Genetic variation in Iranian rice (*Oryza sativa* L.) genotypes using physiological traits and SSR markers

Mehrnesa Qarehkhani [1], Saeid Navabpour [1]*, Hossein Sabouri [2], and Sanaz Ramezanpour [1]

1. Plant Breeding and Biotechnology Department, Gorgan University of Agriculture Sciences and Natural Resources, Gorgan, Iran
2. Plant Production Department., Gonbad Kavous University, Gonbad, Iran

Abstract

In order to assess genetic diversity, 38 Iranian rice cultivars were evaluated using 15 agronomic traits and 14 SSR markers. Rice genotypes were cultivated in a Randomized Complete Block Design in a research farm located in Azadshahr during 2013. According to the analysis of variance (ANOVA) genotypes showed a significant difference for all traits (P<0.01). Regression analysis showed that the weight of filled grains correlated with panicle number and chlorophyll content and explained 93.5% of the filled grains weight variations. The highest and lowest contents of polymorphic information was recorded in RM262 (0.81) and RM241 (0.33), respectively and the mean PIC was 0.51. The average Shannon Index was 0.33. The genotypes were categorized according to morphological traits into three groups but based on molecular and chlorophyll data they were assigned to three and two groups, respectively. Among the microsatellite markers used in this study, the RM142 RM255, RM341, RM262 markers had a high degree of polymorphism. RM297, RM104, and RM274 showed the highest correlation with rice chlorophyll content, which can be promising in marker selection programs.

Keywords: chlorophyll, regression, microsatellite, marker selection programs, polymorphic

Qarehkhani M., S. Navabpour, H. Sabouri, and S. Ramezanpour. 2022. 'Genetic variation in Iranian rice (*Oryza sativa* L.) genotypes using physiological traits and SSR markers*'. Iranian Journal of Plant Physiology 12* (3), 4239- 4251.

Introduction

Rice belongs to Gramineae family and is one of the most important food crops in the world growing throughout the tropical and cold regions. More than half of the world's people feed on rice as a main meal (Jamal et al., 2009). Rice cultivation in Asia is one of the most important farming practices in Southeast Asia and alone producing 96% of the world's total rice (Noormohammadi et al., 2004; Ebadi, 2013; Choudhury, 2013; Lin et al., 2012). The global population's demand for rice is projected to reach about 758 million tons by 2025 (Honar Nezhad, 2002). In 2012, rice production was 719,738,273 tons and its cultivated area was 16,319,909 hectares (FAO, 2010). Providing the world's future needs for rice is based on the use of plant breeding methods and the production of high yield varieties. For this purpose, crop breeding professionals should choose the right parenting choice to use, which is critical. The

* *Corresponding Author*
E-mail Address: s.navabpour@yahoo.com
Received: February, 2020
Accepted: July, 2021

importance of this issue lies in the fact that performance is a complex feature, consisting of several components, each of which positively or negatively affecting performance, which is generally very sensitive to environmental changes and atmospheric fluctuations (Gourbani et al, 2011). Leaf chlorophyll content is one of the most important physiological traits associated with the ability of photosynthesis in rice and has an impact on yield (Ishimaru et al., 2001). Since most of the important morphological and physiological traits affecting performance, such as chlorophyll content, are small and the gene blocks are involved, classical genetic engineering is not able to examine the behavior of the genes controlling quantitative traits as separate genes. With the advancement of the technology of molecular markers, the mapping of continuity, and locating software of quantitative traits controlling genes, powerful tools have been provided for the quantitative characterization of quantitative traits (Sabouri et al., 2008). Molecular markers are a powerful tool for detecting genetic variation and interpreting the diversity of plant species and have the potential to prove genetic variation and to help genetic resources of plants. In the meantime, microsatellites are an important tool for defining genetic variation in germplasm (Das et al., 2013). Simple microsatellite or repetitive sequences (SSRs) played a significant role in identifying the diversity of important plant species such as cotton, rice, wheat, corn, etc. Microsatellite markers include one to six repeat units that are scattered in the genome of most eukaryotes, so that at each 10,000 pairs of openings along the DNA, at least one microsatellite row with two unique rows on both sides can be seen (Weising et al., 2005). SSR markers have advantages such as high speed, simplicity, polymorphism, and stability, so they are very fast used in genetic variations, mapping, genes placement, fingerprinting, genetic purity testing, germplasm analysis, heterosis, and species interpretation and the introduction of related species (Das et al., 2013). In Juan et al. (2016), a collection of 217 rice variet mainly cultivated in temperate regions was generated. The collection encompasses modern elite and old cultivars, as well as traditional landraces covering a wide genetic diversity available for rice breeders. Whole genome sequencing was performed on 14 cultivars representative of the collection and the genomic profiles of all cultivars were constructed using a panel of 2697 SNPs with wide coverage throughout the rice genome, obtained from the sequencing data. The population structure and genetic relationship analyses showed a strong substructure in the temperate rice population, predominantly based on grain type and the origin of the cultivars. Dendrogram also agrees with population structure results in Roodbar. Calari et al. (2001) examined 113 cultivars of indigenous and modified Iranian cultivars. Among them, cultivars with similar local names collected from different regions had a significant genetic variation, so that the genetic distance between these cultivars and the genetic distance from other cultivars did not show any significant difference. Iranian rice in the gene bank had a great variety even if they had similar names. Zang et al. (2004) examined the genetic variation of 33 rice genotypes using 25 SSR markers. A total of 123 alleles with an average of 4.9 alleles per location were found in 33 genotypes. The number of alleles in each location had a painful relationship between two to nine alleles per location. The content of multiform data was 0.66-0.85, with an average of 0.57. Cluster analysis of genotypes showed two groups with a similarity of 10% based on Jaccard's similarity coefficient. Vaniraja et al. (2012) used Indicia and Japonica rice genotypes to use nine SSR markers, the Indiana and Japonica rice genotypes were divided into two distinct groups. Rabiei et al (2013) examined the genetic variation of 8 rice types from 46 markers SSR and 245 RFLP markers. Of the 46 SSR markers, 20 polymorphic markers were given. The number of alleles in each location varied from 2 alleles (RM133, RM215, RM433) to 6 (RM271) alleles. The polymorphic information contents for SSR and RFLP markers was 0.682 and 0.103, respectively. Cluster analysis based on data from both SSR and RFLP markers created two groups. The genetic distance obtained based on the formula and based on SSR markers, a range between 0.339 to 0.682 and the mean genetic distance was 0.554. 30 rice genotypes were examined using SSR markers. Of the 35 SSR markers used, 28 showed polymorphic markers. The content of multiform data (RM274) varied from 0.64 to (RM580) 0.72, with an average

Table 1
Rice genotypes used in this research

Genotype Number	Genotype Name	Genotype Number	Genotype Name
1	Dorfak	20	Domsiyah
2	KMP41	21	Khazar
3	Deylamani	22	Domsepid
4	Sangjo	23	IR72046-B-R-7-2-2-1
5	DOLAR	24	IR50
6	Zireh Bandpey	25	IR24
7	Tarom Amiri	26	Hashemi
8	Gil3	27	Gil1
9	Salari	28	IR43
10	Hassan Saraee	29	IR74099-3R-2-2
11	IR 82639-B-B-118-3	30	sadri
12	Tarom Pakotah	31	Sepidroood
13	binam	32	Gharib
14	Hassan Saraee Atashgah	33	IR28
15	Hasani	34	Champabodar
16	IR 83384-B-B-102-3	35	R69626B
17	IR60	36	IR 82590-B-B-94-4
18	IR64	37	Shahpasand
19	Mohammadi Chaparso	38	Mosa Tarom

of 0.46. Jaccard's likelihood coefficient ranged between 0.42-0.92. The genetic similarity between the five groups in which the genotypes were located was 56%. Analysis of the main components justifies 41.6% of variation (Seetharam et al., 2009). Doku et al. (2013) studied the genetic diversity of 18 African rice cultivars from different parts of Ghana using SSR markers. Of the 24 SSR markers, 23 polymorphic markers (95.83%) were used. Three different formulas were used to estimate genetic variation, which showed all three high varieties (I = 1.178, He = 0.625 and Nei's He = 0.608). The index of consolidation of genotypes collected from 4 regions was 51.5%.

Due to the importance of genetic diversity in rice and also the importance of this plant as a strategic crop, this study was designed to assess the genetic diversity of rice cultivars in the northeastern region of the Iran.

Materials and Methods

In this research, 38 rice genotypes (Table 1) from Gonbad Kavous University were evaluated. Research field located at 45 km south-east of Azadshahr. Rice seedlings were transplanted at the end of June 2013 to the main field. Seedlings were planted in a row with a spacing of 25 cm in plot of 2 m^2 at three replicates. At the end of the growing season, random samples were identified from each row of the experimental unit and chlorophyll contents were measured at five growth stages from the beginning of the planting until the harvest. Ten plants were randomly selected from each replicate and placed in separate envelopes after the harvest and measurements were made on them. The plant characteristics studied in this study were plant height, number of tiller per plant, weight of panicles, straw weight, flag leaf length, flag leaf width, number of panicles per panicle, number of filled grains per plant, and weight of filled grains per plant.

DNA extraction and SSR marker analysis

The DNA was extracted from 21-day-old seedling leaves of aromatic rice genotypes using hexadecyltrimethyl ammonium bromide (CTAB) method. The quality of DNA was determined by running it on 1% agarose gel with 1x TBE buffer (Trizma base with EDTA and boric acid; pH was adjusted to 8.0 with NaOH) at 70 V for 45 minutes. The gel was observed by a UV transilluminator

Table2
Name, chromosome number, and sequence of SSR primers used in the experiments for estimation of genetic diversity

Primer	Chromosome Number	forward Sequences	Backward Sequence
RM302	1	TCATGTCATCTACCATCACAC	ATGGAGAAGATGGAATACTTGC
RM472	1	CCATGGCCTGAGAGAGAGAG	AGCTAAATGGCCATACGGTG
RM297	1	TCTTTGGAGGCGAGCTGAG	CGAAGGGTACATCTGCTTAG
RM104	1	GGAAGAGGAGAGAAAGATGTGTGTCG	TCAACAGACACACCGCCACCGC
RM341	2	CAAGAAACCTCAATCCGAGC	CTCCTCCCGATCCCAATC
RM29	2	CAGGGACCCACCTGTCATAC	AACGTTGGTCATATCGGTGG
RM550	2	CTGAGCTCTGGTCCGAAGTC	GGTGGTGGAAGAACAGGAAG
RM262	2	CATTCCGTCTCGGCTCAACT	CAGAGCAAGGTGGCTTGC
RM434	9	GCCTCATCCCTCTAACCCTC	CAAGAAAGATCAGTGCGTGG
RM257	9	CAGTTCCGAGCAAGAGTACTC	GGATCGGACGTGGCATATG
RM241	4	GAGCCAAATAAGATCGCTGA	TGCAAGCAGCAGATTTAGTG
RM274	5	CCTCGCTTATGAGAGCTTCG	CTTCTCCATCACTCCCATGG
RM142	4	CTCGCTATCGCCATCGCCATCG	TCGAGCCATCGCTGGATGGAGG
RM255	4	TGTTGCGTGTGGAGATGTG	CGAAACCGCTCAGTTCAAC

lamp. The DNA was diluted to 50 ng using TE buffer and stored at 4 ℃ before the commencement of PCR. Fourteen (14) SSR markers belonging to different genomic regions of rice were selected in this study and (Table 2). PCR reactions were carried out in thermal cycler (Bio Red Inc. USA) with the total reaction volume of 15 µl containing, 5 µl of genomic DNA, 1 X assay buffer, 200 µM of DNTPs, 2 µM $MgCl_2$, 0.2 µM of forward and reverse primer and 1 unit of Taq DNA polymerase. PCR cycles were programmed as 95 ℃ for 2 min, 94 ℃ for 1 min, 55 ℃ for 1 min, and 72 ℃ for 10 min for final extension. The amplified products were separated on 2.5% agarose gel prepared in 0.5 X TAE buffer. The amplified product was separated on 6% polyacrylamide gel electrophoresis (PAGE) and visualized with silver-staining. Clearly resolved unambiguous bands were scored visually for their presence or absence with each primer. The scores were obtained in the form of a matrix with '1' and '0', indicating the presence and absence of bands in each variety, respectively (Masood Shah et al., 2013).

Data Analysis

Analysis of morphological data

SPSS software was used for correlations, regression, principal component, and cluster analysis (Ward algorithm as well as Euclidean distance square coefficient). Also, SAS software was used for analysis of variance.

Ward's similarity coefficients were calculated by pair-wise comparisons of varieties using SPSS software. Based on an average linkage algorithm (UPGMA, unweighted pair group method with an arithmetic average), clustering of genotypes was done. To depict the similarity or dissimilarity among groups or individual genotypes, Principal Component Analysis was done using SPSS software.

Analysis of molecular data

Genetic diversity parameters such as percentage polymorphic loci (PPL), effective allele number

Table 3
Analysis of variance of the rice traits under study

Source of Variation	df	Mean of Squares													
		Height	Tiller number	Panicles length	Panicles weight	Straw weight	Number of filled grains in the panicle	Weight of filled grains	Flag leaf length	Flag leaf weight	Number of panicles	Chlorophyll content 20 days after transplant	Chlorophyll content 40 days after transplant	Chlorophyll content 70 days after transplant	Chlorophyll content 90 days after transplant	Chlorophyll content 110 days after transplant
Replication	2	398.254 ns	5.482 **	22.367**	24.397*	137.35**	21138.68 ns	8.769 ns	57.390*	0.0277 ns	7.447*	5.481**	5.909**	*12.975	0.978 ns	3.008 ns
Genotype	37	1226.173**	36.386**	14.571**	116.524**	90.480**	38.298.55**	181.965**	45.033**	0.19**	6.710**	15.658**	42.871 ns	18.592**	3.877*	6.373 **
Error	74	818.818	0.923	2.690	5.198	20.462	13678.25	3.970	15.931	0.045	2.050	1.372	4.814	3.179	2.070	1.261

ns, **, and *Show not significant, significant at 1%, and significant 5% probability levels, respectively

Table 4
Analysis of regression for weight of filled grains as a dependent variable and other traits as independent variables

Variable	Regression Coefficient	Standard Error	R^2	Adjusted R^2	Standard Error
Constant	5.349**	1.569			
Spikelet number	0.021**	0.001	0.911	0.908	2.35749
Chlorophyll content 110 days after transplantation	0.059**	0.016	0.935	0.931	2.03914

(N_e), gene diversity (h), Shannon's information index (I), and gene frequency were computed using POWER MARKER software.

Polymorphism information content (PIC) was computed from the formula given below:

$$PIC = 1 - \sum_{j=1}^{n} P_{ij}^2$$

where P_{ij}^2 is the frequency of the j*th* allele for the i*th* marker and summed over alleles.

Utilizing binary data generated by SSR primers, Jaccard's similarity coefficients (Jaccard, 1908) were calculated between genotypic pairs using NTSYS-pc 2.02 program (Rohlf, 1998). From the similarity coefficient matrix thus generated, the dissimilarity coefficients (JD; Genetic distances = 1 - similarity coefficient) were calculated. The dissimilarity coefficient matrices of the other agronomic traits showed that there was a significant difference between genotypes at 1% level. Thus, the genetic potential of the studied traits in genotype were again subjected to PCoA to explore and establish similarity or dissimilarity among groups or individual genotypes.

Results

Morphological data

Analysis of variance

In Table 3, the results of analysis of variance of 15 quantitative traits were evaluated in rice

Table 5
Correlation coefficients between morphological traits, yield, and yield components

	Height	Tiller number	Panicle length	Panicles weight	Straw weight	Number of filled grains per panicles	Weight of filled grains	Flag leaf length	Flag leaf width	number of panicles	Chlorophyll content 20 days after transplant	Chlorophyll content 40 days after transplant	Chlorophyll content 70 days after transplant	Chlorophyll content 90 days after transplant	Chlorophyll content 110 days after
1	1														
2	-0.424*	1													
3	0.387*	-0.084	1												
4	0.131	-0.115	-0.107	1											
5	0.015	0.426**	0.210	0.111	1										
6	0.060	0.147	-0.268	0.779**	0.281	1									
7	0.213	0.083	-0.262	0.779**	0.249	0.954**	1								
8	0.092	-0.156	0.107	-0.119	-0.136	-0.116	-0.088	1							
9	-0.258	-0.060	-0.010	0.043	0.201	0.086	-0.206	-0.111	1						
10	-0.121	0.250	0.028	0.239	0.463**	0.269	0.158	-0.206	0.526**	1					
11	0.047	-0.252	-0.047	0.152	-0.190	0.074	0.046	-0.530**	-0.134	-0.100	1				
12	-0.202	0.089	-0.136	0.006	0.040	0.085	0.075	-0.022	0.110	0.092	0.035	1			
13	0.005	0.063	-0.009	-0.212	-0.158	-0.178	-0.191	0.100	0.116	0.119	-0.042	0.224	1		
14	0.029	-0.082	0.055	0.082	-0.159	0.019	0.062	0.022	0.228	-0.217	0.020	0.148	0.154	1	
15	-0.237	0.038	-0.226	-0.378	0.031	-0.083	-0.097	-0.180	0.076	0.037	-0.263	0.140	0.120	-0.097	1

* and **: significant 1% and 5%

genotypes. Analysis of variance showed differences between repetitions in the number of tillers, panicle length, straw weight, and chlorophyll contents 20 and 40 days after transplantation (significant at $p<0.01$) and panicles weight, flag leaf length, number of panicles, and chlorophyll contents 70 days after transplantation (significant at $p<0.05$). Differences between genotypes on height, tiller number, panicles length, straw weight, panicles weight, flag leaf length, number of panicles, number of filled grain in the panicle, weight of filled grains, flag leaf weight, and chlorophyll contents 20, 70, and 110 days after transplantation were significant ($p<0.01$) and chlorophyll content 90 days after transplantation trait were significant at the 5% level.

Analysis of regression

Fitting best regression model is conducted due to evaluation and effect of the most important independent variables that have the greatest influence on the dependent variable and the removal of variables that have a negligible effect on the dependent variable. Stepwise method was used to determine the final model and results are depicted in Table 4. In this research, the number of spikelet and the chlorophyll content, 110 days after transplantation were related to the weight of filled grains as dependent variables.

The second trait introduced into the model was chlorophyll content 110 days after transplantation (0.935). In total, 93.5% of the changes in the weight of grains were explained by these two traits.

Correlation

There was a positive correlation between chlorophyll content and most traits 90 days after transplantation. The number of filled grains showed a positive and significant correlation with panicle weight at 1% probability level (Table 5). The weight traits of grains filled had a significant positive correlation between the number of filled grains in the panicle and the weight of the panicle at 1% level. The number of panicles showed a significant positive correlation with flag leaf width and straw weight ($p\leq0.01$). Negative correlations were found between the number of tillers and plant height and local rice varieties had fewer tillers. The straw weight showed a significant positive correlation with the number of tillers

(p≤0.01). The length of the panicles showed a significant positive correlation with plant height (p≤0.05).

The relationship between plant height and the number of panicles per plant was positive. Also, the number of grains per plant positively correlated with grain yield in each plant.

Cluster analysis

Cluster analysis of rice genotypes was divided into three groups (Fig. I). The genotypes of the first group consisted of cultivars Salari, Domsiah, Shahpasand, Hasan Saraee, Hassani, Gil 1, Domsepid, IR74099-3R-2-2, Gharib, Binam, Hassan-Saraee Atashgah, Dorfak, Tarom Amiri, Zireh Bandpey, Hashemi, IR43, Dolor, Sadri, Sang Jo, Sepidrod, Champabodar, IR82639-BB-118-3, and Mohammadi Chaparso. The second group consisted of TaromPakotah, Khazar, Gil 3, KMP41, IR60, R24, Deylamani, R50, IR83384-B-B-102-3, IR72-46-B-R-7-2-2, R28, and IR64. The third group consisted of IR82590-B-B-94-4, and Mosa Tarom, R69626B. Multivariate analysis of variance was used to determine the best cutting point for the dendrogram. The highest amount of F in this analysis was related to the cut point and all genotypes were divided into three groups (Fig. I). The third group of genotypes had the highest number of grains, full grain weight, panicles weight, straw weight, and height. It can be argued

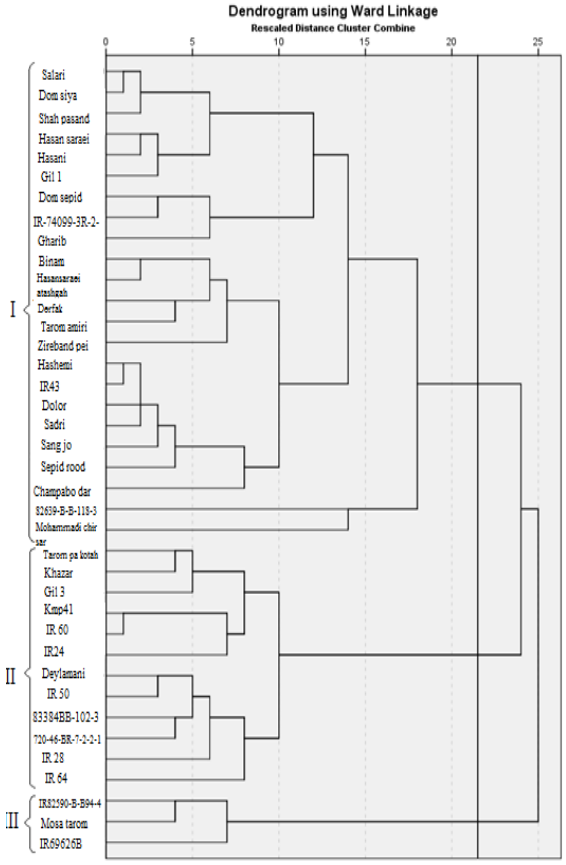

Fig I. Dendrogram of 38 rice genotypes constructed from a similarity matrix based on UPGMA using morphological data

that this group had the highest performance. The genotypes of the first group for the most of the studied traits had an average performance. The

Table 6
Genetic diversity analysis of rice genotypes using PCR-based SSR markers

Marker	Alley Frequency	Number of Alleles	Gene Variation	Polymorphism Information Content	Average Effective Number of Alleles	Average Shannon Index
RM302	0.625	4	0.5566	0.5139	1.43	0.38
RM472	0.4865	5	0.5920	0.5087	1.33	0.36
RM297	0.4900	3	0.5873	0.5001	1.41	0.42
RM104	0.3529	5	0.7318	0.6871	1.24	0.32
RM341	0.2647	11	0.8188	0.7957	1.13	0.19
RM29	0.3421	4	0.7216	0.6700	1.3	0.38
RM550	0.3784	4	0.6954	0.6392	1.33	0.37
RM262	0.2162	7	0.8327	0.8109	1.16	0.26
RM434	0.3947	4	0.7161	0.6663	1.3	0.37
RM257	0.3684	5	0.7258	0.6797	1.24	0.33
RM241	0.6897	2	0.4281	0.3364	1.67	0.57
RM274	0.4167	4	0.6991	0.6462	1.31	0.38
RM142	0.2778	5	0.7731	0.7358	1.24	0.33
RM255	0.2632	5	0.7774	0.7408	1.31	0.36
Average	0.3973	4	0.6897	0.6379	1.27	0.33

second group genotypes had lower weight of the grains and the average number of tillers and number of panicles were higher than the average traits. All three groups were equal in terms of flag leaf length and flag leaf width.

Molecular evaluations

The usefulness of microsatellite markers

According to Table 6, the highest and lowest allele frequencies were related to RM302 (0.62) and RM262 genetic locations (0.21), rspectively. The mean allelic frequency was 0.39. The most effective alleles were in the rank (1.67) RM241 while the lowest ranked (1.13) RM341. The mean of effective alleles in all samples was 1.27. The most effective alleles were in the rank (1.67) RM241 and the lowest ranked (1.13) RM341. The mean of effective alleles in all samples was 1.27. The highest number of alleles was found at RM341 (11) and the lowest number was at RM241 (2). The average number of alleles was 4.85, indicating a high genetic variation by the markers used. The highest gene diversity was found in RM341 (0.81) and RM262 (0.83). The mean genetic diversity was 0.68. The value of PIC was between 0.81 (RM262) and 0.33 (RM241). The mean of polynomial data was 0.63. High PIC, which reflects the diversity and frequency of alleles in genotypes showed the high number of alleles in a gene where the above results confirm this. Shannon index had a range of RM341(0.19)-RM241 (0.57). The average Shannon index for all locations was 0.33.

The mean genetic diversity was 0.244. The total number of observed alleles in the 14 was 63 alleles. The allele frequency ranged from 0.425 (RM252) to 0.975 (RM315), with an average of 0.6472. The value of polymorphism information content was 0.0476 (RM315) – 0.5993 (RM252) and the mean PIC for each indicator was calculated as 0.37885.

Cluster analysis

Fig. II. shows the grouping of people into three groups. The presence of genotypes in a group indicates a high genetic link between them. A large number of subgroups are represented in the form of a high diversity among genotypes.

Table 7
Eigenvalues, explained variance, and cumulative explained variance for extracted component in 38 rice genotypes under study

Principle component	Eigenvalues	explained variance	Cumulative explained variance
1	0.988	12.26	12.26
2	0.680	8.44	20.7
3	0.593	7.36	28.06
4	0.541	6.71	34.77
5	0.477	5.92	40.69
6	0.398	4.94	45.63
7	0.376	4.67	50.3
8	0.353	4.38	54.68
9	0.330	4.09	58.77
10	0.284	3.52	62.29

The first group consists of 3 cultivars (Gil 1, Sadri, and IR69626B) whose placement in a peduncle indicates a high genetic association between them. In the second group, 9 genotypes were placed (Mosa Tarom, Sang Jo, Gil 3, Tarom Pakotah, Domsiyah, KMP41, IR60, Hassani, and IR83384-B-B-102-3). A larger number of varieties was placed in the third group consisting of 26 varieties (Dorfak, Deylamani, IR50, IR64, IR24, Hashemi, Hassan Saraee Atashgah, Zireh Bandpey, Tarom Amiri, IR28, Binam, Shahpasand, IR72046-BR-7-2-2-1, Champabodar, IR82590-BB-94-4, Gharib, Mohammadi Chaparso, Salari, Khazar, IR74099-3R-2-2, IR82639-BB-118-3, Hassan Saraee, Sepidrod, Dolor, IR43, and DomSepid) were placed.

Principle coordinate analysis

The first ten components justify only 62.29% of the variation (Table 7). The first component justifies 12.26 % of the variation, and the second component justifies 8.44 % of the changes that the first component did not justify. Similarly, the process of justifying the changes by components continues, and eventually the tenth component justifies 3.52% of the changes that were not justified by the 9 previous components.

Association analysis

In order to determine the association between morphological traits and alleles, molecular data was processed on each trait (Table 8). Plant height

Table 8
Association analysis for rice morphological traits as dependent variable and amplified alleles as independent variables

Row	Quality	Alleles	R^2 adjusted	R^2	P value
1	Height	RM341-d, RM274-a, RM104-c, RM262-f, RM434-a, RM434-b, RM241-b, RM257-e, RM255-b, RM302-a, RM257-a	0.432	0.9	0.000
2	Tiller number	RM262-b, RM434-a	0.295	0.550	0.012
3	Panicles length	RM257-e, RM472-a, RM297-c, RM434-c, RM255-c	0.319	0.963	0.000
4	Panicles weight	RM104-e, RM341-a, RM262-c, RM257-e	0.463	0.948	0.000
5	Straw weight	RM142-c, RM302-a	0.239	0.496	0.023
6	Number of filled grains	RM142-c, RM104-c, RM257-c, RM262-c	0.724	0.963	0.000
7	The weight of filled grains	RM142-c, RM255-d, RM472-b, RM262-e	0.759	0.984	0.000
8	Flag leaf length	RM434-c, RM257-b, RM241-a, RM274-c	0.5	0.904	0.000
9	Flag leaf width	RM257-d, RM257-e, RM302-d, RM302-c	0.495	0.942	0.000
10	Number of Panicles	RM257d, RM257-e	0.352	0.705	0.001
11	Chlorophyll content 20 days after transplantation	RM142-e, RM104-e, RM241-b	0.516	0.867	0.000
12	Chlorophyll content 40 days after transplantation	RM341-b, RM297-b, RM472-b, RM274-c, RM241-b, RM434-a, RM29-a, RM434-b, RM302-c, RM262-c	0.292	0.9	0.000
13	Chlorophyll content 70 days after transplantation	RM302-c, RM341-b, RM550-a, RM104-e, RM257-b, RM29-d, RM341-e, RM104-c, RM302-a	0.363	0.9	0.000
14	Chlorophyll content 90 days after transplantation	RM142-c, RM297-c	0.255	0.538	0.014
15	Chlorophyll content 110 days after transplantation	RM274-a, RM262-d, RM104-b	0.572	0.879	0.000

and chlorophyll content 40 days after transplantation were the traits with the highest number of positive markers and the coefficient of explanation was 0.9. RM257, RM104, RM434, and RM262 had the highest relationship with the traits studied, which can be useful in subsequent studies as this indicates that these attributes have close proximity or are probably affected by genes (there are several effects). To understand this, it is essential to provide segregate generations and linkage maps. Finally, RM274, RM104, and RM297 showed the highest correlation with rice chlorophyll content.

Discussion

In recent years, increased production of genetically-linked varieties by rice growers has led to an increase in genetic diversity. Information on genetic variation among cultivars is needed to use the genetic resources available to create new genotypes. This helps with the production of new types that require a continuous evaluation of germplasm for useful traits based on morphological data. Morphological markers reflect only the genetic component of the cultivar, but the interaction of the genotype with the environment (G×E) in which it is expressed. Molecular markers have the ability to detect

genetic diversity and help manage plant genetic resources. Compared with morphological traits, molecular markers can show the difference between genotypes at the DNA level and provide direct and reliable tools for identifying, preserving, and managing germplasm (Thenmozhi and Rajasekaran, 2013). Variable refers to the presence of differences between individuals of a species. This is due to the differences in the genetic structure of individual plants or in the environment in which they grow. The success of genetic improvement depends on the nature of the changes in the gene set for the person given. Hence, it is important to evaluate the variations available to each person in the gene set for a plant breeder product to start a recreational planting program. In this study, 38 genotypes of rice were evaluated morphologically due to morphological traits.

Significant differences were detected in the study between genotypes for all traits. Honarnezhad (2002) examined the genetic parameters in six varieties of Iranian rice. The results of analysis of variance of diallel in each of the four methods of Griffing's diallel analysis on grain yield in the plant and other agronomic traits showed that there was a significant difference between genotypes at 1% level. Thus, it can be concluded that the genetic potential of the studied traits is in genotype.

Estimation of correlation coefficient between different traits shows the direction of relationships. The correlation between traits for three reasons is attributed to the genetic causes of correlation due to the genetic phenotypic effect, the changes caused by selection, and natural selection in which the relationship between quantitative traits and fitness is the primary factor that determines the genetic properties in a natural population (Sambrok et al., 1990). The function of a complex character is influenced by a large number of other features. Knowledge about the relationship between performance and its features, as well as component features, help improve choice performance. In the present study, the correlation coefficient (Table 5) was measured between 15 quantitative characteristics. Among the studied traits, panicle weight (0.779) and number of grain per panicle (0.954) had a positive and significant correlation with grain weight, grain number per panicle, and grain weight. This suggests that selection in each of these traits will increase other traits in comparison with other traits and thus increasing grain yield. Salam Khan et al. (2009) showed a positive and significant correlation between grain yield with plant height, panicle length, flag leaf length, and number of grains per panicle. The relationship between plant height and positive tillers was significant. Also, the number of grains per plant had a good correlation with grain yield per plant.

From the results of correlation analysis, it is expected that by increasing the traits, the weight of filled grains, the number of filled grains in the panicle, and the weight of the panicle eventually increase the yield. The results were similar to those of Mahdavi et al. (2005) and Honar Nezhad (2002).

The relationship between chlorophyll content 110 days after transplantation with most negative traits because at the end of the growth season, the chlorophyll content was low compared with the straw weight of the plant and chlorophyll was depleted in most organs which were aging. Salam Khan et al. (2009) indicated that there was a positive and significant correlation between grain yield and plant height, panicle length, flag leaf, and grain number per panicle.

Regression relationships between dependent and direct variables provide a very good understanding of the relationships between different traits. Examining the results of regression analysis showed that the weight of filled grains had a strong positive relationship with the number of panicles and the chlorophyll content 110 days after transplantation, and this corresponded to the results of simple correlations. In this study, the number of panicles (R^2 = 91.1) as the first variable entered into the model, and its regression coefficient remained significant in the model due to its significance. However, 91% of the variations in the weight of grains filled with peduncle numbers were explained. Comparison of the standardized regression coefficients in Table 5 indicated that the number of panicles was more important than the chlorophyll content and increasing the values of this trait that has a

positive regression coefficient increased the weight of filled grains.

The present study showed that the clustering of genotypes was based on morphological traits that caused three clusters (Fig. I). Chakravarti et al. (2006) studied the genotypes in 11 groups using 30 rice genotypes, employing 30 cluster analysis methods. Choudhury et al (2013) examined 29 varieties of rice by using the analysis of variance by input method; they were assigned in three groups. The results of this grouping can be applied to the parent in order to select the best parent in breeding programs. The maximum and minimum number of cultivars were found in clusters (I) and (III). Genotypes were divided into three groups based on the highest F value using multivariate analysis of variance. The third group genotypes had the highest panicle weight, number of filled grains, full grain weight, and height. The genotypes of the first group and the two groups were evaluated for the traits and average yield. The second group had the lowest grain weight and the average number of tillers in the other two groups. Three groups were equal in terms of flag leaf length and width. The results of this grouping can be used to achieve parents with desirable features. SSR marker data were able to divide rice genotypes into three separate clusters (Fig. II). The maximum and minimum number of cultivars were found in cluster 3 (26 cultivars) and cluster 1 (3 cultivars), respectively. We know that modified cultivars such as Gil 1, the results of crossing between local and introduced cultivars or from the direct introduction have probably resulted in the dendrogram being included in two groups of introduced and landrace varieties. In other words, these three genotypes have similar repetitive sequences despite apparent differences and these repetitive units may not be associated with different traits. If the number of primers were used in this experiment, the breakdown of these dendrogram might have been better.

Although these figures were placed in the subgroup of the group, the best reason for them to be in the same group is to have similar repetitive areas that fit these dendrogram into a group. Two varieties of Gil 1 and 3, although of similar names, differ in terms of repeat DNA sequences and are divided into two distinct groups. As can be seen in the figure, many groups have been created, so even with the low number of primers used in this research there is a large variation among the varieties within each group so that the varieties in terms of this indicator can be varied.

In general, the data obtained from molecular evaluations showed that the use of these markers are beneficial in later studies or corrective work due to the high degree of variation in them.

The close proximity of this number to the real number of alleles, namely, two alleles, is a reason for the good effect of high-density polymorphisms on alleles and the estimation of genetic variation. All of the 14 analyzed loci featured a multitude of forms. The high number of alleles represents a high diversity in one locus.

Sherria et al. (2011) reported the average number of observed alleles (na) and the number of effective alleles (ne) in the study of 8 varieties of rice (1.66 and 1.37). Shaha et al. (2013) In a gene diversity study of 40 rice cultivars, using a total of 24 SSR markers distributed throughout the genome, identified a total of 66 alleles with an average of 2.75.

The reason for these differences can be due to the different genetic resources used in these experiments, and that the primers designed for different microsatellite positions are completely dedicated and only replicate the same target position. So this difference may be due to a change in the repeat unit, which may have been altered due to mutation (the mutation level in the microsatellite is high), as a result of these factors, diversity occurs in genome.
Finally, association analysis suggested that RM274, RM104, and RM297 showed the most relevance to chlorophyll content of rice, which can be used to find QTLs related to chlorophyll content in QTL mapping applications.

Acknowledgments

This study was supported by Gonbad Kavous University and the Gorgan University of

Agriculture Sciences and Natural Resources of Iran.

References

Botstein, D., R. L. White., M. Skolnick and **R. W. Davis**. 1980. 'Construction of a genetic linkage map in man using restriction fragment length polymorphisms'. *American journal of human genetic*, 32:314.

Cavalli-Sforza, L. L. and **A. W. F. Edwards**, 1967. 'Phylogenetic analysis. Models and estimation procedures'. *American journal of human genetic*, 19: 233.

Chakravarthi, B. k. and **R. Naravaneni**, 2006. 'SSR marker-based DNA finger printing and diversity study in rice (*Oryza sativa. L*)'. *Journal of African of Biotechnology*, 5(9):684-688.

Choudhuy, B., M. Latifkhan and **S. Dayanandan**, (2013). Genetic structure and diversity of indigenous Rice varieties in the Eastern Himalayan vegion of NorthestIndia. www. SpringerPlus. ir.

Das, H. k., A. Singhand and **S. Datti**, 2013. 'Phylogenetic relationship of wild and cultivated Oryza speciesusing SSR markers'. *Journal of Asian of Contemporary Sciences*, 2(2): 35-44.

Doku, H. A., E. Y. Danquah, A. N. Among, K. Nyalemegbe and **H.M. Amoatey**, 2013. 'Genetic diversity among 18 accessions of African rice (*Oryza* glaberri masteud) using simple sequence repeat (SSR) markers'. *Journal of Cultural*, 8(2): 106-112.

Doyle, J. J. and **J. L. Doyle**, 1987. 'A rapid DNA isolation procedure for small quantities of fresh leaf tissue'. *Phytochemical Bulletim*, 19: 11-15.

Ebadi, M., H. Riahi and **R. Zare**, 2013. 'Genetic diversity of Fusarium semitectum isolates from rice, using RAPD and REP-PCR markers'. *Journal Mycologia Iranica*, 1 (1): 14-20.

FAO. 2010. 'FAO Land and plant nutrition management service'. http://www.fao.org

Gourbani, H. R., H. A. Samiezade lahijani, B. Rabiei and **A. Allah golipour**, 2011. 'Classification of different rice genotypes using factor analysis and cluster analysis'. *Journal Agricultural and stable production*, 21(3): 104-89.

Honarnezhad, R, 2002. 'Estimation of Genetic parameters in Rice using different diluent griving techniques'. *Journal science and technology of agriculture and natural resources*, 11(41): 247-258.

Ishimaru, k., M. Yano, N. Aoki, T. Hirose, Y. Lin, L. Monna, T. Sasaki and **R. Ohsugi**, 2001. 'Toward the mapping of physiologi and agronomic characters on a rice function map: QTL analysis and comparisons between QTLs and expressed sequence tags'. J. *Theor. Appl. Genet*, 102:793-800.

Jamal, K., I. H. Khalil, A. Bari, S. Khan and **I. Zada**, 2009. 'Genetic variation for yield and yield components in rice'. Journal Agriculture & Biotechnology Science, 4(6): 60-64.

Juan, L., V. Reig, V. Juan, S. Ester, M. Luis, T. Javier, G. Marta, D. Sophia, T. Manuel and **D. Concha**, 2016. 'Genetic Diversity and Population Structure of Rice Varieties Cultivated in Temperate Regions'. *Springer link* (http://creativecommons.org/licenses/by/4.0/).

Lin, H. Y., Y. P.Wu, A. L. Hour, H. O. Sw, F. J. Wei, Y. I. Hsing and **Y. R. Lin**, 2012. 'Genetic diversity of rice germplasm used in Taiwan breeding programs'. *Journal of Botanical studies*, 53: 363-376.

Liu, K. and **S. V. Muse**. 2005. 'Power marker: an integrated analysis'. *Bioinformatics*, 21: 2128-2129.

Mahdavi, F., M. A. Esmaeiliy, A. Fallah and **H. A. Pirdashti**, 2005. 'Study of Morphological characteristics of physiological indices yield and grain yield components in native and modified rice cultivars (*Oryza Sativa L*.)'. *Journal Iran Crop Science*, 7(4): 297-280.

Masood Shah, S., A. N. Shahzad and **M. Arif**, 2013. 'Genetic diversity basmati and non-basmati rice varieties based on microsatellite markers'. Journal of *Pakistan Biotechnology*, 45(S1):423-431.

Noormohammadi, G. H., A. A. Seyadat and **A. Kashani**, 2004. *Culture* (first volume): Shahid Chamran: University Publishers. Page: 446.

Rabiei, H., K. H. Salem and **M. Mattar**, 2013. 'The genetic diversity and relatedness of eight rice (*Oryza Sativa L.*) cultivar as revealed by AFLP and SSRs marker'. *Journal of Life Science*, 10(1): 1471-1479.

Roodbar calari, F., A. A. Farshadfar and **B. Gharehyazy**, 2001. 'Evaluation of Genetic Variation in Rice (Oryza sativa L) Cultivars in Iran Using RAPD markers'. *Journal Iran Crop Science*, 3(4):8-15.

Sabouri, H., A. Sabouri, M. Nahavi, A. R. Dadras and **M. Katozi,** 2008. 'Locating QTL controlling chlorophyll content of seedling and reproductive stages in rice under salt stress'. *Journal New Genetics,* 3(3): 30-20.

Salem, K. H and **A. Sallam**, 2016. 'Analysis of population structure and genetic diversity of Egyptian and exotic rice (*oryza sativa* L.) genotypes'. *Journal of Science direct*, 3(39):1-9.

Sambrook, E., F. Fritsch and **T. Maniatis**, 1990. Molecular Ctonlng: A Laboratory Manual. Book review, pp: 17-18.

Seetharam, M., S. Thirumeni and **K. Paramasivam**, 2009. 'Estimatin of genetic diversity in rice (Oryza Sativa L.) genotypes using SSR markers and morphological characters'. *Journal of African of Biotechnology*, 8(10): 2050-2059.

Shaha, S. M., S. A. Naveed and **M. Arif**, 2013. 'Genetic diversity in basmatic and non-basmatic rice varieties based on microsatellite markers'. *Journal of Pakistan Biotechnology,* 45(1): 423-431.

Sherria, R., S. Sen and **P. Muneer**, 2011. 'Analysis of genetic variability in rice varieties (Oryza sativa L.) of Kerala using RAPD marker'. *Journal of Genetic engineering and Biotechnology*, 11:1-9.

Thenmozhi, P and C. Rajasekaran, 2013. 'Genetic diversity and relationship among 40 rice accessions from North-Eastern Zone Tamil Nadu using morphological and SSR markers'. *Research journal of biotechnology*. 8 (9): 32-41.

Vaniraja, C., K. K. Vinod and **A. Pereira**, 2012. 'Molecular evaluation of genetic diversity and association studies in rice (Oryza Sativa L.)'. *Journal of Genetic*, 91(1):9-19.

Weising, K., H. Nybom, K. Wolff and **G. Kahl,** 2005. 'DNA fingerprinting in plants principles methods and application (second edition). CRC press, pp:7-10.

Zang, L., T. R. Kwon, X. Liu, C. Wilson, C. M. Grieve and **G. B. Gregorio**, 2004. 'Genetic diversity analyzed by microsatellite markers among rice (Oryza Sativa L.) genotype with different adaptation to saline soils'. *Journal of Science Direct*, 166(68): 1275-1285

Doyle, J. J. & Doyle, J. L, (1990). Isolation of Plant DNA from fresh tissue. Focus, 12, 13-15

Gower, J. C, (1966). Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika, 63, 315-28

J**accard, P,** (1908). Nouvelles researches sur la distribution florale. Bulletin de la Societe Vaudoise des Sciences Naturelles, 44, 223-270

Rohlf, F. J, (1998). NTSYS-pc Numerical taxonomy and multivariate analysis system, version 2.02. New York: Exeter Publication

Sokal, R. R. & Michener, C. D, (1958). A statistical method for evaluating systematic relationships. The University of Kansas Science Bulletin, 38, 1409-1438

Yeh, F. C., Yang, R. C. and T. Boyle, 1999. Popgene v. 1.31. Microsoft Window-Based Freeware for Population Genetic Analysis, University of Alberta and Tim Boyle, Centre for International Forestry Research, 1999, A Joint Project Development by Francis C. Yeh and Rong-Cai Yang.