E

BMT/15/8
ORIGINAL: English
DATE: April 26, 2016

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

Geneva

WORKING GROUP ON BIOCHEMICAL AND MOLECULAR TECHNIQUES AND DNA-PROFILING IN PARTICULAR

Fifteenth Session

Moscow, Russian Federation, May 24 to 27, 2016

ASSOCIATION MAPPING FOR DUS CHARACTERISTICS IN A PANEL OF BARLEY LANDRACES USING SSR AND SRAP MARKERS

Document prepared by experts from the Islamic Republic of Iran

BACKGROUND AND PURPOSE

1. Technical tests for granting plant breeder's rights (PBRs) to owners of new barley varieties are generally relied on field-based morphological characteristics within two years of the so-called distinctness, uniformity and stability (DUS) tests. However, because of plentiful advantages of molecular markers over these characteristics, the Working Group on Biochemical and Molecular Techniques and DNA-Profiling in Particular (BMT) has considered Guidance on the Use of Biochemical and Molecular Markers in the Examination of Distinctness, Uniformity and Stability (DUS), TGP/15 (UPOV 2013). Markers which are directly linked to traditional characteristics and/or as a predictor of them can gain particular attention compared to markers with neutral functionality.

2. We used different statistical models to find Simple Sequence Repeats (SSR) and Sequence-related amplified polymorphism (SRAP) markers that are associated with DUS characteristics in a panel of 143 barley landraces and cultivars.

MATERIALS AND METHODS

Plant materials

3. A panel of 143 barley landraces and advanced breeding lines including 63 six-rowed and 80 two-rowed accessions with geographical origin from eight countries (Islamic Republic of Iran, Egypt, China, United States of America, United Kingdom, India, Pakistan, and Algeria) used for association study. These accessions were provided by Plant Genetic Resources and Cereals research departments of Seed and Plant Improvement Institute, Islamic Republic of Iran. This collection had records of seasonal growth habit (SGH), in which 62 were spring, 68 winter, and 12 facultative types, leaving one accession not assigned. The vast majority of six-rowed barley samples were winter types (88.8%), on the contrary there were spring types that constituted more than two-thirds (70%) of two-rowed accessions (SM 1).

Field trials and phenotyping

4. The 143 lines were planted in the experimental field at Seed & Plant Certification Institute (SPCRI) under lattice square (12×12) design with two replications in autumn of 2013 and 2014, and harvested in

spring of the following year. In order to alleviate the vernalization requirements of spring and facultative type accessions, their seeds were soaked in water for 24 hours and then stored at 2° C for 40 days before cultivation.

5. A number of 36 morphological characteristics were observed, of which 9 and 27 were quantitative (continuous) and qualitative (categorical) variables. The characteristics were chosen from Guidelines for the conduct of tests for distinctness, uniformity and stability, barley (*Hordeum vulgare* L. *sensu lato*), TG/19/10 (UPOV 1994), Community Plant Variety Office of the European Union (CPVO) Protocol for distinctness, uniformity and stability tests, *Hordeum vulgare* L. *sensu lato*, barley. CPVO-TP/019/3 Final (CPVO 2012), India's DUS test guideline (PPV & FRA 2011), and characteristics used by Wang et al. 2012. The characteristics (DUS characteristics hereafter) were coded as U, C, I, or W (denoting the literature that was taken) plus a number indicating characteristics number of the guideline. Among categorical variables, nine were binary (presence or absence of a characteristic), 13 were ordinal (a visual scale of the expression intensity of a characteristic), and five were nominal (like color or shape of an organ). Moreover, four ordinal variables (KCAL-AR-GACN-GSLN) were also regarded as binary variables. Similarly, one variable (ED) was both measured (continuous) and observed as an ordinal characteristic (SM 2).

Genotyping with SSR and SRAP markers

6. Genomic DNA of materials was extracted using the CTAB method (Saghai Maroof et al. 1984) from a bulk of 15 plants of each accession. A set of 149 Simple Sequence Repeats (SSR) and EST-SSR (derived from expressed sequence tags) markers evenly distributed over seven barley chromosomes were used for amplification. Polymerase chain reaction (PCR) assays were performed in a final volume of 10 µL containing 40 ng template DNA, 1X PCR buffer, 0.025 units Taq DNA polymerase (Cinnagen Co.), 0.2 mmol dNTPs, and 0.8 pmol forward and reverse primers. The amplification consisted of initial denaturation at 94°C for 4 min, followed by 35 cycles of 94°C for 1 min, 58–68°C (depending on primer pairs) for 1 min and 72°C for 1 min, and a final extension at 72°C for 7 min. A laser scanning Gel-Scan™ 3000 electrophoresis system (Corbett Life Science, a Qiagen company) was used for resolving PCR products of SSR loci through ultra-thin (0.2) polyacrylamide gel stained with ethidium bromide.

7. A set of 30 primer combinations (5 × 6 me-em primers) of SRAP markers was used according to Li and Quiros (2001). The PCR conditions followed that of the authors' recommendations. The products were electrophoresed in a dual system (CBS Scientific) through a 10% polyacrylamide gel with a thickness of 1.5 mm and stained with GelRed™ (Biotium company).

8. The SRAP and SSR markers were coded as presence/absence as molecular binary data. Markers with minor allele frequency (MAF) less than 0.05 and markers with more than 15% missing data were removed from analyses. After quality control checking, a total of 684 binary markers remained for further analyses.

Inference of population structure and family-based relatedness

9. The population structure of 143 barley accessions was inferred by the Bayesian clustering method using STRUCTURE v.2.3.4 (Pritchard et al. 2000). We modeled a burn-in period of 50,000 cycles followed by 100,000 Markov Chain Monte Carlo (MCMC) iterations. As no prior information regarding the number of clusters (K) were used, we presumed 1-10 sub-populations and repeated 20 times for each K value for the stability of each model set for no-admixture and correlated allele frequencies. The most probable number of sub-populations (clusters) was determined by plotting the quantity of ΔK as a function of K (Evanno et al. 2005) using the online software STRUCTURE HARVESTER (Earl and von Holdt 2012) to generate membership coefficient (Q) matrix. In order to assign each individual to a particular group a cut-off limit of 60% membership (Q-matrix) was considered (Jacob et al.2014). Accessions with less than this criterion were considered as admixed individuals.

10. Analysis of molecular variance (AMOVA) was performed using GenAlex v.6.41(Peakall and Smouse 2006) using 999 permutations to estimate population differentiation among pre-defined sub-populations i.e. number of ear rows (NER) (two-rowed vs. six-rowed), and SGH (winter type vs. spring type). Principal coordinate analyses (PCoA) were run on Nei genetic distance matrix (Nei 1973). Furthermore, unweighted neighbor joining (NJ) dendrogram was constructed by DARwin V.5.0.158 software (Perrier and Jacquemoud-Collet 2006) based on Jaccard's dissimilarity matrix calculated in NTSYS v.2.02 software (Rohlf 1998). All population structure analyses were carried out on molecular data of 147 SSR marker loci.

Statistical Analysis

11. Shapiro–Wilk normality test was used to analyze the distributions of the nine quantitative DUS characteristics at 0.05 significance level. The association of DUS characteristics with NER groups was tested by using Fisher's exact chi-square test for binary and nominal characteristics, Mann-Whitney U test for ordinal characteristics and Student t test for quantitative characteristics.

Association analysis of DUS characteristics

12. A general linear model (GLM) and a mixed linear model (MLM) (Yu et al. 2006) were used for testing associations between markers and DUS characteristics using TASSEL version 3.0 software (Bradbury et al. 2007). The membership coefficient matrix (Q) was harvested from STRUCTURE analysis where ΔK was at higher value and incorporated into the GLM approach as covariate to correct for the effect of population structure. For MLM analysis, Jaccard's dissimilarity coefficients between individuals were added to the model (Q+KJ) to account for relatedness between individuals. Moreover, kinship values (scaled between 0 and 2) computed by TASSEL was also incorporated into the other MLM method (Q+KT).

13. Furthermore, allelic association tests (AAT) were used for categorical (nominal and binary) characteristics based on allele frequencies between states using PLINK v.1.07 software (Purcell et al. 2007). Moreover, Cochran-Mantel-Haenszel test for 2x2xK stratified tables were used in stratified analysis (SA) using PLINK, in which NER groups was considered as K to consider the population structure in the analysis.

14. The P values of associated markers were adjusted for multiple testing using Bonferroni correction criterion (P value× number of markers), and those were selected that exceeded the significance threshold (α/n) to control false discovery rate (FDR) at 0.05 and 0.01 significance levels. Therefore, two conservative (0.05/684) and stringent (0.01/684) cut-offs were employed in all association analyses.

RESULTS

Phenotyping statistical results

15. Among 27 categorical morphological characteristics observed in the panel of barley accessions, four characteristics had similar (monomorphic) states. All 80 two-rowed varieties had sterile spikelets with full development (ch. EDSS). The grains of all samples were husked (ch. GH), hairless in ventral furrow (ch. GHVF), and bearing clasping lodicules (ch. GDL).The states of remaining categorical (ordinal, binary, and nominal) characteristics were polymorphic between samples (SM 3).

16. Shapiro–Wilk normality test showed that two quantitative characteristics (time of ear emergence and plant length) out of nine were normally distributed ($P = 0.25$ and 0.36 respectively), and the other seven characteristics did not fit a normal distribution ($P < 0.01$).

Genetic structure of barley panel

17. By investigating SSR allele frequencies in NER and SGH groups, we found that just more than half (0.56) of the alleles had frequencies differing more than two fold between two-rowed and six-rowed groups. In that respect, 4 and 9 percent of alleles were only present in respective row groups. Similarly, nearly two-third (0.61) of allele frequencies differed more than two fold between spring and winter groups, with 12 and 9 percent of them existing only in either SGH type respectively. These data suggest strong population structure effect in the panel of 143 barley landraces and advanced breeding lines. The population structure was inferred using 147 SSR markers by the Bayesian clustering method in STRUCTURE software. The higher quantity for ΔK was detected at K=2 (Fig. 1) implying the existence of two sub-populations in the barley samples. The constructed membership (Q) matrix at K=2 made two sub-populations primarily due to the divergence in NER of barley samples (Fig. 2). Similarly, the dendrogram of unweighted neighbor joining algorithm on Jaccard's dissimilarity matrix, clustered the individuals in two major groups (six- and two-rowed samples) and a minor sub-group (consisting 10 two-rowed samples) based on NER (Fig.2). These results show the consistency of hierarchical clustering procedure with that of model-based method through Bayesian statistics.

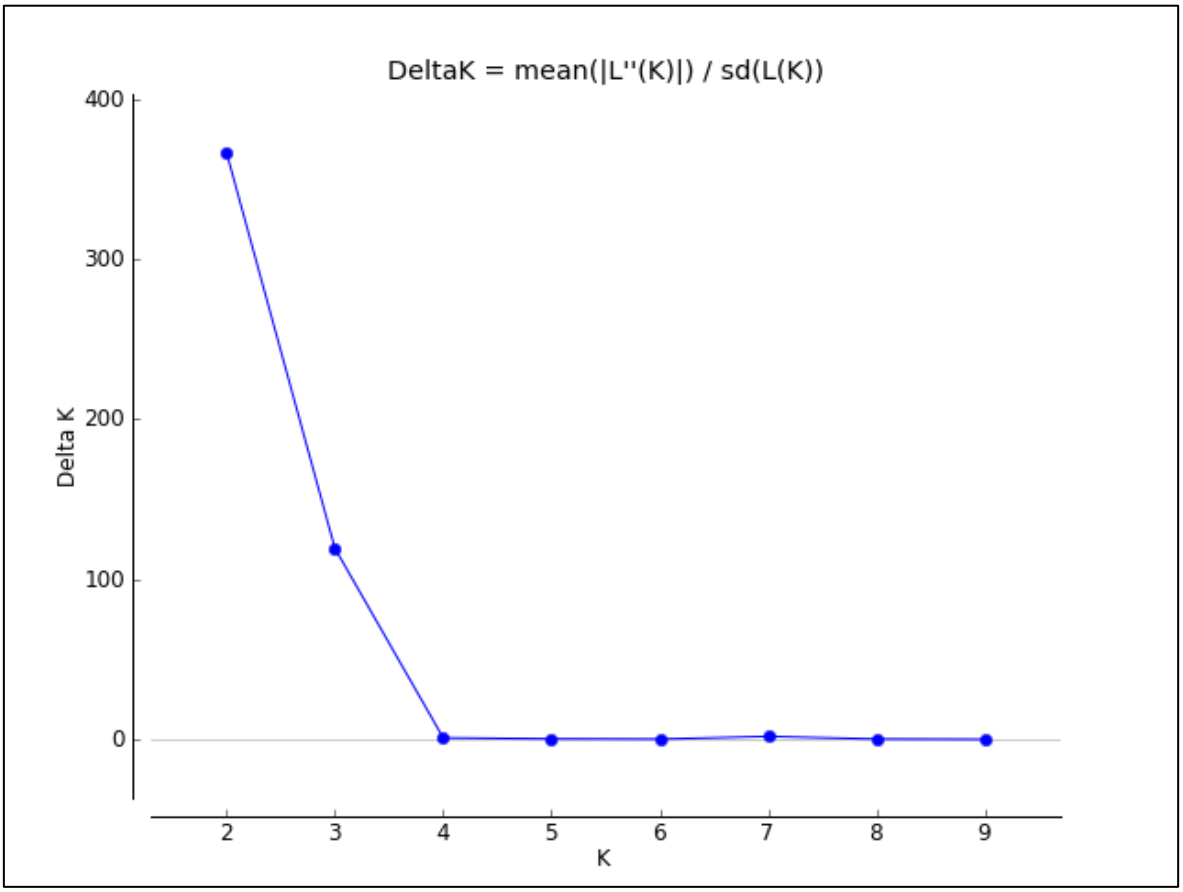

Fig. 1 Delta K values as a function of K, according to Evanno et al. (2005). The higher value (K=2) was used to build Q-matrix, in which two clusters represented the number of ear rows of 143 barley samples.

18. The allocation of each individual to either of Q1 and Q2 groups considering a threshold of 60% membership (Jacob et al.2014) essentially corresponded to NER of barley accessions, leaving two (Nos. 122 and 142) and six (Nos. 40, 41, 125, and 143) samples with admixed structure. In the resulting Q-matrix, all 63 six-rowed samples fell in Q1 sub-population except one individual (No. 32). Interestingly, this individual resided in two-rowed group derived from NJ tree. Besides, 13 individuals of 80 two-rowed samples remained in Q1 (six-rowed) group, of which 10 individuals stayed in a separate group within the dendrogram (Fig. 2).

19. Moreover, the extent of genetic differentiation among predefined sub-populations was measured by AMOVA and PCoA analyses. The percentage of molecular variance among sub-populations (Rst) was 0.24 when considering the NER as sub-populations. This value dropped to 0.12 when the SGH of samples were regarded as sub-populations. In the PCoA, the percentage of variation explained by the first three axes was 33.5%, of which PCoA1 and PCoA2 contributed 20.5 and 8% respectively of the SSR variation among 143 accessions. Therefore, plotting the first two principal coordinates based on the NER gave a better differentiation than SGH grouping (Fig. 2).

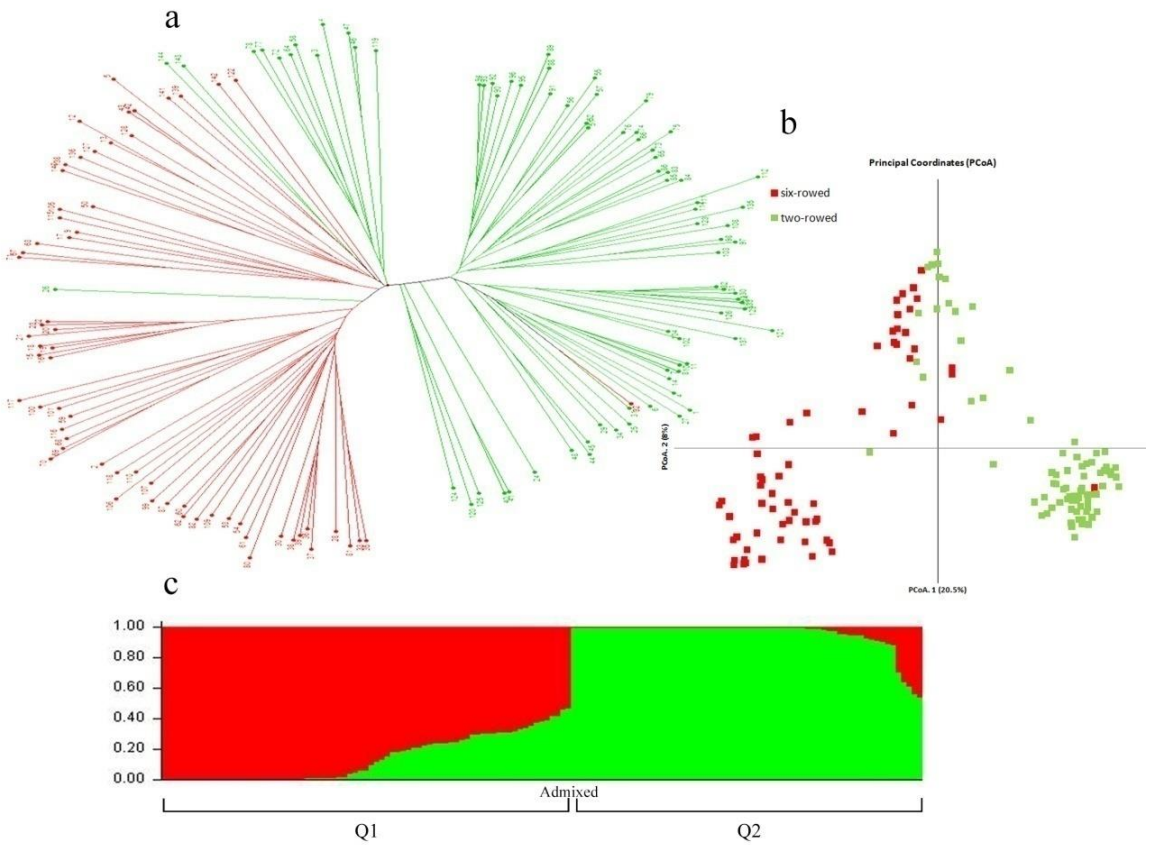

Fig 2. Differentiating six- (red) and two-rowed (green) 143 barley samples with unweighted neighbor joining tree based on Jaccard's dissimilarity (a), principal coordinate analysis on two first axes (b), and model-based clustering method based on Bayesian statistics into Q1 and Q2 clusters (c)

Comparison of association models

20. Association analyses for 36 DUS characteristics (quantitative and categorical) were performed using GLM and MLM models. Using GLM model, a total of 66 marker-characteristic associations were found considering -log10 Bonferroni threshold (0.05/684= 4.13). These associations reduced to 48 when a more stringent threshold (0.01/684=4.83) was considered. By employing two kinship coefficients in MLM models, 61 and 72 marker-characteristic associations in which only one association was found that surpassed the cut-off at both significance levels. These data suggest that the GLM model outperforms two MLM models for controlling spurious associations. Incorporating Q-matrix in GLM was sufficient to correct the effect of population structure, as entering other factors (kinship coefficients) in MLM models might have resulted in over-correction. Moreover, for both binary and nominal (treated as binary) the results of SA and AAT analyses based on allele frequencies between morphological states were consistent with nearly all results of GLM model.

Association of DUS characteristics with NER groups

21. The difference between four quantitative DUS characteristics (ear length, ear density, 1000-seed weight, and coleoptile length) of NER groups was significant at 0.05 level using Student t test. The remaining five continuous characteristics did not differ between six- and two-rowed barley samples (Fig 3.). For all ordinal characteristics, Mann-Whitney U test indicated significant values ($P < 0.05$) between NER groups, except for two characteristics i.e. anthocyanin color of lemma nerves, and awn length. Fisher's exact test for contingency tables of all binary and nominal DUS characteristics showed significance ($P < 0.05$) between NER groups, but not for ear shape and spiculation of inner lateral nerves of lemma.

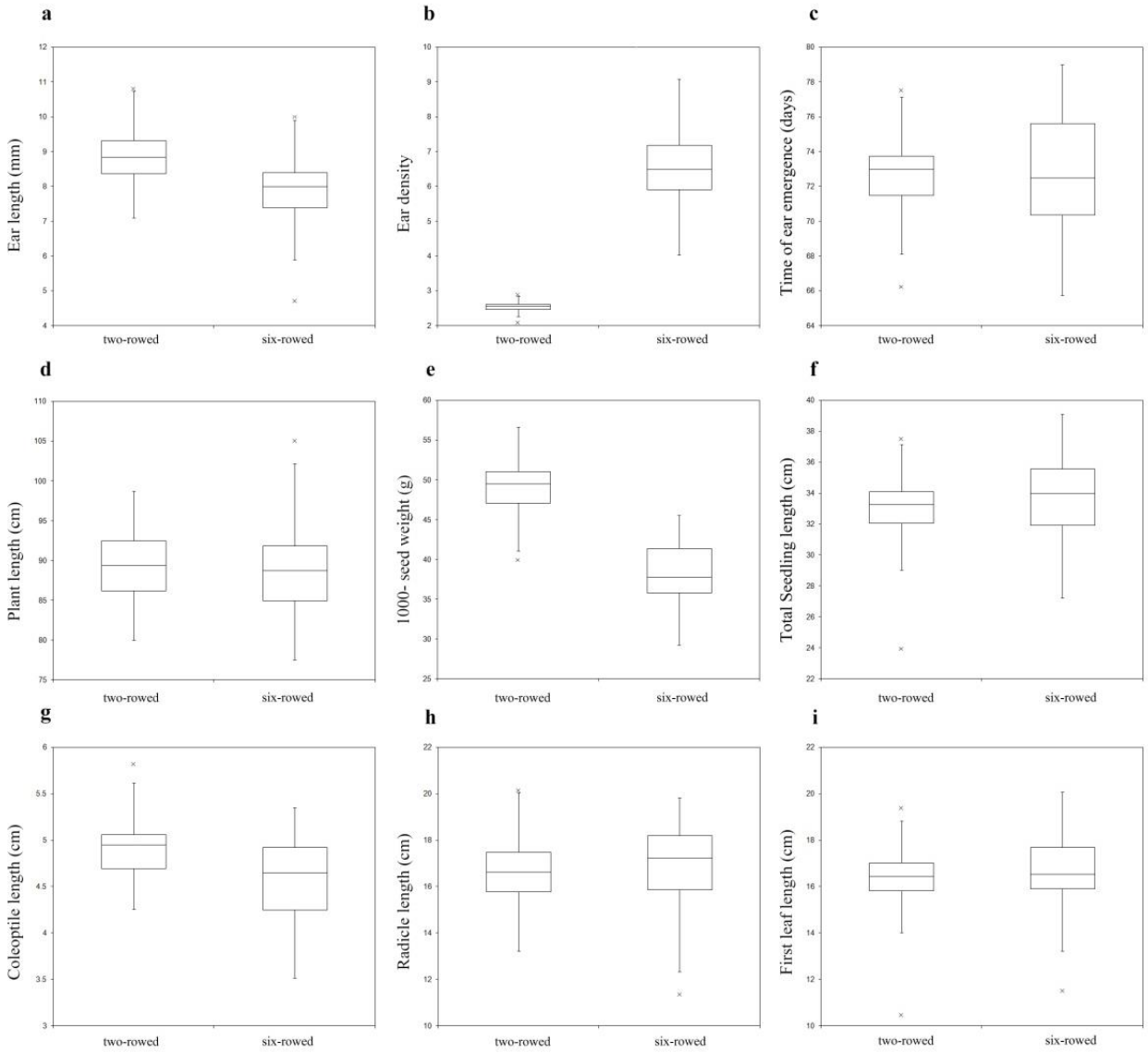

Fig 3. Box plots of nine quantitative DUS characteristics grouped by NER. The 143 inbred lines were differentiated into two- and six-rowed sub-populations by model-based Bayesian clustering and NJ tree. Student t test was used to estimate the significance of difference between NER means across two years. Four characteristics (a, b, e, g) were significantly associated with two- and six-rowed groups. Legends, a: ear length; b: ear density; c: time of ear emergence; d: plant length; e: 1000-seed weight; f: total seedling length; g: coleoptile length; h: radical length; i: first leaf length.

Associations of DUS characteristics and markers

Quantitative DUS characteristics

22. The GLM approach resulted in finding association between four (out of nine quantitative DUS characteristics) and SSR markers. Among these four characteristics, student t test showed that only two quantitative characteristics (ear density and 1000-seed weight) were associated ($P < 0.05$) with NER groups. When considering ear density as ordinal characteristic, the Mann-Whitney U value was also significant between six- and two-rowed groups.

23. The GLM analysis resulted in detection of one associated SSR marker (GBM1400) on chromosome 6 with ear density (ED). This marker showed association (R2=0.20) when phenotypic states of ED was both considered as quantitative (number of grains divided by ear length) and ordinal (very lax, lax, medium, dense, and very dense) in separate analyses. SSR marker GBM1221 on chromosome 4 also showed association (R2=0.16) with ED, when a less conservative Bonferroni threshold (4.13) was considered. Three SSR

markers (EBMAC560, BMAG0518, BMAC0113) were found to be significantly associated with time of ear emergence (TEE). These markers were located on chromosomes 1, 2, and 5, and explained 31, 29, and 24 percent of phenotypic variance. Another SSR marker (GBM1309) on chromosome 2 was also detected when cut-off at 4.13 was considered. The GLM analysis revealed the association of only one SSR marker i.e. GBM1464 (on chromosome 7) with plant length (PL) with contribution of 19 percent of variation. Two SSR markers i.e. GBM1463 (chromosome 5), and GBM1400 (chromosome 6) were associated with 1000-seed weight (SW). These two markers explained 12 % of SW variation. When considering the conservative cut-off (4.13) two other markers (R2=0.09) i.e. GBM1293 (chromosome 5) and GBM1299 (chromosome 4) were also added to previous associated markers (see Annex I, Table 1).

Ordinal DUS characteristics

24. The association of six ordinal DUS characteristics i.e. intensity of anthocyanin coloration of awn tips (AIAC), curvature of first segment of rachis (RCFS), awn roughness (AR), ear density (ED), color of grain aleurone layer (KCAL), and length of glume relative to grain (MSLG) with SSR markers was revealed in results of GLM model. Mann-Whitney U test showed that all these characteristics were significant ($P < 0.05$) between NER groups.

25. By considering stringent Bonferroni threshold (4.83), AIAC associated with BMAG0518 (chromosome 2) and EBMAC0541 (chromosome 3) wherein 22 and 19% of phenotypic variation was explained by this SSR marker respectively. Four markers (GBM1400, GBM1309, GBM1463, and GBM1221) on chromosomes 6, 2, 5, and 4 were associated with AR when phenotypes of barley awns were considered as ordinal (smooth, intermediate, and rough) states. GBM1400 explained 40% of AR variation and also was found in both MLM models, while the other three markers contributed in 21-27 percent of it. When we added AR states as binary (smooth vs. rough) values into the GLM analysis, GBM1309 was common with previous results. Moreover, two other SSR markers (BMAG0345 and GBM1110) showed also association with AR. Similarly, two separate values as ordinal (whitish, weakly colored, and strongly colored) and binary (whitish vs. colored) were incorporated into GLM analysis of KCAL. Five SSR markers (BMAG0345, GBM1461, GMS116, EBMAC0624, and BMAG0807) associated with KCAL when phenotypes were regarded as ordinal values. These markers explained 23-31% of ordinal variation. By considering values of KCAL as binary states, just one marker (BMAG0345) exceeded the 4.83 threshold. Only one marker (SCSSR18076) on chromosome 5 with contribution of 18% of phenotypic variation, associated with MSLG. Furthermore, GBM1400 (chromosome 6) and GBM1299 (chromosome 4) were associated with RCFS while explaining 11 and 9% of characteristic variation respectively (see Annex I, Table 1).

Binary and nominal DUS characteristics

26. For binary and nominal DUS characteristics those associated markers were selected which surpassed the stringent threshold (4.83) in all three analyses i.e. the GLM (using TASSEL software) and AAT and SA (using PLINK software). Therefore, a total of 26 marker-characteristic associations were found for nine characteristics. Fisher's exact test for contingency tables of these characteristics showed significance ($P < 0.05$) between NER groups except for rachilla hair type of grain.

27. One marker (BMAC0310) on chromosome 4 found to be associated with spiculation of inner lateral nerves of lemma (GSLN) which explained 34% of phenotypic variation. No markers exceeded the thresholds when the values of GSLN were considered as ordinal states (absent or very weak, weak, medium, strong, and very strong) in the GLM method. The most number of associated markers (eight SSR and one SRAP markers) were found in analysis of anthocyanin color of awns tips (AACT), with the most of them residing on chromosome 2 of barley. Total variation (R2) explained by these markers ranged 0.17-0.45 with the highest value for BMAG0518 SSR marker. Associations of four markers (HVI3, GBM1408, BMAG0140, and HV13GEIII) with rachilla hair type of grain (GRHT) were revealed in all three analyses while exceeding the stringent cut-off. These markers contributed in 25-34 of total GRHT variation. SSr markers EBMAC0624 (R2=0.18) and BMAG0345 (R2=0.16) on chromosomes 6 and 1 respectively were associated with grain color. There was only one SSR marker (EBMAC0788) on chromosome 4 that showed association with attitude of sterile spikelets (SSA) in two-rowed barley samples. Moreover, in these two-rowed samples four markers (EBMAC0624, BMAG0504, GBM1525, BMAG0807) were found to be associated tip shape of sterile spikelets (SSTS). Furthermore, hairiness of lower leaves sheats (LLHL) was associated with EBMAC679 (chromosome 4) explaining 21% of LLHL variation (see Annex I, Table 2).

CONCLUSION

28. The 143 barley landraces we used in the present study represented a diverse panel in terms of morphological DUS characteristics. On the other hand, SSR and SRAP markers grouped these samples based on the number of ear rows when both hierarchical and model-based clustering methods were employed. The resulting population structure was taken into account in linear models (GLM and MLM) and chi-square-based tests (SA and AAT) for quantitative and categorical DUS characteristics respectively. By considering Bonferroni threshold and controlling false discovery rate, a number of associated markers with characteristics were found. These markers represent valuable functional markers to predict DUS characteristics. These markers are especially cost-effective for test authorities with less access to high-throughput SNP genotyping platforms.

Seyed Hossein JAMALI [a], Seyed Abolghasem MOHAMMADI [b], Leila SADEGHI [a], Maryam NAJAFIAN [a], Behzad SADEGHZADEH [c]

[a] Plant Variety Identification & Registration Dep., Seed & Plant Certification & Registration Institute (SPCRI), Karaj, Islamic Republic of Iran
[b] Plant Breeding and Biotechnology Dep., College of Agriculture, University of Tabriz, Islamic Republic of Iran
[c] Cereals Research Dep., Dryland Agricultural Research Institute, Maragheh, Islamic Republic of Iran

Abbreviations:

SGH:	seasonal growth habit
NER:	number of ear rows
SRAP:	Sequence-related amplified polymorphism
SSR:	Simple Sequence Repeats
DUS:	Distinctness, Uniformity, and Stability
NJ:	neighbor joining
GLM:	General Linear Model
MLM:	Mixed Linear Model
AAT:	allelic association test
SA:	Stratified Analysis

[Annex I follows]

ANNEX I

Table 1. Significant markers exceeding the stringent Bonferroni threshold (4.83) detected in general linear model (GLM), stratified analysis (SA), and allelic association test (AAT)

DUS Characteristic	Marker	Map	Associated Allele	R^2	-log10 adjusted P		
					GLM	SA	AAT
Anthocyanin color of awns tips	BMAG0518_SSR9	2	5	0.45	24.26	17.36	18.43
	SCSSR02306_SSR18	5	1	0.31	13.52	13.78	14.19
	BMAC0213_SSR59	1	8	0.26	9.87	10.38	11.95
	GBM1366_SSR64	2	1,2	0.22	7.77	6.58	9.88
	GBMAS0183_SSR51	3	4	0.23	7.52	5.44	8.42
	EBMAC0854_SSR60	2	3	0.19	6.61	13.02	9.88
	EBMATC0039_SSR33	2	1,2	0.19	5.92	6.01	9.89
	me4-em5_SRAP	–	4	0.17	5.75	6.13	8.24
	GBMS0160_SSR102	2	8	0.17	5.64	5.06	7.27
Awn roughness	GBM1309_SSR2	2	1,2	0.30	10.36	15.08	8.74
	BMAG0345_SSR84	1	3	0.23	7.01	9.99	7.62
	GBM1110_SSR20	3	1	0.22	6.62	10.4	5.47
Grain color	EBMAC0624_SSR30	6	3	0.18	6.07	13.19	8.73
	BMAG0345_SSR84	1	3	0.16	5.15	8.84	6.50
Rachilla hair type of grain	HVI3_SSR137	–	1	0.34	9.86	8.42	8.15
	GBM1408_SSR141	2	1,2	0.25	6.90	8.17	6.01
	BMAG0140_SSR81	2	1	0.25	6.58	8.89	5.90
	HV13GEIII_SSR74	3	1	0.25	5.97	8.82	5.16
Spiculation of lateral nerves of lemma	BMAC0310_SSR108	4	2	0.34	10.12	18.24	7.38
Color of grain aleurone layer	BMAG0345_SSR84	1	3	0.33	10.56	17.73	9.23
Attitude of sterile spikelets	EBMAC0788_SSR120	4	8	0.32	8.68	13.65	7.03
Tip shape of sterile spikelets	BMAG0504_SSR29	1	1,3	0.28	9.53	9.58	9.81
	GBM1525_SSR73	4	1	0.20	5.89	4.07	6.61
	BMAG0807_SSR117	6	1	0.20	5.39	6.63	6.63
	EBMAC0624_SSR30	6	3	0.17	7.08	14.5	10.00
Hairiness of lower leaves sheats	EBMAC679_SSR10	4	2	0.21	4.96	8.89	4.56

Table 2. Significant associated markers detected in general linear model (GLM) and two mixed linear models (MLM) in which Jaccard's (K_J) and kinship coefficients calculated by TASSEL (K_T) were incorporated as co-factors. Markers exceeding the stringent Bonferroni threshold (4.43) by GLM method are denoted in bold font.

Characteristic	Marker	Map	Associated allele	GLM		MLM (Q+K_J)		MLM (Q+K_T)	
				-log10 adjusted P	R^2	-log10 adjusted P	R^2	-log10 adjusted P	R^2
Intensity of anthocyanin color of awns tips	**BMAG0518_SSR9**	2	5	6.55	0.22	–	–	–	–
	EBMAC0541_SSR90	3	2	5.40	0.19	–	–	–	–
	BMAC0213_SSR59	1	1	4.57	0.18	–	–	–	–
Awn roughness	**GBM1400_SSR47**	6	3	13.10	0.40	3.64	0.21	1.89	0.14
	GBM1309_SSR2	2	1	7.73	0.27	0.23	0.08	–	–
	GBM1463_SSR67	5	1	6.11	0.23	0.27	0.09	–	–
	GBM1221_SSR1	4	1	5.53	0.21	–	–	–	–
	GBM1438_SSR66	5	3	4.61	0.20	–	–	–	–
	BMAG0345_SSR84	1	3	4.64	0.19	–	–	–	–
	GBM1461_SSR4	1	5	4.28	0.18	–	–	–	–
Ear density	**GBM1400_SSR47**	6	3	5.58	0.20	–	–	–	–
	GBM1221_SSR1	4	1	4.18	0.16	–	–	–	–
Color of grain aleurone layer	**BMAG0345_SSR84**	1	3	10.37	0.31	–	–	–	–
	GBM1461_SSR4	1	5	8.82	0.28	–	–	–	–
	GMS116_SSR107	3	2	7.30	0.25	–	–	–	–
	BMAG0807_SSR117	6	1	6.43	0.24	–	–	–	–
	EBMAC0624_SSR30	6	3	6.53	0.23	–	–	–	–
	GBM1438_SSR66	5	1	4.82	0.19	–	–	–	–
	HVM40_SSR43	4	2	4.25	0.17	–	–	–	–
length of glume (relative to grain)	**SCSSR18076_SSR37**	5	1,2	7.22	0.18	–	–	–	–
Plant length	GBM1464_SSR35	7	7	4.63	0.19	–	–	–	–
Curvature of first segment of rachis	**GBM1400_SSR47**	6	2,3	7.30	0.11	–	–	–	–
	GBM1299_SSR94	4	2	6.19	0.09	–	–	–	–
1000-seed weight	**GBM1463_SSR67**	5	1	7.45	0.13	–	–	–	–
	GBM1400_SSR47	6	3	5.74	0.12	–	–	–	–
	GBM1293_SSR61	5	2	4.57	0.09	–	–	–	–
	GBM1299_SSR94	4	1	4.37	0.09	–	–	–	–

Table 2. Significant associated markers detected in general linear model (GLM) and two mixed linear models (MLM) in which Jaccard's (K_J) and kinship coefficients calculated by TASSEL (K_T) were incorporated as co-factors. Markers exceeding the stringent Bonferroni threshold (4.43) by GLM method are denoted in bold font.

Characteristic	Marker	Map	Associated allele	GLM		MLM ($Q+K_J$)		MLM ($Q+K_T$)	
				-log10 adjusted P	R^2	-log10 adjusted P	R^2	-log10 adjusted P	R^2
Time of ear emergence	**EBMAC560_SSR100**	1	1,2	9.60	0.31	–	–	–	–
	BMAG0518_SSR9	2	5	8.56	0.29	–	–	–	–
	BMAC0113_SSR7	5	6	6.24	0.24	–	–	–	–
	GBM1309_SSR2	2	1	4.37	0.19	–	–	–	–
Anthocyanin color of nerves of lemma	BMAG0740_SSR86	4	9	4.34	0.20	2.27	0.16	–	–

[Annex II follows]

ANNEX II

REFERENCES

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex characteristics in diverse samples. Bioinformatics 23:2633–2635

CPVO, Community Plant Variety Office (2012) Protocol for distinctness, uniformity and stability tests, *Hordeum vulgare* L. *sensu lato*, barley. CPVO-TP/019/3 Final.

Earl DA, vonHoldt BM (2012) Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. Conser Genet Resour 4:359–361

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.

Jakob SS, Rodder D, Engler JO, Shaaf S, Ozkan H, Shaaf, S., Özkan, H., Blattner, F.R. and Kilian, B. (2014) Evolutionary history of wild barley (*Hordeum vulgare* subsp. *spontaneum*) analyzed using multilocus sequence data and paleodistribution modeling. Genome Biol Evol 6: 685–702.

Nei M (1973) Analysis of gene diversity in subdivided populations. PNAS 70:3321–332

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6(1):288–295

Perrier X and Jacquemoud-Collet JP (2006) DARwin software. http://darwin.cirad.fr/darwin

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus Genotype Data. Genetics 155: 945–959.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. Am J Hum Genet 81:559-575

PPV & FRA, Protection of Plant Varieties and Farmers' Rights Authority of Government of India (2011) Guidelines for the conduct of test for distinctiveness, uniformity and stability on barley (*Hordeum vulgare* L.)

Rohlf FJ (1998) NTSYS-pc Numerical Taxonomy and Multivariate Analysis System, Version 2.02. Exeter Software, New York

Saghai Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA Spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81 (24):8014-8018

UPOV, International Union for the Protection of New Varieties of Plants (1994) Guidelines for the conduct of tests for distinctness, uniformity and stability, barley (*Hordeum vulgare* L. *sensu lato*). TG/19/10.

UPOV, International Union for the Protection of New Varieties of Plants (2013) Guidance on the Use of Biochemical and Molecular Markers in the Examination of Distinctness, Uniformity and Stability (DUS), TGP/15.

Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

[Annex III follows]

ANNEX III

SUPPLEMENTARY MATERIALS (SM)

SM 1. List of 143 barley samples used in present study with number of ear rows (NER) and seasonal growth habit (SGH) of genotypes

No	Name/Origin	NER	SGH
1	England1	2	S
2	England2	6	W
3	Algeria1	2	W
4	Algeria2	2	S
5	Iran43	6	W
6	USA1	2	S
7	Russia1	2	S
8	Russia2	2	S
9	Spain1	6	W
10	Spain2	2	S
11	Egypt1	2	S
12	Egypt2	6	F
13	Egypt3	6	S
14	Egypt4	2	S
15	Egypt5	6	W
16	Egypt6	6	W
17	Egypt7	6	W
18	Egypt8	6	W
19	Egypt9	6	W
20	Egypt10	6	W
21	Egypt11	6	W
22	Egypt12	6	W
23	India	6	W
24	Ethiopia	2	S
25	Russia3	2	S
26	Pakistan1	2	S
27	Pakistan2	2	S
28	China1	6	W
29	China2	2	S
30	China3	6	W
31	China4	2	S
32	China5	6	W
33	China6	6	W
34	China7	2	S
35	China8	6	W
36	China9	6	W
37	China10	6	W
38	China11	6	W
39	China12	6	W
40	China13	2	S
41	China14	2	S
42	China15	6	S
43	China16	6	S

No	Name/Origin	NER	SGH
49	Iran/Miyandoab1	2	W
50	Iran/Unknown1	6	W
51	Iran/Korand	2	S
52	Iran/Unknown2	2	W
53	Iran/Unknown3	6	W
54	Iran/Ghazvin	6	W
55	Iran/Unknown4	2	S
56	Iran/Unknown5	2	S
57	Iran/Unknown6	2	S
58	Iran/Bojnord1	2	W
59	Iran/Bojnord2	6	W
60	Iran/Bojnord3	6	W
61	Iran/Bojnord4	6	W
62	Iran/Bojnord5	6	W
63	Iran/Golpayegan1	6	S
64	Iran/Golpayegan2	2	S
65	USA2	2	S
66	Iran/Azarbaijan1	2	S
67	Iran/Azarbaijan2	6	W
68	Iran/Tehran	6	W
69	Iran/Azarbaijan4	6	W
70	Iran/Kerman1	6	W
71	Iran/Kerman2	6	W
72	Iran/Gorgan1	2	W
73	Iran/Gorgan2	2	S
74	Iran/Gorgan3	2	S
75	Iran/Gorgan4	2	S
76	Iran/Kerman3	2	S
77	Iran/Kerman4	2	W
78	Iran/Unknown7	2	W
79	Iran/Miyandoab2	2	F
80	Iran/Unknown8	6	W
81	Iran/Unknown9	6	W
82	Iran/Unknown10	6	W
83	Iran/Unknown11	2	S
84	Iran/Unknown12	2	S
85	Iran/Unknown13	2	S
86	Iran/Unknown14	2	S
87	Iran/Unknown15	6	W
88	Iran/Unknown16	2	W
89	Iran/Unknown17	2	F
90	Iran/Unknown18	2	F
91	Iran/Unknown19	2	W

No	Name/Origin	NER	SGH
97	Iran/Unknown25	2	F
98	Iran/Unknown26	2	F
99	Iran/Unknown27	6	W
100	Iran/Unknown28	6	W
101	Iran/Unknown29	6	W
102	China22	2	F
103	China23	2	F
104	Iran/Unknown30	2	F
105	Iran/Unknown31	6	W
106	Iran/Torbat-E-Jam	6	W
107	Iran/Azarbaijan3	6	W
108	Iran/Miyandoab3	2	S
109	Iran/Unknown32	6	W
110	Iran/Unknown33	6	W
111	Iran/Unknown34	2	S
112	Iran/Unknown35	2	F
114	Iran/Unknown37	2	S
115	Iran/Unknown38	6	W
116	Iran/Unknown39	6	W
117	Iran/Unknown40	6	W
118	Iran/Unknown41	6	W
119	Iran/Unknown42	2	S
120	CWB117-77	2	S
121	Tokak/Demir-2	2	S
122	Zarjau-80	6	W
123	AZE-ICB	2	S
124	CWB117-5	2	S
125	ICB01-1402	2	S
126	Alpha/Gumhuriyet	2	S
127	Rihane-03	6	W
128	Makoee	6	W
129	Sahand	2	S
130	Abidar	2	S
131	Dayton	2	W
132	Yea-168	2	S
133	Denmark	2	S
134	Obruk-86	2	S
135	Gara-Arpa	2	S
136	Ec-79	6	S
137	Bolbol	2	W
138	Dikto	6	W
139	Radical	6	W
140	Dobrynya	6	W

No	Name/Origin	NER	SGH
44	China17	2	S
45	China18	2	S
46	China19	2	NA
47	China20	2	W
48	China21	2	S

No	Name/Origin	NER	SGH
92	Iran/Unknown20	2	S
93	Iran/Unknown21	2	S
94	Iran/Unknown22	2	F
95	Iran/Unknown23	2	F
96	Iran/Unknown24	2	S

No	Name/Origin	NER	SGH
141	Ec-80	6	W
142	Ec-84	6	S
143	Erb-86	2	S
144	Erb-87	2	S

SM 2. List of 36 DUS characteristics used in present study.

No.	Ch. code	Type of variable	Time of Obs.*	Characteristics	Ch. Abbr.	State of Expression (Note)
1	I-27	N	00	Grain: color	GC	white (1), yellow (2), green (3), black (4)
2	U-28	O	00	Kernel: color of aleurone layer	KCAL	whitish (1), weakly colored (2), strongly colored (3)
3		C	00	1000-seed weight	SW	Measurement
4		C	05	Radicle length	RL	Measurement- 7 days after germination
5		C	07	Coleoptile length	CL	Measurement- 7 days after germination
6		C	10	First leaf length	FLL	Measurement- 7 days after germination
7		C	10	Total seedling length (first leaf plus radicle)	TSL	Measurement- 7 days after germination
8	U-2	B	25-29	Lowest leaves: hairiness of leaf sheats	LLHL	Absent (1), present (9)
9	U-1	O	25-29	plant: growth habit	PGH	Erect (1), semi-erect (2), intermediate (3), semi prostrate (4), prostrate (5)
10	U-3	B	45-49	Flag leaf: anthocyanin coloration of auricles	FLAC	Absent (1), present (9)
11	U-4	O	45-49	Flag leaf: intensity of anthocyanin coloration of auricles	FLIA	very weak (1), weak (3), medium (5), strong (7), very strong (9)
12	C-5	O	49-51	Flag leaf: attitude	FLA	erect (1), semi-erect (3), horizontal (5), semi-drooping (7), drooping (9)
13	U-7	O	50-52	Time of ear emergence	TEE	very early (1), early (3), medium (5), late (7), vary late (9)
14	U-6	O	50-60	Flag leaf: glaucosity of sheet	FLGS	very weak (1), weak (3), medium (5), strong (7), very strong (9)
15	U-8	B	60-65	Awns: anthocyanin coloration of tips	AACT	Absent (1), present (9)
16	U-9	O	60-65	Awns: intensity of anthocyanin coloration of tips	AIAC	very weak (1), weak (3), medium (5), strong (7), very strong (9)
17	I-13	O, B	60-65	Awn: roughness	AR	O: smooth (3), intermediate (5), rough (7) B: smooth (1), rough (9)
18	U-24	O	80-85	Grain: anthocyanin coloration of nerves of lemma	GACN	absent or very weak (1), weak (3), medium (5), strong (7), very strong (9)
19	U-17	O	80-92	Awn: length (compared to ear)	AL	short (3), medium (5), long (7)
20	U-15	C, O	80-92	Ear: density	ED	C: Measurement O: very lax (1), lax (3), medium (5), dense (7), very dense (9)
21	U-16	O	80-92	Ear Length	EL	Very short (1), short (3), medium (5), long (7), very long (9)
22	U-13	B	80-92	Ear: number of rows	ENR	two (1), more than two (9)
23	U-14	N	80-92	Ear: shape	ESh	tapering (3), parallel (5), fusiform (7)
24	U-22	B	80-92	Grain: rachilla hair type	GRHT	short (1), long (2)
25	U-12	O	80-92	Plant Length	PL	Very short (1), short (3), medium (5), long (7), very long (9)
26	W-8	N	92	collar: type	CT	recurred (1), platform (2), cup (3)
27	C-19	B	92	Ear: development of sterile spikelets	EDSS	non or rudimentary (1), full (2)
28	U-27	B	92	Grain: disposition of lodicules	GDL	frontal (1), clasping (2)
29	U-23	B	92	Grain: husk	GH	Absent (1), present (9)
30	U-26	B	92	Grain: hairiness of ventral furrow	GHVF	Absent (1), present (9)

31	U-25	O, B	92	Grain: spiculation of inner lateral nerves of dorsal side of lemma	GSLN	O: absent or very weak (1), weak (3), medium (5), strong (7), very strong (9) B: Absent (1), present (9)
32	U-21	O	92	Median spikelet: length of glume and its awn relative to grain	MSLG	shorter (1), equal (2), longer (3)
33	U-19	O	92	Rachis: curvature of first segment	RCFS	absent or very weak (1), weak (3), medium (5), strong (7), very strong (9)
34	U-18	O	92	Rachis: Length of first segment	RLFS	O: Short (3), medium (5), long (7), very long (9)
35	U-20	N	92	Sterile spikelet: attitude (in mid-third of ear)	SSA	Parallel (1), parallel to weakly divergent (2), divergent (3)
36	W-32	N	92	sterile spikelet: tip shape	SSTS	pointed (1), rounded (2), squared (3)

Abbreviations: U=UPOV's DUS test guideline, I=India's DUS test guideline, C=CPVO's DUS test protocol, B=Binary variable, O=Ordinal variable, N=Nominal Variable, C=Continuous variable, * Time of observation according to Zadoks two-digit growth scale

SM 3. Results of morphological (ordinal, binary, and nominal) variables measured or observed on 143 barley samples

Ch. Abbr.	Type of variable	Characteristics	State of expression (No., Frequency of varieties out of 143)
KCAL	O,B	Kernel: color of aleurone layer	O: whitish (72,0.5), weakly colored (19,0.13), strongly colored (52,0.36) B: whitish (72, 0.5), colored (71, 0.5)
PGH	O	Plant: growth habit	Erect (6,0.04), semi-erect (74,0.52), intermediate (59,0.41), semi prostrate (4, 0.03)
LLHL	B	Lowest leaves: hairiness of leaf sheats	Absent (138, 0.97), present (5,0.03)
FLAC	B	Flag leaf: anthocyanin coloration of auricles	Absent (58,0.41), present (85,0.49)
FLIA	O	Flag leaf: intensity of anthocyanin coloration of auricles	absent or very weak (58,0.41), weak (49,0.34), medium (35,0.24), strong (1,0.01)
FLA	O	Flag leaf: attitude	erect (77,0.54), semi-erect (58,0.41), horizontal (8,0.06)
FLGS	O	Flag leaf: glaucosity of sheet	weak (36,0.25), medium (92,0.64), strong (15,0.1)
AACT	B	Awns:anthocyanin coloration of tips	Absent (28,0.2), present (115,0.8)
AIAC	O	Awns: intensity of anthocyanin coloration of tips	absent or very weak (28,0.2), weak (82,0.57), medium (25,0.17), strong (8,0.06)
GACN	O,B	Grain: anthocyanin coloration of nerves of lemma	O: absent or very weak (126,0.88), weak (11,0.08), medium (5,0.08), very strong (1,0.01) B: Absent (126,0.08), present (17,0.12)
ENR	B	Ear: number of rows	two (80,0.56), more than two (63,0.44)
ESh	N	Ear: shape	tapering (9,0.06), parallel (134,0.94)
ED	O	Ear: density	Very lax (21,0.15), lax (69,0.48), medium (43,0.3), dense (9,0.06), very dense (1,0.01)
GRHT	B	Grain: rachilla hair type	short (7,0.05), long (136,0.95)
RCFS	O	Rachis: curvature of first segment	absent or very weak (55,0.38), weak (88,0.62)
EDSS	B	Ear: development of sterile spikelets	Full (80,1.00), monomorphic in all 80 two-rowed varieties
SSA	N	Sterile spikelet: attitude (in mid-third of ear)	parallel to weakly divergent (5,0.06), divergent (75,0.94)
MSLG	O	Median spikelet: length of glume and its awn relative to grain	shorter (6,0.04), equal (90,0.63), longer (47,0.33)
GH	B	Grain: husk	present (143,1.00), monomorphic in all 143 varieties
GSLN	O	Grain: spiculation of inner lateral nerves of dorsal side of lemma	O: absent or very weak (16,0.11), weak (53,0.37), medium (48,0.34), strong (18,0.13), very strong (8,0.06) B: Absent (16,0.11), present (127,0.89)
GHVF	B	Grain: hairiness of ventral furrow	Absent (143,1.00), monomorphic in all 143 varieties
AL	O	Awn: length (compared to ear)	short (5,0.03), medium (31,0.22), long (107,0.75)
AR	B,O	Awn: roughness	B:smooth (39,0.27), rough (104,0.73) O:smooth (9,0.06), intermediate (30,0.21), rough (104,0.73)
SSTS	N	sterile spikelet: tip shape	pointed (35), rounded (45)
CT	N	collar: type	recurred (66,0.46), cup (77,0.54)
GC	N	Grain: color	yellow (106,0.74), green (3,0.02), black (34,0.24)
RLFS	O	Rachis: Length of first segment	Short (26,0.18), medium (90,0.63), long (26,0.18), very long (1,0.01)
GDL	B	Grain: disposition of lodicules	Clasping (143, 1.00), monomorphic in all 143 varieties

[End of annex and of document]