INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
19–1319/2020/23–5–943–950
DOI: 10.17957/IJAB/15.1373
http://www.fspublishers.org

Full Length Article

DNA Fingerprinting and Population Structure of Date Palm Varieties Grown in Punjab Pakistan using Simple Sequence Repeat Markers

Shakra Jamil, Rahil Shahzad*, Shamsa Kanwal, Erum Yasmeen, Sajid Ur Rahman and Muhammad Zaffar Iqbal
Agricultural Biotechnology Research Institute, Ayub Agriculture Research Institute, Faisalabad 38040, Pakistan
*For correspondence: rahilshahzad91@gmail.com
Received 26 August 2019; Accepted 23 December 2019; Published 03 March 2020

Abstract

Fingerprinting is a powerful tool to distinguish closely related cultivars. In this study 209 SSR markers were used for DNA fingerprinting of 13 date palm varieties of Punjab, Pakistan. 191 SSR markers were amplified and 163 were found polymorphic. A total of 927 alleles were amplified (with 4.84 alleles per locus) among which 597 were polymorphic (with 3.65 polymorphic alleles per locus). Maximum number of alleles and polymorphic alleles per locus were 23 and 16 respectively for marker PDAG1018. Maximum polymorphic information content (0.95) was also recorded for PDAG1018. Genetic similarity coefficients were calculated and dendrogram was constructed using Unweighted Pair Group Method with Arithmetic Averages. Khalas, AseelKhumba & Khupra were assigned to group I whereas Ajwa, Zahidi and Khurma were classified to group III. Group II contains 07 genotypes which were subdivided to 03 subgroups. Shamran and Amber were classified to group IIa, Hallawi, Khudri & Shakri to group IIb and Haleni & Barhi to IIc. Structure analysis results were in line to cluster analysis. The genotypes classified in group II were admixture with mixed genetic backgrounds. Whereas group I and III members have similar genetic backgrounds within group and have common ancestors. Further, DNA fingerprints were available for 12 varieties out of 13 except Hallawi which is identified using 02 step methods. The information generated in this study will be helpful in variety registration process under Plant Breeders Rights and variety protection and right choice of SSRs for genetic diversity studies. © 2020 Friends Science Publishers

Keywords: Alleles; Cluster analysis; Date Palm; Genetic Similarity coefficients; Polymorphic Information Content; Simple Sequence Repeat markers; Structure analysis

Introduction

Phoenix dactylifera L., commonly known as Date Palm, is dioecious, perennial, monocotyledonous ancient fruit tree (2n=36 chromosomes) belonging to Arecaceae (*Coryphoideae*) family and has a great nutritional and socioeconomic importance (Enan and Ahemad 2014). The genome size for date palm is estimated to be 658 Mbp long and is most commonly cultivated in Middle East, Gulf countries and North Africa for fruit, food and ornamental purposes (Al-Dous *et al*. 2011). In Pakistan, the cultivation of date palm has been greatly elevated in previous years, whereas, province Sindh has been considered as the center of date palm biodiversity (Akhtar *et al*. 2014). Globally, Pakistan is considered to be the fifth largest dates producing country (622,000 tons per year) sharing approximately 9% of the world's total production (MirBahar *et al*. 2016).

Humans had significantly influenced date palm through cultivation and selection making them vulnerable to biotic and abiotic stress and severely harmed genetic conservation (Yusuf *et al*. 2015). Previously, various morphological descriptors including leaves, spines, weight, texture and fruit had been used to characterize date palm cultivar but these features are laborious, time-consuming, unreliable and ambiguous as immensely affected by environmental conditions and developmental stages (Elhoumaizi *et al*. 2002; Al-Ruqaishi *et al*. 2008). Furthermore, there is narrow difference in morphological traits among date palm species. Biochemical approaches such as isozyme analysis was employed to discriminate date palm varieties but exhibit low level of polymorphism making them difficult to characterize (Gothwal *et al*. 2013). Hence, there is need to utilize DNA based polymorphic markers to unmask genetic differentiation among date palm germplasm.

The knowledge of genetic diversity is of prime importance to boost up the breeding programs and preserve Pakistani date palm germplasm. Genetic markers are an efficient tool for the identification of cultivars and estimation of relatedness. PCR based markers like RAPD, AFLP, ISSR, RFLP and SNPs are robust, easy and have become promising tool to assess genetic diversity and fingerprinting (Adawy *et al*. 2005; Eid *et al*. 2019; ElKadri *et al*. 2019). Microsatellites or Simple sequence

repeats (SSR) are multi-locus, short tandem repeats that are randomly present across the genome of eukaryotes. These show high rates of mutation and are used to study genetic diversity, fingerprinting, gene tagging and genome mapping because of high degree of polymorphism and independence from environmental effects (Pashley *et al*. 2006; Guichoux *et al*. 2011). They are easy to amplify using PCR, highly reproducible, highly polymorphic and higher PIC value gives them priority over other markers. Polymorphism Information content (PIC) depends on the numbers of alleles detected by a particular marker and its frequency in the given population which indicates the genetic variability in the population (Elmeer *et al*. 2011). Moreover, analysis of structure of a population enables an intensive understanding of genetic diversity, facilitates the association mapping and defines the geographical background of germplasm (Nielsen *et al*. 2014).

To obtain a deeper comprehension of genetic organization, we have employed 209 SSR markers to produce genotypic data providing unique allelic profile in order to discriminate 13 date palm varieties. Main objective of this study is to examine genetic relationship in date palm gene pool available in Punjab Pakistan providing a standardize-able reference based on DNA fingerprints establishing unique genotypic identity. Further degree of polymorphism of the markers using polymorphic information content (PIC) was worked out to explain the informativeness of SSR markers. To assess the genetic diversity on the basis of geography, we inferred the population structure analysis between the markers and the traits using model-based approach indicating the distribution of alleles. Moreover, the genetic relationship among genotypes was inferred using UPGMA cluster analysis representing the closeness and divergence among date palm cultivars. Present study will be helpful for germplasm management in order to improve their conservation and production of elite cultivars.

Materials and Methods

Plant materials

In the present study, a population of thirteen cultivars of Date Palm had been evaluated for varietal identification (Fig. 1). All the experimentation was conducted at Agriculture Biotechnology Research Institute (ABRI), Faisalabad, Pakistan during 2018–2019. The plant material was comprised of young leaflets collected from 12–15 year-old palms plants from the Horticultural Research Station Bahawalpur. Leaves were dried in silica gel and stored at -40ºC.

DNA extraction and PCR

The genomic DNA was isolated using modified Cetyl Trimethyl Ammonium Bromide (CTAB) method Allen *et al*. (2006). DNA was quantified using Nanodrop spectrophotometer (ND 2000, Thermo Scientific, U.S.A.). DNA was considered pure when A_{260}/A_{280} ratio ranged between 1.80 and 2.0. The quality of extracted DNA was also assessed by loading DNA 20 ng/ μL on 0.8% (w/v) agarose gel stained with ethidium bromide. All the DNA extracts were stored at -40°C. 209 SSR markers were selected (Table S1) and synthesized according to information provided by (Elmeer *et al*. 2011; Mathew *et al*. 2014; Elmeer and Mattat 2015; Al-Faifi *et al*. 2016; Racchi and Camussi 2018).

PCR reaction was conducted in thermal cycler with total reaction volume of 25 μL including 20 ng/μL genomic DNA of each variety, 0.6 μM of each forward and reverse primers and 12 μL of green master mix for different SSR markers. The following temperature conditions were applied for amplification: initial denaturation 94ºC for 5 min, 35 cycles of denaturation 94ºC for 1 min, annealing at variable temperatures according to primers for 1 min (Table 1), extension at 72ºC for 1 min. Final extension at 72ºC for 7 min. The amplified products were stored at 4ºC.

Polyacrylamide gel electrophoresis (PAGE) analysis

All the amplified products were resolved on Vertical Gel Electrophoresis System model POWERPRO-3AMP (cleaver scientific limited) using 6% PAGE performed at 16 watts power followed by Silver nitrate staining for visualization. The staining protocol is described in detail by Caetano-Anolles (1997). Images were captured using Syngene trans-illuminator.

Statistical analysis

The data for SSR markers were taken in the form of binary matrix. The presence of band was scored as 1, whereas the absence was scored as 0. To detect polymorphism among thirteen cultivars on the basis of alleles, the distance matrix was computed using Un-weighted pair Group Method of Arithmetic Means (UPGMA) with NTSYSpc 2.0 version and dendrogram was generated. The genetic diversity levels and geographic structure of thirteen date palm genotypes were assessed using model-based Bayesian clustering approach implemented in STRUCTURE v. 2.3.4 (Pritchard *et al*. 2000). Genotyping data of 209 microsatellites were used to determine population's structure of various date palm varieties. Population structure analysis was performed using following Parameters: no admission model; K ranging from 1 to 06; 10,000 Burn-in period; Reps. hypothetical populations' number (k) (03), number of in-iteration burns (10, 000), number of Markov chain Monte Carlo simulations (100000). Most likely number of cluster was determined by plotting LnP(K) values against ΔK values by selecting appropriate K value using Evanno Test (Evanno *et al*. 2005). Further Polymorphic Information contents (PIC) were calculated for all primers along with number of alleles, polymorphic alleles and different allelic diversity parameters (Table 1).

Results

SSR polymorphism

A total of 209 SSR markers were used to evaluate genetic diversity of thirteen date palm cultivars grown in Punjab Pakistan. A sum total of 937 alleles with a mean value of 4.84 alleles per locus were amplified. The microsatellites were highly polymorphic owning 597 polymorphic alleles with an average of 3.65 alleles per locus. The highest number of alleles and polymorphic alleles per microsatellite were 23 (for SSR PDAG1018) and 16 (for SSR mPdIRD05 and PDAG1018) respectively (Fig. 2). The results also indicate that 191 out of 209 markers (91.38%) exhibited amplification, among them, 163 generated polymorphic bands (85.34%) while, 28 appeared to be monomorphic and 18 failed to amplify. The PIC values for 209 markers ranged from 0.0 to 0.95 with a mean value of 0.64. The most informative markers were DPG1195, DPG1199 and PDAG1018 with PIC values of 0.92, 0.90 and 0.95 respectively (Table 1).

Cluster/dendrogram analysis

To examine the organization of genetic diversity within thirteen date palm cultivars, UPGMA (Unweighted Pair Group of Arithmetic Averages) based cluster analysis was conducted based on genomic-SSR data. For this purpose, Similarity matrix was first generated according to SHAN similarity index, followed by the dendrogram construction (Fig. 1a). The population was broken in to three major groups representing strong clustering patterns with similarity co-efficient ranging from 0.79 to 0.88. The group I contained three cultivars *i.e*., Khalas, AseelKhumba and Kkhupra. Group II was sub clustered into subgroups IIa, IIb and IIc. Group IIa consists of two cultivars Shamran and Amber exhibiting 0.83 similarity coefficient regardless of different geographical distribution; former from Iraq and latter from Saudi Arabia (Fig. 1a).

Group IIb has Hallawi, Khudri and Shakri. The highest similarity coefficient was observed between Hallawi and khudri (0.88) indicating that they are genetically closest varieties and are closely regrouped. The members of group IIc consists of Haleni and Barhi, they share 0.85 similarity coefficient. Group II mainly comprised of Saudi Arabian and Iraqi varieties except for Barhi which originated from Pakistan, represents a close relationship among the two origins. Lastly, group III is comprised of the varieties from both Iraqi and Saudia origin: Ajwa, Zahidi and Khurma. Ajwa and Zahidi exhibit a strong similarity coefficient of 0.86 (Fig. 1a). The clustering did not support geographical distribution.

On the other hand, dendrogram on the basis of their geographical origin was also constructed separately for Pakistani, Iraqi and Saudi Arabian varieties (Fig. 1b). According to Zango *et al*. (2017) AseelKhumba, Zahidi,

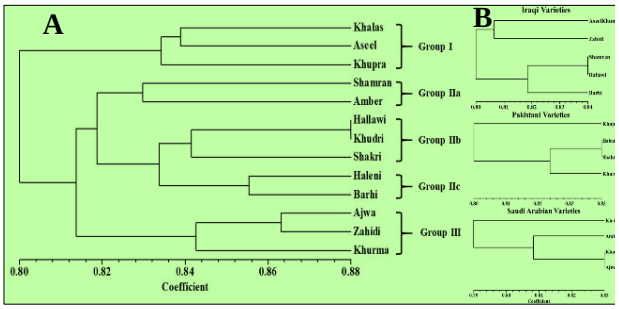

Fig. 1: Cluster analysis of 13 date palm varieties using unweighted Pair Group Method with Arithmetic Averages

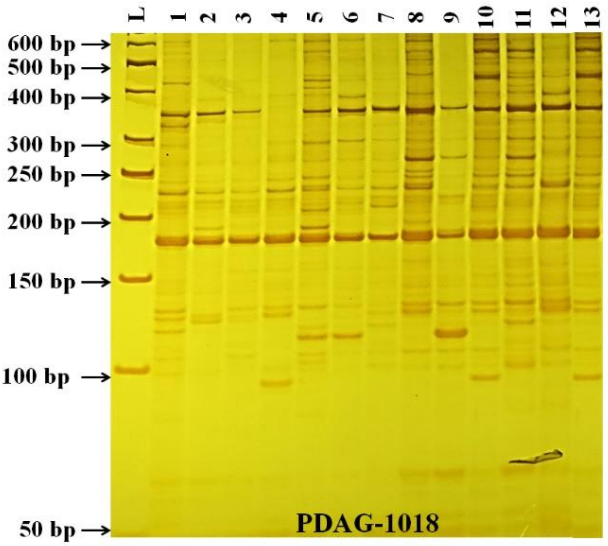

Fig. 2: Segregation pattern of most diverse SSR marker PDAG1018 across 13 date palm varieties run on 6% vertical polyacrylamide gel with 50 bp ladder (L)

Shamran, Hallawi and Barhi are Iraqi varieties, four varieties belong to Pakistan *i.e*., Khupra, Haleni, Shakri and Khurma, whereas, Khalas, Amber, Khurma and Ajwa are of Saudia Arabian origin. The Iraqi varieties showed 79 to 84% similarity wherein, Shamran & Hallawi, and Barhi & Hallawi shared 84% similarity. Pakistani varieties displayed more than 80% genetic similarity. Among four Pakistani varieties, Haleni and Shakri are closely related presenting 0.83 similarity coefficient. Cluster of Saudi Arabian varieties demonstrated 79 to 84% similarity indicating Ajwa and khudri regrouped together with 0.83 similarity coefficients. Overall, on an average, the Saudi Arabian varieties and Iraqi varieties are closely related indicating 81% similarity (Fig. 1b).

Genetic structure of date palm varieties

Model-based cluster analysis based on a Bayesian approach was carried out to infer the population structure among 13 date palm varieties using 209 SSR markers.

Table 1: List of SSR markers used in the study along with polymorphic information content (PIC), number of alleles (NOA) polymorphic alleles (PA) and annealing temperature (Ta)

Sr. No.	Marker Name	PIC	PA	NOA	Ta	Sr. No.	Marker Name	PIC	PA	NOA	Ta
1.	DP150	0.54	3	4	55	52.	DPG1709	0.48	2	2	56
2.	DP152	0.75	4	4	54	53.	DPG1713	0.50	1	2	56
3.	DP153	0.79	1	5	52	54.	DPG2109	0.50	2	2	56
4.	DP154	0.69	3	4	54	55.	DPG2110	0.41	2	2	56
5.	DP155	0.83	1	6	55	56.	DPG2111	0.72	1	4	56
6.	DP156	0.62	4	4	55	57.	DPG2112	0.50	0	2	56
7.	DP157	Not Amplified				58.	DPG2113	0.50	0	2	56
8.	DP158	0.59	2	3	55	59.	DPG2114	0.00	1	1	56
9.	DP159	0.84	8	8	55	60.	DPG2115	0.50	2	2	56
10.	DP160	0.00	1	1	52	61.	DPG2116	0.55	3	3	56
11.	DP162	0.00	1	1	55	62.	DPG2117	0.48	2	3	56
12.	DP163	0.00	0	1	55	63.	DPG2118	0.84	9	9	56
13.	DP164	0.00	0	1	55	64.	DPG2119	0.00	0	1	56
14.	DP165	0.59	1	3	55	65.	DPG2120	0.00	1	1	56
15.	DP166	0.00	0	1	55	66.	DPG2121	0.00	0	1	56
16.	DP167	0.50	2	4	55	67.	DPG2122	0.50	5	5	56
17.	DP168	0.61	2	3	55	68.	DPG2123	0.00	1	1	56
18.	DP169	0.50	0	2	57	69.	DPG2124	0.86	9	11	56
19.	DP170	0.47	2	2	52	70.	DPG2395	0.50	0	2	56
20.	DP171	0.87	4	8	56	71.	DPG2787	Not Amplified			
21.	DP172	0.66	1	3	54	72.	DPG3359	0.87	8	11	56
22.	DP173	0.00	1	1	55	73.	DPG3508	0.50	0	2	56
23.	DP174	0.85	11	11	55	74.	DPG4966	0.50	5	5	56
24.	DP175	0.83	2	4	60	75.	DPG4697	0.49	2	2	56
25.	DP176	0.81	2	6	55	76.	KSU-PDL53	0.61	4	4	55
26.	DP177	0.67	3	4	55	77.	KSU-PDL16	0.78	5	6	55
27.	DP178	Not Amplified				78.	KSU-PDL18	0.81	6	8	55
28.	DP179	0.83	7	7	55	79.	KSU-PDL18-2	0.43	1	2	55
29.	DPG0001	0.63	4	4	56	80.	KSU-PDL25	0.43	1	2	55
30.	DPG0002	0.67	1	3	56	81.	KSU-PDL29	0.00	0	1	55
31.	DPG0003	0.82	6	6	56	82.	KSU-PDL3	0.79	5	6	55
32.	DPG0004	0.83	2	4	56	83.	KSU-PDL39	0.77	5	5	55
33.	DPG0005	0.74	4	5	56	84.	KSU-PDL4	0.76	4	6	55
34.	DPG0006	0.78	7	7	56	85.	KSU-PDL5	0.82	6	6	55
35.	DPG0007	0.79	4	5	56	86.	KSU-PDL58	0.85	4	8	55
36.	DPG0008	Not Amplified				87.	KSU-PDL6	0.79	4	6	55
37.	DPG0009	0.67	0	3	56	88.	KSU-PDL61	0.79	4	6	60
38.	DPG0010	0.67	1	3	56	89.	KSU-PDL73	0.80	5	6	55
39.	DPG1184	0.82	4	6	56	90.	KSU-PDL74	0.67	3	3	55
40.	DPG1195	0.92	11	15	56	91.	KSU-PDL76	0.57	3	3	55
41.	DPG1196	0.50	0	2	56	92.	mPdCIR010	0.86	9	10	55
42.	DPG1197	0.84	3	7	56	93.	mPdCIR015	0.63	2	3	54
43.	DPG1198	0.00	1	1	56	94.	mPdCIR016	0.64	3	3	54
44.	DPG1199	0.90	13	16	56	95.	mPdCIR025	0.73	4	4	54
45.	DPG1202	0.61	1	3	56	96.	mPdCIR032	0.59	2	3	54
46.	DPG1297	0.67	0	3	56	97.	mPdCIR035	0.71	3	4	54
47.	DPG1701	0.67	0	3	56	98.	mPdCIR044	Not Amplified			
48.	DPG1702	0.80	8	8	56	99.	mPdCIR048	0.66	2	3	54
49.	DPG1703	0.50	0	2	56	100.	mPdCIR050	0.78	5	7	55
50.	DPG1704	0.39	2	2	56	101.	mPdCIR057	0.75	4	4	55
51.	DPG1705	0.86	9	9	56	102.	MPdCIR063	Not Amplified			
Sr. No.	**Marker Name**	**PIC**	**PA**	**NOA**	**Ta**	**Sr. No.**	**Marker Name**	**PIC**	**PA**	**NOA**	**Ta**
103.	DPG1706	0.81	5	7	56	157.	mPdCIR070	0.77	1	5	52
104.	DPG1707	0.65	2	4	56	158.	mPdCIR078	0.41	2	2	52
105.	DPG1708	0.64	2	3	56	159.	mPdCIR085	Not Amplified			
160.	mPdCIR090	0.77	5	6	54	160.	PDCAT11	0.61	3	3	54
161.	mPdCIR093	0.81	1	6	52	161.	PDCAT12	0.44	2	2	50
162.	mPdIRD01	0.81	1	6	60	162.	PDCAT13	0.75	4	5	50
163.	mPdIRD013	0.79	4	6	60	163.	PDCAT15	Not Amplified			
164.	mPdIRD03	0.67	3	3	60	164.	PDCAT17	0.73	4	4	54
165.	mPdIRD031	0.75	3	5	60	165.	PDCAT18	0.71	5	5	54
166.	mPdIRD033	0.75	3	5	60	166.	PDCAT2	0.89	5	10	50
167.	mPdIRD040	Not Amplified				167.	PdCAT20	0.62	3	3	54
168.	mPdIRD05	0.80	16	17	60	168.	PDCAT21	0.72	2	4	50
169.	mPdIRD07	0.62	2	4	60	169.	PDCAT3	0.81	6	6	50

Table 1: Continue

Table 1: Continue

170.	mPdIRD08	0.75	1	4	60	170.	PDCAT4	0.77	5	5	50
171.	mPdIRD10	Not Amplified				171.	PDCAT5	0.80	5	6	50
172.	mPdIRD11	0.86	5	8	60	172.	PDCAT8	0.66	5	5	5
173.	mPdIRD13	0.75	0	4	60	173.	PdCUC3-ssr1	0.77	3	6	60
174.	mPdIRD14	0.75	0	4	60	174.	PdCUC3-ssr2	0.59	3	4	60
175.	mPdIRD17	0.83	1	6	60	175.	pd-GSSR18525	0.87	8	9	58
176.	mPdIRD20	0.50	0	2	60	176.	pd-GSSR19852	0.80	0	5	60
177.	mPdIRD22	0.67	3	3	60	177.	pd-GSSR2118	0.75	0	4	58
178.	mPdIRD24	0.64	1	3	60	178.	pd-GSSR4967	0.50	0	2	58
179.	mPdIRD28	0.76	1	5	60	179.	pd-GSSR8157	0.86	3	8	60
180.	mPdIRD29	0.88	0	8	60	180.	DPG150	0.54	3	4	
181.	mPdIRD30	0.88	8	10	60	181.	PDCAT06	Not Amplified			
182.	mPdIRD32	Not Amplified				182.	DP110	Not Amplified			
183.	mPdIRD35	0.50	2	2	60	183.	KSU-PDL64	Not Amplified			
184.	mPdIRD36	0.88	5	6	60	184.	DP161	0.66	1	3	55
185.	mPdIRD37	0.68	2	4	60	185.	KSU-PDL18-2	0.68	4	5	55
186.	mPdIRD42	0.71	3	5	60	186.	DPALM139	0.77	6	6	50
187.	mPdIRD43	0.63	2	3	60	187.	DPALM125	0.84	5	7	50
188.	mPdIRD45	0.80	0	5	60	188.	mPdIRD26	0.80	5	5	60
189.	mPdIRD46	Not Amplified				189.	mPdIRD04	0.65	1	3	60
190.	PDAAG1019	0.59	1	3	59	190.	DPALM112	0.80	0	5	50
191.	PDAAG1020	0.75	6	6	60	191.	DPALM144	0.86	2	7	50
192.	PDAAG1021	0.78	7	7	57	192.	DPALM119	0.48	1	2	50
193.	PDAAG1022	0.86	2	8	58	193.	DPALM110	Not Amplified			
194.	PDAAG1023	0.77	6	6	58	194.	DPALM120	0.81	7	7	50
195.	PDAAG1024	0.48	2	2	59	195.	DPALM121	0.76	4	5	50
196.	PDAAG1025	0.0	1	1	60	196.	DPALM104	0.50	2	2	50
197.	PDAG1001	0.71	3	4	59	197.	DPALM142	0.36	1	2	50
198.	PDAG1002	Not Amplified				198.	DPALM107	0.39	2	2	50
199.	PDAG1003	0.81	2	6	59	199.	DPALM146	0.85	5	7	50
200.	PDAG1004	0.74	2	4	61	200.	DPALM123	0.50	2	3	50
201.	PDAG1005	0.67	0	3	57	201.	DPALM133	0.44	3	4	50
202.	PDAG1006	0.75	4	44	58	202.	DPALM141	0.74	1	4	50
203.	PDAG1007	0.68	4	6	58	203.	PDAG1015	0.86	8	9	58
204.	PDAG1009	Not Amplified				204.	PDAG1016	0.53	2	3	60
205.	PDAG1010	0.89	9	12	58	205.	PDAG1017	0.43	1	2	57
206.	PDAG1011	0.80	0	5	58	206.	PDAG1018	0.95	16	23	58
207.	PDAG1012	0.49	1	2	61	207.	PdAG1-ssr	0.50	2	2	55
208.	PDAG1013	0.56	2	3	58	208.	PdAP3-ssr	0.80	4	6	57
209.	PDAG1014	0.76	4	5	58	209.	PdCAT01	0.83	7	7	54
210.	PDCAT10	0.00	0	1	54						

Structure analysis revealed highest peak value for delta K at optimum K value (K=2) indicating that at least 02 distinct population exists among selected genotypes. Khalas, AseelKhumba and Khupra varieties were placed in sub-population 1 showing similar genetic makeup irrespective of their geographic origin as indicated by green color. Similarly Ajwa, Zahidi and Khurma were placed in sub-population 2 as indicated by red color. Whereas Shamran, Amber, Hallawi, Haleni, Barhi, Shakri and Khudri are admixture. Expected heterozygosity between two populations did not varied significantly among two populations. Sub-population 1 was less heterozygous group with average heterozygosity value of 0.40 however sub-population 2 was relatively more heterozygous with 0.51 heterozygosity. However sub-population 2 showed less genetic diversity with an Fst value 0.013 whereas sub-population 01 showed high genetic diversity with Fst value of 0.18 (Fig. 3).

Cultivar identification key

Forty-five SSR markers were successful for identification of 12 genotypes of date palm out of 13, whereas, one genotype Hallawi was identified using two step identification method. Khalas was identifiable using PDCUC-3SSR1, mpdIRD28, PDAG1023, mpdIRD01, mpdIRD01, mpdIRD05, DP146, DP133, DPG1195 and DPG1702 at 380 bp, 190 bp, 145 bp & 175 bp, 185 bp, 80 bp, 150 bp, 180 bp, 245 bp, and 235 bp alleles, respectively. AseelKhumba exclusively exhibited unique alleles with KSU0PDL64 and KSU-PDL6 at 235 bp and 175 bp, respectively. Similarly, banding pattern obtained from KSU-PDL73, PDCAT8 and DPG2118 at 130 bp, 220 bp, and 210 bp, respectively, discriminated Shamran. Khupra was uniquely identified using KSU-PDL6 and DP120 at 70 and 215 bp respectively. Amber was differentiated from other cultivars using three different SSR markers PDAG1015 (250 bp), mpdIRD05 (225 bp) and DPG2115 (210 bp, 390 bp). Haleni was recognizable using distinct banding patterns *i.e*., DP174, PDCAT8, mpdIRD05, DP133 and DPG2116 at 215 bp, 350 bp, 140 bp, 250 bp, and 185 bp, respectively (Table 2).

Set of five SSR markers were produced DNA fingerprints of Barhi *i.e*., DP174 (175 bp, 500 bp), mpdCIR093 (160 bp), DPG1705 (190 bp), DPG3359 (200 bp) and DPG2118 (200 bp). Distinct DNA fingerprints at

Table 2: Cultivar identification key of 13 date palm varieties

Varieties	DNA Fingerprints Marker name (size in base pairs)
Khalas	PDCUC-3SSR1 (380), mpdIRD28 (190), PDAG1023 (145, 175), mpdIRD01 (185), mpdIRD05 (80), DP146 (150), DP133 (180), DPG1195 (245) and DPG1702 (235)
Aseel Khumba	KSU-PDL64 (235) and KSU-PDL6 (175),
Khupra	KSU-PDL6 (70) and DP120 (215),
Shamran	KSU-PDL73 (130), PDCAT8 (220) and DPG2118 (210)
Amber	PDAG1015 (250), mpdIRD05 (225) and DPG2115 (210, 390)
Hallawi	Two step identification procedure for identification of Hallawi is as follow. 1. KSUPDL-64 gives 145 bp in Hallawi and Khudri 2. DP174 amplified 174 bp band in all genotypes except Hallawi 3. mpdCIR010 amplifies 80 bp in Amber and Hallawi 4. PDAG1021 amplifies 150 bp band in Shamran and Hallawi 5. DPG2110 amplifies 172 bp band in Khupra and Hallawi
Haleni	DP174 (215), PDCAT8 (350), mpdIRD05 (140), DP133 (250) and DPG2116 (185)
Barhi	DP174 (175, 500), mpdCIR093 (160), DPG1705 (190), DPG3359 (200) and DPG2118 (200)
Shakri	DP175 (285), KSUPDL76 (220), mpdCIR090 (240) and mpdIRD05 (90)
Khudri	mpdIRD42 (240, 285, 300) and DPG1702 (200)
Ajwa	KSU-PDL53 (240), DP139 (210, 255), mpdIRD05 (120) and DPG1702 (195)
Zahidi	mpdIRD033 (400), mpdCIR050 (330), PDAG1007 (140, 150)
Khurma	mpdIRD36 (125, 130)

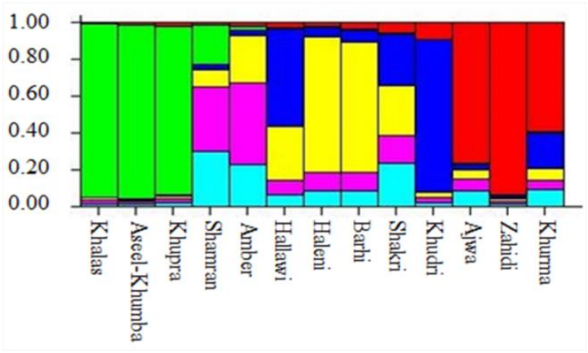

Fig. 3: Structure Analysis of Date Palm varieties grown in Punjab Pakistan. Parameters: no admission model; K = 02; 10,000 Burn-in period; 100000 Rep. Green color indicates sub-population 1, Red indicates sub-population 2 whereas all other genotypes were admixture

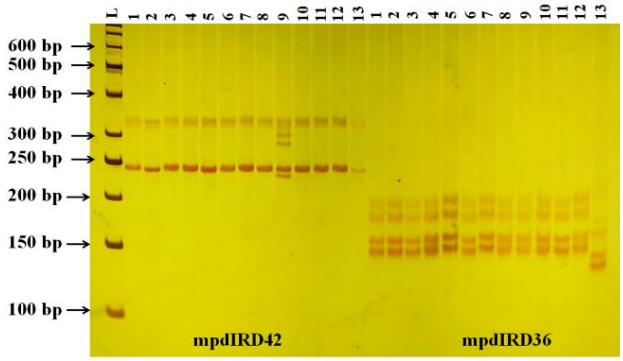

Fig. 4: Segregation pattern of mpdIRD42 and mpdIRD36 SSR markers showing DNA fingerprints for Khudri (240, 285 and 300 bp) and Khurma (125 and 130 bp) respectively run on 6% vertical polyacrylamide gel with 50 bp ladder (L)

285bp, 220bp, 240bp and 90bp using DP175, KSUPDL76, mpdCIR090 and mpdIRD05 were obtained for Shakri. MpdIRD42 amplified 240, 285 and 300 bp alleles (Fig. 4) and DPG1702 amplified 200bp allele as a DNA fingerprints for Khudri. Ajwa generated unique alleles with KSU-PDL53 (240bp), DP139 (210bp & 255bp), mpdIRD05 ((120bp) and DPG1702 (195bp). Zahidi was identified using mpdIRD033 allele at 400bp, mpdCIR050 allele at 330bp and PDAG1007 allele at 140bp & 150bp. Khurma exhibited unique allele at 125 and 130 bp with mpdIRD36 for unique identification (Fig. 4). Exceptionally Hallawi did not produce any unique single band directly but could be identified following two steps identification procedure. KSUPDL-64 amplified 145bp allele with Hallawi and Khudri (Table 2).

Discussion

DNA fingerprinting and population structure of 13 date palm varieties was exploited in present study using 209 polymorphic SSR markers (Elmeer *et al*. 2011; Elmeer and Mattat 2015; Al-Faifi *et al*. 2016; Racchi and Camussi 2018). 163 polymorphic SSR markers amplified 937 alleles and 597 polymorphic alleles with an average 4.84 and 3.65 alleles per locus and polymorphic alleles per locus respectively (Table 1). However, polymorphism reported in this study is much lower than previously reported results *i.e*., 7.7, 9.71 and 8.54 alleles per locus from Qatar, Tunisia and Iraq respectively (Elmeer *et al*. 2011; Khierallah *et al*. 2011; Zehdi *et al*. 2012). Mattat *et al*. (2019) reported 07 average alleles per SSR from Saudi Arbian and Qatar date palm varieties which are also higher as observed in this study.

Results obtained by ElKadri *et al*. (2019) were similar to our study with 05 average alleles per SSR. Previously MirBahar *et al*. (2016) conducted a study for DNA fingerprinting of date palm cultivars from Pakistan. Although they use 25 genotypes but numbers of DNA markers used in their study were only 07 in comparison to our 209 SSR markers. Further they also obtained varieties

from Sindh but our study only covers date palm varieties widely cultivated in Punjab. In our study SSR marker PDAG-1018 amplified 23 alleles whereas as previous study by Arabnezhad *et al*. (2012) reported six alleles for same marker. Date palm varieties used by Arabnezhad *et al*. (2012) were originated from Iran, Tunisia, Morocco and Algeria whereas genotypes used in our study were belonged to Saudi Arabia, Iraq and Pakistan (Fig. 1). Hence it is speculated that variation in the number of alleles may be linked with difference in genetic backgrounds of the varieties as was also found in case of olive SSR marker UDO-28. Abdessemed *et al*. (2015) reported 11 alleles whereas Sakar *et al*. (2016) reported 16 alleles for UDO-28 due to difference in genetic backgrounds of the varieties used in each case.

Genetic similarity coefficients were worked out among genotypes which varied from 80 to 88%. These results once again suggested that date palm cultivars used in this study are very closely linked and have narrow genetic base which may be explained by intensive selection operations. Normally date palm varieties are selected by farmers on the base of fruit shape and adaptation which is mainly controlled by limited number of genes and majority of the genome remain conserved leading to maximum genetic diversity. However, if varieties are undergone rigorous round of selection considering many traits than genetic diversity is lost (Zehdi *et al*. 2004).

Based on their origin genotypes mainly classified to three groups *i.e*. Pakistani (Khupra, Haleni, Shakri and Khurma), Iraqi (AseelKhumba, Zahidi, Shamran, Hallawi and Barhi) and Saudi Arabian (Khalas, Amber, Khudri and Ajwa) (Zango *et al*. 2017). However, dendrogram and structure analysis results do not support the hypothesis that genotypes originated from the same geographic origin or cultivated at a place have same genetic makeup as was also reported by (Zehdi *et al*. 2002).

Khalas, Aseel and Khupra (Group I) although have different origin but have same genetic makeup same as is the case with Ajwa, Zahidi and Khurma (Group III). However, rest of varieties lies in group II and have mixed genetic makeup as is indicated by structure analysis and also proved by cluster analysis as it contains 03 different subgroups (Fig. 1). New cultivars in date palm appear as results of sexual reproduction followed by selection by farmers. Also exchange of propagules which is a mixture of seed and vegetative propagated material is conducted by farmers which results in mixed genome (Elshibli and Korpelainen 2008) as is observed in case of seven varieties of group II in our study.

Genetic similarity coefficients were worked out separately and dendrogram was constructed for varieties belonging to different geographic origin. It was observed that highest genetic similarity (84%) was observed between varieties originated from Iraq followed by 83% genetic similarity among varieties originated from Pakistan and Saudi Arabia (Fig. 1b).

Identification of date palm cultivars on the basis of vegetative and fruiting characteristics is very difficult (Eid *et al*. 2019) even with the help of isozymes markers (Salem *et al.* 2001). As these traits contain limited genetic diversity and are highly influenced by environmental conditions and plant developmental stages. SSR markers used in this study have successfully distinguished 12 out of 13 date palm varieties with direct DNA fingerprints however Hallawi was identifiable following two step method for identification at DNA level. The identified DNA fingerprints are highly reproducible and will be helpful in plant variety registration/protection under Plant Breeders Rights and will help in easy identification of varieties even at seedling stage.

Conclusion

Date palm varieties grown in Punjab Pakistan mainly classified to 03 groups on the basis of structure and cluster analysis. DNA fingerprints for 12 date palm varieties except Hallawi (which is identifiable using two step approaches) was available. Population structure analysis revealed that genotypes have different genetic makeup irrespective of their origin and majority of date palm cultivars (07) have mixed genetic makeup. Polymorphic information contents values and number of alleles and polymorphic alleles of 209 SSR markers was available which will be helpful in designing different genetic studies of date palm in future.

Acknowledgement

Authors are highly thankful to Punjab Agriculture Research Board (PARB) for providing financial support through PARB Project No. 908 entitled “DNA Barcoding / Fingerprinting for the identification of Cotton, Wheat, Maize, Potato, Tomato and Date Palm Varieties” for conductance of this research work. Also to Horticultural Research Institute, Faisalabad for providing plant material.

References

Abdessemed S, I Muzzalupo, H Benbouza (2015). Assessment of genetic diversity among Algerian olive (*Olea europaea* L.) cultivars using SSR marker. *Sci Hortic* 192:10–20

Adawy SS, EH Hussein, SEME Ismail, HA El-Itriby (2005). Genomic diversity in date palm (*Phoenix dactylifera* L.) as revealed by AFLPs in comparison to RAPDs and ISSRs. *Arab J Biotechnol* 8:99–114

Akhtar W, AW Rasheed, ZK Shinwari, SMS Naqvi, T Mahmood (2014). Genetic characterization of different Pakistani date palm varieties. *Pak J Bot* 46:2095–2100

Al-Dous EK, B George, ME Al-Mahmoud, MY Al-Jaber, H Wang, YM Salameh, EK Al-Azwani, S Chaluvadi, AC Pontaroli, J DeBarry, V Arondel, J. Ohlrogge, IJ Saie, KM Suliman-Elmeer, JL Bennetzen, RK Kruegger, JA Malek (2011). *De novo* genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nat Biotechnol* 29:1–8

Al-Faifi SA, HM Migdadi, SS Algamdi, MA Khan, MH Ammar, RS Al-Obeed MI Al-Thamra, EH El-Harty, J Jakse (2016). Development, characterization and use of genomic SSR markers for assessment of genetic diversity in some Saudi date palm (*Phoenix dactylifera* L.) cultivars. *Electr J Biotechnol* 21:18–25

Allen GC, MA Flores-Vergara, S Krasynanski, S Kumar, WF Thompson (2006). A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat Protoc* 1:2320–2325

Al-Ruqaishi IA, M Davey, PG Alderson, S Mayes (2008). Genetic relationships and genotype tracing in date palms (*Phoenix dactylifera* L.) in Oman, based on microsatellite markers. *Plant Genet Resour* 6:70–72

Arabnezhad H, M Bahar, HR Mohammadi, M Latifian (2012). Development, characterization and use of microsatellite markers for germplasm analysis in date palm (*Phoenix dactylifera* L.). *Sci Hortic* 134:150–156

Caetano-Anolles G (1997). Resolving DNA amplification products using polyacrylamide gel electrophoresis and silver staining. *In: Fingerprinting Methods based on Arbitrarily Primed PCR*, pp: 119–134. Springer, Berlin, Heidelberg, Germany

Eid M, S Sabry, M Hussein (2019). DNA Fingerprinting and Characterization of some Egyptian Date Palm Cultivars Using Simple Sequence Repeats (SSRs). *J Plant Prod Sci* 3:31–41

ElKadri N, MB Mimoun, JI Hormaza (2019). Genetic diversity of Tunisian male date palm (*Phoenix dactylifera* L.) genotypes using morphological descriptors and molecular markers. *Sci Hortic* 253:24–34

Elhoumaizi MA, M Saaidi, A Oihabi, C Cilas (2002). Phenotypic diversity of date-palm cultivars (*Phoenix dactylifera* L.) from Morocco. *Genet Resour Crop Evol* 49:483–490

Elmeer K, I Mattat (2015). Genetic diversity of Qatari date palm using SSR markers. *Genet Mol Res* 14:1624–1635

Elmeer K, H Sarwath, J Malek, M Baum, A Hamwieh (2011). New microsatellite markers for assessment of genetic diversity in date palm (*Phoenix dactylifera* L.). *3Biotech* 1:91–97

Elshibli S, H Korpelainen (2008). Microsatellite markers reveal high genetic diversity in date palm (*Phoenix dactylifera* L.) germplasm from Sudan. *Genetica* 134:251–260

Enan MR, A Ahamed (2014). DNA barcoding based on plastid matK and RNA polymerase for assessing the genetic identity of date (*Phoenix dactylifera* L.) cultivars. *Genet Mol Res* 13:3527–3536

Evanno G, S Regnaut, J Goudet (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620

Gothwal RK, R Bhargava, PK Yadav, RR Meghwal, MK Agnihotri, NK Moraniya (2013). Evolutionary relationship study in date palm cultivars using morphological and biochemical parameters. *Bioscan* 8:1251–1254

Guichoux E, L Lagache, S Wagner, P Chaumeil, P Léger, O Lepais, C Lepoittevin, T Malausa, E Revardel, F Salin, RJ Petit (2011). Current trends in microsatellite genotyping. *Mol Ecol Resour* 11:591–611

Khierallah HS, SM Bader, M Baum, A Hamwieh (2011). Genetic diversity of Iraqi date palms revealed by microsatellite polymorphism. *J Amer Soc Hortic Sci* 136:282–287

Mathew LS, M Spannagl, A Al-Malki, B George, MF Torres, EK Al-Dous, EK Al-Azwani, E Hussein, S Mathew, KF Mayer, YA Mohamoud, K Suhre, JA Malik (2014). A first genetic map of date palm (*Phoenix dactylifera*) reveals long-range genome structure conservation in the palms. *BMC Genomics* 15; Articlde 285

Mattat I, A Al-Malki, AG Al-Mamari, K BoJulaia, A Hamwieh, M Baum (2019). Assessing genetic diversity of Shishi date palm cultivars in Saudi Arabia and Qatar using microsatellite markers. *Intl J Hortic Sci Technol* 6:1–9

Mirbahar AA, S Khan, GS Markhand, N Kauser, R Saeed (2016). DNA fingerprinting of some pakistani date palm (*Phoenix dactylifera* L.) cultivars using ISSR markers. *Pak J Bot* 48:2005–2010

Nielsen NH, G Backes, J Stougaard, SU Andersen, A Jahoor (2014). Genetic diversity and population structure analysis of European hexaploid bread wheat (*Triticum aestivum* L.) varieties. *PLoS One* 9; Article e94000

Pashley CH, JR Ellis, DE McCauley, JM Burke (2006). EST databases as a source for molecular markers: lessons from Helianthus. *J Hered* 97:381–388

Pritchard JK, M Stephens, P Donnelly (2000). Inference of population structure using multilocus genotype data. *Genetic* 155:945–959

Racchi ML, A Camussi (2018). The date palms of Al Jufrah-Libya: a survey on genetic diversity of local varieties. *J Appl Res Intellect Disabil* 112:161–184

Salem AOM, M Trifi, H Sakka, A Rhouma, M Marrakchi (2001). Genetic inheritance analysis of four enzymes in date-palm (*Phoenix dactylifera* L.). *Genet Resour Crop Evol* 48:361–368

Sakar E, H Unver, SI Ercisli (2016). Genetic diversity among historical olive (*Olea europaea* L.) genotypes from southern anatolia based on SSR markers. *Biochem Genet* 6:842–853

Yusuf AO, A Culham, W Aljuhani, CD Ataga, AM Hamza, JO Odewale, LO Enaberue (2015). Genetic diversity of Nigerian date palm (*Phoenix dactylifera* L.) germplasm based on microsatellite markers. *J Biol Sci Biotechnol* 7:121–132

Zango O, E Cherif, N Chabrillange, S Zehdi-Azouzi, M Gros-Balthazard, SA Naqvi, A Lemansour, H Rey, Y Bakasso, F Aberlenc (2017). Genetic diversity of Southeastern Nigerien date palms reveals a secondary structure within Western populations. *Tree Genet Genomics* 13:1–11

Zehdi S, E Cherif, S Rhouma, S Santoni, AS Hannachi, JC Pintaud (2012). Molecular polymorphism and genetic relationships in date palm (*Phoenix dactylifera* L.): The utility of nuclear microsatellite markers. *Sci Hortic* 148:255–263

Zehdi S, M Trifi, N Billotte, M Marrakchi, Christophe, J Pintaud (2004). Genetic diversity of Tunisian date palms (*Phoenix dactylifera* L.) revealed by nuclear microsatellite polymprohism. *Hereditas* 141:278–287

Zehdi S, M Trifi, OAM Salem, M Marrakchi, A Rhouma (2002). Survey of inter simple sequence repeat polymorphisms in Tunisian date palms (*Phoenix dactylifera* L.). *J Genet Breed* 56:77–83