Tobacco Research 40 (1) : 21-27, 2014

GENETIC DIVERSITY IN CHEWING TOBACCO GROWN IN GUJARAT AND WEST BENGAL AS REVEALED BY RAPD AND SSR MARKERS

K. SIVA RAJU AND T.G.K. MURTHY

ICAR-Central Tobacco Research Institute, Rajahmundry- 533 105, Andhra Pradesh, India
(Received on 29th January, 2014 and accepted on 1st April, 2014)

Genetic diversity among released varieties of chewing tobacco belonging to *Nicotiana tabacum* and *Nicotiana rustica* grown in Gujarat and West Bengal states was studied using randomly amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) markers. Seventeen out of 19 RAPD primers amplified 82 fragments of which 70 (85.35%) were polymorphic. Twenty SSR primers amplified 81 fragments of which 46 (56.79%) were polymorphic. The extent of pair-wise similarity ranged from 0.29 to 0.93 with an average of 0.61 in RAPD, in contrast to a range of 0.20 to 0.91 with a mean of 0.55 in SSR markers. The formation of two clusters in dendrogram by both the markers was similar where the three varieties of *N. tabacum* were grouped togather in cluster 1 and the 7 varieties of *N. rustica* were grouped in cluster 2 based on species specificity. The cluster formation was based on species and sub-cluster formation was based on parentage. The genetic variation in chewing tobacco varieties of *N. tabacum* grown in Gujarat was very narrow. In *N. rustica* cultivars, the genetic distance was 12 to 52% and 11 to 54% in RAPD and SSR markers respectively indicating a large genetic variability. SSR markers were found more efficient than RAPD in parentage estimation of chewing tobacco cultivars and could be used for development of markers for various traits.

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

INTRODUCTION

The genus *Nicotiana* is a member of the family *Solanaceae*. Out of 72 recognized species (Goodspeed, 1954; Narayan, 1987), only two species *viz.*, *N. tabacum* and *N. rustica* are grown commercially in India and other countries. Different types of tobaccos grown commercially are defined to a large extent by region of production, method of curing and intended use in the manufacture as well as some distinct morphological characters and chemical differences. Hence, breeding and selection of tobacco varieties is location specific right from the choice of parents to the final stages of evaluation and selection.

In Gujarat and West Bengal, the cultivated chewing tobacco varieties belong to both *N. tabacum* and *N. rustica* species. The *rustica* tobacco is short day and low temperature loving crop largely grown in the states of West Bengal, Bihar, Uttar Pradesh, Haryana, Punjab and Gujarat for hookah, chewing and snuff purposes. Due to its high nicotine content, many times, it is also blended with *bidi* and cigarette tobaccos to improve their smoking strength. *Rustica* tobacco, known as calcutti tobacco in Gujarat, the produce of which either lamina flakes (*bhuka* or *chura*) or leaf bundles (chopadia) are mainly used for chewing, hookah, snuff and blending with *bidi* tobacco, is grown in north and middle Gujarat as *rabi* crop. The northern part of West Bengal occupies unique place in cultivation of good quality of *Motihari* tobacco which belongs to *N. rustica*. *Motihari* tobacco is famous for its quality and strong flavour and this tobacco is used for hookah as well as chewing purposes. It occupies 60% of the tobacco growing area in the North Bengal region including Cooch Behar, Jalpaiguri, West Dinajpur, Malda and Murshidabad districts. Though the varietal diversification is apparent among commercial cultivars, an assessment of their genetic diversity is lacking.

Information regarding DNA polymorphism is more elaborative than polymorphism at phenotypic level to properly explain the genetic diversity and to find out more diverse genotypes for crossing programme. DNA fingerprinting in different tobacco types has been done through RAPD and AFLP markers (Ren and Timko, 2001; Siva Raju *et al.*, 2008). Among different molecular markers, SSRs are more abundant, ubiquitous in

presence, hyper variable in nature and have high polymorphic information (Gupta *et al.,* 1996). These markers have been used for genotype identification in many crop plants such as soybean (Song *et al.*, 1999), potato (Barandalla *et al.,* 2006), brinjal (Behera *et al.*, 2006) and rice (Sheetal Yadav *et al.*, 2008), chewing tobacco (Siva Raju, 2011) among several others. Thus these markers are ideal for discriminating individuals and for parentage determinations. Systematic information on the existing diversity pattern in chewing tobacco varieties in Gujarat and West Bengal, which is crucial to define future breeding strategy in this important crop, is not available. Characterization of genotypes is imperative for best utilization in breeding programs aimed at development of varieties with various required traits. Also in the changing context of agriculture and various kinds of rights of plant material, it is necessary to characterize the indigenously developed plant material at molecular level. This paper reports the genetic diversity based on RAPD and SSR analysis among 10 chewing tobacco varieties belonging to *N. tabacum* and *N. rustica* species.

MATERIALS AND METHODS

Ten genotypes of chewing tobacco belonging *to N. tabacum* and *N. rustica* species (Table 1) comprising released varieties were subjected to RAPD and SSR analysis. This collection represents different commercial chewing tobacco varieties defined by their morphological and biochemical characteristics. Thirty-day-old seedlings of each variety were collected, bulked, frozen in liquid nitrogen, and stored at -80^{0}C until used for DNA extraction. DNA was extracted from pulverized frozen seedlings as per the modified protocol by Doyle and Doyle (1990). The DNA concentration was estimated by comparing with known concentration of lambda DNA as standard in agarose gel electrophoresis (0.8% gel containing 0.5 µg/ml ethidium bromide).

Polymerase chain reaction

A total of 20 decamer primers belonging to different series (Operon Technologies Inc, Alameda, CA, USA) were screened for RAPD analysis. Of these, 17 primers (OPP 01, OPP 03,

Table 1: Chewing tobacco varieties of of *N. tabacum* and *N. rustica* sp.

S. no.	Variety	Species	Parentage	Intended use	Growing state
1	GT-6	*N. tabacum*	783-51-1-13-618 x Smyrna	Chewing & Lal and Kala chopadia	Gujarat
2	GT-8	*N. tabacum*	Improved version of Dharmaj a local selection	Chewing and Kala chopadia	Gujarat
3	Anand-145	*N. tabacum*	Local variety	Chewing	Gujarat
4	Coker-1	*N. rustica*	Exotic introduction	Chewing	Gujarat
5	GC-1	*N. rustica*	Local selection	Chewing	Gujarat
6	GCT-2	*N. rustica*	GC1 x Coker 1	Chewing	Gujarat
7	GCT-3	*N. rustica*	Selection from GCT-2	Chewing	Gujarat
8	DD-437	*N. rustica*	Local selection from Hemti bulk	Hookah and Chewing	West Bengal
9	HD 65-40	*N. rustica*	BVM 5 x Selection 47-33	Hookah and Chewing	West Bengal
10	Dharla	*N. rustica*	DD 437 x C 304	Hookah and Chewing	West Bengal

OPP 04, OPP 09, OPP 10, OPP 12, OPB 01, OPB 10, OPB 14, OPC 04, OPC 08, OPC14, OPC 18, OPL 02, OPL 14, OPAB 8, and OPAB 18) 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 ml 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, M.J. Biotech, USA) programmed to run the following thermal profile: initial denaturation at 94°C for 5 min, 40 cycles of 94°C x 1 min, 37°C x 1 min, 72°C x 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, U.K). The size of the fragments was estimated using Gene Ruler 100 bp DNA ladder plus (MBI Fermentas, Lithuania) marker.

Eleven SSR primers out of 18 were used in the present study based on reproducibility and their NCBI accession numbers are; DQ865412, DQ865416, DQ865427, DQ865430, DQ865431, DQ865432, DQ865433, DQ865436, DQ865437, DQ865441, DQ865442, DQ865444, DQ865445, DQ865446, DQ865450, DQ865459, EF375960, EF375967, EF375968 and EF375969. Amplifications were carried out in a 25 micro liter 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 X 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 Gene Ruler as in RAPD.

Data analysis

All the genotypes were scored for presence or absence of the RAPD, SSR and isoenzyme bands. Only intense bands and those showing repeated amplification 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 'Microsoft 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.

RESULTS AND DISCUSSION

Out of 19 primers used in RAPD to characterize chewing tobacco varieties belonging to *N. tabacum* and *N. rustica* grown in Gujarat and West Bengal, reproducible patterns were obtained with 17 primers, which generated a total of 82 highly reproducible bands in the size ranged from 0.206 to 2.64 kb. The total number of fragments amplified per primer ranged from 2 to 14 with an average of 4.82 (Table 2). Sixteen primers (94%) were found to be polymorphic while the primer OPAB8 was monomorphic. The level of polymorphism detected with different primers ranged from 0 to 100%. Out of all the 82 amplified fragments, 70 (85%) were polymorphic. The observed level of polymorphism was very much higher than that reported by Del Piano *et al.* (2000), Yang *et al.* (2006) and Siva Raju *et al.* (2008; 2012), as two species were involved in the present study and also may be due to difference in genetic architecture. Besides, the primers used in the present study being different in sequence, possibly assayed different genome regions. A total of 12 bands were common to all the cultivars indicating the conserved nature in genome among tobacco genotypes belonging to two species. The primer OPAB8 produced fragments which are common to all the cultivars. The amplified fragment with size 1500 bp by the primer OPC 14 was specific to the variety HD65-40 whereas the band with 370 bp by primer OPC 4 was specific to the variety Anand-145. The combined profiles based on the polymorphic primers, however, provide variety specific patterns and thus could

be distinguished some of the varieties. Photograph of the gel by RAPD primer OPC14 was given in Fig. 1.

Tobacco specific SSR marker analysis was carried out using 25 primers, out of which 20 primers were chosen for analysis based on reproducibility. The number of fragments amplified with these primes ranged from 1 to 7 with an average of 4.05. Out of 81 fragments amplified by 20 primers, 46 (56.79%) were polymorphic (Table 3). The primers DQ865441 and DQ865459 produced one monomorphic band each in all the varieties whereas all other primers DQ865412, DQ865427, DQ865430, DQ865433 and DQ865437 gave the highest level of polymorphism (100%). 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

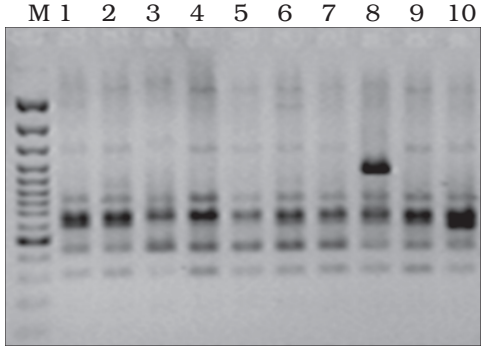

Figure 1: Gel photograph of RAPD amplification by OPC14. 1. GT-6, 2. GT-8, 3. Anand-145, 4. Cocker-1, 5. GC-1, 6. GCT-2, 7. DD437, 8. HD65-40, 9. GCT-3, 10. Dharla and M= DNA marker

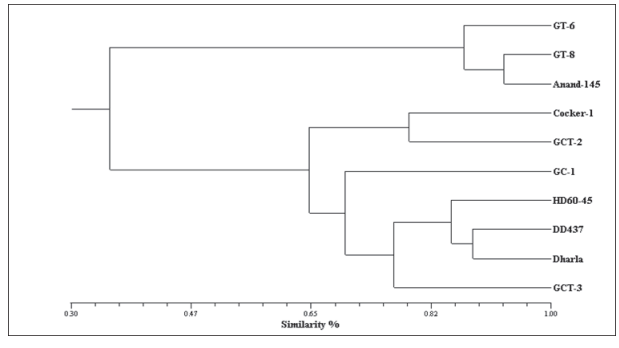

Figure 2: Dendrogram based on all the RAPD markers revealing genetic relationship among the varieties.

Table 2: Data on RAPD amplification

Primer	No. of bands	No. of polymorphic bands	Polymo-rphism (%)	Specific markers
OPP-1	4	4	100	
OPP-3	6	6	100	
OPP-4	2	1	50	
OPP-9	3	2	75	
OPP-10	2	1	50	
OPP-12	4	4	100	
OPB-1	5	5	100	
OPB-10	6	6	100	
OPB-14	5	5	100	
OPC-4	7	5	71	700 bp specific Anad-145
OPC-8	6	5	83	
OPC-14	7	5	71	1500 bp specific to HD65-40
OPC-18	14	13	92	1500 bp specific to Dharla, 2000 bp specific to GCT-3
OPL-2	3	3	100	
OPL-14	3	3	100	
OPAB-8	2	0	0	
OPAB-18	3	2	66	

collected from different countries based on inter simple sequence repeats (ISSR) and inter retrotransposon amplification polymorphism (IRAP) markers. The size of the amplified fragment varied between 185 to 600 bp. 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. Photograph of the gel by SSR primer DQ865412 is given in Fig. 3.

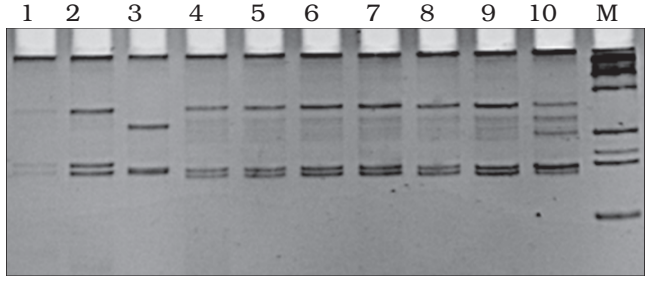

Figure 3: Gel photograph of amplification by SSR primer DQ865412. 1. GT-6, 2. GT-8, 3. Anand-145, 4. Cocker-1, 5. GC-1, 6. GCT-2, 7. DD437, 8. HD65-40, 9. GCT-3, 10. Dharla and M = DNA marker

Genetic relationship among the chewing tobacco genotypes

The genetic relationship among chewing tobacco varieties belonging to *N. tabacum* and *N. rustica* grown in Gujarat and West Bengal was established separately with the two DNA marker systems. In case of RAPD, the Jaccard's coefficient of similarity for all varieties, pair-wise comparisons ranged from 0.29 to 0.93 with an average of 0.61 (Fig. 2). Between the species, the variety GT-8 of *N. tabacum* and GCT-2 of *N. rustica* showed minimum genetic similarity (29 %) whereas the varieties GT-8 and GT-6 of *N. tabacum* showed maximum genetic similarity of 41% with the variety GCT-3 of *N. rustica*. Among the varieties of *N. tabacum*, the genetic similarity varied between 84 and 93% while it ranged form 48 to 88% in the

Table 3: Data on SSR amplification

Primer	No. of bands	No. of polymorphic bands	Polymorphism (%)
DQ865412	7	7	100
DQ865416	6	3	50
DQ865427	3	3	100
DQ865430	5	5	100
DQ865431	6	4	66
DQ865432	3	1	33
DQ865433	4	4	100
DQ865436	3	2	66
DQ865437	3	3	100
DQ865441	1	0	0
DQ865442	4	1	25
DQ865444	4	0	0
DQ865445	5	3	60
DQ865446	5	4	80
DQ865450	2	0	0
DQ865459	1	0	0
EF375960	4	1	25
EF375967	5	2	40
EF375968	5	0	0
EF375969	5	3	60

varieties of *N. rustica*. In the dendrogram constructed based on RAPD (Fig. 3), the 10 genotypes were divided into two main clusters (1 and 2). The 3 varieties of *N. tabacum* were grouped in the cluster 1 where, the variety GT-8 showed 93% of genetic similarity with the variety Anand -145. The second cluster was formed with 7 varieties of *N. rustica* with two sub-clusters. The sub-cluster 2a was formed by Cocker-1 and GCT-2 where, Cocker 1 was one of the parents to GCT-2 and showed 79% of genetic similarity. In the sub-cluster 2b, the rustica tobacco varieties grown in West Bengal were grouped together along with GCT-3 and GC-1 which are grown in Gujarat and the genetic similarity among the varieties grown in West Bengal varied between 81 and 88%. The variety DD 437 was one of the parents to the variety Dharla and grouped together with 88% of genetic similarity. The varieties GCT-3 and GCT-2, which are grown in Gujarat were linked to the group. Thus the clustering pattern was based on species specificity and sub-clustery pattern was based on parentage.

Based on the SSR profiles, the pair-wise similarity coefficients ranged from 0.20 to 0.91 per cent with an average 0.55. Between the species, the variety HD 65-40 of *N. rustica* and the variety Anand-145 showed minimum genetic similarity whereas the variety Anand-145 of *N. tabacum* and Cocker-1 of *N. rustica* showed a maximum genetic similarity of 34%. Among the cultivars of *N. tabacum,* the genetic similarity varied between 88 and 91% whereas the genetic similarity among the varieties of *N. rustica* ranged from 46 to 89% with an average genetic similarity of 67.5%. The average genetic similarity obtained in the present study was similar to the average genetic similarity in chewing tobacco varieties grown in Tamil Nadu and Bihar (Siva Raju, 2011). In the dendrogram constructed based on SSR (Fig. 4), the 10 genotypes were divided into two main clusters (1 and 2). The 3 varieties of *N. tabacum* formed cluster 1. The second cluster was formed by the 7 varieties of *N. rustica* species with two subclusters and the variety GCT-3 was linked independently to the main cluster 2. The sub-cluster 2a was formed by the varieties Cocker-1, GCT-2 and GC-1 whereas the sub-cluster 2b was formed by the varieties grown in West Bengal. As in the case of RAPD the main clustering pattern

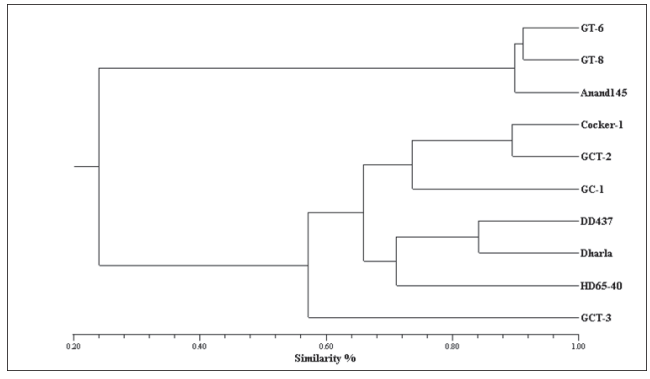

Figure 4: Dendrogram based on SSR markers revealing genetic relationship among the varieties.

was based on species and the sub-clustering pattern was based on parentage.

The average of amplified fragments/primer was higher in RAPD (4.82) than SSR markers (4.02), it could be due to the fact that SSR primers used in this study were tobacco specific. The cluster analysis from RAPD and SSR data produced similar but not identical phylogenetic relationships; it could be due to the fact that probability of scoring of non-homologous RAPD fragments as identical is higher in RAPD than SSR. The formation of two clusters in dendrogram by the both markers was similar where the three varieties of *N. tabacum* were grouped togather in cluster 1 and the 7 varieties of *N. rustica* were grouped in cluster 2 based on species specificity. In RAPD, the sub-cluster 2a was formed by Cocker 1 and GCT-2, where Cocker 1 was one of the parents to GCT-2, but the variety GC-1 which was one of the parents to GCT-2 and GCT-3, which was a selection from GCT-2 were grouped along with the varieties grown in West Bengal. Whereas, in the dendrogram farmed by SSRs, the sub-cluster 2a was formed by GC-1, Cocker 1 and GCT-2 where the first two were parents to GCT-2. The variety GCT-3 was linked to the main cluster 2. Thus the cluster formation was based on species and sub-cluster formation based on parentage. It has clearly emerged from the cluster analysis based on RAPD and SSR markers that the genetic variation in chewing tobacco varieties of *N. tabacum* grown in Gujarat was very narrow. In *N. rustica* cultivars, the genetic distance was 12 to 52% and 11 to 54% in RAPD and SSR markers, respectively indicating that there is a possibility

to select as parents in breeding programme to develop improved chewing varieties of *N. rustica.*

Thus, in general RAPD and SSR markers are efficient in detecting polymorphism even among closely related chewing tobacco varieties of *N. tabacum* and *N. rustica.* However, SSR markers were found to be more efficient than RAPD in estimation of parentage of the chewing tobacco varieties of *N. tabacum* and *N. rustica.*

REFERENCES

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

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

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

Doyle, J.J. and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. **Focus** 12: 13-5.

Goodspeed, T.H. 1954. *The Genus Nicotiana.* Chronica Bot Wallham, Mass. USA.

Gupta, P.K., H.S. Balyan, P.C. Sharma and B. Ramesh. 1996. Microsatellites in plants: A new class of molecular markers. **Curr. Sci.** 70: 45-54.

Narayan, R.K. 1987. Nuclear DNA changes, genome differentiation and evolution in *Nicotiana* (*Solanaceae*). **Plant Syst. Evol**. 157: 161-80.

Ren, N. and M.P. Timko. 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. **J. Plant Biochem. Biotech**. 17: 1-8.

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

Siva Raju, K. 2011. Genetic diversity in Indian chewing tobacco (*Nicotiana tabacum*) as revealed by RAPD and SSR markers. **Indian J. Agric. Sci**. 81: 12-7.

Siva Raju, K., S. Amarnath and V. Krishnamurthy. 2012. Genetic Diversity in *Jati* and *Motihari* tobacco types grown in West Bengal by RAPD markers**. Indian J. Agric. Biochem**. 25: 31-5.

Song, Q.J., C.V. Quigley, R.L. Nelson, T.E. Carter, H.R. Boerma and J.R. Strachan. 1999. A selected set of trinuclotide simple sequence repeat markers for soybean cultivar identification. **Plant Var. Seeds** 12: 207-20.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski and S.V. Tingey. 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Res.**18: 6531-5.

Yang, B.C., B.G. Xiao, X.J. Chen and C.H. Shi. 2006. Assessing the genetic diversity of tobacco germplasm using intersimple sequence repeat and inter-retrotransposon amplification polymorphism markers. **Ann. Appl. Biol.** 150: 393-401.