Indian Journal of Agricultural Sciences **81** (1): 15–9, January 2011

Genetic diversity in Indian chewing tobacco (*Nicotiana tabacum*) as revealed by RAPD and SSR markers

K SIVA RAJU

Central Tobacco Research Institute, Rajahmundry, Andhra Pradesh 533 105

Received: 6 November 2009; Revised accepted: 15 October 2010

ABSTRACT

Randomly amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) markers were employed to assess genetic diversity among released varieties of chewing tobacco cultivars grown in Tamil Nadu and Bihar. In RAPD, 15 out of 25 primers amplified 79 fragments of which 43 (54.43%) were polymorphic. Twelve SSR primers amplified 52 fragments of which 35 (67.3%) were polymorphic. Pair-wise genetic similarity based on RAPD assay ranged from 0.48 to 0.94 with an average of 0.71 while it ranged from 0.37 to 0.95 with a mean of 0.66 based on SSR marker assay. From the cluster analysis based on RAPD and SSR markers, the chewing tobacco varieties grown in these two states were found genetically divergent. The clustering pattern of the cultivars corresponded more with parentage. SSR markers were more efficient than RAPD markers in differentiating chewing tobacco cultivars and could be used for development of markers for various traits. The SSR primers TbM 27 and TbM 29 produced variety-specific markers.

Key words: Chewing tobacco, Genetic diversity, RAPD, SSR

Assessment of genetic diversity is critical for any successful breeding programme. Usually, breeders employ morphological markers for genetic diversity estimation. Chewing tobacco (*Nicotiana tabacum* L.) belongs to genus *Nicotiana* of family Solanaceae. India is the only country where several types of tobacco, viz Flue-Cured Virginia (FCV), chewing, burley, *natu*, cigar filler, cigar wrapper, cheroot, *hookah*, HDBRG and *bidi* are grown under different agro-climatic conditions (Gopalachari 1984). The qualities of tobacco grown in different locations differ considerably depending on the genotype, climatic conditions and cultural practices adopted. Chewing tobacco is an important cash crop in Bihar, Gujarat, Uttar Pradesh and Tamil Nadu producing nearly 83 million kg of cured leaf annually (Krishnamurthy and Dev Singh 2005). Chewing tobacco fetches sizeable income to the farmers, traders and government and provides employment to millions of people involved in its cultivation, curing, grading, processing and marketing.

Tobacco cultivars are currently distinguished by morpho-physio-chemical characters. Cultivar identification is rather slow and complicated by cultivar×environmental interaction (Nielsen 1985). Low level of genetic diversity within and among cultivated tobacco types was reported **(**Ren and Timko 2001). Cultivars that are closely related cannot be readily distinguished by morphological markers. Identification and differentiating the cultivars has become increasingly important for protecting intellectual property associated with the cultivars.

DNA marker-based genetic diversity assessment is more reliable as it is environment-independent and crop-stage non-specific. DNA markers such as RAPD and AFLP have been used to fingerprint different tobacco types. Among different molecular markers, SSRs are more abundant, ubiquitous in presence, hyper variable in nature and have high polymorphic information (Gupta *et al.* 1996). SSR markers have been used for genotype identification in many crop plants/cultivars, such as soybean (Song *et al*. 1999), potato (Barandalla *et al.* 2006), brinjal (Behera *et al*. 2006) and rice (Sheetal Yadav *et al*.2008). SSR markers are almost twice as informative as dominant markers such as RAPD and AFLP (Powell *et al*. 1996). SSR markers are considered as ideal for discriminating individuals and for parentage determinations. Systematic information on the diversity pattern among the released chewing tobacco varieties is limited. Hence genetic diversity among 12 chewing tobacco varieties released by Central Tobacco Research Institute were assessed based on RAPD and SSR markers.

MATERIALS AND METHODS

Twelve genotypes of chewing tobacco (Table 1) comprising 5 and 7 released high-yielding varieties from

[1]Principal Scientist (Plant Biochemistry) (e mail: rajuks2002@yahoo.co.in)

Table 1 Details of chewing tobacco varieties used

Variety	Parentage	Year of release	Growing state
'Vairam'	I 64×VTK I	1977	Tamil Nadu.
'Bhagyalakshmi'	I 64×I 735	1980	Tamil Nadu
'Maragadam'	PV 7×I 115	1981	Tamil Nadu.
'Meenakshi'	PV 7×HV67 9	1992	Tamil Nadu
'Abirami'	Chemical mutant of I 64	2006	Tamil Nadu
'DP 401'	Local selection	1958	Bihar
'Gandak Bahar'	Local selection	1976	Bihar
'Sona'	'Borimalingar Tohra' ×DP 401	1977	Bihar
'Prabha'	'Desila madhata'× 'Hingla kiratpur'	1981	Bihar
'Pusa Tobacco 76'	'HP 60-1'×'Bori'	1990	Bihar
'Vaishali Special'	'Bandi'×'Sona'	1993	Bihar
'Lichchavi'	'Boritepari'×'Sona'	2001	Bihar

Tamil Nadu and Bihar, respectively were subjected to RAPD and SSR analysis. These varieties differ for morphological and biochemical characteristics. They were grown in pots under glasshouse during 2008–09 season. Thirty-day-old seedlings of each variety were collected, bulked, frozen in liquid nitrogen, and stored at –80°C until used for DNA extraction. DNA was extracted from pulverized frozen seedlings as per the modified protocol by Doyle and Doyle (1990). The quantity of the extracted DNA was examined under UV light following agarose gel electrophoresis (0.8% gel containing 0.5 µg/ml ethidium bromide). The DNA concentration was estimated by comparing with known concentration of lambda DNA as standard.

A total of 25 decamer primers belonging to different series (Operon Technologies Inc, Alameda, CA, USA) were screened for RAPD analysis. Of these, 15 primers (OPP04, OPP05, OPP06, OPP09, OPP10, OPP16, OPP17, OPAB01, OPAB02, OPL16, OPL17, OPB05, OPM03, OPM06 and OPM12) were selected for analysis based on repeatability. The PCR amplification conditions were as described by Williams *et al*. (1990) with minor modifications. Amplifications were carried out in a 25 µl reaction mixture containing 25 ng template DNA, 0.5 units of Taq DNA polymerase, 0.2 mM of each dNTP and 20 ng of each primer. Amplification was carried out in a DNA thermal cycler (PTC–100, Peltier thermal cycler, MJ Biotech USA) programmed to run the following thermal profile : initial denaturation at 94°C for 5 min, 40 cycles of 94°C×1 min., 37°C×1 min., 72°C×2 min., followed by final extension at 72°C for 10 min. Amplification products were resolved on 1.4% agarose gel at 50 V and gels were documented in gel documentation system (Gene genius, Syngene bio-imaging system, UK). The size of the fragments was estimated using Gene Ruler 100 bp DNA ladder plus (MBI Fermentas, Lithuania) marker.

Twelve SSR primers were used in the present study and their NCBI accession numbers are DQ865416, DQ865430, DQ865431, DQ865433, DQ865434, DQ865435, DQ865437, DQ865439, EF375960, EF375967, EF375968 and EF375969. Amplifications were carried out in a 25 µl reaction mixture containing 15 ng template DNA, 0.5 units of Taq DNA polymerase, 0.2 mM of each dNTP and 10 ng of each of forward and reverse primers, 2.5 ml of 10×buffer and 2.5 mM $MgCl_2$. PCR cycles consisted of one cycle at 94°C for 5 min., followed by 35 cycles at 94°C for 1 min., 55°C for 1 min. and 72°C for 2 min. Final primer extension for 7 min. was carried out at 72°C. The PCR products were electrophoresed on 10% polyacrylamide gels to achieve better resolution of the bands. The gels were stained in ethidium bromide for 20 min. and documented. The size of the fragments was estimated using 50 bp Gene Ruler as in RAPD.

All the genotypes were scored for presence or absence of the RAPD/SSR bands. Only intense bands and those that were repeatedly amplified in RAPD were scored visually. The data were entered into a binary matrix as discrete variables, 1 for presence and 0 for absence of the character and this data matrix was subjected to further analysis. The Excel file containing the binary data was imported into NT Edit of NTSYS-pc software version 2.02 (Rohlf 1998). The 0/1 matrix was used to calculate similarity as DICE coefficient using SIMQUAL subroutine in Similarity routine. The resultant similarity matrix was employed to construct dendrogram using Sequential Agglomerative Hierarchical Nesting (SAHN) based Unweighted Pair Group Method with Arithmetic Means (UPGMA) to infer genetic relationships and phylogeny. To compare the efficiency of the two marker system, parameters such as expected heterozygocity (H_{ep}), fraction of polymorphic loci (ß), effective multiplex ratio (E) and marker index (MI) were estimated as described in Maras *et al.* (2008).

RESULTS AND DISCUSSION

Extent of genetic variation in chewing tobacco

RAPD and SSR markers were employed to characterize Indian chewing tobacco (*Nicotiana tabacum* L.) varieties released by CTRI research centres. Out of 25 primers used in RAPD, reproducible patterns were obtained with 15 primers, which generated a total of 79 highly reproducible bands in the size ranged from 0.2 to 2.9 kb. The total number of fragments amplified per primer ranged from 2 to 9 with an average of 5.26 (Table 2). Fourteen (93.3%) primers were found to be polymorphic while the primer OPM12 was monomorphic. The level of polymorphism detected with different primers ranged from 0 to 100%. Out of all the 79 amplified fragments, 43 (54.43%) were polymorphic. The

Table 2 Data on RAPD and SSR markers

Primer	Total bands	No. of polymorphic bands	Polymorphism (%)
OPP4	5	3	60
OPP5	5	2	40
OPP6	4	1	25
OPP9	2	2	100
OPP10	3	3	100
OPP16	6	4	66.6
OPP17	8	3	37.5
OPAB	5	3	60
OPAB2	5	3	60
OPL16	5	4	80
OPL17	8	7	87.5
OPB5	4	2	50
OPM3	9	5	55.55
OPM6	6	1	16.66
OPM12	4	0	0
TbM12	4	3	75
TbM26	4	2	50
TbM27	4	3	75
TbM29	3	2	66.6
TbM30	3	2	66.6
TbM31	6	6	100
TbM33	4	2	50
TbM35	4	2	50
TbM38	4	4	100
TbM45	4	3	75
TbM46	4	2	50
TbM47	8	4	50

observed level of polymorphism in chewing tobacco was very much higher than that reported by Del Piano *et al*. (2000) in 12 tobacco genotypes and was nearer to that reported by Siva Raju *et al*. (2008), who reported 59.4% polymorphism in 10 types of tobacco by RAPD, it may be due to different tobacco types. Besides, the primers used in the present study being

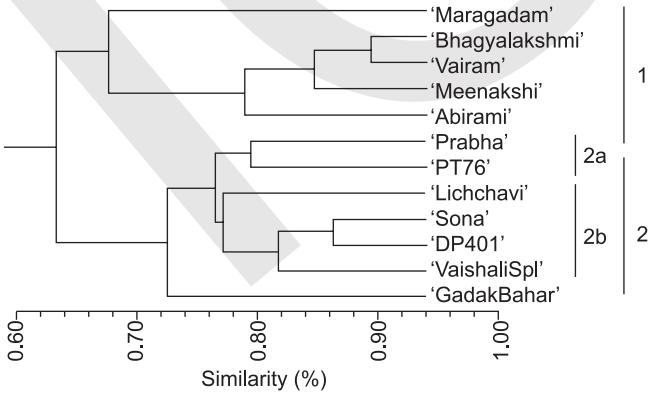

Fig 1 Dendrogram constructed using an UPGMA and SAHN algorithm from Jaccard's similarity data from all the RAPD marker profiles revealing genetic relationship among the varieties used in the study

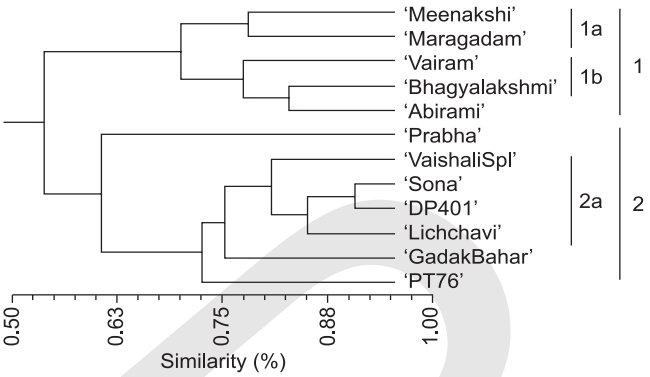

Fig 2 Dendrogram constructed using an UPGMA and SAHN algorithm from Jaccard's similarity data from all the SSR marker profiles revealing genetic relationship among the varieties used in the study

different in sequence, possibly assayed different genome regions. Four primers, namely OPP 9, OPP 10, OPL 16 and OPL 17 used in this study were found to be highly polymorphic. Therefore, these primers could be employed in the future RAPD-based diversity analysis of chewing tobacco lines and germplasm.

SSR analysis was carried out using 12 primers which were locus-specific tobacco primers (Fig 3). The number of fragments amplified with these primes ranged from 2 to 8 with an average of 4.3 (Table 3). Out of 52 fragments amplified from 12 primers, 35 (67.3%) were polymorphic. The primer TbM 31 and TbM 38 gave the highest level of polymorphism (100%), followed by TbM 12, TbM 27 and TbM 45. The number of polymorphic bands ranged from 2 to 6 with an average of 3.18. The observed level of polymorphism was very much higher than that reported by Yang *et al*. (2006), who reported 24.43% and 38.56% polymorphic loci in different tobacco types collected from

Table 3 Level of polymorphism and comparison of the discriminating capacity of RAPD and SSR markers in chewing tobacco varieties

Indexes	Marker system	
	RAPD	SSR
Number of assay units	15	12
Number of polymorphic bands	43	35
Number of monomorphic bands	36	17
Average no. of polymorphic bands/ assay unit	2.86	2.91
Number of loci	79	52
Number of loci/assay unit	5.26	4.33
Expected heterozygosity of the polymorphic loci	0.324	0.398
Fraction of polymorphic loci	0.54	0.67
Effective multiplex ratio	2.84	2.91
Marker index	0.92	1.158
Mean genetic similarity	0.71	0.66

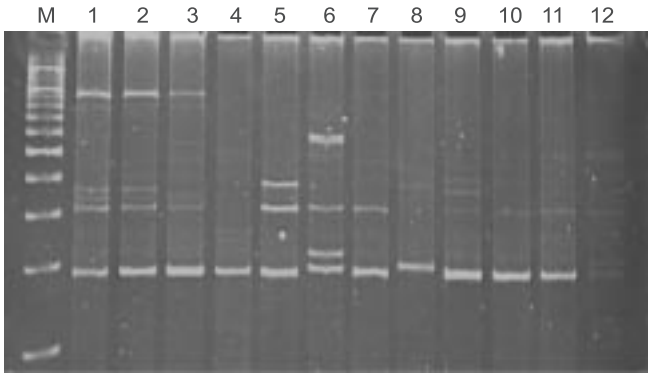

Fig 3 Amplified products of 12 chewing tobacco cultivars using SSR primer TbM31. Lanes *1*. 'Bhagyalakshmi', *2*. 'Maragadam', *3*. 'Vairam', *4*. 'Meenakshi', *5*. 'Abirami', *6*. 'DP 401', *7*. 'Gandak Bahar', *8*. 'Sona', *9*. 'Prabha', *10*. 'PT 76', *11*. 'Vaishali Special' and *12*. 'Lichachavi. M. DNA Ladder'

different countries based on inter-simple sequence repeats (ISSR) and inter retrotransposon amplification polymorphism (IRAP) markers. The SSR primers TbM 27 and TbM 29 produced variety-specific markers to Lichachavi and Abirami, respectively. No single primer individually differentiated all the genotypes but genotype-specific profile was observed with some primers. The combined profiles based on the polymorphic primers, however provide variety-specific patterns and thus could be distinguished some of the varieties.

Genetic relationship among the genotypes

The genetic relationship among chewing tobacco varieties was established separately with the two marker systems. In case of RAPD, the Jaccard's coefficient of similarity ranged from 0.48 to 0.94 with an average of 0.71. The similarity index values among varieties cultivated in Tamil Nadu ranged from 0.62 to 0.94, whereas it ranged from 0.69 to 0.91 among varieties grown in Bihar. The varieties grown in Tamil Nadu remained together in cluster 1 with an average genetic similarity of 0.78 while those of Bihar remained in the cluster 2 with an average genetic similarity of 0.8 (Fig 1). In cluster 1, the grouping of the varieties was based on parentage. 'Maragadam' variety was found most divergent and thus remained isolated from the other four varieties. High degree of similarity (94%) between 'Bhagyalakhmi' and 'Vairam' showed could be due to the common parental line I 64 (Fig 1). The variety 'Gandak Bahar', a local selection was found most divergent and thus remained isolated from the rest of the 6 varieties in the cluster 2. The varieties 'Prabha' and 'Pusa' Tobacco 76 (PT 76) which shared dissimilar parents were included in sub-cluster 2a. Varieties 'DP401', 'Sona', 'Vaishali Special' and 'Lichchavi' were included in sub-cluster 2b with very close genetic similarity due to close parental relation among them.

Based on the SSR profiles, the pair-wise Jaccards's similarity coefficients ranged from 0.37 to 0.95 with an average 0.66. The similarity indices among varieties bred from Tamil Nadu ranged from 0.64 to 0.86, whereas it ranged from 0.59 to 95 among the varieties bred from Bihar. The varieties of Tamil Nadu remained in the cluster 1 with an average genetic similarity 0.75, while those of Bihar remained in the cluster B with an average similarity of 0.77 (Fig 2).

The average of amplified fragments/primer was higher in RAPD (5.26) than SSR markers (4.33), it could be due to the SSR primers were specific to tobacco. Marker index and effective multiplex ratio were higher for SSR markers than RAPD markers (Table 3). The mean similarity index was higher in RAPD (0.71) than SSR markers (0.66) and cluster analysis from RAPD and SSR data produced similar but not identical clustering pattern, which could be due to higher probability of scoring of non-homologous RAPD amplicons as identical than that of SSR amplicons. It clearly emerged from the cluster analysis based on RAPD and SSR markers that the chewing tobacco grown in two states was distinctly different and genetically divergent. The clustering pattern of the cultivars corresponded more with the SSR markers than RAPD markers as it evident from the formation of sub-cluster 1a with the varieties 'Maragadam' and 'Meenakshi', which had a common parent (line PV 7) and the sub-cluster 1b with varieties 'Abirami', 'Bhagyalakshmi' and 'Vairam' with background of common ancestry (line I 64). Similarly the varieties 'Vaishali Special', 'Sona', 'DP 401' and 'Lichchavi' were grouped together (sub-cluster 2a) where 'DP 401' was one of the parents to the variety 'Sona' and in turn 'Sona' was one of the parents to varieties 'Vaishali Special' and 'Lichchavi'.

Both RAPD and SSR markers were efficient in detecting polymorphism among the chewing tobacco cultivars showing their utility in validation/characterization of chewing tobacco cultivars. However, SSR markers were more efficient than RAPD markers in differentiating chewing tobacco cultivars. However, limited genetic diversity among chewing tobacco varieties warrants enhanced use of genetic resources in future breeding programmes to diversify the genetic base of newly released cultivars.

REFERENCES

Barandalla L, Ruiz de Galarreta, Rios D and Ritter E. 2006. Molecular analysis of local potato cultivars from Tenerife Island using microsatellite markers. *Euphytica* **152**: 283–91.

Behera T K, Sharma P, Singh B K, Kumar G, Kumar R, Mohapatra T and Singh N K. 2006. Assessment of genetic diversity and species relationship in eggplant (*Solanum melongena* L.) using STMS markers. *Scientia Horticulture* **107**: 352–7.

Del Piano L, Abet M, Sorrenlino C, Acanfora F, Cozzolino E, Acanfora F, Cozzolino E and DiMuro A. 2000. Genetic variability in *Nicotiana tabacum* and *Nicotiana* species as revealed by RAPD procedure. *Beitrege Zur Tabac International* **19**: 1–15.

Doyle J J AND Doyle J L. 1990. Isolation of plant DNA from fresh tissue. *Focus* **12**: 13–5.

Gopalachari N C. 1984. *Tobacco*. Indian Council of Agricultural Research. New Delhi.

Gupta P K, Balyan H S, Sharma P C and Ramesh B. 1996. Microsatellites in plants: a new class of molecular markers., *Current Science* **70**: 45–54.

Krishnamurthy V and Dev Singh K. 2005. *Non-FCV Tobacco Soils of India, their Fertility and Management*. *Bulletin*, pp 50–7. Central Tobacco Research Institute, Rajahmundry, Andhra Pradesh.

Nielsen G. 1985. The use of isozymes as a probe to identify and lable plant varieties and cultivars. (*in*) *Isozymes: Current Topics in Biological and Medical Research*, Vol 12. pp 1–32. Rattazzi M C, Scandalios J G & Whitt G S (Eds). Alan R. Liss, New York.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalski A. 1996. The comparision of RFLP, RAPD, AFLP and SSR markers for germplasm analysis. *Molecular Breeding* **2**: 225–38.

Ren N and Timko M P.2001. AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild *Nicotiana* species. *Genome* **44**: 559–71.

Rohlf F J. 1998. NTSYS-pc: Numerical taxonomy and multivariate analysis system. Version 2.02, Exter Software, Setauket, New York.

Sheetal Yadav, Poonum Rana, Navinder Saini, Sunita Jain and Rajinder Jain. 2008. Assessment of genetic diversity among rice genotypes with different adaptations to salinity using physio-morphological and molecular markers. *Journal of Plant Biochemistry and Biotechnology* **17**: 1–8.

Siva Raju K, Madhav M S, Sharma R K, Murthy T G K, Singh N K, Bansal K C, Koundal K R and Mohapatra T. 2008. Molecular diversity in Indian tobacco types as revealed by Randomly Amplified DNA polymorphism. *Journal of Plant Biochemistry and Biotechnology* **17**: 51–6.

Williams J G K, Kubelik A R, Livak K J, Rafalski J A and Tingey S V.1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**: 6531–35.