Journal of Biosafety
Volume 14, Number 4, Winter, 2022
pISSN 2717-0632, eISSN 2716-9804

Assessment of Genetic Diversity and Population Structure among Barley Breeding Lines and Landraces Using SNP Markers

20.1001.1.27170632.1400.14.4.6.9

Farhad Ahakpaz[1], Iraj Bernosi[*2], Babak Abdollahi Mandoulakani[3], Saber Golkari[4], Jafar Jafarzadeh[5], Sripada Udupa[6]

1. PhD Student and Scientific Member, Dryland Agricultural Research Institute (DARI), Agriculture Research, Education and Extension Organization (AREEO), Maragheh, Iran.

2. Associate professor, Department of Plant Production and Genetics, Uromieh University, Uromieh, West Azerbaijan, Iran.

3. Professor, Department of Plant Production and Genetics, Uromieh University, Uromieh, West Azerbaijan, Iran.

4. Associate Professor, Agriculture Biotechnology Research Institute of Iran (ABRII), Karaj, Iran.

5. Assistant Professor, Dryland Agricultural Research Institute (DARI), Agriculture Research, Education and Extension Organization (AREEO), Maragheh, Iran.

6. Senior Scientist, International Center for Agricultural Research in the Dry Areas (ICARDA), Rabat, Morocco.

i.bernosi@urmia.ac.ir

Resived Date: 2022/01/16, Accepted Date: 2022/01/31
Page: 113-134

Abstract

Genetic diversity is critical for developing broadly adapted cultivars with desirable traits. This research aimed to investigate the genetic diversity, population structure, and genome-wide linkage disequilibrium (LD) of 108 barley genotypes, including rainfed and irrigated breeding lines, cultivars, and landraces with the 50K IlluminaTM iSelect single nucleotide polymorphism (SNP) array. After quality control, 14,943 polymorphic SNPs were subjected to estimate Shannon's information index (*I*), Nei's gene diversity (*H*), polymorphic information content (PIC), fixation index (*Fst*), and principal component analysis (PCA). The *I* and *H* indices were 0.74 and 0.336, respectively. The PIC values were 0.367 and 0.178 for six and two-rowed barley, respectively. PCA using Nei's genetic distance identified three subpopulations. Subpopulations 1 and 2 had 0.38, and 0.29 PIC values, respectively, and were more diverse than subpopulation 3 (0.09). The Fst value was 0.381. Subpopulations 1 and 3 indicated the highest (0.379) and the lowest (0.040) genetic diversity within subpopulations, respectively. The average LD for two-rowed genotypes and subpopulation 3 was more than that for six-roweded genotypes. A high level of genetic variation and genetic differentiation among subpopulations was observed in this panel. The irrigated six-rowed types and landraces indicated higher genetic diversity, whereas rainfed two-rowed barley indicated the highest LD and the lowest genetic variation due to high selection intensity. This panel could be used for genome-wide association studies to identify the trait-marker associations in future genetic improvement programs for barley.

Keywords: Barley Genetic Resource, Population Structure, Genome-Wide Linkage Disequilibrium, Polymorphic Information Content, Single Nucleotide Polymorphism.

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

Introduction

Barley is one of the most important cereal crops broadly used for human consumption, animal feeding, and malt production for beer brewing worldwide, after wheat, rice, and maize (Ferreira et al. 2016). Barley is a low-demand crop planted in many areas due to its ability to grow in marginal and low-input lands, well adaptation to drought and salinity (Ferreira et al. 2016). Barley's planting area and production rates were 52 million hectares and 158 million tons worldwide, while it was 3.6 million hectares and 1.7 million tons, respectively, in Iran (FAO. 2020).

Diploid barley includes wild and cultivated types, while all the tetraploid and hexaploid barley are wild types (Blattner. 2018). Cultivated barley has winter, facultative, and spring growth habits, row type of spike (two and six), awn and awnless, hulled, and naked grain (Paulitz and Steffenson, 2010). Barley genome has 5.1 gigabases (Gb) with 80% repetitive sequences (Schulte et al. 2009; Wicker et al. 2008). In spite, barley was widely used for different genetic studies due to its few numbers of chromosomes, easy hybridization, high rate of self-pollination, a close relative of barley genome with rice and wheat, and short growth life (Hill et al. 2021; Sreenivasulu et al. 2008).

Long-term domestication and cultivar development decreased genetic variation in cereal crops such as barley and wheat. Although breeding programs in different crops lead to the development of high-yielding and quality cultivars, their genetic diversity is less than that in landraces and wild relatives related to other agronomic traits (Al-Abdallat et al. 2017; Pasam et al. 2014). Effect of long-term breeding on genetic variation of Minnesota's barley (from 1958 to 1998) using 71 simple sequence repeat (SSR) markers revealed that genetic diversity of breeding lines decreased extensively so that the average allelic variation per SSR locus reduced from 5.89 for parents to 2.39 for advanced lines . A reduction in genetic diversity of breeding germplasms could slow or reduce future genetic gains (Condón et al. 2009). Genetic bottlenecks cause that for more agronomic traits, barley elite lines and cultivars had less genetic variation than their wild relatives and early domesticates. Narrow genetic diversity in breeding populations leads to develop more homogeneous cultivars, which could be susceptible to different diseases, pathogens, and harsh environmental conditions.

Furthermore, allele frequency has been changed due to genetic bottlenecks through domestication and cultivar development, resulting in different levels of genetic variation and linkage disequilibrium (LD). Hence, LD levels increased from wild relatives to landraces and developed cultivars, while the genetic variation levels decreased from wild relatives to developed cultivars (Pasam et al. 2014). Thus, exploiting the genetic variation of wild relatives, landraces, accessions, and local varieties is crucial to the genetic diversity of the barley gene pool and high-yielding varieties development with good resistance to biotic and abiotic stresses and better-adapted cultivars (Dotlačil et al. 2010; Pasam et al. 2014).

Single nucleotide polymorphisms (SNP) markers are a more common marker that covers the most frequent type of genetic polymorphism across the whole genome (Lai et al. 2015).

These markers were efficiently used in marker-assisted breeding, genome-wide association study, genetic diversity, and population structure analysis in most crops and barley (Kalinowski, 2002;

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

Smith et al. 2007; Varshney et al. 2012). Kanazin et al. (2002) reported that there was one SNP per 189 base pairs in barley. There were also common alleles in *Hordeum vulgar* L and *Hordeum spontaneum* L, which show gene exchanging between two species in the evolutionary process (Jehan and Lakhanpaul, 2006). A genome-wide association study (GWAS) was conducted to identify genomic regions controlling drought stress-tolerant traits on 185 cultivated barley (*H. vulgare* L.) and 38 wild barley (*H. spontaneum* L.) from 30 different countries using 816 molecular markers including SSR, Diversity Array Technology (DArT), and SNP (Varshney et al. 2012).

This study classified barley genotypes into five clusters: origin, domesticated, and wild type. Furthermore, the population structure of 224 spring barley was studied using 1536 SNP markers and revealed that morphology of spike (row type) and origin were the main reasons for population structure (Pasam et al. 2012).

The polymorphic information content (PIC), Wright's Fst statistic (*Fst*), and Nei's genetic distance indices have been used to study population structure and genetic diversity among and within populations using SNPs. The genetic diversity of 60 barley genotypes was studied using 40 markers (32 SSR and 8 SNP), of which 33 were polymorphic markers, and PIC values ranged from 0.06 to 0.77 with an average of 0.36 (Elakhdar et al. 2016). Also, they used principal component analysis (PCA) and cluster analysis to classify and visualize sub-populations of the barley panel, which Fst value was 0.235 between the two subpopulations. Genetic diversity, population structure, and LD was estimated for 180 spring barley breeding lines and cultivars using 48 SSR and 6208 polymorphic SNPs markers by PIC, Shannon's diversity index, Nei's genetic distance, and principal coordinates analysis (PCoA) (Bengtsson et al. 2017a).

The average PIC value was 0.46 and 0.28 for the SSR and the SNP markers, respectively. Furthermore, the genetic diversity index was 0.514 and 0.359 for the SSR and the SNP markers, respectively (Bengtsson et al. 2017a). Diversity in plant genetic resources is crucial for plant breeders to develop new broadly adapted cultivars with desirable characteristics. So, the objectives of this study were i: to determine genetic variation among a collection of barley genotypes, including landraces, rainfed and irrigated breeding lines, and cultivars using SNP markers. ii: to assess the population structure and compare the level of polymorphism among subpopulations according to row type (two and six), growth habit, rainfed and irrigated, landraces, and breeding lines/cultivars.

Materials and methods

Germplasm

The Dryland Agricultural Research Institute (DARI) is the main public breeding institute that works on rainfed wheat and barley to develop new high-yielding breeding lines and cultivars for all rainfed areas of Iran, including cold, moderate, and warm regions. In this regard, a collection of 108 barley genotypes was used for genotyping and assessment of genetic diversity. This panel included 49 rainfed barley varieties and elite breeding lines for cold and moderate cold areas from DARI, Iran, and ICARDA-DARI breeding program, 25 irrigated barley varieties, and advanced breeding lines from Seed and Plant Improvement Institute (SPII), Iran, and 34 landraces from England, Algeria, Spain, Egypt, India, Russia, China, Azerbaijan, and Iran (received from gene bank, Karaj, Iran) (Table 1). Barley genotypes had winter (W), spring (S), and facultative (F) growth habits with two-rowed and six-rowed types. All two-rowed types were rainfed barley adapted for cold and moderate cold rainfed areas, whereas 84% of six-rowed types were irrigated barley (Table 1).

Table 1. General information for origin, growth habit, row-type and the adapted environment of barley genotypes.

No.	Genotypes	Origin	Growth habit	Row type	Rf/Irr
1	Makouee (Star)	SPII	W	6	Irr
2	Bahman (WA 2196-68/NY6005-18, F1//Scotia I)	SPII	W	6	Irr
3	Bereke 54	SPII	F	6	Irr
4	Radical/Birgit//Pamir-154	SPII	F	6	Irr
5	Michailo/Dobrinya	SPII	F	6	Irr
6	Bahtim 7DL/79-W40762//Deir Alla106	SPII	F	6	Irr
7	Michailo/K-096M3	SPII	W	6	Irr
8	Pamir-168	SPII	W	6	Irr
9	Torsh/Legia	SPII	W	6	Irr
10	Pamir 013/Sonata	SPII	F	2	Irr
11	Ste/L.640//Hml-02/Arabi Abiad*2/3/1-BC-80593	SPII	F	6	Irr
12	Bereke-54/Alanda	SPII	F	6	Irr
13	L.1242/ZARJOW//LB.Iran/Una8271//Gloria"S"/Com"S	SPII	F	6	Irr
14	Makouee/C.C89//Rihane"s"/3/Roho/Mazurka	SPII	W	6	Irr
15	L.527/MB2367//(CI7117-9/DeirAlla106)/3/Beecher	SPII	W	6	Irr
16	Zarjow/CM67/4/Schuyler/3/M.Rnb86.80/NB2905/L.527 LPD 92	SPII	W	6	Irr
17	Schuyler/3/M.Rnb86.80/NB2905//L.527	SPII	W	6	Irr
18	Roho//Alger/Ceres362-1-1/3/CWB117-77-9-7/4/Alpha/Durra//Antares/K2y63	SPII	W	2	Irr
19	Bolgali/4/Roho//Alger/Ceres362-1-1/3/CWB117-77-9-7	SPII	F	2	Irr
20	Gara arpa	DARI-IRAN	SP	2	Rf
21	Sahand	ICARDA-DARI	F	2	Rf
22	Abidar	ICARDA-DARI	F	2	Rf
23	Ansar	ICARDA-DARI	F	2	Rf
24	Kuban-06	ICARDA-DARI	F	2	Rf
25	PAMIR-158/ZDM1454 (Artan)	ICARDA-DARI	F	2	Rf
26	TOKAK//STEPTO/ANTARES (Qaflan)	ICARDA-DARI	F	2	Rf
27	Uzno-Kazakastan/3/CWB117-77-9-7//Alpha/Durra	ICARDA-DARI	F	2	Rf
28	Orza-96/4/Tokak/3/CWB117-77-9-7//Alpha/Durra (Arda)	ICARDA-DARI	F	2	Rf
29	Roho/Masurka//ICB-103020/3/Alpha/Durra//Slr	ICARDA-DARI	F	2	Rf
30	ChiCm/An57//Albert/3/ICB-102379/4/GkOmega/5/Tokak	ICARDA-DARI	F	2	Rf
31	Ste/Antares//YEA762-2/YEA605-5/3/Slr//Alpha/Durra	ICARDA-DARI	F	2	Rf
32	Sahand/C-25041	ICARDA-DARI	F	2	Rf
33	ARM-ICB-123199/3/Zarjau/80-5151//Skorohod	ICARDA-DARI	F	2	Rf
34	Yea168.4/Ywa605.5//Radical	ICARDA-DARI	F	2	Rf

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

35	Unknown-F6-88-9	ICARDA-DARI	F	2	Rf
36	Sahand/Radical	ICARDA-DARI	F	2	Rf
37	Schuyler//Alpha/Durra/3/Radical	ICARDA-DARI	F	2	Rf
38	Uzno-Kazakestan/3/CWB117-5-9-5//YEA389-3/YEA475-4	ICARDA-DARI	F	2	Rf
39	Fasih	ICARDA-DARI	SP	2	Rf
40	Taram	ICARDA-DARI	F	2	Rf
41	Reyhan-03	SPII	SP	6	Irr
42	Bulbul	ICARDA-DARI (Turkey)	F	2	Rf
43	Dubrinia	SPII	F	6	Irr
44	Dictoo	SPII	F	6	Irr
45	Dayton/Runney	ICARDA-DARI	F	2	Rf
46	Sararood-1	ICARDA-DARI	F	2	Rf
47	Nader	ICARDA-DARI	F	2	Rf
48	Radical	SPII	W	6	Irr
49	Cumra 2001	ICARDA-DARI (Turkey)	SP	2	Rf
50	Efes98	ICARDA-DARI (Turkey)	SP	2	Rf
51	Cumhariyet50	ICARDA-DARI (Turkey)	SP	2	Rf
52	Catalhuyuk2001	ICARDA-DARI (Turkey)	SP	2	Rf
53	Keser	ICARDA-DARI (Turkey)	SP	2	Rf
54	Sahand / Obruk-86	ICARDA-DARI	F	2	Rf
55	Antares/Ky36-1294//Slrlcbh-0383 /3/ Sahand	ICARDA-DARI	F	2	Rf
56	Yea168.4/Ywa605.5 Yea206-4A-3 // Dictoo	ICARDA-DARI	F	2	Rf
57	Yea168.4/Ywa605.5 Yea206-4A-3 // Denmark	ICARDA-DARI	F	2	Rf
58	ORZA96	ICARDA-DARI	F	2	Rf
59	R018 UK	ICARDA-DARI	F	2	Rf
60	PI-549081 -	ICARDA-DARI	W	6	Irr
61	PI-560331 -	ICARDA-DARI	W	6	Irr
62	C-25041//Yea168.4/Ywa605.5 Yea206-4A-3	ICARDA-DARI	F	2	Rf
63	Dayton/Ranney/4/K-88 M1/3/Rhn-03/lignee 640//ICB-107766	ICARDA-DARI	F	2	Rf
64	Zarjau/80-5151//DZ-40-66/3/Alanda	SPII	F	6	Irr
65	Hispanic/Sararood	SPII	W	2	Irr
66	Sahand/3/Alpha/Gumhuriyet//Sonja	ICARDA-DARI	F	2	Rf
67	Abidar/4/K-88 M1/3/Rhn-03/lignee 640//ICB-107766	ICARDA-DARI	F	2	Rf
68	Icb-100059/3/Tipper/ICB-102854//Alpha/Durra	ICARDA-DARI	F	2	Rf
69	71411	England	F	2	Rf
70	71411	England	W	6	Irr-Rf
71	71426	Algeria	W	2	Irr
72	71538	Spain	W	6	Irr
73	71576	Egypt	F	2	Rf
74	71608	Egypt	W	6	Irr
75	71657	Egypt	W	6	Irr-Rf
76	71663	India	W	6	Irr-Rf
77	71850	Russia	F	2	Rf
78	72113	China	W	6	Irr-Rf
79	72295	China	W	6	Irr-Rf
80	72322	China	W	6	Irr-Rf
81	72368	China	W	6	Irr-Rf

82	72439	China	W	2	Irr
83	72466	Iran	W	2	Irr
84	72488	Iran	W	6	Irr-Rf
85	72494	Iran	W	6	Irr-Rf
86	72522	Iran	W	6	Irr
87	72524	Iran	W	6	Irr-Rf
88	72550	USA	F	2	Rf
89	72557	Azerbaijan	W	6	Irr
90	72562	Iran	W	6	Irr
91	72602	Iran	W	6	Irr-Rf
92	72646	Iran	W	6	Irr-Rf
93	72647	Iran	F	2	Rf
94	72653	Iran	W	6	Irr
95	72703	Iran	W	6	Irr-Rf
96	72726	Iran	W	6	Irr-Rf
97	72472	Iran	W	6	Irr
98	72482	Iran	W	6	Irr-Rf
99	72553	Azerbaijan	W	6	Irr-Rf
100	72704	Iran	W	6	Irr
101	72744	Iran	W	6	Irr-Rf
102	72747	Iran	F	2	Irr
103	Obruk-86	ICARDA-DARI (Turkey)	F	2	Rf
104	Star/4/M25-84/Attiki//Cr366-15-2/NopalS/3/Skorohod	ICARDA-DARI	F	6	Irr
105	Viringa'S'/Radical//Mattina	ICARDA-DARI	W	6	Irr
106	Kozir/Regina	ICARDA-DARI	W	2	Irr
107	Alanda/Regina	ICARDA-DARI	F	6	Irr
108	Pamir-147/Sonata/8/Alpha/Durra/7/P101/5/3896/1-15/3/3896/28//584/28/4/5050/6/Tipper	ICARDA-DARI	W	2	Irr

W: winter, S: spring, F: facultative, Rf: rainfed, and Irr: irrigated.

SNP genotyping

Genomic DNA for barley genotypes was extracted from young leaf seedlings collected from a single plant for each line using the modified CTAB (cetyltrimethylammonium bromide) method (Saghai-Maroof et al. 1984), digested with the restriction enzymes PstI and MspI (Poland et al. 2012), and quantified using Picogreen. The genotyping of samples was performed using 50K IlluminaTM iSelect SNP array (IPK-Gatersleben, Germany) as described in detail in Bayer et al., (2017). A total of 36,864 SNP markers were used for genotyping of 108 barley genotypes. However, genotype number 47 (Nader variety (Table 1)) had an issue in genotyping process and was discarded, so all genetic analysis was done for the final number of 107 genotypes. Marker data were filtered for missing data (<20%), heterozygosity rate 1%, redundant SNP markers, and minor allele frequency (MAF) (<5%) for a final number of 14,943 SNP markers selected for genome diversity analysis. The filtering process was done using TASSEL 5 (Bradbury et al. 2007) and in-house script R statistical software (v3.5.2; R Core Team, 2021).

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

Genetic diversity and population structure analysis

The following genetic diversity parameters were computed using the BIO-R program (Pacheco et al. 2016): the number of effective alleles per locus (Ne); Ne=1/1-He, where He is expected heterozygosity; $He = 1 - \sum_{i=1}^{k} Pi^2$ where p_i is the frequency of the ith allele in locus, Shannon's information index (I); $I = -\sum_{i=1}^{k} p_i ln p_i$ where *ln* is the natural logarithm of p_i (Tomar et al. 2021). Polymorphic information content (PIC) value that shows the amount of polymorphism within a population was estimated for each SNP according to (Botstein et al. 1980) for assessment of genetic diversity of barley germplasms using in-house script R as follows:

$$PIC = 1 - \sum_{i=1}^{k} P_i^2 - \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} 2P_i^2 P_j^2$$

Where p_i is the frequency of the i^{th} allele, p_i is the frequency of j^{th}, and k is the number of alleles per marker (Botstein et al. 1980; Kumar et al. 2020).

To account for the population structure, PCA method was performed using 14,943 SNPs and the first three principal components were used for scattering distribution to identify subpopulations in the R environment (R Core Team, 2021). Also, the multidimensional scaling (MDS) analysis is applied to compute the pairwise genetic distance matrix between the barley genotypes following Nei et al. (1983) index using the BIO-R v3.0 (Pacheco et al. 2016). The genetic distance matrix was used to construct a dendrogram using the Ward's method to identify the subpopulation structure of barley genotypes based on the whole genome marker information of 14,943 SNPs in BIO-R v3.0 (Pacheco et al. 2016).

The fixation index (*Fst*) which indicates the variance of allele frequencies among populations (genetic differentiation) was estimated based on Nei's genetic distance using BIO-R (Pacheco et al. 2016) as follows:

$$Fst = \frac{D_{ST}}{H_e} \qquad D_{ST} = H_e - H_s$$

Where D_{ST} indicates diversity among individuals within the subpopulation and H_e is the expected heterozygosity. The H_s is the mean of diversity within subpopulations as:

$$H_s = \frac{1}{L} \sum_{j=1}^{L} Hslj$$

Where L is the total loci, and j is the number of loci, and Hsl is the diversity within subpopulations and computed as:

$$Hsl = 1 - \sum Psi^2$$

Where Psi is the frequency in the i^{th} allele in one locus in the s^{th} subpopulation (Pacheco et al. 2016).

The pairwise r2 for SNP markers on each chromosome for whole genotypes, two-rowed, and six-rowed barley genotypes, and genotypes within subpopulations were calculated for linkage disequilibrium (LD) using the statistical program R (R Core Team, 2021). The loci were considered to be in significant LD when $P < 0.01$, the rest of r2 values was not considered as informative.

Results

The distribution of allele frequencies was classified in five categories (Fig. 1) indicating the proportion of SNP markers with MAF values that fell within the following ranges: 0.05 to <0.1, 0.1 to <0.2, 0.2 to <0.3, 0.3 to <0.4, and 0.4 to ≤0.5. The SNPs whose MAF value was less than 0.05 were discarded. A higher proportion of SNPs (51%) fell in the two middle categories (0.1≤MAF<0.3), while 35% of the SNPs showed a MAF≥0.3 and fell in the last two categories. The lower proportion of SNPs (14%) had MAF<0.1 (Fig.1).

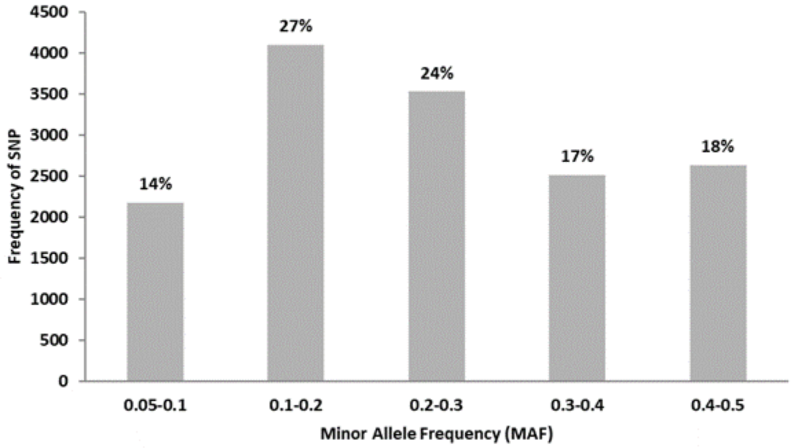

Figure 1. Distribution of SNP frequencies based on MAF categories within the population of 107 barley.

The distribution of SNPs varied on chromosomes and chromosome 5H had the highest SNPs (2658 SNP) while chromosome 1H had the lowest SNPs (1571 SNP) (Table 2). The average of PIC value for each chromosome ranged from 0.320 for chromosome 7H to 0.349 for chromosome 6H. Furthermore, the distribution of PIC value of 14,943 SNP for all genotypes (Fig. 2) indicated that PIC value ranged from 0.09 to 0.50, and 52% of SNPs had PIC value higher than the average PIC value (0.34), representing higher polymorphisms in the studied SNPs.

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

Table 2. Distribution of SNP markers on each chromosome and average PIC value for each chromosome.

Chromosome	SNPs number	Ave. PIC	SD PIC
1H	1571	0.338	0.13
2H	2533	0.336	0.13
3H	2472	0.337	0.12
4H	1744	0.336	0.12
5H	2658	0.337	0.13
6H	1918	0.349	0.12
7H	2047	0.320	0.13
Average	2134	0.336	

SNP: single nucleotide polymorphism, Ave. PIC: average polymorphic information content, and SD: standard deviation

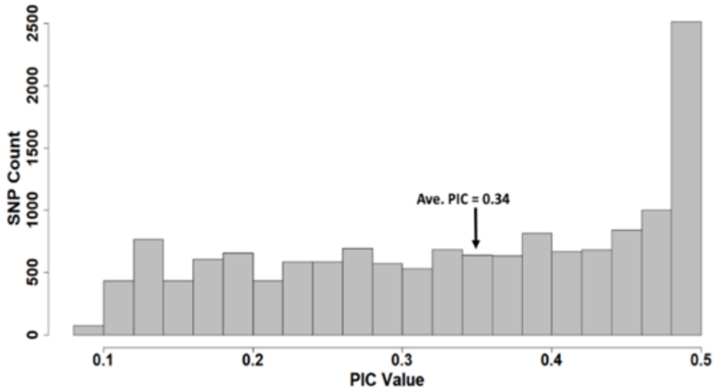

Figure 2. Distribution of PIC values of 14943 SNPs using entire barley genotypes.

The PIC values were computed for two-rowed and six-rowed types using the whole-genome information (Fig. 3) and each chromosome (Table 3). The average PIC value for the six-rowed and two-rowed types were 0.367, and 0.178, respectively (Fig. 3 A, and B). For the six-rowed types, the PIC value ranged from 0.0377 to 0.5 and 57.16% of SNPs indicated PIC values greater than average (0.367) while for the two-rowed types the range of PIC values was between 0.0345 to 0.484 and 44.31% of SNPs showed PIC values greater than average (0.0.178). For six-rowed type, average PIC value ranged from 0.361 (chromosome 7H) to 0.375 (chromosome 4H) while it ranged from 0.164 (chromosome 2H) to 0.191 (chromosome 6H) for two-rowed types (Table 3).

Table 3. The average PIC value for each chromosome for six-rowed and two-rowed barley.

	Six-rowed types		Two-rowed types	
Chromosome	Ave. PIC	SD PIC	Ave. PIC	SD PIC
1H	0.366	0.13	0.184	0.10

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

2H	0.365	0.12	0.164	0.09
3H	0.373	0.12	0.178	0.09
4H	0.375	0.12	0.167	0.09
5H	0.366	0.12	0.179	0.10
6H	0.365	0.13	0.191	0.09
7H	0.361	0.13	0.180	0.10
Average	0.367		0.178	

Ave. PIC: average polymorphic information content, and SD: standard deviation.

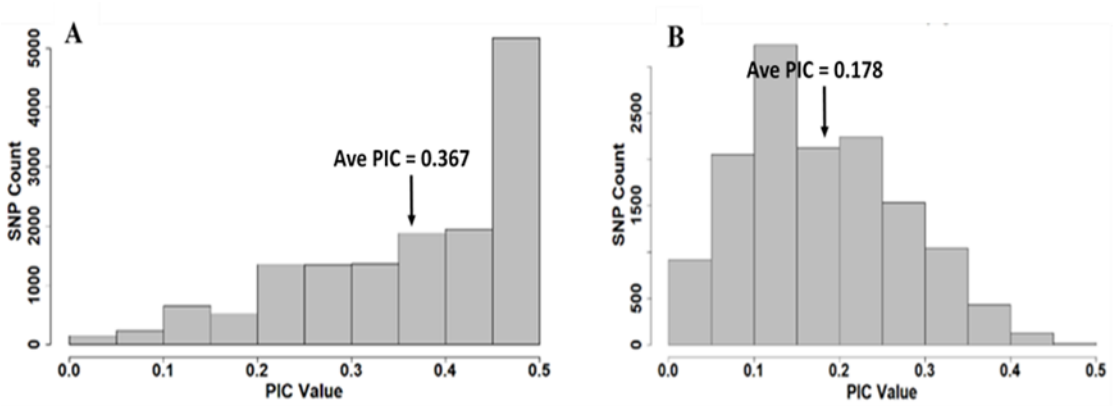

Figure 3. Distribution of PIC values for six-rowed (A) and two-rowed barley (B).

Population structure

The PCA and Ward's method was used to uncover the population structure of 107 barley genotypes using 14,943 whole-genome SNPs which led to identifying three subpopulations with 42, 20, and 45 genotypes for subpopulations 1, 2, 3, respectively (Fig. 4 A). Moreover, the hierarchical cluster analysis with Ward's method and Nei's genetic distance was used to analyze population structure, resulting in three distinctive subpopulations with 41, 21, and 45 genotypes (individuals) (Fig. 4 B). According to PCA and the Ward's clustering, subpopulation 1 comprised all 25 barley genotypes from SPII, eight lines from the ICARDA-DARI breeding program, nine genotypes from national and international landraces (Fig. 4 A, blue cluster). Genotypes of this subpopulation included one spring (2%), 26 winters (62%), and 15 facultative (36%) barley genotypes. Also, 33 of 42 genotypes in subpopulation 1 had a six-rowed types, and all of 42 genotypes except of "Sararood 1" cultivar were irrigated barley. However, the "Sararood 1" cultivar was adapted for supplementary irrigation conditions. Subpopulation 2 (Fig. 4 A, red cluster) included national and international landraces (20 genotypes) which all of them except "landrace 102" were winter and six-rowed types. In this cluster, there were three irrigated and 17 Irr-Rf barley genotypes (Fig. 4 A, red cluster). Subpopulation 3 included 45 barley genotypes in which 40 originated from the ICARD-DARI breeding program and five were from national and international landraces (Fig. 4 A green cluster). All genotypes in this cluster were rainfed with two-rowed types, and 39 (87%), and six genotypes (13%) had facultative and spring growth habits, respectively (Fig. 4 A green cluster, and Table 1). Grouping of genotypes into subpopulations by Nei's genetic distance (population structure) (Fig. 4 B) was very similar to PCA result, except for two genotypes (genotype 19 and 61 from SPII and ICARDA-DARI breeding program, respectively) which differed in subpopulation 1, and fell into subpopulation 2. In subpopulation 2, only one genotype (102 from Iranian landraces) differed which fell into subpopulation 1. All genotypes in subpopulation 3 were similar for both methods (Fig. 4 B).

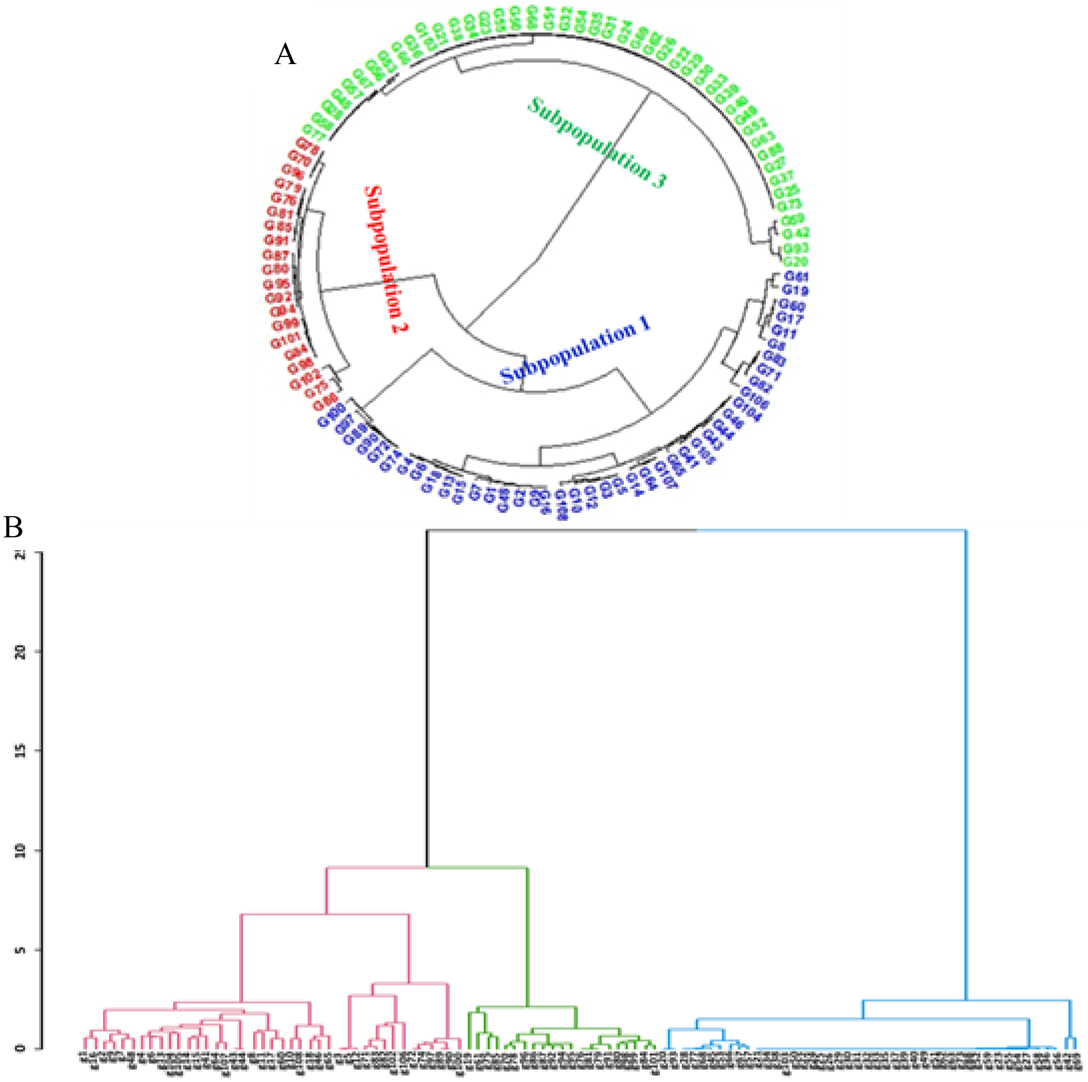

Figure 4. Dendrogram of grouping 107 barley genotypes using 14943 SNPs into three subpopulations (clusters) based on PCA and Ward method (A), and Nei's genetic distance and Ward method (B).

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

The distribution of PIC values was different for three subpopulations so that subpopulation 1 with an average PIC of 0.38 was more diverse than subpopulation 2 (Ave. PIC = 0.29), and subpopulation 3 (Ave. PIC = 0.09) (Fig. 5). In subpopulation 1, 59.86% of SNPs had a PIC value greater than the average PIC, representing higher polymorphic SNPs in this subpopulation, by contrast average PIC was respectively calculated 48.30% and 28.99% for subpopulations 2 and 3. Furthermore, the range of PIC value was 0.023–0.5.00, and 0.022−0.411, 0.051−0.500 for subpopulation 1, 2, and 3, respectively. The average PIC value for each chromosome ranged from 0.370 for 2H to 0.395 for 3H in subpopulation 1, while this value ranged from 0.255 for 1H to 0.313 for 7H in subpopulation 2. For subpopulation 3, the average PIC value for each chromosome was smaller than those in subpopulations 1 and 2 and ranged from 0.077 for 2H to 0.116 for 3H (Table 4) representing the lowest genetic diversity.

Table 4. Average PIC values using 14943 SNPs for three subpopulations in each chromosome.

Chromosome	Ave. PIC Subpopulation 1	SD	Ave. PIC Subpopulation 2	SD	Ave. PIC Subpopulation 3	SD
1H	0.388	0.11	0.255	0.15	0.091	0.09
2H	0.370	0.12	0.296	0.15	0.077	0.08
3H	0.395	0.11	0.301	0.15	0.116	0.12
4H	0.384	0.11	0.289	0.15	0.100	0.10
5H	0.380	0.11	0.275	0.15	0.083	0.08
6H	0.384	0.12	0.292	0.15	0.090	0.09
7H	0.374	0.12	0.313	0.14	0.100	0.10
Average	0.382		0.289		0.094	

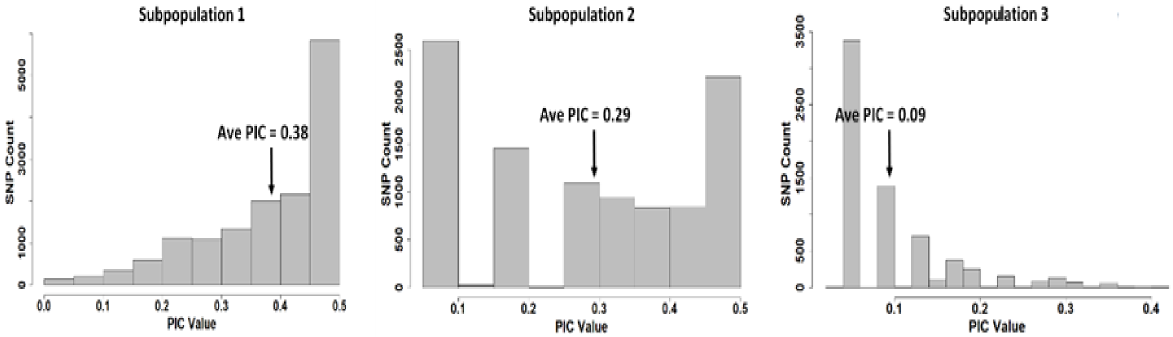

Figure 5. Distribution of PIC values by 14943 SNPs for subpopulation 1, subpopulation 2, and subpopulation 3.

There was a good agreement between Nei's genetic distance classification and PCA visualization so that the three clusters classified by Nei can be identified in the PCA plot. The first three axes of the PC analysis explained 48.12% of the total genetic variation of 107 genotypes, and the proportion of the 1st, 2nd, and 3rd PCs was 31.55%, 11.52% and 5.06%, respectively (Fig. 6).

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

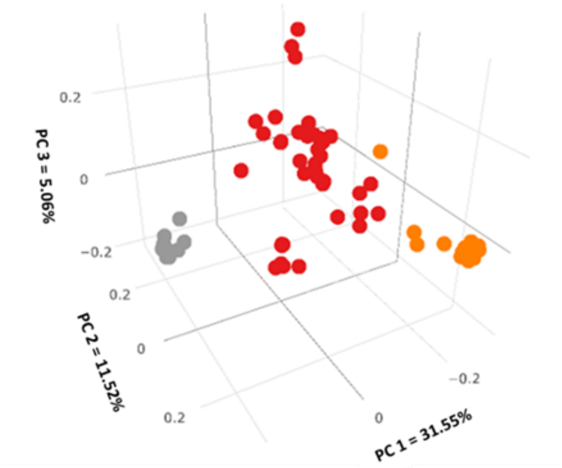

Figure 6. Grouping of 107 barley genotypes to three subpopulations with 14943 SNPs by three first principal components (PCA). Red color: subpopulation 1, Orange color: subpopulation 2, and Gray color: subpopulation 3. The first three PCs explained 48.12% of variation and the 1st PC had the highest value which representing a correct relationship among subpopulations. Lower PIC (Table 4), *Hsl* (Table 5), and higher LD amounts for subpopulation 3 representing higher familial relationship within genotypes which lead to the highest value for PC1 compared to the previous studies (Bengtsson et al. 2017a; Jabbari et al. 2018; Melchinger et al. 1994).

Genetic differentiation among populations

The *Fst* measures the degree of genetic differentiation among populations, according to the allele frequencies (Pacheco et al. 2016). Pacheco et al. (2016) classified Fst as following: 0–0.05; small genetic differentiation, 0.05–0.15; middle genetic differentiation, 0.15–0.25; big genetic differentiation, and *Fst*≥0.25; very big genetic differentiation. Fixation index (*Fst*) was computed to evaluate genetic differentiation among subpopulations with Nei's genetic distance matrix. In this study, Fst was 0.381 indicating a large level genetic differentiation among three subpopulations. The mean of diversity within subpopulations (*Hs*) was 0.208 while the *Hsl* values revealed that subpopulations 1 and 3 had the highest (0.379) and lowest (0.040) genetic diversity, respectively.The mean value of the effective number of alleles (*Ne*) was 1.56. The value of Shannon's information index (*I*) was 0.74 and the average of gene diversity based on Nei's gene diversity (*He*) was 0.336. The observed heterozygosity (*Ho*) and average expected heterozygosity (*He*) was 0.004, and 0.336, respectively (Table 5).

Table 5. Genetic diversity parameters for entire population and subpopulations based on 14943 SNPs.

He = Nei	*Ho*	*Ne*	*I*	*Hsl1*	*Hsl2*	*Hsl3*	*Hs*	*Dst = He - Hs*	*Fst = Dst/He*
0.336	0.004	1.56	0.74	0.379	0.204	0.040	0.208	0.128	0.381

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

He: expected heterozygosity, *Nei*: Nei genetic distance index, *Ho*: observed heterozygosity, Ne: number of effective alleles, I: Shannon information index, *Hsl1*: diversity within subpopulation1, *Hsl2*: diversity within subpopulation2, *Hsl3*: diversity within subpopulation3, *Hs*: mean of diversity within subpopulations, *Dst*: diversity among individuals within a subpopulation, and *Fst*: fixation index.

The pairwise LD measurement r^2 of 14943 SNPs related to physical distance indicated that LD varied in different chromosomes. For whole-genome LD (entire population), the r^2 value of 16,477,702 intra-chromosomal SNP markers was estimated with average r^2= 0.0957 in which r^2 values of 6,734,162 pairwise (40.1%) was significant (*P<0.01*). The chromosome 5H comprised the highest pairwise SNPs (3531153) (21.43%) in which the r2 value of 1472144 pairwise SNPs (21.9%) was significant (*P<0.01*) while the chromosome 1H with the 1233235 pairwise SNPs (7.48%) included the lowest pairwise SNPs in which r^2 value of 502460 pairwise markers (7.46%) was significant (*P<0.01*). The highest average r^2 value was for the chromosomes 2H and 4H with 0.1060 and 0.1020, respectively, while the lowest average r^2 value belonged to chromosome 7H (0.0773) (Table 6).

Table 6. Total pairwise SNPs, pairwise SNPs with significant r^2, average pairwise LD measurement r^2 for whole-genome, and each chromosome with 14943 SNPs and 107 barley genotypes.

Chromosome	Total pairwise SNPs	Ave. r^2	$r^2 = 1$	Pairwise SNPs with r^2 (*P*<0.01)	r^2 and *P*<0.01
1H	**1233235**	0.0954	35	**502460**	0.202
2H	3206778	**0.1060**	63	1403937	**0.215**
3H	3054156	0.0908	74	1203018	0.197
4H	1519896	**0.1020**	43	648644	**0.209**
5H	**3531153**	0.0988	64	**1472144**	0.206
6H	1838403	0.0997	53	785836	0.203
7H	2094081	0.0773	48	718123	0.184
Sum	16477702	0.0957	380	6734162	0.202

r^2: average correlation coefficient square, r^2= 1 shown the pairwise SNPs that are in complete LD, and r^2 and *P<0.01*: pairwise SNPs that their r^2 was significant at *P<0.01* (r^2 is average value for each chromosome).

The pairwise LD measurement r^2 was computed on each chromosome for two and six-rowed type barley, subpopulations 1 and 3 separately to assess population structure effects on LD pattern (Table 7 and 8). For all chromosomes, the proportion of pairwise SNPs with significant p-value in the whole population was more than that in two and six-rowed types and also genotypes within subpopulations 1 and 3 (Table 7 and 8). But the average r^2 of pairwise SNPs with significant p-value for two and six-rowed types and genotypes in subpopulation 1 was more than that in the whole population on different chromosomes (Table 7 and 8). Furthermore, the average r^2 for two-rowed genotypes and subpopulation 3 (mainly comprised rainfed two-rowed barley) was more than that for six-rowed type genotypes (which 84% of them was irrigated barley) (Table 7 and 8).

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

Table 7. Total pairwise SNPs, pairwise SNPs with significant r^2, average r^2 for two- and six-rowed type barley genotypes in each chromosome.

	Two-rowed barley					Six-rowed barley			
Chr.	Total pairwise SNPs	Ave. r^2	$r^2 = 1$	$P<0.01$	Ave. r^2 and $P < 0.01$	Ave. r^2	$r^2 = 1$	$P<0.01$	Ave. r^2 and $P<0.01$
1H	1233235	0.187	2966	276683	0.39	0.076	333	180041	0.272
2H	3206778	0.193	14533	723309	0.412	0.082	874	531320	0.273
3H	3054156	0.172	10464	708915	0.398	0.068	443	398093	0.255
4H	1519896	0.212	8062	301832	0.446	0.08	428	246871	0.266
5H	3531153	0.212	12598	913632	0.415	0.069	558	467557	0.256
6H	1838403	0.184	5547	537153	0.388	0.065	401	220642	0.26
7H	2094081	0.159	5065	395836	0.381	0.056	329	193705	0.249
Sum	16477702	0.158	59235	3857360	0.404	0.071	3366	2238229	0.262

SNP: single nucleotide polymorphism, Ave. r2: average correlation coefficient square, and P: p-value.

Table 8. Total pairwise SNPs, pairwise SNPs with significant r^2, average r^2 for subpopulations 1 and 3 in each chromosome.

	Subpopulation 1					Subpopulation 3			
Chr.	Total pairwise SNPs	r^2	$r^2 = 1$	$P<0.01$	r^2 and $P<0.01$	r^2	$r^2 = 1$	$P<0.01$	r^2 and $P<0.01$
1H	1233235	0.063	329	96212	0.297	0.277	21499	40303	0.586
2H	3206778	0.065	827	258241	0.298	0.184	14991	9758	0.570
3H	3054156	0.058	514	208908	0.286	0.202	13924	35966	0.642
4H	1519896	0.063	419	116152	0.296	0.355	1239	3476	0.652
5H	3531153	0.056	742	211791	0.288	0.158	11129	19424	0.531
6H	1838403	0.184	508	129794	0.300	0.214	24167	50384	0.568
7H	2094081	0.060	468	111520	0.289	0.191	14422	40789	0.566
	16477702	0.078	3807	1132618	0.293	0.226	101371	200100	0.588

SNP: single nucleotide polymorphism, r^2: correlation coefficient square, and P: p-value.

Discussion

This study was aimed to assess genetic diversity and population structure of a set of breeding lines, cultivars, and landraces using SNP markers. In this study the number of SNPs per chromosome was different and chromosome 5H had the highest SNPs while chromosome 1H had the lowest SNPs. This finding was similar to the result of a previous study on Nordic spring barley panel with SNP markers (Bengtsson et al. 2017b) and research on a set of winter and spring barley accessions with polymorphic SNPs (Xu et al. 2018). However, our finding was in disagreement with the results of a study on a set of six-rowed barley from the USA and Kazakhstan with SNPs for which chromosomes 1H and 4H comprised the lowest number of SNPs while 3H included the highest SNPs (Almerekova et al. 2019).

These differences could be due to in the amount of genomic coverage of different sets of SNPs that were used in previous studies. The genetic diversity parameters were measured for whole barley genotypes and each subpopulation. The Nei's and Shannon's indices as reliable parameters for assessing genetic diversity highlighted high genetic variation among this panel (Feng et al. 2018; Nei. 1978; Tomar et al. 2021; Yu et al. 2021). Structure analysis of the panel was carried out using PCA, Nei's genetic distance among genotypes, and hierarchical cluster analysis based on the Ward's algorithm which lead to similar results and 107 barley genotypes placed into three subpopulations in which the breeding lines and cultivars were well grouped based on row type, rainfed and irrigated (subpopulation 1 included irrigated and six-rowed barley and subpopulation 3 comprised rainfed and two-rowed barley) while landraces distributed in all three subpopulations due to high genetic diversity. Therefore, strong population structure effect in this panel was related to row type (two-rowed and six-rowed), and adaptation to irrigated, and rainfed environment (Ataei et al. 2018).

Results of earlier studies indicated that growth habit (winter and spring type), row type (two and six), and geographical origin were the main factors leading to population structure in barley populations (Bengtsson et al. 2017a; Comadran et al. 2009; Hamblin et al., 2010; Malysheva-Otto et al. 2006; Tondelli et al. 2013; Zhang et al. 2009). In a study LD and genetic diversity pattern of 192, Mediterranean barley with SSR and DArT markers grouped similar genotypes into five subpopulations according to growth habit, row type, and geographical origin (Comadran et al. 2009). The PCA showed that the three first PCs explained a higher amount of the total genetic variation of the panel that was higher than that was reported in the previous studies representing a correct relationship among subpopulations and genotypes (Bengtsson et al. 2017a; Jabbari et al. 2018; Melchinger et al. 1994). The highest value for PC1 was due to lower PIC values (Table 4) and Hsl (Table 5) for subpopulation 3 (included all two-rowed barley) which led to higher familial relationship within genotypes of this subpopulation. A genetic diversity study was conducted on 100 six-rowed winter barley genotypes using 3964 SNPs and indicated that the 1st and 2nd PCs explained 13.8 and 8.97% of whole genetic variation of the panel, respectively (Ataei et al. 2018).

In this study the PC1 value was lower than PC1 value of our finding which was due to using diverse panel in their study. The results of fixation index, *Fst*, indicated a higher genetic differentiation existing among three subpopulations. In accordance with the Pacheco et al. (2016), *Hsl* values indicated that subpopulation 1 (six-rowed, and irrigated barley) had the highest and subpopulation 3 (two-rowed and rainfed barley) had the lowest genetic variation. Ataei et al. (2018) was reported that genetic diversity in six-rowed barley cultivars was greater than that in two-rowed barley cultivars. The average PIC of 0.34 demonstrated a high genetic divergence of the panel in this study. Furthermore, 52% of SNPs had a PIC value greater than the average indicating that the SNPs used in this study were very informative markers for studying the genetic diversity of the barley population (Kumar et al. 2020; Tomar et al. 2021).

Similar results were reported for six-rowed winter barley genotypes in which average PIC value was 0.39 (0.19-0.5) (Ataei et al. 2018) whereas our finding for six-rowed types was 0.37 and for entire genotypes was 0.34 (0.05-0.50).

Hill et al. (2021) estimated a low polymorphism of 0.17 PIC for 632 accessions in a barley panel. Furthermore, Varshney et al. (2010)

obtained average PIC value for SSR and SNP markers as 0.63 and 0.38, respectively for a set of ICARDA barley germplasm collection including 185 cultivated (*H. vulgare* L.) and 38 wild (*H. spontaneum* L.) genotypes originated from 30 countries. The average and distribution of PIC values of three subpopulations showed a different trend of PIC. In agreement with the Hsl values, the trend increased toward 0.5 of PIC (maximum diversity) in subpopulation 1 while it decreased in subpopulation 3 indicating very high genetic variation in subpopulations 1. Decreasing trend of PIC value in sub-population 3 could be due to selection intensity for specific traits under rainfed conditions, and common parents that were used in DARI's barley breeding program for line and cultivar development. Regarding this result, genetic diversity of European two-rowed spring barley showed low PIC values at regions on chromosome 1H (47.8–55.4), 6H (30.2–53.6) and 7H (29.8–47.6), which high probably was due to selection for malting quality traits and yield (Tondelli et al. 2013).

For six-rowed winter type barley, Ataei et al., (2018) indicated very similar results in which the average of PIC was varied from 0.37 (2H and 5H chromosomes) to 0.42 (3H and 7H chromosomes) (Table 3; PIC values for six-rowed barley). There was no specific trend for PIC values in subpopulation 2 (Iranian and international landraces) and its genetic variation was lower than that for subpopulation 1 and much higher than that for subpopulation 3. The observed average PIC for this panel (0.34) was comparable to previous studies and PIC values also varied among chromosomes (Jabbari et al., 2018). The highest average PIC value was 0.349 for chromosome 6H, which resembles the results of a study on a set of European barley cultivars (Jabbari et al. 2018; Roy and Shil, 2020). The PIC variations among chromosomes differed for six and two-rowed barley in which for six-rowed types, the upper and lower level of PIC was for chromosomes 4H, and 7H, which was in disagreement with the results of a study on a set of six-rowed type barley genotypes from the USA and Kazakhstan with SNPs for which the range of PIC values was between 0.28 for 2H and 0.34 for 3H (Almerekova et al. 2019).

These differences could be due to different sets of SNPs that were used in these studies. For two-rowed barley the highest and lowest average PIC value was for chromosomes 2H and 6H, respectively, which corresponds with the result of Jabbari et al. (2018). The distribution of whole-genome and average PIC value for each chromosome indicated that six-rowed type barley were more diverse than the two-rowed types which was in agreement with the finding of Pasam et al. (2012) and Ataei et al. (2018). The pairwise LD measurement (r^2) was computed for whole-genome, each chromosome, two and six-rowed type barley, subpopulations 1 and 3 separately to assess population structure effects on LD pattern. The whole-genome LD (entire population) results indicated that LD amount was different on chromosomes and the highest amount of LD was on chromosomes 2H and 4H while chromosome 7H had the lowest amount of LD. The LD amount for two and six-rowed types and subpopulation 1 was more than whole-genome LD on different chromosomes. Furthermore, the LD amount of two-rowed genotypes and subpopulation 3 (rainfed two-rowed types) was more than that for six-rowed types. The higher LD within subpopulations could be due to population structure resulting from small size and higher familial relationships among individuals compared to the

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

entire population. Results revealed that the LD amounts for rainfed barley were higher than that for irrigated barley. Additionally, lower genetic variation (the lowest PIC and *Hsl*) within subpopulation 3 (rainfed barley) leads to higher LD which could be due to higher selection intensity for traits under rainfed conditions (Stracke et al. 2007). In agreement with our result and conclusion, Bengtsson et al. (2017) found slower LD decay within the two-rowed lines from the southern part compared to the six-rowed lines and the two-rowed lines from the northern part, that may be a result of strong selection for higher malting quality and yield in the southern parts. Also, there were more common parents in the pedigree of rainfed breeding lines and cultivars resulting in higher familial relationships among rainfed genotypes, and consequently higher LD. On the other hand, most of the rainfed barley used in this study were elite, advanced, promising breeding lines and cultivars which were selected for important agronomic traits simultaneously leading to narrow down genetic variation, and consequently increasing familial relationships and LD in this germplasm (Gupta et al. 2005).

Conclusion

A high level of genetic diversity was observed among barley breeding lines, cultivars, and landraces based on informative SNP markers. Irrigated winter six-rowed types were more diverse than rainfed two-rowed type barley.

Strong population structure in this paned was related to row-type and adaptation to irrigated and rainfed conditions. The LD amount of rainfed two-rowed barley was more than that for irrigated six-rowed types, which could be the result of strong selection for traits under rainfed conditions.

The potential genetic variation of this pane could be used in barley improvement programs to extend genetic diversity of germplasms specially rainfed two-rowed types resulting in to develop new lines with desirable traits under rainfed conditions. Also, landraces could be used in the pre-breeding program to introgression novel genetic diversity to the background of breeding lines and cultivars. The final goal of exploring genetic diversity will be improving productivity and stability under rainfed conditions.

Acknowledgments

This work was conducted in the Dryland Agricultural Research Institute (DARI) as a public research institute in Iran. The results of the present study will support DARI's barley breeding program to extend the genetic variation of barley germplasms.

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

References منابع

1. Al-Abdallat AM., Karadsheh A., Hadadd NI., Akash MW., Ceccarelli S., Baum M., Hasan M., Jighly A. and Abu Elenein JM. (2017). Assessment of genetic diversity and yield performance in Jordanian barley (*Hordeum vulgare* L.) landraces grown under Rainfed conditions. BMC Plant Biology. 17 1–13. https://doi.org/10.1186/s12870-017-1140-1
2. Almerekova S., Sariev B., Abugalieva A., Chudinov V., Sereda G., Tokhetova L., Ortaev A., Tsygankov V., Blake T., Chao S., Genievskaya Y., Abugalieva S. and Turuspekov Y. (2019). Association mapping for agronomic traits in six-rowed spring barley from the USA harvested in Kazakhstan. PLoS One. https://doi.org/10.1371/journal.pone.0221064
3. Ataei R., Gholamhoseini M. and Ahmadi H. (2018). Study of genetic dand population structure in barley (*Hordeum vulgare* L.) based on SNP markers. Iranian Journal of Field Crop Science (Iranian Journal of Agricultural Sciences). 49 (2): 35-59. DOI: 10.22059/ijfcs.2017.233380.654321
4. Bayer MM., Rapazote-Flores P., Ganal M., Hedley PE., Macaulay M., Plieske J., Ramsay L., Russell J., Shaw PD., Thomas W. and Waugh R. (2017). Development and evaluation of a barley 50k iSelect SNP array. Front. Plant Science. 8: 1–10. https://doi.org/10.3389/fpls.2017.01792
5. Bengtsson T., Åhman I., Bengtsson T., Manninen O., Veteläinen M., Reitan L., Alsheikh M., Gertsson B., Tuvesson S., Jalli M., Jahoor A., Jensen JD., Orabi J., Backes G., Krusell L., Hjortshøj RL., Helgadóttir Á., Göransson M., Sveinsson S., Manninen O., Jahoor A. and Orabi J. (2017a). Genetic diversity, population structure and linkage disequilibrium in Nordic spring barley (*Hordeum vulgare* L. subsp. vulgare). Genetic Resource and Crop Evolution. 64: 2021–2033. https://doi.org/10.1007/s10722-017-0493-5
6. Bengtsson T., Åhman I., Bengtsson T., Manninen O., Veteläinen M., Reitan L., Alsheikh M., Gertsson B., Tuvesson S., Jalli M., Jahoor A., Jensen JD., Orabi J., Backes G., Krusell L., Hjortshøj RL., Helgadóttir Á., Göransson M., Sveinsson S., Manninen O., Jahoor A. and Orabi J. (2017b). Genetic diversity, population structure and linkage disequilibrium in Nordic spring barley (*Hordeum vulgare* L. subsp. vulgare). Genetic Resource and Crop Evolution. 64: 2021–2033. https://doi.org/10.1007/s10722-017-0493-5
7. Blattner FR. (2018). Taxonomy of the Genus Hordeum and Barley. pp. 11–23. https://doi.org/10.1007/978-3-319-92528-8_2
8. Botstein D., White RL., Skolnick M. and Davis RW. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. The American Journal of Human Genetics. 32: 314–331.
9. Bradbury PJ., Zhang Z., Kroon DE., Casstevens TM., Ramdoss Y. and Buckler ES. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics. 23: 2633–2635. https://doi.org/10.1093/bioinformatics/btm308
10. Comadran J., Thomas WTB., van Eeuwijk FÁ., Ceccarelli S., Grando S., Stanca AM., Pecchioni N., Akar T., Al-Yassin A., Benbelkacem A., Ouabbou H., Bort J., Romagosa I., Hackett CA. and Russell JR. (2009). Patterns of genetic diversity and linkage disequilibrium in a highly structured *Hordeum vulgare* association-mapping population for the Mediterranean basin. Theoritical Applied Genetics. 119: 175–187. https://doi.org/10.1007/s00122-009-1027-0
11. Condón F., Rasmusson DC., Schiefelbein E., Velasquez G. and Smith KP. (2009). Effect of advanced cycle breeding on genetic gain and phenotypic diversity in barley breeding germplasm. Crop Science. 49: 1751–1761. https://doi.org/10.2135/cropscience.2008.10.0585
12. Dotlačil L., Hermuth J., Stehno Z., Dvořáček V., Bradová J. and Leišová L. (2010). How can wheat landraces contribute to present breeding? Czech Journal of Genetics and Plant Breeding. 46. https://doi.org/10.17221/1519-cjgpb
13. Elakhdar A., EL-Sattar MA., Amer K., Rady A. and Kumamaru T. (2016). Population structure and marker–trait association of salt tolerance in barley (*Hordeum vulgare* L.). Comptes Rendus Biologies. 339: 454–461. https://doi.org/10.1016/j.crvi.2016.06.006
14. FAO. (2020). FAOSTAT. ROM. www.faostat.fao.org.
15. Feng Y., Ryan UM. and Xiao L. (2018). Genetic diversity and population structure of cryptosporidium. Trends in Parasitology. 34: 997–1011. https://doi.org/10.1016/j.pt.2018.07.009

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

16. Ferreira JR., Pereira JF., Turchetto C., Minella E., Consoli L. and Delatorre CA. (2016). Assessment of genetic diversity in Brazilian barley using SSR markers. Genetics and Molecular Biology. 39: 86–96. https://doi.org/10.1590/1678-4685-GMB-2015-0148

17. Gupta PK., Rustgi S. and Kulwal PL. (2005). Linkage disequilibrium and association studies in higher plants: Present status and future prospects. Plant Molecular Biology. https://doi.org/10.1007/s11103-005-0257-z

18. Hamblin MT., Close TJ., Bhat PR., Chao S., Kling JG., Abraham KJ., Blake T., Brooks WS., Cooper B., Griffey CA., Hayes PM., Hole DJ., Horsley RD., Obert DE., Smith KP., Ullrich SE., Muehlbauer GJ. and Jannink JL. (2010). Population structure and linkage disequilibrium in U.S. barley germplasm: Implications for association mapping. Crop Science. 50: 556–566. https://doi.org/10.2135/cropsci2009.04.0198

19. Hill CB., Angessa TT., Zhang XQ., Chen K., Zhou G., Tan C., Wang P., Westcott S. and Li C. (2021). A global barley panel revealing genomic signatures of breeding in modern Australian cultivars. The Plant Journal. 106: 419–434. https://doi.org/10.1111/tpj.15173

20. Jabbari M., Fakheri BA., Aghnoum R., Nezhad NM. and Ataei R. (2018). GWAS analysis in spring barley (*Hordeum vulgar*e L.) for morphological traits exposed to drought. PLoS One. https://doi.org/10.1371/journal.pone.0204952

21. Jehan T. and Lakhanpaul S. (2006). Single nucleotide polymorphism (SNP) - methods and applications in plant genetics: A review. Indian Journal of Biotechnology. 5: 435-456.

22. Kalinowski S. (2002). How many alleles per locus should be used to estimate genetic distances? Heredity (Edinb). 88: 62–65. https://doi.org/10.1038/sj.hdy.6800009

23. Kanazin V., Talbert H., See D., DeCamp P., Nevo E. and Blake T. (2002). Discovery and assay of single-nucleotide polymorphisms in barley (*Hordeum vulgare*). Plant Molecular Biology. 48: 529–537. https://doi.org/10.1023/A:1014859031781

24. Kumar P., Banjarey P., Malik R., Tikle AN. and Verma RPS. (2020). Population structure and diversity assessment of barley (*Hordeum vulgare* L.) introduction from ICARDA. Journal of Genetics. 99. https://doi.org/10.1007/s12041-020-01226-6

25. Lai K., Lorenc MT., Lee HC., Berkman PJ., Bayer PE., Visendi P., Ruperao P., Fitzgerald TL., Zander M., Chan CKK., Manoli S., Stiller J., Batley J. and Edwards D. (2015). Identification and characterization of more than 4 million intervarietal SNPs across the group 7 chromosomes of bread wheat. Plant Biotechnology Journal. 13(1): 97-104. https://doi.org/10.1111/pbi.12240

26. Malysheva-Otto LV., Ganal MW. and Röder MS. (2006). Analysis of molecular diversity, population structure and linkage disequilibrium in a worldwide survey of cultivated barley germplasm (*Hordeum vulgare* L.). BMC Genetics. 7. https://doi.org/10.1186/1471-2156-7-6

27. Melchinger AE., Graner A., Singh M. and Messmer MM. (1994). Relationships among European Barley Germplasm: I. genetic diversity among winter and spring cultivars revealed by RFLPs. Crop Science. 34(5): 1191-1199. https://doi.org/https://doi.org/10.2135/cropsci1994.0011183X003400050009x

28. Nei M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 89: 583–590. https://doi.org/10.1093/genetics/89.3.583

29. Nei M., Tajima F. and Tateno Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. Journal of Molecular Evolution. 19: 153–170. https://doi.org/10.1007/bf02300753

30. Pacheco Á., Alvarado G., Rodríguez F. and Burgueño J. (2016). BIO-R (Biodiversity analysis with R for Windows) Version 3.0. https://doi.org/hdl:11529/10820

31. Pasam RK., Sharma R., Malosetti M., van Eeuwijk FA., Haseneyer G., Kilian B. and Graner A. (2012). Genome-wide association studies for agronomical traits in a world wide spring barley collection. BMC Plant Biology. 12. https://doi.org/10.1186/1471-2229-12-16

32. Pasam RK., Sharma R., Walther A., Özkan H., Graner A. and Kilian B. (2014). Genetic diversity and population structure in a legacy collection of spring barley landraces adapted to a wide range of climates. PLoS One. 9: 1–29. https://doi.org/10.1371/journal.pone.0116164

33. Paulitz TC. and Steffenson BJ. (2010). Biotic stress in Barley: Disease problems and solutions. Barley, Wiley Online Books. https://doi.org/https://doi.org/10.1002/9780470958636.ch11

34. Poland JA., Brown PJ., Sorrells ME. and Jannink JL. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One. 7: e32253. https://doi.org/10.1371/journal.pone.0032253

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

35. R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

36. Roy SC. and Shil P. (2020). Assessment of genetic heritability in rice breeding lines based on morphological traits and caryopsis ultrastructure. Scientific Reports. 10. https://doi.org/10.1038/s41598-020-63976-8

37. Saghai-Maroof MA., Soliman KM., Jorgensen RA. and Allard RW. (1984). Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proceedings of the National Academy of Sciences of the United States of America. 81: 8014–8018. https://doi.org/10.1073/pnas.81.24.8014

38. Schulte D., Close TJ., Graner A., Langridge P., Matsumoto T., Muehlbauer G., Sato K., Schulman AH., Waugh R., Wise RP. and Stein N. (2009). The international barley sequencing consortium - At the threshold of efficient access to the barley genome. Plant Physiology. 149(1): 142-7. https://doi.org/10.1104/pp.108.128967

39. Smith CT., Antonovich A., Templin WD., Elfstrom CM., Narum SR. and Seeb LW. (2007). Impacts of marker class bias relative to locus-specific variability on population inferences in Chinook salmon: A comparison of single-nucleotide polymorphisms with short tandem repeats and allozymes. Transactions of the American Fisheriries Society. 136: 1674–1687. https://doi.org/10.1577/t06-227.1

40. Sreenivasulu N., Graner A. and Wobus U. (2008). Barley genomics: An overview. International Journal of Plant Genomics. https://doi.org/10.1155/2008/486258

41. Stracke S., Presterl T., Stein N., Perovic D., Ordon F. and Graner A. (2007). Effects of introgression and recombination on haplotype structure and linkage disequilibrium surrounding a locus encoding bymovirus resistance in barley. Genetics. 175(2): 805-17. https://doi.org/10.1534/genetics.106.063800

42. Tomar V., Dhillon GS., Singh D., Singh RP., Poland J., Joshi AK., Tiwari BS. and Kumar U. (2021). Elucidating SNP-based genetic diversity and population structure of advanced breeding lines of bread wheat (*Triticum aestivum* L .). p 9: e11593. https://doi.org/10.7717/peerj.11593

43. Tondelli A., Xu X., Moragues M., Sharma R., Schnaithmann F., Ingvardsen C., Manninen O., Comadran J., Russell J., Waugh R., Schulman AH., Pillen K., Rasmussen SK., Kilian B., Cattivelli L., Thomas WTB. and Flavell AJ. (2013). Structural and temporal variation in genetic diversity of European spring two-row barley cultivars and association mapping of quantitative traits. Plant Genome. 6: 1–14. https://doi.org/10.3835/plantgenome2013.03.0007

44. Varshney RK., Baum M., Guo P., Grando S., Ceccarelli S. and Graner A. (2010). Features of SNP and SSR diversity in a set of ICARDA barley germplasm collection. Molecular Breeding. 26: 229–242. https://doi.org/10.1007/s11032-009-9373-9

45. Varshney RK., Paulo MJ., Grando S., van Eeuwijk FA., Keizer LCP., Guo P., Ceccarelli S., Kilian A., Baum M. and Graner A. (2012). Genome wide association analyses for drought tolerance related traits in barley (*Hordeum vulgare* L.). Field Crop Research. 126: 171–180. https://doi.org/10.1016/j.fcr.2011.10.008

46. Wicker T., Narechania A., Sabot F., Stein J., Vu GTH., Graner A., Ware D. and Stein N. (2008). Low-pass shotgun sequencing of the barley genome facilitates rapid identification of genes, conserved non-coding sequences and novel repeats. BMC Genomics. 9: 1–15. https://doi.org/10.1186/1471-2164-9-518

47. Xu X., Sharma R., Tondelli A., Russell J., Comadran J., Schnaithmann F., Pillen K., Kilian B., Cattivelli L., Thomas WTB. and Flavell AJ. (2018). Genome-wide association analysis of grain yield-associated traits in a pan-european barley cultivar collection. Plant Genome. 11: 170073. https://doi.org/10.3835/plantgenome2017.08.0073

48. Yu H., Yang J., Cui H., Abbas A., Wei S. and Li X. (2021). Distribution, genetic diversity and population structure of aegilops tauschii coss. In major wheat-growing regions in China. Agriculture. 11. https://doi.org/10.3390/agriculture11040311

49. Zhang LY, Marchand S, Tinker NA, Belzile F. 2009. Population structure and linkage disequilibrium in barley assessed by DArT markers. Theoretical Applied Genetetics. 119, 43–52. https://doi.org/10.1007/s00122-009-1015-4

[DOR: 20.1001.1.27170632.1400.14.4.6.9]

"Journal of Biosafety; Volume 14, Number 4, Winter 2022"

مجله ایمنی زیستی

دوره ۱۴، شماره ۴، زمستان ۱۴۰۰

ISSN 2716-9804 الکترونیکی، ISSN 2717-0632 چاپی

ارزیابی تنوع ژنتیکی و ساختار جمعیت توده‌های بومی و لاین‌های اصلاحی جو با استفاده از نشانگرهای SNP

فرهاد آهک‌پز[1]، ایرج برنوسی[2]،بابک عبداللهی مندولکانی[3]، صابر گلکاری[4]، جعفر جعفرزاده[5] و سریپادا یودوپا [6]

۱. مربی پژوهشی، موسسه تحقیقات کشاورزی دیم کشور، سازمان تحقیقات آموزش و ترویج کشاورزی، مراغه، ایران

۲. دانشیار، گروه اصلاح و بیوتکنولوژی، دانشکده کشاورزی، دانشگاه ارومیه، آذربایجان غربی، ایران

۳. استاد، گروه اصلاح و بیوتکنولوژی، دانشکده کشاورزی، دانشگاه ارومیه، آذربایجان غربی، ایران

۴. دانشیار، موسسه تحقیقات کشاورزی بیوتکنولوژی ایران، کرج، ایران

۵. استادیار، موسسه تحقیقات کشاورزی دیم کشور، سازمان تحقیقات آموزش و ترویج کشاورزی، مراغه، ایران

۶. محقق ارشد، مرکز بین المللی تحقیقات کشاورزی در مناطق خشک (ایکاردا)، مراکش

i.bernosi@urmia.ac.ir

تاریخ دریافت: ۱۴۰۰/۱۰/۲۶، تاریخ پذیرش: ۱۴۰۰/۱۱/۱۱

صفحه ۱۳۴-۱۱۳

چکیده

بهره‌گیری از تنوع ژنتیکی برای تولید ارقام جدید ضروری است. دراین تحقیق تنوع ژنتیکی، ساختار جمعیت و عدم تعادل پیوستگی در سطح ژنوم (LD) ۱۰۷ لاین، رقم و توده بومی دیم و آبی با نشانگرهای SNP بررسی شد. تعداد ۱۴۹۴۳ نشانگر برای برآورد شاخص تنوع ژنی شانون (I)، نئی (H)، محتوای اطلاعات چندشکلی (PIC)، شاخص تثبیت (Fst) و تجزیه مؤلفه‌های اصلی (PCA) استفاده شدند. مقادیر I و H، ۰/۷۴ و ۰/۳۳۶ بود. میانگین کل PIC، ۰/۳۴ و برای جو شش و دو ردیفه به ترتیب ۰/۳۶۷ و ۰/۱۷۸ بود. تجزیه PC سه زیرجمعیت را شناسایی کرد که زیرجمعیت‌های ۱ و ۲ با میانگین مقدار PIC، ۰/۳۸ و ۰/۲۹ به ترتیب نسبت به زیرجمعیت ۳ (۰/۰۹) متنوع‌تر بودند. میانگین Fst، ۰/۳۸۱ بود. زیرجمعیت‌های ۱ و ۳ به ترتیب بیشترین (۰/۳۷۹) و کمترین (۰/۰۴۰) تنوع ژنتیکی زیرجمعیت‌ها را داشتند. میانگین LD برای جوهای دو ردیفه بیشتر از شش ردیفه بود. جوهای شش ردیفه آبی و توده‌های بومی بیشترین تنوع ژنتیکی و دو ردیفه‌های دیم کمترین تنوع ژنتیکی و بیشترین LD را نشان دادند. ساختار جمعیت مربوط به تعداد ردیف سنبله و آبی و دیم بودن جوها بود. این ژنوتیپ‌ها می‌توانند در یافتن ارتباط صفت - نشانگر در برنامه اصلاحی جو استفاده شوند.

واژه‌های کلیدی: منابع ژنتیکی جو، ساختار جمعیت، عدم تعادل لینکاژی در سطح ژنوم، محتوای اطلاعات چندشکلی و نشانگرهای چندشکلی تک نوکلئوتیدی.

[DOR: 20.1001.1.27170632.1400.14.4.6.9]