Genetic diversity assessment in *Nicotiana tabacum* L. with iPBS-retrotransposons

Gülsüm YALDIZ[1], **Mahmut ÇAMLICA**[1], **Muhammad Azhar NADEEM**[1], **Muhammad Amjad NAWAZ**[2], **Faheem Shehzad BALOCH**[1,]*

[1]Department of Field Crops, Faculty of Agricultural and Natural Science, Abant İzzet Baysal University, Bolu, Turkey
[2]Department of Biotechnology, Chonnam National University, Chonnam, Republic of Korea

Abstract: Knowledge of genetic diversity could be beneficial by contributing important information in the selection of breeding material. The objective of this study was to explore the genetic diversity and relationship in the Turkish tobacco germplasm with the iPBS-retrotransposons marker that emerged as a universal method. A total of 90 landraces and 6 commercial cultivars collected from different geographical regions of Turkey were used in this study. Eleven most polymorphic iPBS-retrotransposons primers yielded a total of 119 scorable bands and 98 of these bands were found to be polymorphic (82.35%), with an average of 8.91 polymorphic fragments for each primer. The mean polymorphism information contents were found to be 0.33, ranging from 0.10 to 0.80. The unweighted pair group method with arithmetic mean revealed that the genotypes belonging to the same geographical regions were often present in the same group and very close to each other. Based on the Bayesian clustering model, the genetic structure of Turkish tobacco germplasm was divided into 2 main groups. This is first study to explore the genetic diversity of Turkish tobacco by the iPBS-retrotransposons and we believe that in future the results of this study will serve as a foundation for the development of new and improved tobacco varieties in Turkey and the rest of world.

Key words: *Nicotiana tabacum* L., Solanaceae, iPBS-retrotransposons, genetic diversity, Turkey

1. Introduction

Tobacco (*Nicotiana tabacum* L.) is an economically important nonfood crop cultivated and consumed largely all over the world (Moon et al., 2009). Tobacco belongs to the family Solanaceae, which contains more than 64 species, and *N. tabacum* is the most cultivated and commercially consumed species of this family (Ren and Timko, 2001). Tobacco is a nonobligatory selfing amphidiploid (2n = 48) crop with a genome size of 4.5 Gb (Renny-Byfield et al., 2011). *N. tabacum* was derived through an interspecific hybridization event between *Nicotiana sylvestris* (2n = 24) as a maternal donor and *Nicotiana tomentosiformis* (2n = 24) as a paternal donor (Leitch et al., 2008). Traditionally this plant was used as a medicine to cure insect bites and cuts as it contains good concentrations of nicotine tartrate that is used in the manufacturing of different medicines (Mackay and Eriksen, 2002). Tobacco is used for cigarettes, water pipe smoking, medicine, creamy snuff, and chewing tobacco (Darvishzadeh et al., 2013). The tobacco industry is not only playing a vital role in the tobacco-producing countries, but countries lacking cigarette factories are also earning a significant amount of annual capital only through its distribution (Ren and Timko, 2001). In 2014 tobacco was cultivated on an area of

3,963,630 ha and production was 7,176,650 t. China, India, and Brazil are the top tobacco-producing countries with a share of 74.57%, 20.39%, and 18.89% in the world tobacco production (http://faostat.fao.org/site/339/defaultaspx).

Turkey is present at a very important geographical position of the world and it served as the source of origin, distribution, and diversity for different crops (Karagoz, 2001; Baloch et al., 2017). Tobacco is native to South America, while Venetian and English sailors were the ones responsible for the introduction of tobacco in Turkey in the 16th century, during the rule of the Ottomans (Esendal et al., 1994). In Turkey tobacco is mainly used as cigarettes, and Turkey shares the 1.75% of the world production. In 2014 in Turkey, this crop was cultivated on an area of 99,000 ha, and production was recorded as 74,696 t (http://faostat.fao.org/site/339/defaultaspx). The Aegean region is the primary tobacco-producing region of Turkey, while the Western Black Sea, Western Marmara, and Southeastern Anatolian regions also contribute an important role in the tobacco production (Güler and Demirbaş, 2016).

The importance of molecular markers in plant breeding is now universally established. Different types of molecular markers have been developed and have been used in plant breeding (Nadeem et al., 2017; Nawaz et

---

* Correspondence: balochfaheem13@gmail.com

al., 2017). Retrotransposons are genetic elements capable of changing their locations and copy numbers and a very important component of the structural evolution of the plant genome (Finnegan, 1989). Retrotransposon contents vary based on plant species and normally they constitute 50% of the plant genome; however, they can also be present up to 90% (SanMiguel et al., 1996). Retrotransposons are grouped as long terminal repeat (LTR) and non-LTR retrotransposons. The plant genome contains higher contents of LTR retrotransposons as compared to non-LTR. However, limiting factors among both LTR and non-LTR retrotransposon marker systems resulted in the development of a new marker system. Inter-primer binding site (iPBS) retrotransposons emerged as a universal marker system because it is a PCR-based marker system depending on the presence of tRNA as a reverse transcriptase primer binding site (Kalender et al., 2010). The iPBS-retrotransposon markers have been successfully applied in various crops like pea (Baloch et al., 2015a), chickpea (Andeden et al., 2013), *Lens* (Baloch et al., 2015b), Turkish okra (Yıldız et al., 2015), and common bean (Nemli et al., 2015). Genetic diversity is very important for the breeders to develop new and improved cultivars having favorable characteristics according to both the farmer's and the breeder's points of interest (Govindaraj et al., 2015; Baloch et al., 2017). Very little is known about genetic diversity and relatedness within the Turkish tobacco germplasm. Knowledge in this area could be beneficial in the management of future germplasm maintenance, helpful in the selection of breeding material for this crop, and could provide essential information for breeding and genetics activities of tobacco. The objective of this study was to investigate the genetic diversity with iPBS-retrotransposon markers among 90 tobacco landraces and 6 commercial cultivars collected from various geographical regions of Turkey.

2. Materials and methods

2.1. Plant material

A total of 96 Turkish tobacco accessions (90 landraces and 6 commercial cultivars) were collected from different geographical regions of Turkey. Sixty-seven landraces were collected from 12 different provinces (Samsun, Hatay, Amasya, Çanakkale, Balıkesir, Bursa, Bitlis, Muş, Aydın, Uşak, Adıyaman, Manisa) of Turkey and the other 23 landraces were obtained from the Aegean Agricultural Research Institute, İzmir, Turkey. Six commercial cultivars (Akhisar 97, Akhisar 97_95, Kılıcemmi, Burley 94, A. Virginia 90, and Islambey) were also used in this study. Brief information about the plant material is described in Table 1 and shown in Figure 1. The seeds of the tobacco landraces were sown into pots under greenhouse conditions and young leaves were harvested for DNA isolation.

2.2. DNA isolation

Young leaves from 10 plants of each landrace and cultivar were bulked and frozen at -80 °C before use. The genomic DNA was isolated from bulked young, healthy, and fresh leaves following the CTAB protocol (Doyle and Doyle, 1990) with some modifications (Baloch et al., 2016). The DNA concentrations were measured with NanoDrop (DeNovix DS-11 FX, USA) and were adjusted to 5 ng/µL for further use in PCR. These samples were stored at -20 °C until PCR amplification. All chemicals used in this study were purchased from Thermo Scientific (USA).

2.3. iPBS-retrotransposon analysis

Initially, a total of 83 iPBS-retrotransposons primers were screened on 8 randomly selected tobacco landraces for PCR amplifications, designed by Kalender et al. (2010). Fifteen iPBS-retrotransposons primers failed to produce any PCR product, and 10 primers did not produce any polymorphic fragment. For the further analysis, 11 iPBS-retrotransposons primers were selected, which produced strong and polymorphic bands (Table 2). These 11 primers produced perfect banding profiles, which were selected for fingerprinting the tobacco landraces and cultivars. The PCR amplifications contained reactions of 20.9 µL with 5 ng/µL template DNA, 2 mM dNTPs, 0.2 U Taq DNA polymerase, 3.2 µM primer, 2 mM 1X PCR buffer, and 8.5 mM distilled water. The PCR amplification was carried out under the following conditions: initial denaturation at 95 °C for 3 min, followed by 30 cycles of denaturation at 95 °C for 15 s and annealing temperature of 50-65 °C (depending on the primers used) for 1 min, and a final extension at 72 °C for 5 min (Kalendar et al., 2010). The PCR products were detected by electrophoresis on 1.2% (w/v) agarose gel using 1X TAE buffer for 2.5 h; the gel was stained with ethidium bromide after the electrophoresis, visualized under the UV Imager Gel Doc XR+ system (Bio-Rad, USA) light, and later photographed. A 250-bp ladder (Thermo Scientific) was used as a molecular weight marker.

2.4. Data analysis

Only the strong, clear, and reproducible amplifiable products were considered for analysis. We scored the iPBS-retrotransposons bands according to a binary system as present (1) or absent (0). The gene diversity and Shannon information index among and within Turkish tobacco landraces and commercial cultivars were measured with PopGene ver. 1.32 (Yeh et al., 2000). The mean polymorphism information contents (PICs) for each selected primer were calculated as previously done (Baloch et al., 2015a). A pairwise genetic distance (GD) matrix between 96 Turkish tobacco accessions was calculated by applying Jaccard's coefficient (Jaccard, 1987) using R statistical software. To visualize the pattern of genetic diversity among the 96 Turkish tobacco accessions,

Table 1. Passport data of 90 landraces and 6 commercial cultivars of Turkish tobacco germplasm.

No.	Landraces	Geographical province	Latitude (N)	Longitude (E)
1	Samsun_1	Samsun-Merkez-Kahyalı	41°10′21.5508″	36°32′21.6564″
2	Samsun_2	Samsun-Terme-Akçay	41°8′ 0.7332″	37°9′22.2264″
3	Samsun_3	Samsun-Canik-Alibeyli	41°12′27.4284″	36°22′0.1992″
4	Samsun_4	Samsun-Tekkeköy-Gökçe	41°10′41.0952″	36°29′36.0492″
5	Samsun_5	Samsun-Tekkeköy-Hamzalı	41°12′15.3576″	36°31′54.6852″
6	Samsun_6	Samsun-Bafra-Gökçeağaç	41°32′45.9204″	35°45′41.4324″
7	Samsun_7	Samsun-Tekkeköy-Çınaralan	41°11′9.6180″	36°26′32.5932″
8	Samsun_8	Samsun-Çarşamba-Cumhuriyet	41°12′20.4372″	36°41′10.6620″
9	Samsun_9	Samsun-Tekkeköy-Balcalı	41°9′18.5436″	36°34′6.6144″
10	Samsun_10	Samsun-Bafra-Merkez1	41°33′43.0488″	35°54′20.4660″
11	Samsun_11	Samsun-Bafra-Merkez2	41°33′43.0488″	35°54′20.4660″
12	Samsun_12	Samsun-Bafra-Merkez3	41°33′43.0488″	35° 54′20.4660″
13	Samsun_13	TR 49184*		
14	Samsun_14	TR 49188*		
15	Samsun_15	TR 49224*		
16	Samsun_16	Samsun-Bafra-Merkez4	41°33′43.0488″	35°54′20.4660″
17	Samsun_17	Samsun-Bafra-Merkez5	41°33′43.0488″	35°54′20.4660″
18	Samsun_18	Samsun-Bafra-Merkez6	41°33′43.0488″	35°54′20.4660″
19	Samsun_19	Samsun-Bafra-Merkez7	41°33′43.0488″	35°54′20.4660″
20	Samsun_20	Samsun-Bafra-Merkez8	41° 33′ 43.0488″	35° 54′ 20.4660″
21	Samsun_21	Samsun-Bafra-Keresteci	41°30′41.6232″	35°44′48.3000″
22	Samsun_22	Samsun-Çarşamba-Cumhuriyet	41°12′20.4372″	36°41′10.6620″
23	Samsun_23	Samsun-Bafra-Merkez9	41°33′43.0488″	35°54′20.4660″
24	Samsun_24	Samsun-Bafra-Merkez10	41°33′43.0488″	35°54′20.4660″
25	Samsun_25	Samsun-Bafra-Merkez11	41°33′43.0488″	35°54′20.4660″
26	Samsun_26	Samsun-Bafra-Göltepe	41°35′12.1″	35°46′21.22″
27	Samsun_27	Samsun-Tekkeköy-Balcalı	41°9′18.5436″	36°34′6.6144″
28	Samsun_28	Samsun-Bafra-Paşaşeyh	41°28′50.8260″	35°44′18.0888″
29	Samsun_29	Samsun-Bafra-Merkez12	41°33′43.0488″	35°54′20.4660″
30	Samsun_30	Samsun-Bafra-Merkez13	41°33′43.0488″	35°54′20.4660″
31	Samsun_31	Samsun-Tekkeköy-A. Çinik Yavuzlar	41°12′6.7860″	36°28′6.0456″
32	Samsun_32	Samsun-Tekkeköy-Hamzalı	41°12′15.3576″	36°31′54.6852″
33	Samsun_33	Samsun-Bafra-Merkez14	41°33′43.0488″	35°54′20.4660″
34	Samsun_34	Samsun-Bafra-Azay	41°32′5.4636″	35°46′22.8936″
35	Samsun_35	Samsun-Bafra-İkiztepe	41°36′44.6796″	35°52′13.3284″
36	Akhisar97	Commercial Cultivar		
37	Akhisar97_95	Commercial Cultivar		
38	Manisa_1	TR 64062*		
39	Manisa_2	TR 64078*		
40	Manisa_3	TR 64093*		
41	Manisa_4	Manisa-Merkez-Karayenice	38°45′48.5856″	27°26′51.3780″
42	Manisa_5	Manisa-Akhisar	38°55′3.5904″	27°50′11.8320″

Table 1. (Continued).

43	Manisa_6	Manisa-Saruhanlı-Halitpaşa	38°42′21.2472″	27°41′4.3836″
44	Manisa_7	Manisa-Kula-Gökçeören	38°33′57.1140″	28°28′ 53.0436″
45	Manisa_8	Manisa-Salihli-Kale köy	38°43′18.6780″	28°8′17.9700″
46	Manisa_9	Manisa-Akhisar	38°55′3.5904″	27°50′11.8320″
47	Manisa_10	Manisa-Saruhanlı-Koldere	38°39′26.5392″	27°38′4.1604″
48	Manisa_11	Manisa-Salihli-Poyrazdamları	38°37′23.8764″	28°8′24.2232″
49	Manisa_12	A Manisa-Akhisar-Sarılar	39°6′26.9100″	28°0′8.7084″
50	Manisa_13	Manisa-Saruhanlı-Gözlet	38°50′1.4784″	27°36′59.0796″
51	Manisa_14	Manisa-Selendi-Yıldız	38°44′40″	28°52′4″
52	Manisa_15	Manisa-Saruhanlı-Büyükbelen	38°38′27.0456″	27°49′5.1996″
53	Manisa_16	Manisa-Akhisar-Kavakalan	38°49′52.1688″	28°1′16.5252″
54	Manisa_17	Manisa-Selendi-Turpcu	38°46′4.8″	28°42′8.3″
55	Balıkesir_1	TR 64073*		
56	Balıkesir_2	TR 42526*		
57	Balıkesir_3	Balıkesir-Sındırgı-Taşköy	39°19′58.8540″	28°14′42.7668″
58	Balıkesir_4	Balıkesir-Sındırgı-Gölcük	39°18′50.1624″	27°59′5.7768″
59	Islambey	Commercial cultivar		
60	Kılıcemmi	Commercial cultivar		
61	Uşak_1	TR 64049*		
62	Uşak_2	Uşak-Eşme-Güllü köyü	38°16′15.0636″	29°6′20.7936″
63	Uşak_3	Uşak-Eşme-Konak köyü	38°03′3928″	40°89′5316″
64	Bursa_1	TR 78215*		
65	Bursa_2	TR 42808*		
66	Bursa_3	Bursa-İnegöl-Çavuşköy	40°10′17.9796″	29°28′17.8068″
67	Bursa_4	Bursa-İnegöl-Çavuşköy	40°10′17.9796″	29°28′17.8068″
68	Hatay_1	Hatay-Yayladağı-Sebenoba	36°2′52.2924″	36°1′12.0144″
69	Hatay_2	TR 42127*		
70	Hatay_3	TR 42129*		
71	Hatay_4	TR 78221*		
72	Hatay_5	TR 42132*		
73	Hatay_6	TR 42130*		
74	Hatay_7	Hatay-Yayladağı	35°4′9.0000″	36°3′38.1600″
75	Hatay_8	Hatay-Yayladağı-Sebenoba	36°′52.2924″	36°1′12.0144″
76	Muş_1	TR 42094*		
77	Muş_2	TR 42076*		
78	Muş_3	Mus-Kızılağaç-Suluca	38°45′23.9760″	41°22′6.8592″
79	Çanakkale_1	TR 42523*		
80	Çanakkale_2	TR 78220*		
81	Çanakkale_3	Çanakkale-Agonya-Yaris köyü	39°47′1.6332″	27°15′51.7320″
82	Çanakkale_4	Çanakkale-Yenice-Reşadiye köyü	39°50′27.5424″	27°12′11.4012″
83	Çanakkale_5	Çanakkale-Yenice-Bağlı köyü	39°48′5.2848″	27°15′4.6620″
84	Çanakkale_6	Çanakkale-Yenice-Çukuroba köyü	40°8′48.1956″	26°24′30.9132″
85	Aydın_1	TR 57502*		

Table 1. (Continued).

86	Aydın_2	Aydın-Karacasu	37°43′44.6484″	28°36′26.0316″
87	Bitlis_1	TR 80110*		
88	Bitlis_2	TR 80111*		
89	Bitlis_3	Bitlis-Mutki-Erler köyü	38°28′46.7472″	41°43′56.5536″
90	Bitlis_4	Bitlis-Mutki-Erler köyü	38°28′46.7472″	41°43′56.5536″
91	Bitlis_5	Bitlis-Yolalan-Düzmahalle	38°24′22.6512″	42°6′20.9664″
92	Amasya	Amasya-Gümüşhacıköy-Kuzalan köyü	40°55′7.2984″	35°8′13.0668″
93	Eski tutun 98	Hatay population	36°3524	36°2935
94	Burley 94	Commercial cultivar		
95	A. Virginia 90	Commercial cultivar		
96	B. Çelikhan	Adıyaman-Çelikhan	38°2′4.1028″	38°14′48.1632″

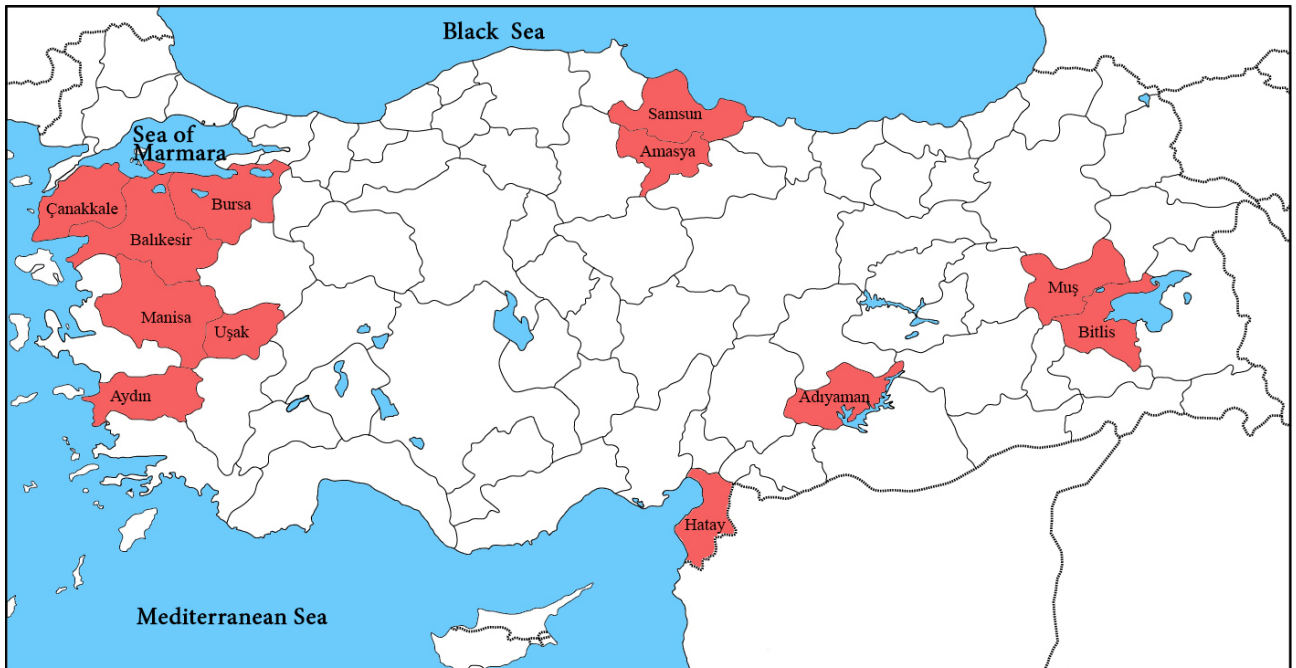* Material taken from the Aegean Agricultural Research Institute, İzmir, Turkey.

Figure 1. Collection sites of 90 landraces and 6 commercial cultivars of Turkish tobacco germplasm.

the unweighted pair group method with arithmetic mean (UPGMA) was performed. The Bayesian clustering model was applied in STRUCTURE in order to obtain a brief understanding of the genetic structure of the Turkish tobacco germplasm. For the determination of a suitable number of clusters (number of K; number of subpopulations) in the STRUCTURE analysis, we followed the criteria suggested by Evanno et al. (2005) and plotted the number of clusters (K) against logarithm probability relative to standard deviation (ΔK).

3. Results

Within the investigated population of 90 landraces and 6 commercial cultivars, 11 most polymorphic iPBS-retrotransposons primers resulted in a total of 119 scorable bands. Among these 119 bands, a total of 98 bands (82.35%) were found to be polymorphic with an average of 8.91 polymorphic fragments per primer (Table 3). The average number of bands per primer was 10.82 and the minimum numbers of bands (7) was produced by primer iPBS2388, while primer iPBS2087 resulted in a maximum

Table 2. Primers names, their sequence, and annealing temperature.

iPBS primer names	Sequence	Annealing temp., °C
iPBS2394	GAGCCTAGGCCA	56,5
iPBS2388	TTGGAAGACCCA	51
iPBS2298	AGAAGAGCTCTGATACCA	60
iPBS2377	ACGAAGGGACCA	53
iPBS2274	ATGGTGGGCGCCA	65,8
iPBS2224	ATCCTGGCAATGGAACCA	55,4
iPBS2243	AGTCAGGCTCTGTTACCA	53,8
iPBS2087	GCAATGGAACCA	52,5
iPBS2230	TCTAGGCGTCTGATACCA	52,9
iPBS2238	ACCTAGCTCATGATGCCA	56
iPBS2241	ACCTAGCTCATCATGCCA	55

Table 3. Primer name, number of total bands, polymorphic bands, and some diversity parameters of the iPBS-retrotransposon primers used during this study.

Markers	Number of bands			Diversity parameters				
	Total bands	Polymorphic bands	P%*	ne*	h*	I*	Ht*	PIC*
iPBS2394	9	8	88.89	1.36	0.23	0.36	0.22	0.36
iPBS2388	7	4	57.14	1.11	0.09	0.17	0.01	0.40
iPBS2298	10	6	60.00	1.17	0.12	0.21	0.11	0.19
iPBS2377	9	6	66.67	1.27	0.19	0.30	0.15	0.31
iPBS2274	10	5	50.00	1.06	0.05	0.12	0.04	0.10
iPBS2224	13	11	84.62	1.41	0.27	0.42	0.21	0.31
iPBS2243	10	10	100.0	1.41	0.27	0.42	0.22	0.80
iPBS2087	16	15	93.75	1.47	0.30	0.46	0.22	0.37
iPBS2230	10	9	90.00	1.20	0.16	0.28	0.10	0.19
iPBS2238	13	12	92.31	1.28	0.20	0.33	0.16	0.30
iPBS2241	12	12	100.0	1.28	0.21	0.35	0.19	0.32
Total	119	98	-	-	-	-	-	-
Average	10.82	8.91	80.31	1.27	0.19	0.31	0.15	0.33

*P%: Polymorphism percentage, ne: effective number of alleles, H: gene diversity, I: Shannon information index, Ht: genetic dispersion index, PIC: polymorphism information contents.

number (16) of bands among all these 11 primers. These 11 iPBS-retrotransposons primers resulted in a higher level of polymorphism with an average of 80.31 in 96 Turkish tobacco accessions (landraces and cultivars) collected from various geographical areas of Turkey (Figure 1; Table 1). The average gene diversity per iPBS-retrotransposons primers was 0.19 with the maximum gene diversity (0.30) observed in iPBS2087, and minimum 0.05 in iPBS2274.

The mean Shannon information index per primer ranged from 0.12 in iPBS2274 to 0.46 in iPBS2087 with an average of 0.31. The maximum value of PIC was 0.80 for iPBS2243, while iPBS2274 resulted in the minimum PIC value (0.10) and the average PIC value per 11 iPBS-retrotransposons was 0.33.

The pairwise genetic distance (GD) matrix between 96 Turkish tobacco accessions was calculated by applying

Jaccard's coefficient using R statistical software (Figure 2). All accessions were mainly grouped into 2 main groups: A (red) and B (green) (Figure 2). Most of the accessions were grouped into Group A. However, as compared to Groups A and B, some genotypes were also found scattered and away from the axis. The minimum genetic distance (0.353) was observed between the landrace Samsun-16 and the commercial cultivar Burley-94, followed by landrace Çanakkale-1 and commercial cultivar Burley-94, having a genetic distance of 0.357. The maximum genetic distance (0.61) was observed between landraces Samsun-6 and Manisa-17, followed by landraces Samsun-6 and Balıkesir-4, having a genetic distance of 0.57.

The UPGMA-based clustering divided the 96 Turkish tobacco accessions into 2 groups: A (red) and B (green) (Figure 3). Group A is smaller than Group B and it contains only 2 landraces, while Group B contains 94 landraces and commercial cultivars. Samsun-29 and Manisa-17 were the 2 landraces clustered in Group A. The main Group B was further subdivided in to 2 subgroups, B1 and B2. B1 was found to be a smaller group as compared to B2, and B2 was further subgrouped into B2I and B2II. This B2II was again further subgrouped into many subgroups according to geographical position. All 6 commercial cultivars were clustered in Group B, and only Islambey and Kılıcemmi were clustered in the same subgroup. The maximum

observed ΔK value was 2. The STRUCTURE analysis also divided the 96 Turkish tobacco accessions into 2 groups: A (red) and B (green) (Figure 4).

4. Discussion

Diversity in plant genetic resources helped plant breeders develop new and improved cultivars having favorable characteristics according to the farmer's and the breeder's points of interest (Govindaraj et al., 2015). Landraces that serve as a source of genetic diversity for any crop are plant populations having specific geographical or ecological representations, which are developed through the influence of the cultural and local environment (Hagenblad et al., 2012). These landraces act as a source of new genes and represent a higher level of inter- and intravariations, which increase their importance for breeding programs (Karaköy et al., 2012). The improvement in plant for traits of interest leads humans to develop different breeding methods. Molecular markers have been successfully applied in plant breeding for various traits of interest. Different molecular technologies have been developed and applied for the improvement of plants. The iPBS-retrotransposons emerged as a universal method and have been applied to explore the genetic diversity and relationships in various crops (Andeden et al., 2013; Mehmood et al., 2013; Baloch et al., 2015a,
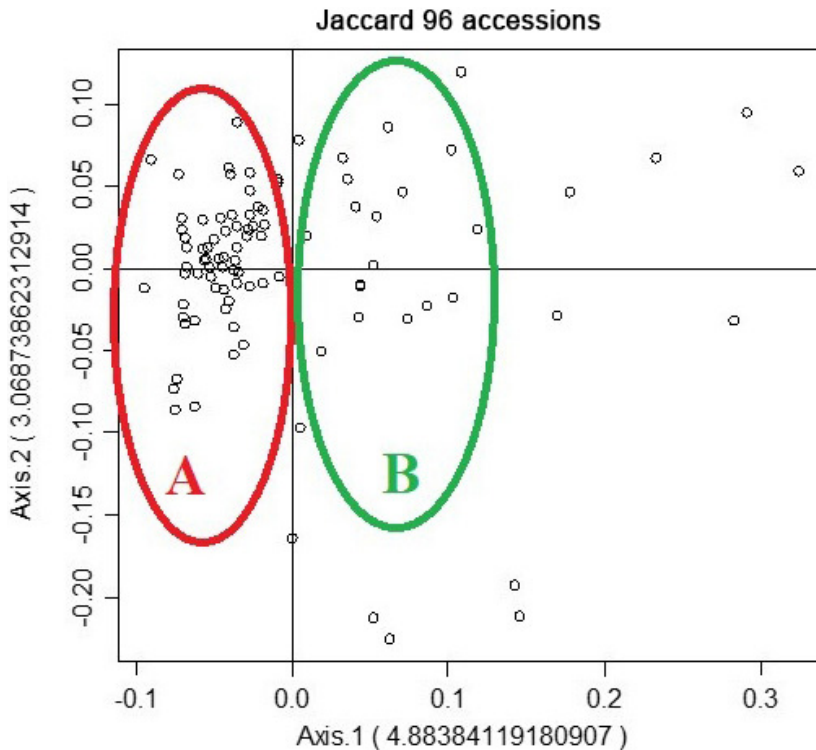

Figure 2. Jaccard's coefficient analysis for the determination of the genetic distance among 96 Turkish tobacco accessions.
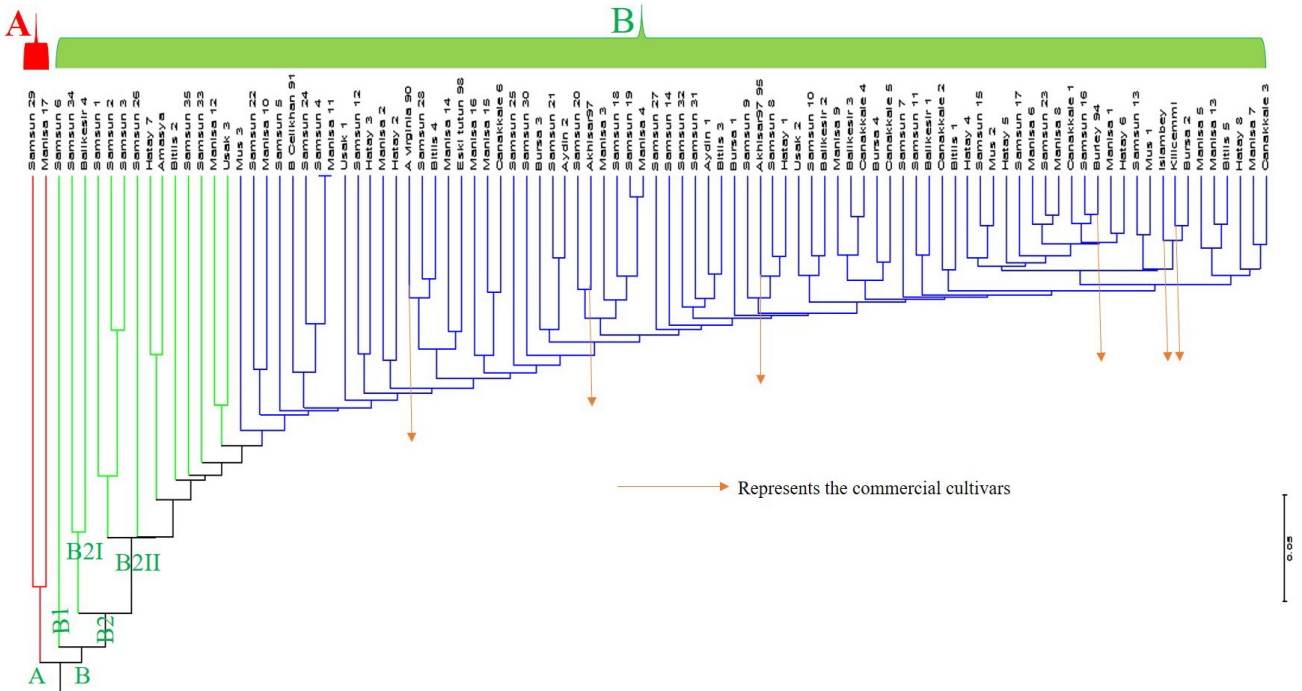
Figure 3. UPGMA-based clustering of 90 landraces and 6 commercial cultivars of Turkish tobacco germplasm Turkish tobacco accessions.
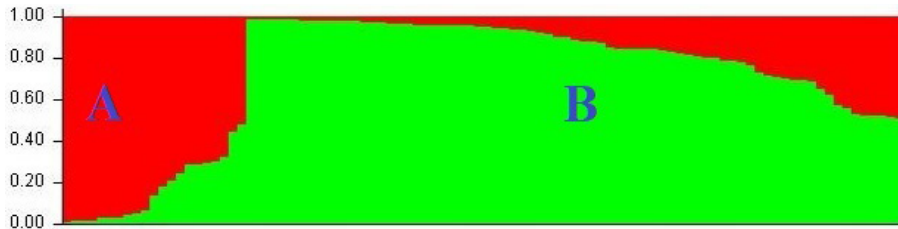

Figure 4. Population structure analysis of the Turkish tobacco germplasm reflecting 2 groups: Group A (red), Group B (green).

2015b; Nemli et al., 2015; Yıldız et al., 2015; Borna et al., 2017; Demirel et al., 2018).

In this study, 90 Turkish tobacco landraces and 6 commercial cultivars collected from various geographical regions of Turkey were used as the study material. The iPBS-retrotransposon markers were used to explore the genetic diversity in the Turkish tobacco germplasm. Very few studies have been conducted on tobacco to investigate the genetic diversity using retrotransposon-based markers (Yang et al., 2007). To our knowledge, this is the first study to explore the genetic diversity and relationship among the Turkish tobacco landraces and cultivars using iPBS-retrotransposon markers. In the past decade different types of molecular markers have been applied to investigate genetic diversity and relatedness in tobacco (Moon et al., 2008; Bindler et al., 2011; Lu et al., 2013; Gong et al.,

2016). Bindler et al. (2011) reported the development of 5119 novel and functional microsatellites markers and constructed the genetic map of the tetraploid tobacco genome using F_2 mapping populations. Recently, Kalivas et al. (2016) performed the genotyping of 34 tobacco cultivars with 10 microsatellite primers and identified 63 new alleles with an average of 6.3 alleles per primer.

As compared to the number of bands obtained in this study, the PIC value provides a clear picture for diversity assessment because it takes account of the relative frequencies of each present band (Cömertpay et al., 2012). Hence, results obtained during this study lead to the selection of more polymorphic markers in order to reduce the number of required loci for precise genotype distinction. During this study, 11 iPBS-retrotransposon markers resulted in an average of 0.33 PIC for 96 Turkish

tobacco germplasms. Dadras et al. (2014) obtained a mean PIC value of 0.26 using a panel of 50 tobacco genotypes with AFLP markers, which is lower than the value of 0.33 obtained in our study. Lu et al. (2013) investigated 330 high-quality DArTseq markers from a panel of 121 tobacco accessions and they resulted in an average PIC value of 0.34, which is close to our resulted PIC value. The average PIC value obtained during this study was lower as compared to that of Xia et al. (2014), who reported a 0.63 average PIC value of 28 SSR markers for 78 tobacco accessions from China. The average gene diversity in the study material was 0.19 and ranged between 0.05 in iPBS2274 and 0.30 in iPBS2087. These results showed a lower polymorphism level and confirmed the previously reported low level of polymorphism in tobacco (Rossi et al., 2001; Julio et al., 2006).

To explore the phylogenetic relatedness among 96 Turkish tobacco accessions, a pairwise genetic distance coefficient score was calculated. The UPGMA-based clustering and principal coordinate analysis grouped the studied material into 2 main groups: A (red) and B (green), on the basis of geographical regions (Figure 3). Group A is smaller than Group B and contains only 2 landraces. Samsun-29 and Manisa-17 were the 2 most diverse landraces clustering in Group A. As both of these landraces made their separate groups by reflecting the variations from other landraces of the same provinces, these landraces can be used as candidate parents for the development of new tobacco cultivars. Main Group B (green) was further subgrouped into B1 and B2 (Figure 3). Subgroup B1 was the smaller group, clustering only 1 landrace, Samsun 6. Subgroup B2 was further grouped into B2I and B2II. B2I was a smaller group than B2II, containing only 2 landraces (Samsun 34 and Balıkesir-4). Subgroup B2II was the bigger subgroup, which was further subgrouped and contained a total of 85 landraces and 6 commercial cultivars. Among the 6 commercial cultivars, none clustered into Group A. Group B, which is the larger group, contained a total of 94 landraces and commercial cultivars. Among the 6 commercial cultivars, 4 commercial cultivars (Akhisar 97, Akhisar 97_95, Burley 94, and A. Virginia 90) were clustered in Group B and they were present near the landraces from Samsun Province. However, all of these 4 cultivars were not clustered within the same subgroup, and Islambey and Kılıcemmi were the only 2 commercial cultivars that were clustered closer to each other in the same subgroup. Most landraces were clustered on the basis of their geographic location. In Group B, some landraces from Manisa, Hatay, and Bursa provinces were also mixed with Samsun landraces. Arslan and Okumus (2006) used RAPD markers in Turkish tobacco germplasm and they also reported that geographically closer populations were present in a single cluster. Denduangboripant et al. (2010)

used AFLP markers in tobacco and they also found that their studied material showed grouping on the basis of respective geographical regions. The grouping obtained in our study can be supported through this possible assumption, as Samsun Province is considered among the main tobacco-growing areas. There is also a possibility that diffusion of seeds may occur from Samsun to these other provinces through farmers and trade within the country, and later selection and hybridization may have resulted in the development of these landraces clustering close to the Samsun Province landraces. Other possible reason is that a big portion of studied material was also collected from Samsun Province, which leads to the clustering of these landraces close to each other. Yang et al. (2007) used ISSR and IRAP markers and found that similar tobacco accessions belonging to the same type were grouped in the same subgroup. In our results there is another possibility that those genotypes clustering close together may belong to the same tobacco type, like flue-cured tobacco, burley tobacco, and oriental tobacco.

In this study, the minimum genetic distance was observed between landrace Samsun-16 and commercial cultivar Burley-94, as well as between Samsun-23 and the commercial cultivar Burley-94. We found the maximum genetic distance between Samsun-6 and Manisa-17, followed by the Samsun-6 and Balıkesir-4 landraces, reflecting the presence of a higher level of diversity that can be used in the near future as a source of novel variation for breeding activities of tobacco. Furthermore, all the commercial cultivars were grouped with the landraces from Samsun Province, except the Islambey and Kılıcemmi cultivars. These 2 commercial cultivars were clustered with the landrace from Bursa Province (Bursa-2) and reflect possible closeness with this landrace. As the other 4 commercial cultivars clustered with the landraces from Samsun Province, there is a possibility that Samsun Province landraces were possibly used as source material for the development of these commercial cultivars. As Samsun-6 and Manisa-17 landraces showed a maximum genetic distance and reflected a higher level of genetic diversity, such landraces with higher genetic diversity are suggested as promising breeding resources for tobacco breeding stratagems aimed to achieve higher yields with better resistances against biotic and abiotic stresses. The population structure analysis grouped these 96 Turkish tobacco accessions into 2 groups, A (green) and B (red), at K = 2 (Figure 4). Clearly, the Bayesian clustering model supported the UPGMA tree and grouped the studied material into 2 main groups. These results show the homogeneous nature of the landraces, as most of the landraces and all cultivars clustered in Group B.

In conclusion, Turkey is acting as a bridge between Europe, Asia, and Africa and contains a great level of

diversity of different crops. During this study, a significant level of diversity was observed in a panel of 96 Turkish tobacco accessions (landraces and cultivars) with the iPBS-retrotransposons. During this study, significant numbers of landraces were collected from Samsun Province. To explore the higher level of diversity and relationship in the Turkish germplasm in the future, it is very important to collect landraces from other parts of Turkey, and the iPBS-retrotransposons could be very beneficial to draw a clear picture of this. In the future, there is also a need to perform genome-wide association studies using this studied material in order to identify the genes associated with different traits of interest. For a long-term breeding program of tobacco, there is a need to convert the identified genes into the Kompetitive Allele Specific PCR

assay. Results of this study provide the genetic relatedness among the Turkish tobacco germplasms and we strongly believe that our results will be very beneficial for the breeders and researchers not only in Turkey but also in other parts of the world who are interested in the Turkish tobacco germplasm.

References

Andeden EE, Baloch FS, Derya M, Kilian B, Özkan H (2013). iPBS-Retrotransposons-based genetic diversity and relationship among wild annual *Cicer* species. J Plant Biochem Biotechnol 22: 453-66.

Arslan B, Okumus A (2006). Genetic and geographic polymorphism of cultivated tobaccos (*Nicotiana tabacum*) in Turkey. Genetika 42: 818-823.

Baloch FS, Alsaleh A, Andeden EE, Hatipoğlu R, Nachit M, Özkan H (2016). High levels of segregation distortion in the molecular linkage map of bread wheat representing the West Asia and North Africa region. Turk J Agric For 40: 352-364.

Baloch FS, Alsaleh A, de Miera LE, Hatipoğlu R, Çiftçi V, Karaköy T, Yıldız M, Özkan H (2015a). DNA based iPBS-retrotransposon markers for investigating the population structure of pea (*Pisum sativum*) germplasm from Turkey. Biochem Syst Ecol 61: 244-252.

Baloch FS, Alsaleh A, Shahid MQ, Çiftçi V, de Miera LE, Aasim M, Nadeem MA, Aktaş H, Özkan H, Hatipoğlu R (2017). A whole genome DArTseq and SNP analysis for genetic diversity assessment in durum wheat from central Fertile Crescent. PLoS One 12: e0167821.

Baloch FS, Derya M, Andeden EE, Alsaleh A, Comertpay G, Kilian B, Ozkan H (2015b). Inter primer binding site (iPBS) retrotransposon and ISSR diversity among wild lens species. Biochem Sys Ecol 58: 162-168.

Bindler G, Plieske J, Bakaher N, Gunduz I, Ivanov N, Van der Hoeven R, Ganal M, Donini P (2011). A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. Theor Appl Genet 123: 219-230.

Borna F, Luo NS, Ahmad NM, Nazeri V, Shokrpour M, Trethowan R (2017). Genetic diversity in populations of the medicinal plant *Leonurus cardiaca* L. revealed by inter-primer binding site (iPBS) markers. Genet Resour Crop Evol 64: 479-492

Cömertpay G, Baloch FS, Kilian B, Ülger AC, Özkan H (2012). Diversity assessment of Turkish maize landraces based on fluorescent labelled SSR markers. Plant Mol Biol Report 30: 261-274.

Dadras AR, Sabouri H, Nejad GM, Sabouri A, Shoai-Deylami M (2014). Association analysis, genetic diversity and structure analysis of tobacco based on AFLP markers. Mol Biol Rep 41: 3317-3329.

Darvishzadeh R, Gholizadeh S, Maleki HH, Abdollahi B, Bernousi I (2013). Study on genetic diversity among Iranian water pipe's tobacco (*Nicotiana* spp.) varieties by using simple sequence repeat markers. Bulg J Agric Sci 19: 557-562.

Demirel U, Tindas İ, Yavuz C, Baloch FS, Çalışkan ME (2018). Assessing genetic diversity of potato genotypes using inter-PBS retrotransposon marker system. Plant Genet Resour 16:137-145.

Denduangboripant J, Piteekan T, Nantharat M (2010). Genetic polymorphism between tobacco cultivar-groups revealed by amplified fragment length polymorphism analysis. J Agr Sci 2: 41-48.

Doyle J, Jeff J (1990). Isolation of plant DNA from fresh tissue. Focus 12: 13-15.

Esendal E (1994). Broad leaf tobacco growth and its importance in Turkey. In: Field Plant Congress, İzmir, Turkey (in Turkish).

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

Finnegan DJ (1989). Eukaryotic transposable elements and genome evolution. Trends Genet 5: 103-107.

Gong D, Huang L, Xu X, Wang C, Ren M, Wang C, Chen M (2016). Construction of a high-density SNP genetic map in flue-cured tobacco based on SLAF-seq. Molecular Breed 36: 100.

Govindaraj M, Vetriventhan M, Srinivasan M (2015). Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genetics Research International 2015: 431487.

Güler D, Demirbaş N (2016). Examination of the developments in the tobacco sector in Turkey. In: IBANESS Conference Series, Plovdiv, Bulgaria.

Hagenblad J, Zie J, Leino MW (2012). Exploring the population genetics of genebank and historical landrace varieties. Genet Resour Crop Evol 59:1185–1199.

Jaccard P (1987). Nouvelles recherces sur la distribution florale. Bull Soc Vaudoise Sci Nat 44: 103-107.

Julio E, Verrier JL, De Borne FD (2006). Development of SCAR markers linked to three disease resistances based on AFLP within *Nicotiana tabacum* L. Theor Appl Genet 112: 335-346.

Kalendar R, Antonius K, Smýkal P, Schulman AH (2010). iPBS: a universal method for DNA fingerprinting and retrotransposon isolation. Theor Appl Genet 121: 1419-1430.

Kalivas A, Ganopoulos I, Bosmali I, Tsaliki E, Osathanunkul M, Xanthopoulou A, Moysiadis T, Avramidou E, Grigoriadis I, Zambounis A et al (2016). Genetic diversity and structure of tobacco in Greece on the basis of morphological and microsatellite markers. Crop Sci 56: 2652-2662.

Karagoz A (2001). Plant genetic resources conservation in Turkey. Acta Hort 598: 17-25.

Karaköy T, Erdem H, Baloch FS, Toklu F, Eker S, Kilian B, Özkan H (2012). Diversity of macro-and micronutrients in the seeds of lentil landraces. Sci World J 2012: 710412.

Leitch IJ, Hanson L, Lim KY, Kovarik A, Chase MW, Clarkson JJ, Leitch AR (2008). The ups and downs of genome size evolution in polyploid species of *Nicotiana* (Solanaceae). Ann Bot-London 101: 805-814.

Lu XP, Xiao BG, Li YP, Gui YJ, Wang Y, Fan LJ (2013). Diversity arrays technology (DArT) for studying the genetic polymorphism of flue-cured tobacco (*Nicotiana tabacum*). J Zhejiang Univ-Sc B 14: 570-577.

Mackay J, Eriksen MP (2002). The Tobacco Atlas. Geneva, Switzerland: World Health Organization.

Mehmood A, Jaskani MJ, Ahmad S, Ahmad R (2013). Evaluation of genetic diversity in open pollinated guava by iPBS primers. Pak J Agr Sci 50: 591-597.

Moon HS, Nicholson JS, Lewis RS (2008). Use of transferable *Nicotiana tabacum* L. microsatellite markers for investigating genetic diversity in the genus *Nicotiana*. Genome 51: 547-559.

Moon HS, Nifong JM, Nicholson JS, Heineman A, Lion K, Van der Hoeven R, Hayes AJ, Lewis RS (2009). Microsatellite-based analysis of tobacco (*Nicotiana tabacum* L.) genetic resources. Crop Sci 49: 2149-2159.

Nadeem MA, Nawaz MA, Shahid MQ, Doğan Y, Comertpay G, Yıldız M, Hatipoğlu R, Ahmad F, Alsaleh A, Labhane N et al. (2017). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. Biotechnol Biotechnol Equip.

Nawaz MA, Yang SH, Rehman HM, Baloch FS, Lee JD, Park JH, Chung G (2017). Genetic diversity and population structure of Korean wild soybean (*Glycine soja* Sieb. and Zucc.) inferred from microsatellite markers. Biochem Sys Ecol 71: 87-96.

Nemli S, Kianoosh T, Tanyolaç MB (2015). Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) accessions through retrotransposon-based interprimer binding sites (iPBSs) markers. Turk J Agric For 39: 940-948.

Ren N, Timko MP (2001). AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild *Nicotiana* species. Genome 44: 559-571.

Renny-Byfield S, Chester M, Kovařík A, Le Comber SC, Grandbastien MA, Deloger M, Nichols RA, Macas J, Novák P, Chase MW et al. (2011). Next generation sequencing reveals genome downsizing in allotetraploid *Nicotiana tabacum*, predominantly through the elimination of paternally derived repetitive DNAs. Mol Biol Evol 28: 2843-2854.

Rossi L, Bindler G, Pijnenburg H, Isaac PG, Giraud-Henry I, Mahe M, Orvain C, Gadani F (2001). Potential of molecular marker analysis for variety identification in processed tobacco. Plant Var Seeds 14: 89-101.

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z et al. (1996). Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765-768.

Xia YS, Guo PG, Li RH, Lü YH, Qiu MW, Zhao WC, Carfi L, He QF, Yu YW (2014). Analysis of genetic diversity and population structure using SSR markers in tobacco. Adv Mat Res 850-851: 1243-1246.

Yang BC, Xiao BG, Chen XJ, Shi CH (2007). Assessing the genetic diversity of tobacco germplasm using inter simple sequence repeat and inter-retrotransposon amplification polymorphism markers. Ann Appl Biol 150: 393-401.

Yeh FC, Yang R, Boyle TJ, Ye Z, Xiyan JM (2000). Pop Gene 32, Microsoft Windows-Based Freeware for Population Genetic Analysis. Version 1.32. Edmonton, Canada: University of Alberta.

Yıldız M, Koçak M, Baloch FS (2015). Genetic bottlenecks in Turkish okra germplasm and utility of iPBS retrotransposon markers for genetic diversity assessment. Genet Mol Res 14: 10588-10602.