Population structure of rice varieties used in Turkish rice breeding programs determined using simple-sequence repeat and inter-primer binding site-retrotransposon data

G. Cömertpay[1], F.S. Baloch[2], M. Derya[3], E.E. Andeden[3], A. Alsaleh[3], H. Sürek[4] and H. Özkan[5]

[1]Eastern Mediterranean Agricultural Research Institute, Doğankent, Adana, Turkey
[2]Department of Field Crops, Faculty of Agricultural and Natural Sciences,
Abant Izzet Baysal University, Bolu, Turkey
[3]Department of Biotechnology, University of Çukurova, Adana, Turkey
[4]Thrace Agricultural Research Institute, Edirne, Turkey
[5]Department of Field Crops, Faculty of Agriculture, University of Çukurova, Adana, Turkey

Corresponding author: H. Özkan
E-mail: hozkan@cu.edu.tr

ABSTRACT. Effective breeding programs based on genetic diversity are needed to broaden the genetic basis of rice (*Oryza sativa* L.) in Turkey. In this study, 81 commercial varieties from seven countries were studied in order to estimate the genomic relationships among them using nine inter-primer binding site (iPBS)-retrotransposon and 17 simple-sequence repeat (SSR) markers. A total of 59 alleles for the SSR markers and 96 bands for the iPBS-retrotransposon markers were detected, with an average of 3.47 and 10.6 per locus, respectively. Each of the varieties could be unequivocally identified by the SSR and iPBS-retrotransposon profiles. The iPBS-retrotransposon- and SSR-based clustering were identical

and closely mirrored each other, with a significantly high correlation (r = 0.73). A neighbor-joining cluster based on the combined SSR and iPBS-retrotransposon data divided the rice varieties into three clusters. The population structure was determined using the STRUCTURE software, and three populations (K = 3) were identified among the varieties studied, showing that the diversity harbored by Turkish rice varieties is low. The results indicate that iPBS-retrotransposon markers are a very powerful technique to determine the genetic diversity of rice varieties.

Key words: Rice variety; Rice diversity; SSR; iPBS-retrotransposon; Turkey

INTRODUCTION

In Turkey, rice (*Oryza sativa* L.) was introduced rather recently, and was only marginally important in Turkish agriculture because it was not a staple Turkish food. However, in the last 30-40 years it has become crucially important, because rice consumption in Turkey has steadily increased. Currently, it is an integral part of the Turkish diet in the form of pilaf. In Turkey, the principal staple food is wheat, but the consumption of wheat per capita has been gradually decreasing, while the per capita consumption of rice is increasing (Sürek, 1997). The two main rice producers in Europe are Italy (224,000 ha) and Spain (117,000 ha). These two countries together contribute more than 80% of the total rice production in Europe. The other major non-EU rice-producing countries in Europe and the Mediterranean region are Egypt and Turkey. According to the Food and Agricultural Organization of the United Nations, rice production in Turkey has increased tremendously in the last 40 years from 266,000 to 750,000 tons, and it has become the fifth most produced cereal crop after wheat, barley, maize, and rye. It is produced in micro- and macroclimatic regions of Turkey, and has achieved great socio-economic importance for family farming systems as a profitable source of income. Rice cultivation is practiced in almost all regions of Turkey; however, it is commercially grown mainly in the western and northern parts of the Marmara and Black Sea regions (approximately 92% of the total cultivation and production area), which have the best climate and water availability. Nearly all of the rice varieties grown in Turkey are the japonica ecotype, as it is the best-adapted to the temperate climate.

In Turkey, rice research and breeding activities were initiated in 1970 by Thrace Agricultural Research Institute. Until 1982, the institute mainly concentrated on regional rice issues, but in 1982 the National Rice Release and Training Project was created, and the Thrace Institute became the coordinating center. Since then, many rice varieties have been introduced from various foreign countries, specifically targeting the yield and quality preferred by local farmers and consumers, ultimately resulting in a rice boom in Turkey. Total milled rice production varied between 150,000 and 200,000 tons annually, which was insufficient to meet domestic demand. Since 1992, due to increasing demand for rice, Turkey started to import rice at more than twice its domestic production. Since then, rice imports have continuously increased.

Knowledge of the genetic relationships between the elements of an improved gene pool is vital for the exact planning and execution of a targeted breeding program (Baloch et al., 2015). There is a remarkably rich genetic diversity in cultivated rice; however, a number of biotic and abiotic stressors limit its productivity. Therefore, there is an urgent need to identify diverse sources

of genes in order to increase tolerance to various stressors, and to broaden the rice gene pool (Joshi et al., 2000). Several studies have investigated the genetic diversity of rice varieties using morphological traits, biochemical markers, and more recently, DNA molecular markers (Lu et al., 2005). DNA molecular markers, such as simple-sequence repeats (SSRs), are widely used in investigating rice germplasm diversity due to their co-dominant nature, abundance in the genome, high rate of polymorphism, informativeness, repeatability, and reliability. They have been particularly useful in investigating rice germplasm from different countries, such as India, Japan, USA, China, and Italy (Hashimoto et al., 2004; Saini et al., 2004; Lu et al., 2005; Bao et al., 2006; Cai et al., 2013). However, there is no comprehensive information available about the population structure and diversity of rice varieties used in Turkish breeding programs.

Among the various DNA markers, retrotransposon markers have been widely applied in numerous evolutionary and genetic studies due to their general applicability, simplicity of implementation, and genotype resolution systems. Some of the important limiting factors in the development of retrotransposon markers for a new species are product size variation and the lack of cloning and sequence information for designing primers that match the flanking genomic DNA at each specific site. Recently, Kalendar et al. (2010) described a novel polymerase chain reaction (PCR)-based technique that overcomes these difficulties, and can both isolate long-terminal repeat (LTR) retrotransposons in virtually any organism as well as serve as a universal marker system in its own right. The inter-primer binding site (iPBS) method has proven to be a powerful DNA fingerprinting technique without the need for previous knowledge of a sequence, and is referred to as a universal marker system because iPBS-retrotransposons comprise the only retrotransposon-based marker system that has allowed polymorphism visualization throughout the plant kingdom. iPBS-retrotransposon DNA markers have been successfully employed in evaluating the genetic diversity of *Saussurea esthonica*, apricot, grapevine, *Cicer* species, guava, Chinese bayberry (*Myrica rubra*), and cultivated and wild *Lens* species (Gailite et al., 2011; Baránek et al., 2012; Guo et al., 2012; Andeden et al., 2013; Mehmood et al., 2013; Fang-Yong and Ji-Hong, 2014; Baloch et al., 2015). However, iPBS-retrotransposon markers have not yet been used in genetic diversity studies of rice.

Thomson et al. (2007) stated that at the global level the population structure of the rice germplasm has been well characterized. However, more detailed analyses of regional or country-specific differences are lacking, and have only just begun in some countries. Turkey is an excellent place to study country-specific differences in the rice germplasm, because it has a wealth of rice diversity. However, Turkey is not the origin or center of diversity for rice. Throughout history, Turkey has been at the intersection of cultures between Europe, Asia, and Africa; consequently, Turkey might have played an important role in spreading the rice germplasm from Asia, the center of rice diversity, to Europe and Africa. The genetic diversity of Turkish rice has not been studied systematically, and an intraspecific classification has not been conducted; indeed, only a few studies have been conducted using Turkish rice germplasm. Buyukunal Bal and Bay (2010) characterized Turkish rice varieties using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis and random amplified polymorphic DNA (RAPD) markers, but the varieties were not differentiated. The molecular characterization of the genetic resources of rice from Turkey has not been attempted. The accessions evaluated here represent the majority of varieties released in Turkey, along with rice varieties from different countries used as parent lines in the rice improvement program. The specific objectives of the present study were as follows: 1) to test the utility of iPBS-retrotransposon markers for phylogenetic analysis of the rice germplasm; 2) to evaluate the population structure of Turkish rice varieties based on iPBS-retrotransposon

markers; and 3) to compare iPBS-retrotransposon data with SSR marker data in order to determine the degree of similarity in the relationships predicted by each DNA marker.

MATERIAL AND METHODS

Plant materials and DNA extraction

Eighty-one commercial rice varieties were used as plant materials, which consisted of 39 commercial rice varieties that had been bred in Turkey and used in production during the last four decades. Forty-two varieties that were mostly used as parent lines in Turkish rice breeding programs were collected from foreign countries. These foreign varieties consisted of 25 from Italy, 4 from Bulgaria, 4 from France, 4 from the USA, 3 from the International Rice Research Institute (IRRI) in the Philippines, and 2 from Russia. The names, origins, release years, and grain types of the different varieties are summarized in Table 1. Ten seeds from each variety were grown in pots, and leaf samples were collected from five genotypes of each variety and combined. Genomic DNA was extracted from young leaf tissue according to the CTAB protocol described by Doyle and Doyle (1990), with minor modifications as described by Baloch et al. (2010).

SSR analysis

In total, 17 SSR primers distributed over the whole rice genome were selected from the published rice microsatellite framework (http://archive.gramene.org; Temnykh et al., 2000) (Table 2). Of the 17 SSR markers, four (RM452, RM489, RM124, and RM277) exhibited a monomorphic pattern. M13 tailed-primer PCR amplification of the SSRs was performed in a 12-µL PCR mix containing 1X buffer, 0.125 mM dNTP, 0.4 pmol M13 sequence tailed forward primer, 0.3 pmol reverse primer, 3.0 pmol universal M13 primer labeled with one of four (6-FAM, VIC, NED, or PET) fluorescent dyes, 0.12 U/µL *Taq* DNA polymerase, and approximately 50 ng genomic DNA (Cömertpay et al., 2012).

iPBS-retrotransposon analysis

A total of 83 iPBS-retrotransposon primers developed by Kalendar et al. (2010) were initially screened on eight randomly selected rice varieties to determine the polymorphism status of the primers. Nine iPBS-retrotransposon primers that produced excellent banding profiles were selected for fingerprinting the entire set of rice varieties. Primer names, sequences, and annealing temperatures are described in Table 3. PCR amplifications were performed in an Eppendorf DNA Thermal Master Gradient Cycler (Eppendorf Netheler Hinz, Hamburg, Germany), and were conducted in a 25-µL reaction mixture containing 25 ng template DNA, 1X Dream *Taq* Green Buffer (Fermentas), 0.2 mM dNTP (Fermentas), 1 µM primer for 12- to 13-nt primers or 0.6 µM for 18-nt primers, 1.5 U Dream *Taq* DNA polymerase (Fermentas), and 0.04 U PFU DNA polymerase (Fermentas) (Baloch et al., 2015). The PCR thermal cycling profile was as follows: initial denaturation at 95°C for 3 min, 30 cycles at 95°C for 15 s, 50-65°C annealing temperature (depending upon the primer) for 1 min, 68°C for 1 min, and a final extension at 72°C for 5 min (Andeden et al., 2013). All of the PCR products were separated by 1.7% (w/v) agarose gel electrophoresis with 0.5X TBE buffer for 2 h, stained with ethidium bromide, and visualized using an ultraviolet transilluminator (Infinity ST5, Vilber Lourmat, France).

Table 1. Material used in this study, with cultivar name, their origin, release year, and grain type.

No	Cultivar Name	Origin	Release Year	Grain type
1	Osmancık-97	Turkey	1997	Medium
2	Kıral	Turkey	2000	Medium
3	Neğiş	Turkey	2002	Medium
4	Trakya	Turkey	1990	Medium
5	Demir	Turkey	2000	Medium
6	Ergene	Turkey	1990	Medium
7	Halilbey	Turkey	2004	Medium
8	Şumnu	Turkey	2006	Medium
9	Gönen	Turkey	2002	Medium
10	Sürek-95	Turkey	1995	Medium
11	Tunca	Turkey	2009	Medium
12	Kızıltan	Turkey	2007	Medium
13	Kırkpınar	Turkey	2004	Medium
14	Durağan	Turkey	2007	Medium
15	Beşer	Turkey	2006	Medium
16	Kargı	Turkey	2002	Medium
17	Edirne	Turkey	2004	Medium
18	Serhat-92	Turkey	1992	Medium
19	Karacadağ	Turkey	1992	Medium
20	Karadeniz	Turkey	2003	Medium
21	Diyarbakır yerli	Turkey	Local variety	Medium
22	Akçeltik	Turkey	Local variety	Short
23	Koral	Turkey	2005	Medium
24	Yavuz	Turkey	2000	Medium
25	Ece	Turkey	2004	Medium
26	Meriç	Turkey	1990	Medium
27	Sarıçeltik	Turkey	Local variety	Medium
28	Gala	Turkey	2009	Short
29	Ipsala	Turkey	1990	Medium
30	TR-1764	Turkey	Under registration	Medium
31	TR-1765	Turkey	Under registration	Medium
32	TR-1830	Turkey	Under registration	Medium
33	TR-1905	Turkey	Under registration	Medium
34	TR-1943	Turkey	Under registration	Medium
35	KA 173	Turkey	Breeding line	Medium
36	İz-160-2	Turkey	Breeding line	Medium
37	İz-300-1	Turkey	Breeding line	Medium
38	TR-1232	Turkey	Breeding line	Medium
39	TR-1324	Turkey	Breeding line	Medium
40	Kros-424	Russia	1983	Short
41	Jubilenni	Russia	-	Short
42	Ribe	Italy	1961	Medium
43	Rocca	Italy	1967	Medium
44	Maratelli	Italy	1919	Medium
45	Veneria	Italy	1975	Medium
46	Agusto	Italy	2002	Medium
47	S. Andrea	Italy	1968	Medium
48	Rialto	Italy	-	Medium
49	Padona	Italy	1968	Medium
50	Roma	Italy	1962	Medium
51	ItalicoRancorollo	Italy	1971	Short
52	Europa	Italy	1969	Medium
53	RossaMarchetti	Italy	1964	Medium
54	Ringo	Italy	-	Medium
55	Bonni	Italy	1969	Medium
56	Rubino	Italy	1975	Short

Table 1. Continued

No	Cultivar Name	Origin	Release Year	Grain type
57	Sirrio	Italy	1978	Medium
58	CRM-3 Cripto	Italy	1990	Medium
59	Lido	Italy	1985	Medium
60	Volono	Italy	1968	Medium
61	Baldo	Italy	1964*	Medium
62	Titanio	Italy	1979	Medium
63	Balilla	Italy	1924*	Short
64	Star	Italy	-	Medium
65	Panda	Italy	1988	Medium
66	Arborio	Italy	1946	Medium
67	Taichung 65	IRRI	-	Short
68	Akihikari	IRRI	-	Medium
69	Chianan 8	IRRI	-	Short
70	Arlaton	France	-	Short
71	Romenica	France	-	Medium
72	Balilla-28	France	1980	Medium
73	Calendal	France	-	Medium
74	Ranballi	Bulgaria	-	Short
75	Plovdiv	Bulgaria	-	Short
76	Rodina	Bulgaria	-	Short
77	İskra	Bulgaria	-	Short
78	Katy (Arkansas)	USA	1989	Long
79	LGRU (Lagrue)	USA	1993	Long
80	Kaybonnet	USA	1994	Long
81	Calrose	USA	1948	Medium

*These varieties were released in Turkey in 1986 and 1980, respectively.

Table 2. Information for 17 SSR marker loci used in this study, including name, repeat units, chromosome number, number of alleles (N), size range in bp, effective number of alleles (NE), gene diversity (h), Shannon's information index (I), and PIC values for 81 rice varieties.

SSR locus	Repeat unit	Chrom.	Size range	N	NE	h	I	PIC
RM5	$(GA)_{14}$	1	123-132	5	1.33	0.25	0.53	0.24
RM154	$(GA)_{21}$	2	196-210	8	3.44	0.71	1.52	0.67
RM338	$(CTT)_6$	3	194-197	2	1.12	0.11	0.22	0.10
RM161	$(AG)_{20}$	5	182-210	5	3.27	0.69	1.52	0.63
RM133	$(CT)_8$	6	243-245	2	1.03	0.03	0.07	0.03
RM162	$(AC)_{20}$	6	222-264	9	2.48	0.61	1.33	0.58
RM125	$(GCT)_8$	7	139-144	3	1.05	0.05	0.13	0.05
RM455	$(TTCT)_5$	7	148-150	2	1.08	0.05	0.16	0.05
RM284	$(GA)_8$	8	164-167	2	1.98	0.50	0.69	0.37
RM316	$(GT)_8-(TG)_9(TTTG)_4(TG)_4$	9	215-231	4	1.65	0.53	0.75	0.47
RM474	$(AT)_{13}$	10	271-279	4	2.47	0.60	1.09	0.54
RM144	$(ATT)_{11}$	11	237-284	7	2.75	0.63	1.25	0.56
RM19	$(ATC)_{10}$	12	231-242	2	1.34	0.25	0.42	0.22
Mean			55/4.23		1.92	0.39	0.74	0.35

PCR amplification was performed with an initial denaturation at 94°C for 5 min, 30 cycles at 94°C for 1 min, 55° to 67°C (the annealing temperature was marker-dependent) for 1 min, 72°C for 1 min, eight cycles at 94°C for 30 s, 53°C for 45 s, and 72°C for 45 s, and a final extension at 72°C for 10 min. A set of four PCR products (each 0.75 μL) labeled with a different dye was combined with 0.14 μL GeneScan-500 LIZ® size standards (Applied Biosystems) and 6.86 μL Hi-Di™ formamide (Applied Biosystems), denatured at 94°C for 5 min, chilled on ice, and separated on an ABI 3130xl Genetic Analyzer (Applied Biosystems).

Data analysis

The SSR amplicon fragments were analyzed using the GeneMapper software v3.7

(Applied Biosystems), as described in the user manual. The reproducibility of the DNA profiles for the retrotransposon marker systems was tested by repeating the PCR amplification with each of the primers selected. Only strong and clear bands were considered for analysis. The iPBS-retrotransposon bands were scored using a binary scoring system based on the presence or absence of bands as 1 and 0, respectively. The typical bands scored were bright and well separated from other bands, and faint bands were not considered in order to avoid the scoring of artificial bands. Gene diversity (*h*) and Shannon's information index (*I*) (Lewontin, 1972) were calculated using the software program POPGENE v1.32 (Yeh et al., 1999). Polymorphism information content (PIC) values for the SSR markers were calculated using the PowerMarker software. For the iPBS-retrotransposon markers, the PIC was calculated using the PICcalc web tool (http://w3.georgikon.hu/pic/english/default.aspx) according to the description by Nagy et al. (2012). Genetic distance coefficients between varieties were estimated according to the Jaccard (1908) index for pairwise comparisons based on the proportion of shared alleles for all primers, and a neighbor-joining cluster was constructed using the genetic distance coefficients with the R software (Paradis, 2012), for both marker systems separately and combined. In order to compare the marker systems, the genetic distances obtained by the iPBS-retrotransposon and SSR markers were compared using Mantel statistics (Mantel, 1967) in R.

Table 3. Primer sequences for iPBS-retrotransposon analysis used in this study.

Primer iPBS	Sequence	Annealing temperature (°C)	Number of bands			Diversity (mean)		
			Total	Polymorphic	%	*h*	*I*	PIC
2232	AGAGAGGCTCGGATACCA	55.4	13	12	0.92	0.31	0.32	0.37
2237	CCCCTACCTGGCGTGCCA	55.0	10	9	0.90	0.27	0.33	0.38
2277	GGCGATGATACCA	52.0	13	11	0.85	0.22	0.35	0.37
2376	TAGATGGCACCA	52.0	7	4	0.57	0.22	0.18	0.38
2382	TGTTGGCTTCCA	50.5	9	6	0.67	0.27	0.28	0.37
2383	GCATGGCCTCCA	50.5	10	6	0.60	0.14	0.22	0.28
2391	ATCTGTCAGCCA	52.6	13	10	0.77	0.21	0.32	0.36
2395	TCCCCAGCGGAGTCGCCA	52.8	12	11	0.92	0.29	0.45	0.34
2399	AAACTGGCAACGGCGCCA	52.0	9	8	0.89	0.15	0.27	0.27
Total			**96**	**77**	-	-	-	-
Average			**10.6**	**8.5**	**0.79**	**0.23**	**0.30**	**0.35**

h = gene diversity; *I* = Shannon's information index (Lewontin 1972).

The population structure and inference of admixture ancestry were assessed with a model-based clustering method implemented in STRUCTURE 2.1. In the structure program, the numbers of populations (K) were assumed; each was characterized by a set of allele frequencies at each locus. Different K values were tested (from 1 to 10), with 10 independent runs for each value, 50,000 burn-in periods, and 150,000 Monte Carlo Markov Chain iterations (Evanno et al., 2005). The change in the log probability of data between successive K values obtained from the STRUCTURE output was used to determine the true number of subpopulations. Subpopulations were assembled by grouping accessions with a membership probability higher than 0.70, while accessions with a lower membership probability were assigned to mixed groups.

RESULTS

Microsatellite diversity

Thirteen SSR markers were used to fingerprint all of the rice varieties. Fifty-five bands were detected at 13 polymorphic microsatellite markers distributed over the 12 chromosomes of

the rice genome across the 81 commercial rice varieties. The number of alleles per locus ranged from two (RM338, RM133, RM455, RM284, and RM19) to nine (RM162), with an average of 4.23 alleles across 13 loci (Table 2). The overall size of the amplified products ranged from 123 bp (RM5) to 284 bp (RM144), and the *h* ranged from 0.03 (RM133) to 0.71 (RM154), with a mean value of 0.39 among the 81 varieties. *I* values for the SSR markers ranged from 0.07 (RM133) to 1.52 (RM154 and RM161). The highest PIC value was 0.67 for the RM154 primer, whereas the lowest was 0.03 for RM133, with an average value of 0.35 per locus (Table 2). A Jaccard genetic distance matrix was used to construct a dendrogram for identifying relationships between the varieties. The neighbor-joining analysis clearly split all of the rice cultivars into three main groups, A, B, and C. Group A was the smallest and contained only four cultivars, while group B had 24 cultivars. Group C harbored 53 cultivars and was the largest group, and was divided into two subgroups, C1 and C2 (Figure 1).
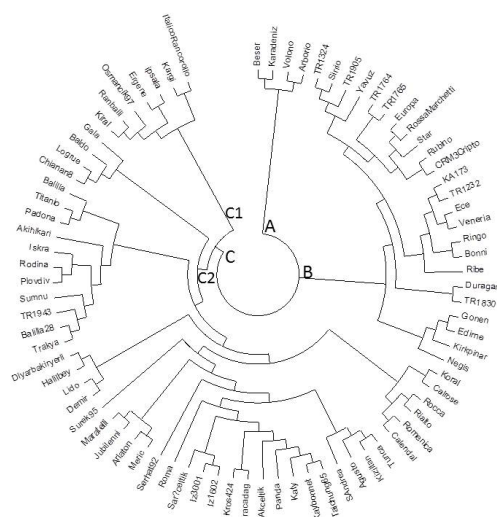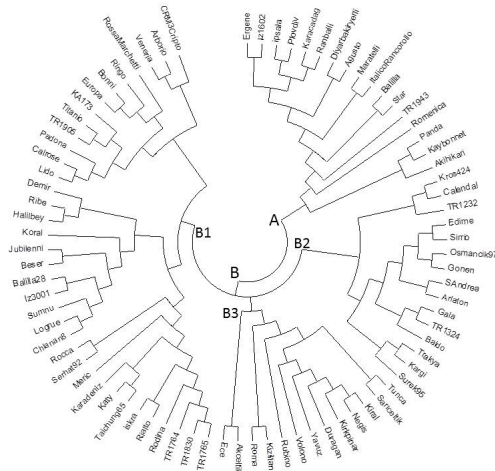

Figure 1. Neighbor-joining cluster of 81 rice varieties based on data of SSR marker.

Retrotransposon diversity

A total of 83 iPBS primers were screened among the 81 rice varieties. Nine primers yielding clear and reproducible banding profiles were selected for further analysis, and produced 96 scorable bands, 19 of which were monomorphic and 77 polymorphic. The number of bands per primer varied from 7 (2376) to 13 (2232, 2277, and 2391), with an average of 10.6 bands per primer (Table 3). The maximum number of polymorphic bands (12) was obtained using the 2232 primer, while the minimum (4) was recorded in 2376, with an average of 8.5 polymorphic bands per primer (Table 3). The polymorphism percentage ranged from 57 to 92%, with an average of 79% across all of the varieties. The *h* values ranged from 0.14 (2383) to 0.31 (2232), with a mean value of 0.23. Mean *I* values ranged from 0.18 (2376) to 0.45 (2395). The average PIC value was 0.35, and ranged from 0.27 (2399) to 0.38 (2376 and 2237) (Table 3). The neighbor-joining dendrogram revealed two main groups (A and B). Group A consisted of 17 cultivars of different origins, while group B was subdivided into the subclusters B1, B2, and B3 (Figure 2). B1 possessed 36 cultivars, B2 harbored 15, and B3 had 13.
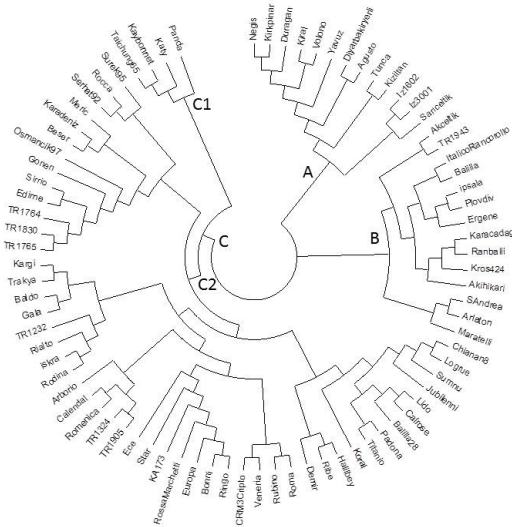
Figure 2. Neighbor-joining cluster of 81 rice varieties based on data of iPBS-retrotransposons marker.

Combined analysis of iPBS-retrotransposon and SSR marker data

To gain a clearer picture of the genetic relationships between the rice varieties, the two DNA marker datasets (iPBS-retrotransposons and SSRs) were combined, and the genetic similarity between the rice varieties was measured. The average genetic distance between the rice varieties was 0.295, and the Logrue and Chianan varieties were the most similar (D = 0.071), whereas the lowest similarity was between Kaybonet and Kiral (D = 0.495). The neighbor-joining analysis divided the rice varieties into three main groups (A, B, and C). Cluster A harbored 13 varieties, cluster B consisted of 14 varieties, and cluster C was the largest group with 54 varieties. Cluster C was differentiated into two subclusters, C1 and C2 (Figure 3).

Figure 3. Neighbor-joining cluster of 81 rice varieties based on combined data of iPBS-retrotransposons and SSR marker.

Using a permutation randomization approach, we found a statistically significant correlation in genetic distance values between the marker systems (r = 0.73, P = 0.008). A combined dataset (SSRs and iPBS-retrotransposons) was used for the STRUCTURE analysis, where the number of groups (K) varied from 1 to 10. The estimated ∆K value (Evanno et al., 2005) was 302.29 for the 81 rice accessions, which represented the three subpopulations (Figure 4).
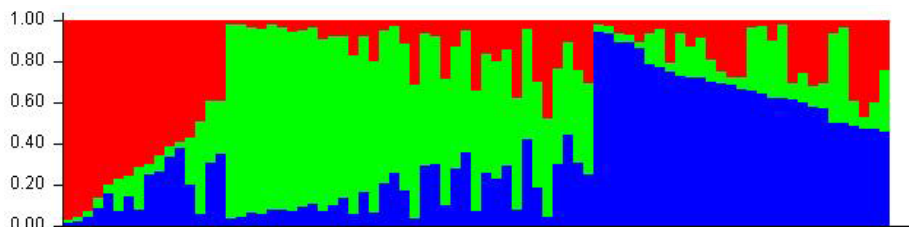
Figure 4. Population structure of 81 rice varieties.

DISCUSSION

Until the 1980s, very little research had been conducted on rice breeding in Turkey. Several rice varieties were introduced from various temperate countries (Manners, 2013), and farmers began to grow rice in most regions of Turkey. Italian rice varieties such as Baldo, Rocca, and Ribe are still grown in Turkey, but the amount is decreasing with the release of Turkish rice varieties. A National Rice Research and Breeding Project was started in 1970, and quickly produced new rice varieties such as Ergene, Meriç, Serhat, and Osmancik. Currently, nearly 100% of the rice varieties grown in Turkey were developed by local Turkish rice breeding programs. Turkish rice breeding programs focus on short growth duration, cold tolerance, blast resistance, and adaptation to local ecological conditions. Besides these traits, rice varieties in Turkey need to exhibit grain characteristics that are preferred by local Turkish consumers, such as medium-thick grains.

We measured the diversity of the rice varieties using SSR markers. Genetic diversity at the gene level can be measured by the number of alleles per locus. The 17 SSR markers used for this study were mapped previously, and four markers were monomorphic. The average number of alleles in our study was 4.23, and ranged from 2 to 9. The average number and range of alleles measured by SSR markers were lower than those obtained by previous studies in different countries (Garland et al., 1999; Ravi et al., 2003; Yu et al., 2003; Jain et al., 2004; Giarrocco et al., 2007; Thomson et al., 2007; Zhang et al., 2009). There are several explanations for the higher genetic diversity detected in rice from other countries using SSR markers. Firstly, the rice germplasm used in those studies was different from that in our study, e.g., Asian rice has a higher diversity than European rice. Secondly, the number of SSR markers varied between studies, which could have resulted in different levels of diversity being measured.

In the present study, we demonstrated the utility of iPBS-retrotransposon markers in investigating the genetic relationships between rice varieties, including Turkish rice varieties and others introduced from Italy, France, Bulgaria, Russia, the IRRI, and the USA, which are commonly used as parent lines in Turkish rice breeding programs. iPBS-retrotransposon markers have been developed only recently, and they have not been previously utilized for diversity studies in rice. The percentage of polymorphisms found in our rice germplasm using iPBS-retrotransposon markers was lower than that reported by Branco et al. (2007), who discovered approximately

96% polymorphisms among 51 rice varieties from Brazil, Japan, and the Philippines using inter-retrotransposon amplified polymorphism markers. Ravi et al. (2003) detected 90% polymorphisms among 40 cultivated varieties and five wild relatives from India using RAPD markers.

It is crucially important to know which type of marker should be used for reliable estimates of diversity, and to know if marker systems other than SSRs can be effectively used for the identification of rice varieties. To study the potential of both markers in detecting genetic polymorphisms among different rice varieties, h, I, and PIC values were calculated. The h and I values for the retrotransposon markers were lower than those for the SSR markers, which demonstrate that microsatellites are more powerful markers for the identification of variety variation. PIC values reflect the level of genetic diversity in a population, and provide a somewhat better estimator of diversity than the raw number of alleles, because they take into account the relative frequencies of each allele present. Both marker techniques differ in their nature, information content, and mode of inheritance. There are many relevant questions about the utility of markers for the identification of polymorphisms, such as how many and what types of polymorphic markers should be used in genetic diversity studies. These questions can be answered by calculating the PIC. In this study, the average PIC value measured using retrotransposon markers was 0.35, and ranged from 0.27 to 0.38, while the average PIC value using SSR markers was 0.35, with a range of 0.03 to 0.67. The average PIC value was almost identical using both marker systems, showing the discriminatory power of iPBS-retrotransposon markers. The average PIC value based on the SSR data used in this study was comparable with values from other rice genetic diversity studies, but lower than those reported in studies on rice varieties from Argentina (0.69) that were performed using 28 SSRs (Giarrocco et al., 2007), from India (0.61) using 30 SSRs (Jain et al., 2004), from Australia (0.74) using 10 SSRs (Garland et al., 1999), or from India using 23 SSRs among 91 landraces (Das et al., 2013).

Only a small set of retrotransposon markers can be evaluated for genetic diversity and cultivar identification in comparison with SSR markers, particularly in developing countries where resources are scarce due to a lack of infrastructure and funds (Nagy et al., 2012). The clustering of the rice cultivars was almost the same, but there were minor differences, as the marker systems amplify different genomic regions. We found a statistically significant correlation between the marker systems.

The neighbor-joining method divided the rice varieties into three main groups (A, B, and C). Group A consisted of 13 varieties, including Volono and Agusto from Italy and all of the varieties from Turkey. These two Italian varieties are used as parent lines in Turkish rice breeding programs. Group B included 14 varieties from all of the countries except the USA, while group C harbored the remaining 54 varieties. Group C was divided into two main subgroups (C1 and C2). Subgroup C1 consisted of four varieties: two varieties from the USA and two from Italy and the IRRI. Subgroup C2 had the highest number, with 50 varieties and genotypes. Turkish varieties always clustered with Italian varieties, followed by French varieties.

The short breeding history of Turkish rice (around 40 years) makes it possible to analyze rice breeding during different periods, and the breeding history of rice in Turkey can be divided into four stages. From 1970 to 1980, rice-breeding efforts were established by the introduction of rice varieties from foreign countries. During that time, four rice varieties from Italy (Baldo, Rocca, Veneria, and Ribe), three from Bulgaria (Plovdiv, Rodina, and Ranballi), and one from Russia (Krasnodarsky-424) were introduced, and experiments were conducted to attempt to adapt these foreign rice varieties to the different agro-ecological conditions of Turkey. The foreign varieties then started to be grown. During the second phase (after 1980), some of the introduced varieties were registered as varieties.

In the third phase (after 1990), extensive crosses were made using the introduced varieties as parent lines in local Turkish rice breeding programs. The pedigrees of some rice varieties from Turkey demonstrate that at least one of the parents in their pedigree was primarily from Italy, followed by France, Bulgaria, and Russia. In order to increase diversity further, more emphasis should be placed on introducing new germplasm from different geographical locations, particularly Asia and South America. In 1990, Turkish rice production was dominated by introduced varieties. In the late 1990s, Turkish-bred varieties (mostly bred from crosses between introduced varieties) were released. The country now meets more than 90% of its domestic demand for rice. Research remains a key factor behind the impressive increases in Turkey's rice production. Fifteen years ago, Turkey grew mainly foreign varieties, and now nearly 100% of the rice grown in Turkey is Turkish.

The population structure analysis revealed that all of the varieties used in this study were derived from three subpopulations, or three gene pools. Only a few varieties were used as parent lines in the Turkish Rice Improvement Program, and these primarily originated from temperate countries. Therefore, a renewed emphasis should be placed on new sources of germplasm in the breeding pool in order to widen the genetic diversity for developing elite varieties. Future parent selection for crosses should consider using genotypes with large genetic distances between them, and avoid genotypes with common backgrounds. Introductions should be made from countries with great rice genetic diversity, such as China, the Philippines, India, Pakistan, and Bangladesh.

The genomic analysis of many crops, including rice, has demonstrated that genetic relationships between different varieties or species using different marker systems can vary significantly. Different marker technologies may amplify different genomic regions. In the present study, we used two different marker systems (iPBS-retrotransposon and SSR markers). These markers differ in the nature of the evolutionary mechanism underlying their variation, and their distribution in the plant genome. iPBS-retrotransposon markers are dominant, while SSR markers are co-dominant. In summary, the patterns of diversity revealed by the iPBS-retrotransposon markers were almost identical to those found using SSR markers, as indicated by the high association found by the Mantel test. The iPBS-retrotransposon technique was successfully applied to rice genetic analysis, and iPBS-retrotransposon fingerprinting was found to be rapid, efficient, cost effective, and had the added advantages of requiring only one primer and a single PCR amplification step. Because this technique is easy to perform, requires a small amount of DNA, is highly reproducible, yields a high number of polymorphisms, and is not species-specific, it could be effectively used for any crop. Therefore, iPBS-retrotransposon fingerprinting of rice genotypes could be used to rapidly characterize a large number of rice accessions, in order to increase diversity and manage breeding populations more effectively. Furthermore, the grouping of rice varieties based on their agronomic traits may serve as a resource for future studies, and selecting diverse parent lines based on molecular data could lead to the improvement of rice.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

Andeden EE, Baloch FS, Kilian B and Özkan H (2013). iPBS-retrotransposons based genetic diversity and relationship among wild annual *Cicer* species. *J. Plant Biochem. Biotechnol.* 22: 453-466. http://dx.doi.org/10.1007/s13562-012-0175-5

Baloch FS, Kurt C, Arıoğlu H and Özkan H (2010). Assaying of diversity among Soybean (*Glycin max* L.) and peanut (*Arachis hypogaea* L.) genotypes at DNA level. *Turk. J. Agric. For.* 34: 285-301.

Baloch FS, Derya M, Andeden EE, Alsaleh A, et al. (2015). Inter-primer binding site retrotransposon and inter-simple sequence repeat diversity among wild *Lens* species. *Biochem. Syst. Ecol.* 58: 162-168. http://dx.doi.org/10.1016/j.bse.2014.12.002

Bao Y, Zhou HF, Hong DY and Ge S (2006). Genetic diversity and evolutionary relationships of *Oryza* species with the B- and C-genomes as revealed by SSR markers. *J. Plant Biol.* 49: 339-347. http://dx.doi.org/10.1007/BF03178809

Baránek M, Mészáros M, Sochorová J, Čechová J, et al. (2012). Utility of retrotransposon-derived markers systems for differentiation of presumed clones of the apricot cultivar Velkopavlovická. *Sci. Hortic. (Amsterdam)* 143: 1-6. http://dx.doi.org/10.1016/j.scienta.2012.05.022

Branco CJS, Vieira EA, Malone G, Kopp MM, et al. (2007). IRAP and REMAP assessments of genetic similarity in rice. *J. Appl. Genet.* 48: 107-113. http://dx.doi.org/10.1007/BF03194667

Buyukunal Bal EB and Bay S (2010). Genetic analysis of Turkish rice varieties (*Oryza sativa* L.) using seed storage proteins and RAPD markers. *Eur. Food Res. Technol.* 230: 609-617. http://dx.doi.org/10.1007/s00217-009-1201-2

Cai X, Fan J, Jiang Z, Basso B, et al. (2013). The puzzle of Italian rice origin and evolution: determining genetic divergence and affinity of rice germplasm from Italy and Asia. *PLoS One* 8: e80351. http://dx.doi.org/10.1371/journal.pone.0080351

Cömertpay G, Baloch FS, Kilian B, Ülger AC, et al. (2012). Diversity assessment of Turkish maize landraces based on fluorescent labelled SSR markers. *Plant Mol. Biol. Rep.* 30: 261-274. http://dx.doi.org/10.1007/s11105-011-0332-3

Das B, Sengupta S, Parida SK, Roy B, et al. (2013). Genetic diversity and population structure of rice landraces from Eastern and North Eastern States of India. *BMC Genet.* 14: 71. http://dx.doi.org/10.1186/1471-2156-14-71

Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh leaf tissue. *Focus* 12: 13-15.

Evanno G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620. http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x

Fang-Yong C and Ji-Hong L (2014). Germplasm genetic diversity of *Myrica rubra* in Zhejiang Province studied using inter-primer binding site and start codon-targeted polymorphism markers. *Sci. Hortic. (Amsterdam)* 170: 169-175. http://dx.doi.org/10.1016/j.scienta.2014.03.010

Gailite A, Ievinsh G and Rungis D (2011). Genetic diversity analysis of Latvian and Estonian *Saussurea esthonica* populations. *Environ. Exp. Biol* 9: 115-119.

Garland SH, Lewin L, Abedinia M, Henry, et al. (1999). The use of microsatellite polymorphisms for the identification of Australian breeding lines of rice (*Oryza sativa* L.). *Euphytica* 108: 53-63. http://dx.doi.org/10.1023/A:1003688612179

Giarrocco LE, Marassi MA and Salerno GL (2007). Assessment of the genetic diversity in Argentine rice cultivars with SSR markers. *Crop Sci.* 47: 853-858. http://dx.doi.org/10.2135/cropsci2005.07.0198

Guo DL, Zhang JY and Liu CH (2012). Genetic diversity in some grape varieties revealed by SCoT analysis. *Plant Mol. Biol. Rep.* 39: 5207-5313.

Hashimoto Z, Mori N, Kawamura M, Ishii T, et al. (2004). Genetic diversity and phylogeny of Japanese sake-brewing rice as revealed by AFLP and nuclear and chloroplast SSR markers. *Theor. Appl. Genet.* 109: 1586-1596. http://dx.doi.org/10.1007/s00122-004-1794-6

Jaccard P (1908). Nouvelles recherché sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat.* 44: 223-270.

Jain S, Jain RK and McCouch SR (2004). Genetic analysis of Indian aromatic and quality rice (Oryza sativa L.) germplasm using panels of fluorescently-labeled microsatellite markers. *Theor. Appl. Genet.* 109: 965-977. http://dx.doi.org/10.1007/s00122-004-1700-2

Joshi SP, Gupta VS, Aggarwal RK, Ranjekar PK, et al. (2000). Genetic diversity and phylogenetic relationship as revealed by inter-simple sequence repeat (ISSR) polymorphism in the genus *Oryza. Theor. Appl. Genet.* 100: 1311-1320. http://dx.doi.org/10.1007/s001220051440

Kalendar R, Antonius K, Smýkal P and Schulman AH (2010). iPBS: a universal method for DNA fingerprinting and retrotransposon isolation. *Theor. Appl. Genet.* 121: 1419-1430. http://dx.doi.org/10.1007/s00122-010-1398-2

Lewontin RC (1972). Testing the theory of natural selection. *Nature* 236: 181-182. http://dx.doi.org/10.1038/236181a0

Lu H, Redus MA, Coburn JR, Rutger N, et al. (2005). Population structure and breeding patterns of 145 U.S. rice cultivars based on SSR marker analysis. *Crop Sci.* 45: 66-76. http://dx.doi.org/10.2135/cropsci2005.0066

Manners G (2013). Rice booms in Turkey. *Rice Today* 12: 1.

Mantel N (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209-220.

Mehmood A, Jaskani MJ, Ahmad S and Ahmad R (2013). Evaluation of genetic diversity in open pollinated guava by iPBS

primers. *Pak. J. Agr. Sci.* 50: 591-597.

Nagy S, Poczai P, Cernák I, Gorji AM, et al. (2012). PICcalc: an online program to calculate polymorphic information content for molecular genetic studies. *Biochem. Genet.* 50: 670-672. http://dx.doi.org/10.1007/s10528-012-9509-1

Paradis E (2012). Analysis of Phylogenetics and Evolution with R. 2 nd edn. Springer New York Dordrecht Heidelberg London.

Ravi M, Geethanjali S, Sameeyafarheen F and Maheswaran M (2003). Molecular marker based genetic diversity analysis in rice (*Oryza sativa* L.) using RAPD and SSR markers. *Euphytica* 133: 243-252. http://dx.doi.org/10.1023/A:1025513111279

Saini N, Jain N, Jain S and Jain RK (2004). Assessment of genetic diversity within and among Basmati and non-Basmati rice varieties using AFLP, ISSR and SSR markers. *Euphytica* 140: 133-146. http://dx.doi.org/10.1007/s10681-004-2510-y

Sürek H (1997). Rice production and research activities in Turkey. In: Activités de recherche sur le riz en climat méditerranéen (Chataigner J, ed.). CIHEAM, Montpellier, 165-173.

Temnykh S, Park WD, Ayers N, Cartinhour S, et al. (2000). Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 100: 697-712. http://dx.doi.org/10.1007/s001220051342

Thomson MJ, Septiningsih EM, Suwardjo F, Santoso TJ, et al. (2007). Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. *Theor. Appl. Genet.* 114: 559-568. http://dx.doi.org/10.1007/s00122-006-0457-1

Yeh FC, Yang RC and Boyle T (1999). POPGENE version 1.32: Microsoft window-based freeware for population genetics analysis. University of Alberta, Edmonton.

Yu SB, Xu WJ, Vijayakumar CHM, Ali J, et al. (2003). Molecular diversity and multilocus organization of the parental lines used in the International Rice Molecular Breeding Program. *Theor. Appl. Genet.* 108: 131-140. http://dx.doi.org/10.1007/s00122-003-1400-3

Zhang D, Zhang H, Wang M, Sun J, et al. (2009). Genetic structure and differentiation of *Oryza sativa* L. in China revealed by microsatellites. *Theor. Appl. Genet.* 119: 1105-1117. http://dx.doi.org/10.1007/s00122-009-1112-4