AcademicPres

Available online at www.notulaebotanicae.ro

Print ISSN 0255-965X; Electronic 1842-4309

Not Bot Horti Agrobo, 2012, 40(1): 247-253

Notulae Botanicae Horti Agrobotanici Cluj-Napoca

Molecular Characterization and Similarity Relationships among Flue-Cured Tobacco (*Nicotiana tabacum* L.) Genotypes Using Simple Sequence Repeat Markers

Soheila GHOLIZADEH[1], Reza DARVISHZADEH[1*], Babak ABDOLLAHI MANDOULAKANI[1]
Iraj BERNOUSI[1], Seyed Reza ALAVI[2], Ardashir KHARABIAN MASOULEH[3]

[1]*Urmia University, Department of Agronomy and Plant Breeding, Urmia, Iran; r.darvishzadeh@urmia.ac.ir (*corresponding author)*

[2]*Urmia Tobacco Research Center, Department of Genetic, Urmia, Iran*

[3]*Southern Cross University, Southern Cross Plant Science, Lismore NSW 2480, Australia*

Abstract

Characterization of genetic diversity has long been a major goal in tobacco breeding programs. Information on genetic diversity is essential for a rational use of genetic resources. In the present study, the genetic variation among 72 flue-cured tobacco genotypes was evaluated using microsatellite markers (SSRs). A set of 104 alleles was generated at 30 SSR loci. The mean number of alleles per locus (n_a) and the effective allele number (n_e) were 3.467 and 2.358, respectively. The expected heterozygosity ranged from 0.29 to 0.75 with average of 0.54. Several methods were used to construct the similarity matrices and dendrograms. The co-phenetic correlation coefficient, which is a measure of the correlation between the similarities represented on the dendrograms and the actual degree of similarity, was calculated for each dendrogram. Among the different methods, the highest value (r=0.76368) was observed for the UPGMA created based on Jaccard's similarity coefficients. The genetic similarity among the tobacco genotypes calculated by using Jaccard's similarity coefficient ranged from 0.08 to 0.84, suggesting the presence of high molecular genetic variability among the studied tobacco genotypes. Based on UPGMA clustering method all studied flue-cured tobacco genotypes, except for 'Glustinusa Rasht', were placed in three distinct groups. We observed an obvious heterotic pattern in the studied flue-cured germplasm corresponding to genetic distances and classification dendrogram, which persuades exploitation of heterosis in flue-cured tobaccos.

Keywords: DNA polymorphism, effective allelic number, genetic variability, male sterile, molecular markers, UPGMA clustering method

Introduction

Nicotiana belongs to family *Solanaceae* and has been divided into three subgenera (*Rustica, Tabacum* and *Petunioides*) which contain more than 64 recognized species. Only two natural amphidiploid species, *Nicotiana tabacum* L. and *Nicotiana rustica* L. with 2n=48 chromosomes (Raju *et al.,* 2008) have widely been cultivated as a medicinal herb, trade commodity and as crop plant in many different cultures for thousands of years (Yang *et al.,* 2007; Zhang and Liu, 2008). This plant is widely utilized as a model system in plant cell-cultures and genetic engineering researches (Zhang and Liu, 2008). Recently, use of tobacco seed oil as a renewable and potential source of energy was reported by authors (Giannelos *et al.,* 2002). Because of its economic importance and value in biological researches, numerous investigations have been undertaken to examine its evolutionary origin and genome structure and organization (Zhang and Liu, 2008).

Numerous types of tobacco are defined by different criteria such as region of production, intended use in cigar (i.e., filler, binder and wrapper) and cigarette manufacturing, method of curing (flue-, air-, sun- and fire-cured tobacco) as well as morphological and biochemical characteristics (i.e., aromatic fire-cured, bright leaf tobacco, Burley tobacco, Turkish or oriental tobacco) (Ren and Timko, 2001). Flue-cured tobacco (*Nicotiana tabacum* L.) is one of the most important commercial types of tobacco in the world (Liu *et al.,* 2009) and uses as a main component in American Blend type cigarettes.

Tobacco breeding aims to develop varieties with wide adaptability, higher yield potential and suitable chemical constituents for cigarette industry. To explore the genetic potential and select suitable parents, it is necessary to study genetic diversity of tobacco germplasm (Yang *et al.,* 2007). Characterization and quantification of genetic diversity has long been a major goal in tobacco breeding programs. Information on genetic diversity is essential for a rational use of genetic resources. Morphological, karyotypic, and physiological characters have already been used to study the genetic diversity of tobacco germplasm (Goodspeed, 1945; Zhang, 1994; Zhang *et al.,* 2005). However, morphological characters usually vary with environments. The number of karyotypical characters is limited, and study of

genotypic diversity based on isozyme variation is restricted to a few polymorphic enzyme systems encoded by a small number of loci (Zhang *et al.,* 2006).

During the last two decades, DNA-based molecular markers have been extensively used for a variety of purposes in many animal and plant systems. DNA markers such as RFLP (Restriction Fragment Length Polymorphism; Botstein *et al*., 1980), RAPD (Random Amplified Polymorphic DNA; Williams *et al*., 1990), SSRs (Simple Sequence Repeats; Rafalski and Tingey, 1993), AFLP (Amplified Fragments Length Polymorphism; Vos *et al*., 1995), as well as SNP (Single Nucleotide Polymorphism; Haff and Smirnov, 1997) has accelerated and facilitated the genetic diversity and evolutionary studies in plants. Molecular markers have become useful tools to provide a relatively unbiased method of quantifying genetic diversity in plants. Studies on the genetic diversity were performed by several researchers in the genus *Nicotiana* using different molecular markers. Zhang *et al.* (2008) used RAPD and AFLP markers to assess the genetic similarity among selected flue-cured tobacco accessions. Arslan and Okumus (2006) studied the genetic polymorphism of cultivated tobaccos (*Nicotiana tabacum* L.) in Turkey by RAPD markers. Qi *et al.* (2006) used ISSR markers to study genetic diversity in Chinese tobacco germplasm and reported that the genetic basis was comparatively narrow in studied germplasm. The recent development of several hundred microsatellite markers for tobacco (Bindler *et al*., 2007) has opened the ways of analysing molecular genetic diversity in this crop. Microsatellite markers, mainly due to their high polymorphism, random distribution and co-dominant Mendelian inheritance, are the most reliable markers for cultivar identification and genetic diversity studies. They were applied in tobacco researches for identification of inbred lines and cultivars (Davalieva *et al.,* 2010; Siva Raju, 2011). Davalieva *et al.* (2010) classified 10 tobacco genotypes into three groups using 24 SSR markers. SSR markers have been successfully employed to reveal genetic variation in chewing tobacco genotypes (Siva Raju, 2011). With the recent advent of high-density SSR maps for tobacco (Bindler *et al*., 2007), it is feasible to estimate genetic variation with a large number of SSR markers that are well distributed across the tobacco genome. The main advantage of using markers with known map positions is ability to create sufficient coverage over the genome. It is thus possible to avoid overrepresentation of certain regions of the genetic map that can produce inaccurate estimates of genetic similarities among individuals.

Iran, due to its geographical situation, is one of the most favourable regions for flue-cured tobacco cultivation. Little is known about the genetic variability of the available germplasm for flue-cured tobacco in Iran. The present study aimed to determine and assess the genetic variability of different local and exotic flue-cured tobacco using tobacco-specific simple sequence repeat (SSR) markers. This is the first report about the genetic diversity of flue-cured tobacco in Iran.

Materials and methods

Plant material and DNA extraction

Seventy-two flue-cured tobacco genotypes provided by Urmia Tobacco Research Center were investigated in the present study (Tab. 1). Seeds of genotypes were cultivated in pot and grown in growth chamber at 25±2°C. Genomic DNA was extracted from the leaves of 50-days-old seedlings using the CTAB-based method (Reichardt and Rogers, 1994). Genomic DNA was re-suspended in 100 µL TE (10 mM Tris, 1 mM EDTA). The concentration of each DNA sample was determined spectrophotometrically at 260 nm (BioPhotometer 6131; Eppendorf, Hamburg, Germany). The quality of the DNA was checked by running 2 µl DNA in 1% (w/v) gels in 0.5X TBE buffer (45 mM Tris base, 45 mM boric acid, 1 mM EDTA pH 8.0). The DNA samples with smear in the gel were rejected.

Thirty SSR markers out of 278 from the tobacco SSR database (Blinder *et al.,* 2007) were used for DNA fingerprinting (Tab. 2). The choice of SSR markers was based on clarity of produced bands and their genetic locations in order to give a uniform coverage of the tobacco genome (Blinder *et al.,* 2007).

Polymerase chain reaction (PCR)

Amplifications were performed in a volume of 25 µl containing 10 µM of each primer, 0.5 unit of Taq DNA polymerase (CinnaGen, Tehran, Iran), 10 mM of each dNTP (BioFluxbiotech, http://biofluxbiotech.com), 2 µl of 10X PCR buffer, 50 mM $MgCl_2$ (CinnaGen, Tehran, Iran), ddH2O and 50 ng DNA by using a 96-well Eppendorf Mastercycler Gradient (Type 5331, Eppendorf AG, Hamburg, Germany). Thermal cycles were programmed for 36 cycles as follows: 1 min at 94°C, 1.5 min at 55°C, and 1.5 min at 72°C, with an initial melting of 4 min at 94°C, and a final extension of 10 min at 72°C. The reaction products were mixed with 5 µl of formamide dyes (98% formamide, 10M m EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol) and resolved in a 3% (w/v) agarose gel in 0.5X TBE buffer, stained with 1.0 µg ml^{-1} ethidium bromide and photographed under UV light using a Gel-Doc image analysis system (Gel Logic 212 PRO, USA).

Data analysis

The amplification products were scored for the presence (1) and absence (0) of bands across the 72 genotypes to construct a binary data matrix. Mean number of allele per locus (n_a), effective allele number (n_e), allele frequency, observed and expected heterozygosities (H_o and H_e, respectively) were estimated using the GenAlEx software version 6.41 (Peakall and Smouse, 2006). $n_a = \sum_{i=1}^{n} n_{ai}/n$, where n_{ai} is the number of alleles at i^th^ locus. $n_e = \sum_{i=1}^{n} n_{ei}/n = \sum_{i=1}^{n} \left(1 - \sum_{j=1}^{m_i} q_{ij}^2\right)/n$ (Brown and Weir, 1983), where n_{ei} is the effective al-

Tab. 1. Name of the flue cured tobacco genotypes and their special cluster group based on un-weighted pair-group method using arithmetic average (UPGMA) algorithm and Jaccard's similarity coefficient on simple sequence repeat data

Code	Flue-cured tobacco genotype	Cluster group	Code	Flue-cured tobacco genotype	Cluster group	Code	Flue-cured tobacco genotype	Cluster group
g1	'MontCalm Brum'	1	g25	'NC95×Hicks'	2	g49	'MC101'	2
g2	'Bell'	2	g26	'K394'	2	g50	'Vo3/15'	2
g3	'PfaTzer'	2	g27	'Kutsaga51E'	2	g51	'NC11-15'	2
g4	'Hawana142'	2	g28	'Bel71-500'	2	g52	'Coker258'	2
g5	'VirRee'	4	g29	'PBD6'	2	g53	'Previ stammV3'	2
g6	'Lassomption'	2	g30	'Vir63'	2	g54	'SpeightG28'	2
g7	'Hawana'	2	g31	'Virree48'	2	g55	'VirE1'	2
g8	'Petrich84'	2	g32	'Kutsaga513'	2	g56	'Vo3/10'	2
g9	'Vir53'	2	g33	'South Carolina'	2	g57	'Tirtash19'	2
g10	'Sumatra9'	2	g34	'Vo3/3'	2	g58	'Bel61-9'	2
g11	'Hicks'	2	g35	'Asombon'	2	g59	'Vir31'	2
g12	'Bel61-10'	4	g36	'Kut110'	2	g60	'Geaderthermer'	2
g13	'RoseCanell'	2	g37	'Vir37'	2	g61	'E1'	2
g14	'AllPurpase'	2	g38	'FixedA1'	2	g62	'Glustinusa Rasht'	3
g15	'Parafum'	2	g39	'Vo3/4'	2	g63	'Coker342'	2
g16	'VirHerzegorina'	4	g40	'PeeDee'	2	g64	'Florida513'	2
g17	'VirH-R'	2	g41	'Harrison'	2	g65	'Perega'	2
g18	'ComSock Spanish'	4	g42	'By4'	2	g66	'Manilla'	2
g19	'GewonGrone'	2	g43	'VirAurea'	2	g67	'MC1'	2
g20	'Ludoyoretz'	4	g44	'Virgold'	2	g68	'Coker176'	2
g21	'NCTG52'	4	g45	'Amers'	2	g69	'VirAmutant'	1
g22	'Vir R.P37'	2	g46	'Vo3/5'	2	g70	'Vo3/11'	2
g23	'NC95'	2	g47	'Coker55'	2	g71	'Coker254'	2
g24	'Hicks26-110'	4	g48	'Coker319'	2	g72	'Vo3/6'	2

lelic number at ith locus, and q_j the frequency of the jth allele. $\text{Allele frequency} = \frac{2N_{XX} + N_{XX}}{2N}$, calculated locus by locus (Hartl and Clark, 1997); where N_{XX} is the number of homozygotes for allele, *X(XX)* N_{XY} is the number of heterozygotes containing the allele *X* (*Y* can be any other allele), N = the number of samples. $H_0 = \sum_{i=1}^{n} H_{oi}/n = \sum_{i=1}^{n}\left(1-\sum_{j=1}^{mi} q_{ij}^2\right)/n$ (Hartl and Clark, 1997), where H_{oi} represents the observed heterozygosity of the ith locus, and q_{ij} is the frequency of the jth allele at i^{th} locus. $H_e = \sum_{i=1}^{n} Hi/n = \sum_{i=1}^{n}\left(1-\sum_{j=1}^{mi} q_{ij}^2\right)/n$ (Lynch and Milligan, 1994), where H_i is the expected heterozygosity of the ith locus, and q_{ij} is the frequency of the jth allele at i^{th} locus. Different methods were used for calculating similarity matrices and constructing dendrograms. The efficiency-of-clustering algorithms and their goodness-of-fit were determined based on co-phenetic correlation coefficients. Data analyses were performed using the NTSYS-pc version 2.11 software (Rohlf, 1998).

Results and discussion

Out of 278 SSR primer pairs tested (Blinder *et al.,* 2007), 30 primer pairs were selected according to their polymorphism (Tab. 2). One hundred and four alleles were generated at 30 SSR loci. The size of amplified PCR products ranged from 140 to 322 bp (Tab. 2). Number of allele per locus ranged from 2 to 6 (PT30285 locus) suggesting the presence of high molecular genetic variability among the studied tobacco genotypes which are in agreement with the finding of Davalieva *et al.* (2010), Denduangboripant *et al.* (2010), and Moon *et al.* (2008; 2009) by means of SSR, AFLP and SSR markers, respectively. In the current study, the mean number of allele per locus was 3.46, which was similarly reported in Macedonian tobaccos with average of 3 alleles per locus (Davalieva *et al.,* 2010). The effective allelic number (n_e) was 2.36 on average and ranged from 1.398 to 4.01. This parameter takes into account both the number of alleles and their frequencies. It allows us to compare genotypes where the number and distributions of alleles differ drastically. The observed heterozygosity ranged from 0.00 to 0.94, with an average of 2.36. Six loci did not show any heterozygosity. The expected heterozygosity ranged from 0.29 to 0.75 with average of 0.54. The informativeness of SSR markers were measured by expected heterozygos-

Tab. 2. Names, sequences and linkage groups of the 30 simple sequence repeat (SSR) primer pairs applied to 72 flue-cured tobacco genotypes

Primer name	Size (bp)	Sequence of primer F 5'→3'	Sequence of primer R 5'→3'	Repeated motif	AT	LG
PT30021	224	CATTTGAACATGGTTGGCTG	CTCAACTCTCGTCGCTCTTG	TA	55	4
PT30132	216	CCTAACAGCATTTGCTACCCA	GATGGACAAGAGTGGCCTTT	TA	55	10
PT30202	225	TCGAAACCTCGAGGACAGTT	TATCCAAATCTCCAAAGCCC	GA	55	7
PT30159	197	GCATGCATATGAACATGGGA	TTTGACATCTCTACTCTTCCGTTT	TA	55	14b
PT30175	229	TTAGGCGGCGGTATTCTTAT	TATGCCTCAATCCCTTACGC	TA	55	14a
PT30285	177	CATCATGGCAAGTCACCATC	TGCTGGAAATTAGCGAGGTT	TA	55	18
PT30324	151	TGCTCTGCGTTAGAACAGGA	CGACGAGAGAAGATTAGTGAAAGA	TAA	55	12
PT20343	322	GGAACACCACCACCATAA	GGAGCTCAGGTTCCAATG	AC/AG/AT	55	4
PT30075	195	CGATCGGGTCGTTACACAAT	CCCATCAGGTTGTTGGGTTA	TA	55	11
PT30241	199	AAGTCTCGTGTGGTTGCTTT	AAAGGGCAATGTGTCTAGCTC	GA	55	15
PT30061	182	TCGTCCATTTCTTTCTCTCTCA	CATAAATAGTTGCTCATTCAATCG	TA	55	11
PT30144	266	TGATTTGTATTGACAGCGTGAAG	TTGTTTAGTTACCCTATTTGACTTGC	TA	55	16
PT30332	230	AAACCGAACCGAACTGATTT	TCAAATTTATGATTCTTGTAGCGAA	TA	55	16
PT30124	228	TCCTCCAACCAAACTCAAGC	TTTCTGTTCGCGTTTCAAAT	TA	55	4
PT30110	213	TTGTACGTTCCTCGCTGATG	GGCCGACAATAAAGTGGCT	TA	55	21
PT20275	184	GTTCTATTTGATCGCCCC	AACAGCACCAACAGCATT	CTT	55	5
PT30260	225	GGTAGGGTGGAACAAATTTATCA	AATATGGTCTATGCCCGCAA	TA	55	8a
PT30067	204	AAGCCTGGTCAGTTATCCCA	ATTCGCACCACTTAATCCCA	TA	55	2
PT30126	208	GTGATTCCAGCGGAAGACAT	TTCGAAATAAGTACCTAGAGTCGG	TA	55	10
PT30034	216	GACGAAACTGAGGATATTCCAAA	TGGAAACAAAGCCATTACCC	TAA	55	22
PT30008	192	CGTTGCTTAGTCTCGCACTG	GGTTGATCCGACACTATTACGA	TA	55	11
PT30165	224	ACCTCTGTGGCCGTAAGCTA	CCTCTACTTCAACAGGGTAAGAAA	TAA	55	19
PT30014	205	TGCCGTGTAAATTTCATTTGG	AGGATTCCTAACGTGTATTATGTTCT	TA	55	11
PT30272	140	GAACCTAACCTCGCTCCACA	AAATGGTAGCTGCGAGGAGA	GA	55	4
PT30171	218	CCCATGCATGCCTAATTTCT	CCCAGAAGCCCTTATACAACC	TA	55	24
PT30172	216	AAACAACGTCGAAGCATTTG	ACGCATGAAATTGTAAGGGC	GAA	55	4
PT30205	193	GGTCGATCCACAATTTAAACG	GCACTTGCTCCTTTGTACCC	TA	55	3b
PT20287	164	CGCCACAACAACTCACCTTA	TCATGCATGTTTCTCCTCCTT	AAG	55	3a
PT30250	177	GAACACACGTTCGTCATTGG	ATAAGTCCCTTTAATTTAATTGCG	TAG	55	10
PT30292	156	AAGACAGATTGGTGCGGAAC	AGCACTTGGACAGGCGAATA	TA	55	7

AT: Annealing temperature, LG: linkage group

ity. SSR markers with high heterozygosity values such as PT30126 and PT30285 could be effectively used in genetic diversity studies. There is a considerable difference between observed and expected heterozygosity for several analysed SSRs, indicating lack of Hardy-Weinberg equilibrium. This situation could be consequence of selection or gene flow or even the genetic drift. Wright fixation index (F) ranged from 0.00 (PT30202, PT30324, PT30285, PT30124, PT30205) to 1.00 (PT30241, PT30332, PT30110, PT20275, PT30171, and PT30250) with an average of 0.64 (Tab. 3).

Different methods were used to construct the similarity matrices and dendrograms (Tab. 4). The co-phenetic correlation coefficient, a measure of the correlation between the similarities represented on the dendrograms and the actual degree of similarity, was calculated for each dendrogram (Tab. 4). Among the different methods, the highest value (r=0.76368) was observed for the UPGMA method based on Jaccard's similarity coefficients (Tab. 4). Therefore, the dendrogram constructed using this method was used to depict the genetic diversity of the flue-cured tobacco genotypes (Fig. 1). The genetic similarity among the tobacco genotypes varied from 0.08 (between 'Gead-erthermer' and 'Hawana142' genotypes) to 0.84 (between 'MC101' and 'Vo3/15' genotypes) based on Jaccard's similarity coefficient, which indicates the high level of genetic variation among studied flue-cured tobacco genotypes. This result also validates that this collection is a valuable tobacco germplasm that has not been exposed to degradation yet. Davalieva *et al.* (2010) indicated a wide range of genetic diversity among the selected tobacco varieties using SSR markers. In a SSR marker study, it was demonstrated that the most of the US *Nicotiana* germplasm collection are considerably distinct from each other (Moon *et al.,* 2009) that is in agreement with the present results. In contrast to the present results, Yang *et al.* (2007) by using inter-simple sequence repeat and inter-retrotransposon amplification polymorphism (IRAP) markers, Arsalan

Tab. 3. Number of alleles (Na), effective allelic number (Ne), observed heterozygosity (Ho) and expected heterozygosity (He) and allele frequency of the 30 simple sequence repeat (SSR) loci applied to 72 flue-cured tobacco genotypes

Primer name	Na	Ne	Ho	He	F	Allel frequency					
						A	B	C	D	E	F
PT30021	3	2.925	0.21	0.658	0.681	0.371	0.371	0.258	-	-	-
PT30132	5	2.901	0.617	0.655	0.059	0.025	0.158	0.525	0.125	0.167	-
PT30202	4	2.895	0.896	0.655	0	0.343	0.104	0.455	0.097	-	-
PT30159	4	3.28	0.183	0.695	0.736	0.242	0.183	0.442	0.133	-	-
PT30175	3	2.697	0.263	0.629	0.582	0.175	0.412	0.412	-	-	-
PT30285	6	4.012	0.91	0.751	0	0.03	0.119	0.284	0.299	0.254	0.015
PT30324	4	2.949	0.871	0.661	0	0.114	0.221	0.5	0.164	-	-
PT20343	3	2.851	0.075	0.649	0.885	0.239	0.425	0.336	-	-	-
PT30075	5	3.296	0.329	0.697	0.528	0.086	0.071	0.421	0.321	0.1	-
PT30241	2	1.958	0	0.489	1	0.574	0.426	-	-	-	-
PT30061	4	1.718	0.058	0.418	0.861	0.051	0.072	0.746	0.13	-	-
PT30144	3	1.833	0.17	0.454	0.626	0.179	0.708	0.113	-	-	-
PT30332	3	2.214	0	0.548	1	0.393	0.541	0.066	-	-	-
PT30124	4	2.697	0.657	0.629	0	0.06	0.127	0.507	0.306	-	-
PT30110	5	2.483	0	0.597	1	0.015	0.029	0.074	0.382	0.5	-
PT20275	2	1.6	0	0.375	1	0.75	0.25	-	-	-	-
PT30260	3	1.443	0.083	0.307	0.729	0.083	0.094	0.823	-	-	-
PT30067	4	2.786	0.069	0.641	0.892	0.017	0.302	0.207	0.474	-	-
PT30126	4	3.551	0.078	0.718	0.891	0.324	0.157	0.353	0.167	-	-
PT30034	3	2.449	0.015	0.592	0.975	0.118	0.493	0.39	-	-	-
PT30008	3	1.94	0.066	0.484	0.865	0.189	0.68	0.131	-	-	-
PT30165	2	1.654	0.086	0.396	0.783	0.271	0.729	-	-	-	-
PT30014	4	1.683	0.148	0.406	0.637	0.041	0.139	0.066	0.754	-	-
PT30272	2	1.747	0.175	0.427	0.592	0.69	0.31	-	-	-	-
PT30171	3	2.09	0	0.522	1	0.424	0.545	0.03	-	-	-
PT30172	2	1.411	0.231	0.291	0.208	0.177	0.823	-	-	-	-
PT30205	4	2.318	0.939	0.569	0	0.015	0.068	0.402	0.515	-	-
PT20287	2	1.398	0.031	0.285	0.89	0.172	0.828	-	-	-	-
PT30250	3	1.574	0	0.365	1	0.042	0.775	0.183	-	-	-
PT30292	5	2.383	0.092	0.58	0.841	0.031	0.069	0.177	0.608	0.115	-
Mean	3.467	2.358	0.242	0.538	0.642						

and Okumus (2006) using RAPD markers and Ren and Timko (2001) using AFLP markers reported the low levels of genetic diversity among Chinese flue-cured tobacco collections, cultivars planted in eastern Anatolia of Turkey and some world tobacco collection, respectively. Differences in results reported by various studies could be due to differences in number of studied genotypes, their genetic background, and number of markers used as well as techniques applied to detect polymorphism.

Tab. 4. Comparison of different methods for constructing similarity matrices and dendrograms in 72 flue-cured tobacco genotypes

Similarity matrices	Algorithm	Co-phenetic coefficient
Jaccard	UPGMA	0.76368
Jaccard	Complete linkage	0.56817
Jaccard	Single linkage	0.71573
Dice (Nie and Li)	UPGMA	0.74314
Dice (Nie and Li)	Complete linkage	0.5381
Dice (Nie and Li)	Single linkage	0.71897
Simple matching	UPGMA	0.74279
Simple matching	Complete linkage	0.57483
Simple matching	Single linkage	0.70047

UPGMA= un-weighted pair-group method using arithmetic average

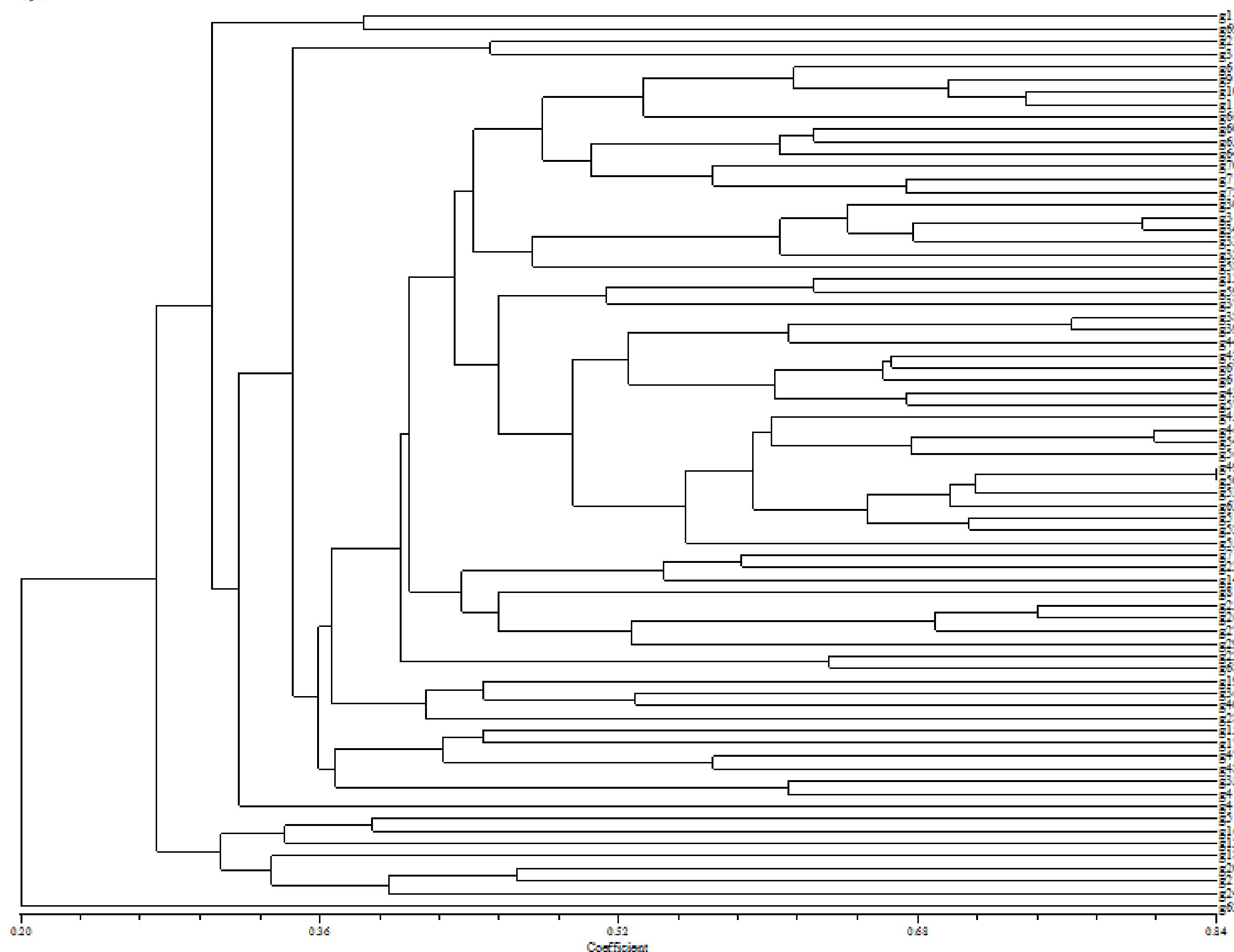

Fig. 1. Dendrogram of 72 flue-cured tobacco genotypes generated by un-weighted pair-group method using arithmetic average (UPGMA) clustering based on Jaccard's coefficients of similarity on the simple sequence repeat (SSR) data

Based on UPGMA clustering method the studied flue-cured tobacco genotypes, except for genotype 'Glustinusa Rasht' were separated into three distinct groups. The majority of genotypes clustered in group 2. Seven male sterile genotypes ('Vo3/3', 'Vo3/4', 'Vo3/5', 'Vo3/6', 'Vo3/10', 'Vo3/11', and 'Vo3/15') were included in the present study as well. They all displayed a similar DNA fingerprint. These genotypes situated in group 2. The genotype 'Vo3/15' showed high genetic similarity (0.84) with 'MC101' genotype. Genotype 'Vo3/3' showed low similarity (0.38) to 'Virree48'. Genotype 'Vo3/4' located in the same sub-cluster with 'FixedA1' genotype. Other male sterile genotypes ('Vo3/11', 'Vo3/6', and 'Vo3/10') located near to each other in cluster 2. Male sterile lines persuade the application of line × tester design in flue-cured tobacco breeding programs. In this study, genotype 'Glustinusa Rasht' was placed into a separate cluster, emphasizing the existence of high level of difference between this genotype and others. Regarding to genetic distances and classification dendrogram, there is obvious heterotic pattern in this flue-cured germplasm which persuade exploitation of heterosis in flue-cured tobaccos. Duplications of genetic material were not observed in the studied germplasm collection.

Conclusions

Identification and quantification of genetic diversity has long been a major aim in plant breeding programs. Information on genetic diversity is essential for use of genetic resources in breeding programs. It is especially useful in characterizing individual, accessions and cultivars as well as in detecting duplications of genetic material in germplasm collections. As a consequence, it is important to select appropriate parents for hybridization in breeding programs and in developing informative mapping populations for QTL identification. Hybridization between any distantly related individuals is expected to yield more heterosis and vigorous plants constituting much of the different traits contained in the two parental lines. The present study shows that development of SSR markers is a valuable issue for genetic diversity studies within the genus *Nicotiana*.

Acknowledgments

We would like to thank the Institute of Biotechnology, Urmia University, Urmia, Iran, and Urmia Tobacco Research Center for financial support of this work.

References

Arslan B, Okumus A (2006). Genetic and geographic polymorphism of cultivated tobacco (Nicotiana tabacum) in Turkey. Russ J Genetics 42:667-671.

Bindler G, van der Hoeven R, Gunduz I, Plieske J, Ganal M, Rossi L, Gadani F, Donini P (2007). A microsatellite marker based linkage map of tobacco. Theor Appl Genetics 114:341-349.

Botstein D, White RL, Skolnick M, Davis RE (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Human Genetic 32:314-331.

Brown AHD, Weir BS (1983). Measuring genetic variability in plant populations, 219-239 p. In: Tanksley SD, Orton L (Eds.). Isozymes in plant genetics and breeding, Part A. Elsevier Science Publishers BV, Amsterdam.

Davalieva K, Maleva I, Filiposki K, spiroski O, Efremov GD (2010). Genetic variability of Macedonian tobacco varieties determined by microsatellite marker analysis. Diversity 2:439-449.

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

Giannelos PN, Zannikos F, Stournas S, Lois E, Anastopoulos G (2002). Tobacco seed oil as an alternative diesel fuel: physical and chemical properties. Ind Crops Prod 16:1-9.

Goodspeed TH (1945). Chromosome number and morphology in Nicotiana. VII. Karyotypes of fifty-five species in relation to a taxonomic revision of the genus. Univ California Pub in Bot 18:345-370.

Haff LA, Smirnov IP (1997). Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectroscopy. Genome Res 7:378-388.

Hartl DL, Clark AG (1997). Principles of Population Genetics, 3rd Ed. Sunderland, Massachusetts: Sinauer Associates, 542 p.

Liu XZ, Shen HeC, Yang YM, Zhang H (2009). Genetic diversity among flue-cured tobacco cultivars on the basis of AFLP markers. Czech J Genetics Plant Breed 4:155-159.

Lynch M, Milligan BG (1994). Analysis of population genetic structure with RAPD markers. Molec Ecol 3:91-99.

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.

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

Qi JM, Wang T, Chen SH, Zhou DX, Fang PP, Tao AF, Liang JX, Wu WR (2006). Genetic diversity and genetic relatives analysis of tobacco germplasm based on inter-simple sequence repeat (ISSR). Acta Agron Sinica 32:373-378.

Rafalski JA, Tingey SV (1993). Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genet 9:275-279.

Raju KS, Madhav MS, Sharma RK, Murthy TGK, Mohapatra T (2008). Genetic polymorphism of Indian tobacco types as revealed by amplified fragment length polymorphism. Current Sci 94:633-639.

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

Rohlf FJ (1998). NTSYSpc: Numerical Taxonomy and Multivariate Analysis System version 2.02, Exeter Software, Setauket, NY.

Reichardt M, Rogers S (1994). Preparation of plant DNA using CTAB, 2.3.3-2.3.7 p. In: Ausubel F, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (Eds.). Current Protocols in Molecular Biology, John Wiley & Sons, NY.

Siva Raju K (2011). Genetic diversity in Indian chewing tobacco (*Nicotiana tabacum*) as revealed by RAPD and SSR markers. The Ind J Agric Sci 81:15-19.

Vos P, Hogers R, Bleeker R, Reijans M, Van DE, Lee T, Hornes M (1995). AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407-4414.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JI, Tingey SV (1990). DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucl Acids Res 18:6531-6535.

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.

Zhang HY, Liu XZ (2008). Advance of molecular marker application in the tobacco research. Afr J Biotechnol 7:4827-4831.

Zhang HY, Liu XZ, Sheng HeC, Yang YM (2008). Genetic diversity among flue-cured tobacco cultivars based on RAPD and AFLP markers. Braz Arch Biol Technol 6:1097-1101.

Zhang HY, Liu XZ, Li TS, Yang YM (2006). Genetic diversity among flue-cured tobacco (*Nicotiana tabacum* L.) revealed by amplified fragment length polymorphism. Molec Biol 47:223-229.

Zhang HY, Liu XZ, He CS, Zheng CM (2005). Random amplified DNA polymorphism of nicotiana tabacum L. cultivars. Biol Plant 49:605-607.

Zhang DM (1994). Study on surface structure scanning of some tobacco seeds with electrical-glass observing. Chinese Tobacco Sinica 2:12-15 (In Chinese).