description of complex traits. In addition, genetic relationships can be used to choose the desirable parents for breeding programs, if the genetic pedigree data is not available. The information produced here can be used to design breeding programs tailored to regional and local needs. The presence of many SilicoDArT and SNP markers, the cost-effectiveness, and the high relatively polymorphism information content (PIC) have provided excellent aspects for widespread screening of the whole genome for genetic diversity, linkage disequilibrium studies, and genome-wide association analysis purposes. The genetic diversity of the durum wheat gene pool should be further explained to facilitate the classification, proper conservation, protection, and use of these valuable genetic resources.

Declarations

Acknowledgments:

We would like to express our gratitude to the Islamic Azad University of Sanandaj branch and the Dryland Agricultural Research Institute, Sararud, Kermanshah, for their cooperation to carry out the present research as best as possible.

Author's contribution:

All authors contributed to the study's conception and design. Material preparation and Methodology have done by Alireza Etminan and Reza Mohammadi; and data collection and analysis have performed by Peyman Ebrahimi, Ezzat Karami, and Reza Talebi. The manuscript has been written by Ezzat Karami and all authors commented on previous versions of it. All authors read and approved the final manuscript, which has been submitted to this Journal with the consent of all authors.

Conflict of Interest:

The authors declare that they have no conflict of interest.

Funding:

This project did not have a financial supporter.

References

1. Akbari M, Wenzl P, Caig V, et al (2006) Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Theor Appl Genet 113:1409–1420. https://doi.org/10.1007/S00122-006-0365-4
2. Al-Maskri A., Sajjad M, Khan SH (2012) (PDF) Association Mapping: A Step Forward to Discovering New Alleles for Crop Improvement. Int J Agric Biol 14:153–160
3. Alam M, Neal J, O'Connor K, et al (2018) Ultra-high-throughput DArTseq-based silicoDArT and SNP markers for genomic studies in macadamia. PLoS One 13:e0203465. https://doi.org/10.1371/JOURNAL.PONE.0203465
4. Alemu A, Feyissa T, Maccaferri M, et al (2021) Genome-wide association analysis unveils novel QTLs for seminal root system architecture traits in Ethiopian durum wheat. BMC Genomics 22:1–16. https://doi.org/10.1186/S12864-020-07320-4/FIGURES/5
5. Ávila CM, Dolores Requena-Ramírez M, Rodríguez-Suárez C, et al (2021) Genome-wide association analysis for stem cross section properties, height and heading date in a collection of spanish durum wheat landraces. Plants 10:1–13. https://doi.org/https://doi.org/10.3390/plants10061123
6. Baloch FS, Alsaleh A, Andeden EE, et al (2016) High levels of segregation distortion in the molecular linkage map of bread wheat representing the West Asia and North Africa region. Turkish J Agric For Pagination 352-364 40:352–364. https://doi.org/10.3906/TAR-1508-27
7. Baloch FS, Alsaleh A, Shahid MQ, et al (2017) A whole genome DArTseq and SNP analysis for genetic diversity assessment in durum wheat from central fertile crescent. PLoS One 12:. https://doi.org/10.1371/journal.pone.0167821
8. Batley J, Edwards D (2007) SNP Applications in Plants. In: Association Mapping in Plants. Springer, New York, NY, pp 95–102
9. Beres BL, Rahmani E, Clarke JM, et al (2020) A Systematic Review of Durum Wheat: Enhancing Production Systems by Exploring Genotype, Environment, and Management (G × E × M) Synergies. Front Plant Sci 11:1665. https://doi.org/10.3389/FPLS.2020.568657/BIBTEX
10. Breseghello F, Sorrells ME (2006) Association Mapping of Kernel Size and Milling Quality in Wheat (Triticum aestivum L.) Cultivars. Genetics 172:1165. https://doi.org/10.1534/GENETICS.105.044586
11. Briñez B, Blair MW, Kilian A, et al (2012) A whole genome DArT assay to assess germplasm collection diversity in common beans. Mol Breed 30:181–193. https://doi.org/10.1007/S11032-011-9609-3
12. Bukhari T, Rana RM, Hassan MU, et al (2022) Genetic diversity and marker trait association for phytophthora resistance in chilli. Mol Biol Rep 49:5717–5728. https://doi.org/10.1007/S11033-022-07635-3
13. Chao S, Dubcovsky J, Dvorak J, et al (2010) Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). BMC Genomics 11:1–17. https://doi.org/10.1186/1471-2164-11-727/FIGURES/6
14. CRP-WHEAT (2016) "Wheat Agri-Food Systems Proposal 2017-2022"Research Program on Wheat, (CGIAR).
15. Dastfal M, Aghaee Sarbarzeh M, Zali H (2022) Genetic diversity and selection of durum wheat pure lines with desirable agronomy traits using SIIG index. Iran J F Crop Sci 53:161–174
16. Dehghani H, Omidi H, Sabaghnia N (2008) Graphic Analysis of Trait Relations of Rapeseed Using the Biplot Method. Agron J 100:1443–1449. https://doi.org/10.2134/AGRONJ2007.0275
17. Deschamps S, Campbell MA (2010) Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. Mol Breed 25:553–570. https://doi.org/10.1007/S11032-009-9357-9
18. Doyle J., Doyle J. (1990) "Isolation of Plant DNA from Fresh Tissue,." Focus (Madison) 12:13–15

19. Etminan A, Pour-Aboughadareh A, Mohammadi R, et al (2018) Applicability of CAAT box-derived polymorphism (CBDP) markers for analysis of genetic diversity in durum wheat. Cereal Res Commun 46:1–9. https://doi.org/10.1556/0806.45.2017.054
20. Etminan A, Pour-Aboughadareh A, Mohammadi R, et al (2016) Applicability of start codon targeted (SCoT) and inter-simple sequence repeat (ISSR) markers for genetic diversity analysis in durum wheat genotypes Applicability of start codon targeted (SCoT) and inter-simple sequence repeat (ISSR) markers for genetic d. Biotechnol Biotechnol Equip 30:1075–1081. https://doi.org/10.1080/13102818.2016.1228478
21. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/J.1365-294X.2005.02553.X
22. Fayaz F, Aghaee Sarbarzeh M, Talebi R, Azadi A (2019) Genetic Diversity and Molecular Characterization of Iranian Durum Wheat Landraces (Triticum turgidum durum (Desf.) Husn.) Using DArT Markers. Biochem Genet 57:98–116. https://doi.org/10.1007/S10528-018-9877-2
23. Flint-Garcia SA, Thornsberry JM, Edward IV SB (2003) Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54:357–374. https://doi.org/10.1146/ANNUREV.ARPLANT.54.031902.134907
24. Fufa TW, Abtew WG, Amadi CO, Oselebe HO (2022) DArTSeq SNP-based genetic diversity and population structure studies among taro [(Colocasia esculenta (L.) Schott] accessions sourced from Nigeria and Vanuatu. PLoS One 17:e0269302. https://doi.org/10.1371/JOURNAL.PONE.0269302
25. Gibson G, Muse S. (2009) A Primer of Genome Science - Greg Gibson; Spencer V. Muse - Oxford University Press. Oxford University Press
26. Giunta F, Pruneddu G, Motzo R (2019) Grain yield and grain protein of old and modern durum wheat cultivars grown under different cropping systems. F Crop Res 230:107–120. https://doi.org/10.1016/J.FCR.2018.10.012
27. Grzebelus D, Iorizzo M, Senalik D, et al (2014) Diversity, genetic mapping, and signatures of domestication in the carrot (Daucus carota L.) genome, as revealed by Diversity Arrays Technology (DArT) markers. Mol Breed 33:625–637. https://doi.org/10.1007/S11032-013-9979-9/FIGURES/3
28. Gupta PK, Varshney RK (2005) Cereal genomics: an overview. In: Cereal Genomics. Springer Netherlands, pp 1–18
29. Hao C, Wang L, Ge H, et al (2011) Genetic Diversity and Linkage Disequilibrium in Chinese Bread Wheat (Triticum aestivum L.) Revealed by SSR Markers. PLoS One 6:e17279. https://doi.org/10.1371/JOURNAL.PONE.0017279
30. Hao CY, Perretant MR, Choulet F, et al (2010) Genetic diversity and linkage disequilibrium studies on a 3.1-Mb genomic region of chromosome 3B in European and Asian bread wheat (Triticum aestivum L.) populations. Theor Appl Genet 121:1209–1225. https://doi.org/10.1007/S00122-010-1382-X
31. Horvath A, Didier A, Koenig J, et al (2009) Analysis of diversity and linkage disequilibrium along chromosome 3B of bread wheat (Triticum aestivum L.). Theor Appl Genet 119:1523–1537. https://doi.org/10.1007/S00122-009-1153-8
32. Hu X, Ren J, Ren X, et al (2015) Association of Agronomic Traits with SNP Markers in Durum Wheat (Triticum turgidum L. durum (Desf.)). PLoS One 10:e0130854. https://doi.org/10.1371/JOURNAL.PONE.0130854
33. Laidò G, Mangini G, Taranto F, et al (2013) Genetic Diversity and Population Structure of Tetraploid Wheats (Triticum turgidum L.) Estimated by SSR, DArT and Pedigree Data. PLoS One 8:e67280. https://doi.org/10.1371/JOURNAL.PONE.0067280
34. Li YC, Korol AB, Fahima T, et al (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol 11:2453–2465. https://doi.org/10.1046/J.1365-294X.2002.01643.X
35. Mantovani P, Maccaferri M, Sanguineti MC, et al (2008) An integrated DArT-SSR linkage map of durum wheat. Mol Breed 22:629–648. https://doi.org/10.1007/S11032-008-9205-3
36. Maulana F, Ayalew H, Anderson JD, et al (2018) Genome-wide association mapping of seedling heat tolerance in winter wheat. Front Plant Sci 9:1272. https://doi.org/10.3389/FPLS.2018.01272/BIBTEX
37. Mehrabi A., Pour-Aboughadareh A., Mansouri S, Hosseini A. (2020) Genome-wide association analysis of root system architecture features and agronomic traits in durum wheat | Enhanced Reader. Mol Breed 40–55
38. Mérida-García R, Bentley AR, Gálvez S, et al (2020) Mapping Agronomic and Quality Traits in Elite Durum Wheat Lines under Differing Water Regimes. Agron 2020, Vol 10, Page 144 10:144. https://doi.org/10.3390/AGRONOMY10010144
39. Mohammadi R, Dehghani H, Karimzadeh G (2014) Graphic analysis of trait relations of cantaloupe using the Biplot method . J Plant Prod Res 21:
40. Mohammadi SA, Prasanna BM (2003) Analysis of Genetic Diversity in Crop Plants—Salient Statistical Tools and Considerations. Crop Sci 43:1235–1248. https://doi.org/10.2135/CROPSCI2003.1235
41. Moragues M, Moralejo M, Sorrells ME, Royo C (2007) Dispersal of durum wheat [Triticum turgidum L. ssp. turgidum convar. durum (Desf.) MacKey] landraces across the Mediterranean basin assessed by AFLPs and microsatellites. Genet Resour Crop Evol 54:1133–1144. https://doi.org/10.1007/S10722-006-9005-8
42. Mwadzingeni L, Shimelis H, Rees DJG, Tsilo TJ (2017) Genome-wide association analysis of agronomic traits in wheat under drought-stressed and non-stressed conditions. PLoS One 12:e0171692. https://doi.org/10.1371/JOURNAL.PONE.0171692
43. Nantongo JS, Odoi JB, Agaba H, Gwali S (2022) SilicoDArT and SNP markers for genetic diversity and population structure analysis of Trema orientalis; a fodder species. PLoS One 17:e0267464. https://doi.org/10.1371/JOURNAL.PONE.0267464
44. Ndjiondjop MN, Semagn K, Gouda AC, et al (2017) Genetic variation and population structure of Oryza glaberrima and development of a mini-core collection using DArTseq. Front Plant Sci 8:1748. https://doi.org/10.3389/FPLS.2017.01748/BIBTEX
45. NGBT NGBT descriptors for durum wheat 1.0. https://www.genesys-pgr.org/descriptorlists/b6544d11-be40-4638-be35-1b3f5349c5c9. Accessed 26 Mar 2023
46. O'Connor K, Kilian A, Hayes B, et al (2019) Population structure, genetic diversity and linkage disequilibrium in a macadamia breeding population using SNP and silicoDArT markers. Tree Genet Genomes 15:. https://doi.org/10.1007/S11295-019-1331-Z

47. Ouaja M, Bahri BA, Aouini L, et al (2021) Morphological characterization and genetic diversity analysis of Tunisian durum wheat (Triticum turgidum var. durum) accessions. BMC Genomic Data 22:1–17. https://doi.org/10.1186/S12863-021-00958-3/FIGURES/4
48. Pour-Aboughadareh A, Mahmoudi M, Moghaddam M, et al (2017) Agro-morphological and molecular variability in Triticum boeoticum accessions from Zagros Mountains, Iran. Genet Resour Crop Evol 64:545–556. https://doi.org/10.1007/S10722-016-0381-4
49. Powell W, Morgante M, Andre C, et al (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2:225–238. https://doi.org/10.1007/BF00564200/METRICS
50. Pozniak CJ, Clarke JM, Clarke FR (2012) Potential for detection of marker-trait associations in durum wheat using unbalanced, historical phenotypic datasets. Mol Breed 30:1537–1550. https://doi.org/10.1007/S11032-012-9737-4
51. Rahmati M, Hoseinpour T, Ahmadi A (2020) Assessment of interrelationships between agronomic traits of wheat genotypes under rain-fed conditions using double and triple biplots of genotype, trait and yield. Iran J Dryl Agric 9:1–20
52. Ranieri R (2015) Geography of the durum wheat crop -. In: Pastaria Open Fields. Collecchio,Italy
53. Reich DE, Lander ES (2001) On the allelic spectrum of human disease. Trends Genet 17:502–510. https://doi.org/10.1016/S0168-9525(01)02410-6
54. Ren J, Sun D, Chen L, et al (2013) Genetic Diversity Revealed by Single Nucleotide Polymorphism Markers in a Worldwide Germplasm Collection of Durum Wheat. Int J Mol Sci 14:7061. https://doi.org/10.3390/IJMS14047061
55. Salgotra RK, Gupta BB, Bhat JA, Sharma S (2015) Genetic Diversity and Population Structure of Basmati Rice (Oryza sativa L.) Germplasm Collected from North Western Himalayas Using Trait Linked SSR Markers. PLoS One 10:e0131858. https://doi.org/10.1371/JOURNAL.PONE.0131858
56. Sansaloni C, Franco J, Santos B, et al (2020) Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. Nat Commun 2020 111 11:1–12. https://doi.org/10.1038/s41467-020-18404-w
57. Saura M, Fernandez A, Rodriguez MC, et al (2013) Genome-Wide Estimates of Coancestry and Inbreeding in a Closed Herd of Ancient Iberian Pigs. PLoS One 8:e78314. https://doi.org/10.1371/JOURNAL.PONE.0078314
58. Schlötterer C (2003) Hitchhiking mapping–functional genomics from the population genetics perspective. Trends Genet 19:32–38. https://doi.org/10.1016/S0168-9525(02)00012-4
59. Shahid Masood M, Javaid A, Ashiq M, et al (2005) PHENOTYPIC DIVERSITY AND TRAIT ASSOCIATION IN BREAD WHEAT (TRITICUM AESTIVUM L.) LANDRACES FROM BALUCHISTAN, PAKISTAN. Pak J Bot 37:949–957
60. Sodedji FAK, Ryu D, Choi J, et al (2022) Genetic Diversity and Association Analysis for Carotenoid Content among Sprouts of Cowpea (Vigna unguiculata L. Walp). Int J Mol Sci 23:3696. https://doi.org/10.3390/IJMS23073696/S1
61. 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. Front Plant Sci 9:81. https://doi.org/10.3389/FPLS.2018.00081/BIBTEX
62. Trebbi D, Maccaferri M, de Heer P, et al (2011) High-throughput SNP discovery and genotyping in durum wheat (Triticum durum Desf.). Theor Appl Genet 123:555–569. https://doi.org/10.1007/S00122-011-1607-7
63. Wang S, Xu S, Chao S, et al (2019) A genome-wide association study of highly heritable agronomic traits in durum wheat. Front Plant Sci 10:919. https://doi.org/10.3389/FPLS.2019.00919/BIBTEX
64. Yan W (2001) GGEbiplot—A Windows Application for Graphical Analysis of Multienvironment Trial Data and Other Types of Two-Way Data. Agron J 93:1111–1118. https://doi.org/10.2134/AGRONJ2001.9351111X
65. Yan W, Frégeau-Reid J (2008) Breeding Line Selection Based on Multiple Traits. Crop Sci 48:417–423. https://doi.org/10.2135/CROPSCI2007.05.0254
66. Yan W, Kang MS (2002) GGE Biplot Analysis : A Graphical Tool for Breeders, Geneticists, and Agronomists. GGE Biplot Anal. https://doi.org/10.1201/9781420040371
67. Yan W, Rajcan I (2002) Biplot Analysis of Test Sites and Trait Relations of Soybean in Ontario. Crop Sci 42:11–20. https://doi.org/10.2135/CROPSCI2002.1100
68. Zeven AC (2000) Traditional maintenance breeding of landraces: 1. Data by crop. Euphytica 116:65–85. https://doi.org/10.1023/A:1004089816030/METRICS
69. Zhu C, Gore M, Buckler ES, Yu J (2008) Status and Prospects of Association Mapping in Plants. Plant Genome 1:. https://doi.org/10.3835/PLANTGENOME2008.02.0089

Figures

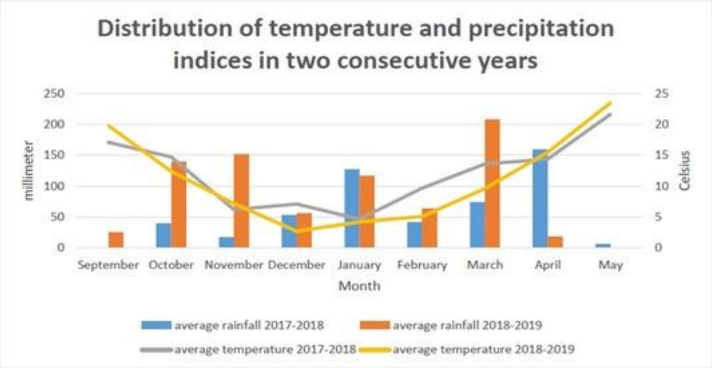


Figure 1

Distribution of temperature and precipitation indices in two consecutive years

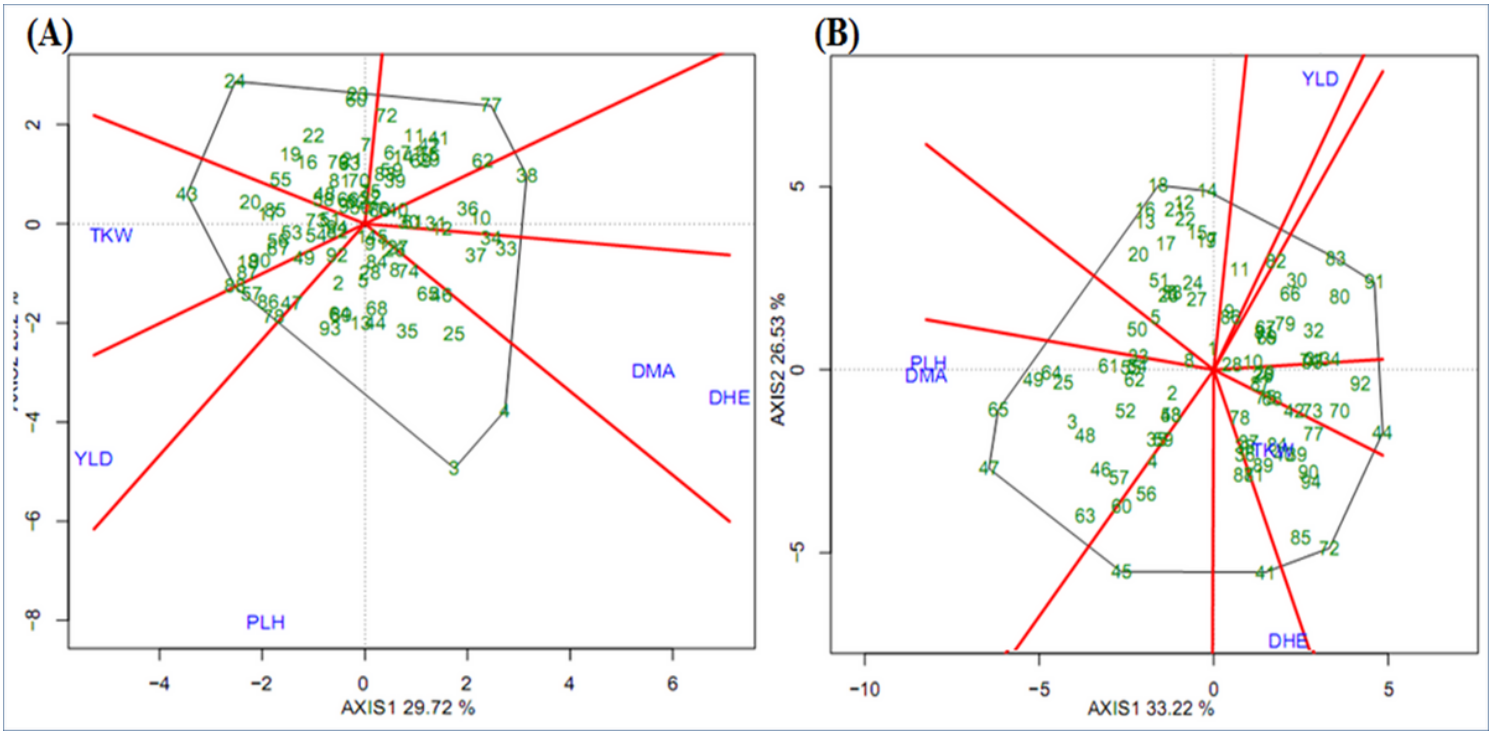


Figure 2

Polygon bi-plot drawn based on the first two components for the measured traits in the first (A) and second (B) years of the experiment. DHE, DM, PLH, TKW and YLD represent the number of days until the emergence of the spike, the number of days to the physiological maturity, the plant height, the thousand kernel weight and the grain yield, respectively.

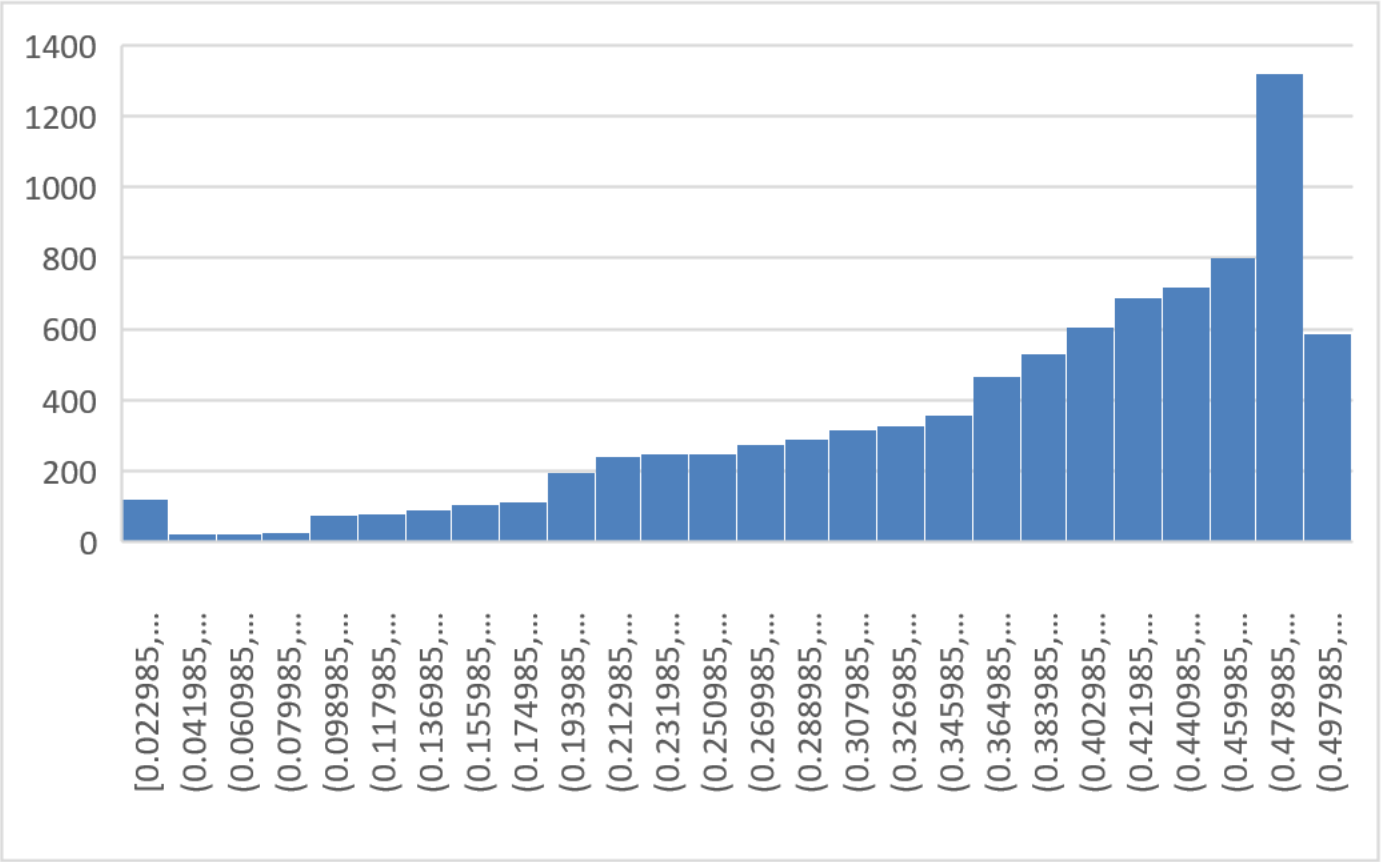


Figure 3

Frequency distribution of PIC values (the X axis) of SilicoDArT markers (the Y axis).

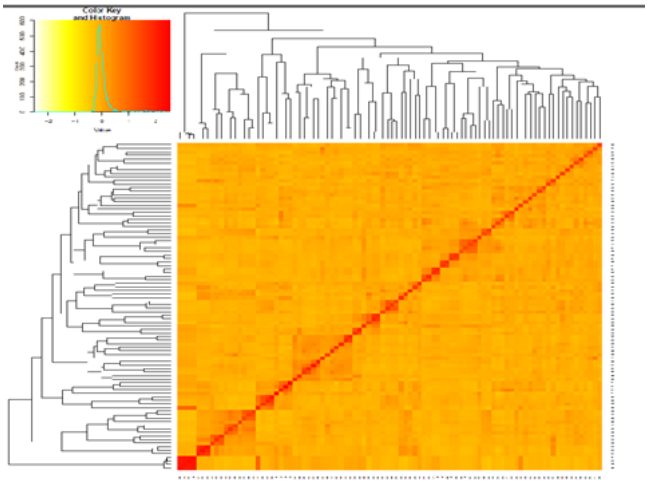


Figure 4

Grouping of genotypes based on relative kinship coefficients between pairs of genotypes

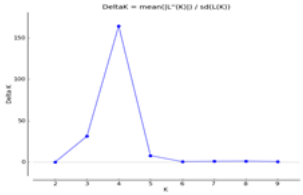


Figure 5

Bivariate diagrams to determine the appropriate number of subpopulations in the studied genotypes (K=4).

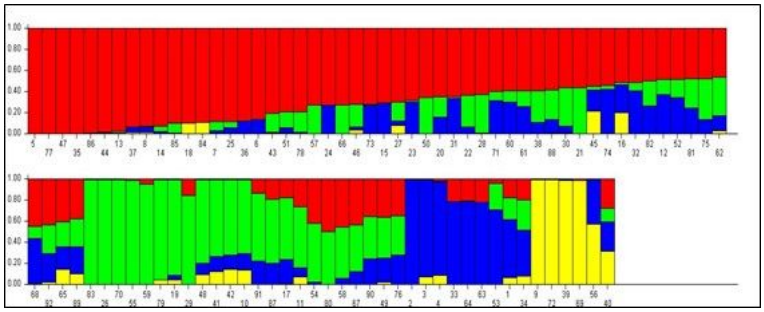


Figure 6

Distribution of the 92 durum wheat genotypes using SilicoDArT markers in subpopulations according to the structure analysis (k = 4). The individuals were represented in vertical bars, each color associated with a different group.

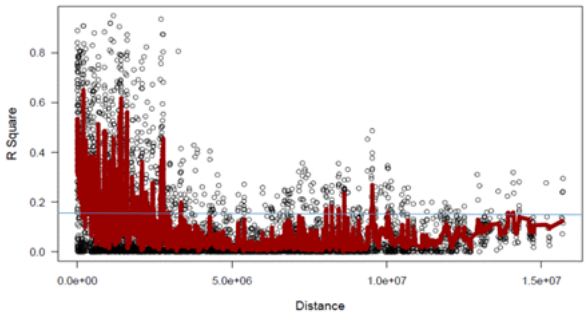


Figure 7

Reduction of linkage disequilibrium at the genome level

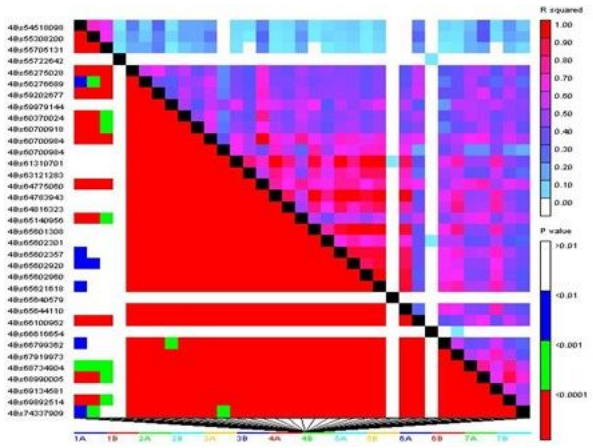


Figure 8

Distribution of some DArTseq markers on chromosome 4B

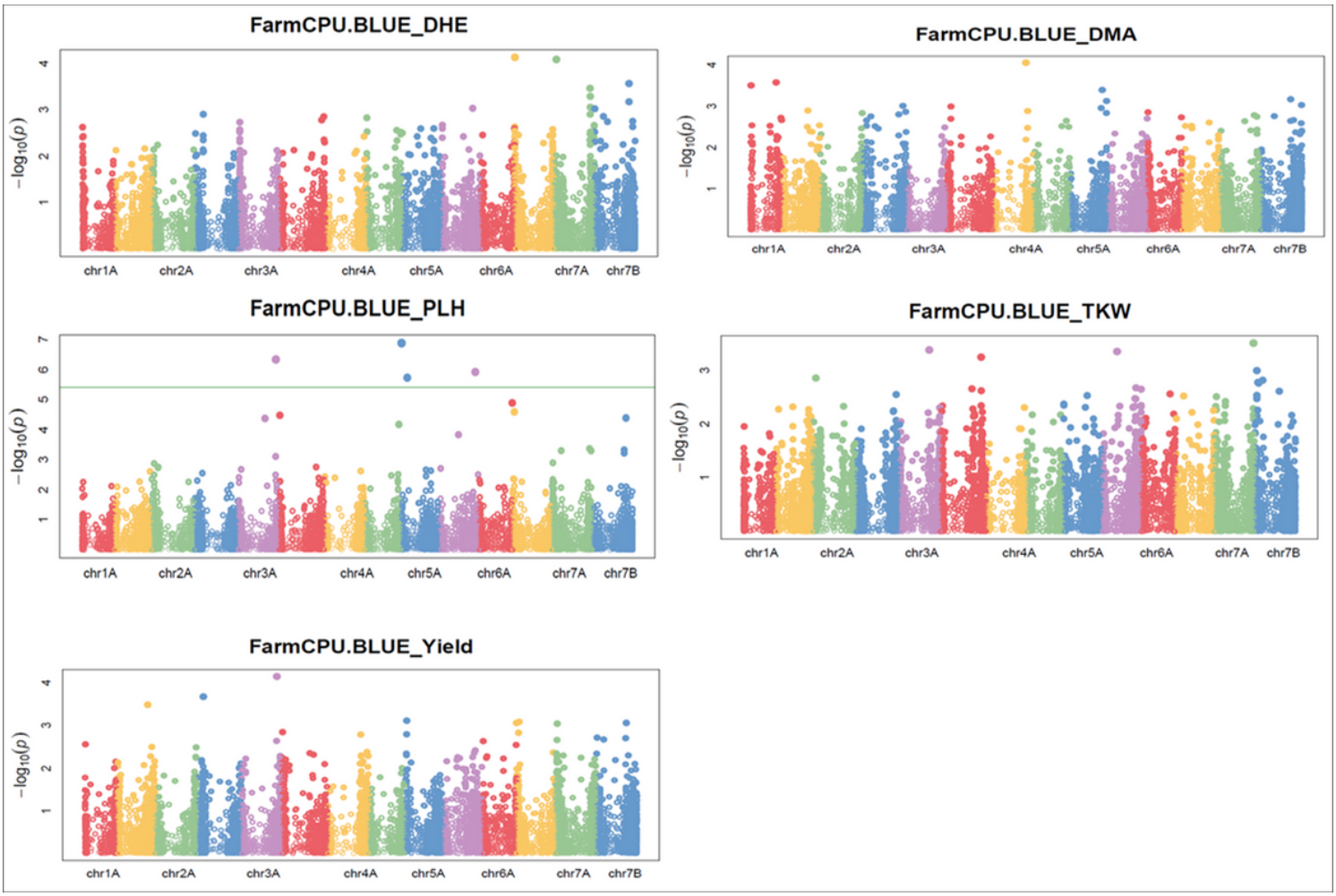


Figure 9

Manhattan plot resulting from linkage location along the length of durum wheat chromosomes in the crop year 2016 for the traits of day to flowering (DHE), day to maturity (DMA), plant height (PLH), thousand grain weight (TKW) and grain yield

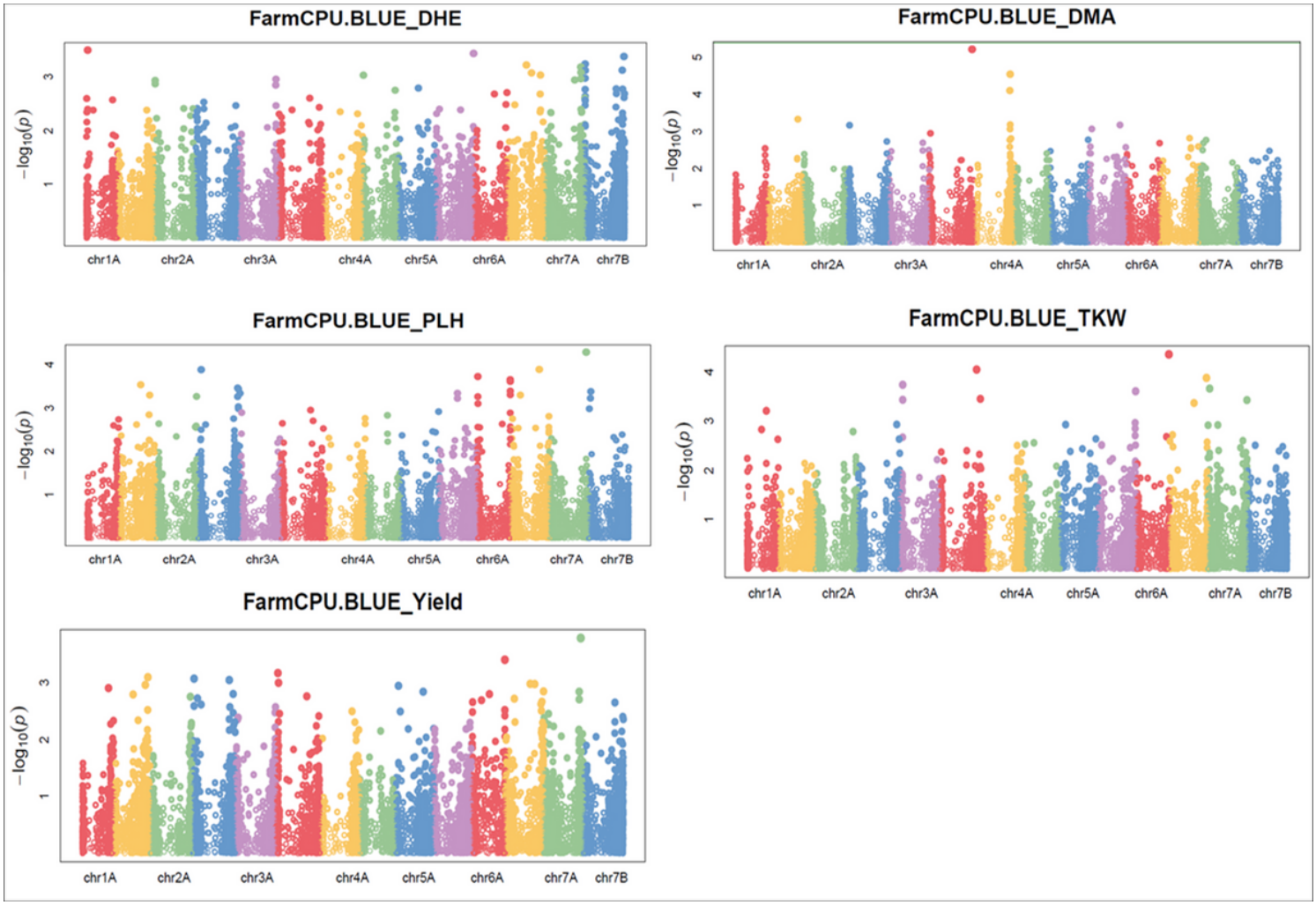


Figure 10

Manhattan plot resulting from linkage location along the length of durum wheat chromosomes in 2016 crop year for the traits of day to flowering (DHE), day to maturity (DMA), plant height (PLH), thousand grain weight (TKW) and grain yield