World Journal of Agricultural Research, 2013, Vol. 1, No. 6, 124-129
Available online at http://pubs.sciepub.com/wjar/1/6/6

DOI:10.12691/wjar-1-6-6

Genetic Variability of *Nicotiana tabacum* (Linn.) Using SSR Marker

Mukund Chandra Thakur[1,*], Sunil Agrawal[1], Mitesh Patel[1], Arif Khan[1], Hiren V. Doshi[2]

[1]Department of Biotechnology, Ashok & Rita Patel Institute of Integrated Study and Research in Biotechnology and Allied Sciences (ARIBAS), New Vallabh Vidyanagar, Gujarat, India
[2]Department of Pharmaceutical Chemistry, Ashok & Rita Patel Institute of Integrated Study and Research, New Vallabh Vidyanagar, Gujarat, India
*Corresponding author: thakurmukund@rocketmail.com

Received December 31, 2012; Revised October 20, 2013; Accepted November 14, 2013

Abstract In present study a total ten varieties of tobacco seeds Gujarat Tobacco-9 (GT-9), Mosaic Resistance Gujarat Tobacco Hybrid – 1 (MRGTH-1), Gujarat Tobacco-7 (GT-7), Gujarat Tobacco-4 (GT-4), Anand-2 (A-2), Anand BT-10 (ABT-10), Anand BD-101 (ABD-101), Anand BD-118 (ABD-118) and Anand-119 (A-119) were procured and grown on simple agar medium. The plants were harvested for further analysis after 5 to 7 days incubation in dark condition. SSR primers tested in present investigation produced fragments of different length. The minimum (150bp) sized fragment and the maximum 600 bp sized amplified fragment was obtained by primer SSR4 & SSR8 respectively. The highest (100%) polymorphism was exhibited by primer SSR1, SSR2, SSR3, SSR5, SSR6, SSR7, SSR9, and SSR10, while the lowest polymorphism (82.30%) was observed with primer SSR4. The maximum scorable bands (34) were generated by primer SSR5, SSR9, & SSR10, whereas the minimum scorable bands (5) generated by primer SSR2. In the SSR analysis, 10 primers were used for polymorphism screening. A total of 18 bands in the size range of 200 to 600 bp were produced by examining across genotypes with 10 SSR primers, with an average polymorphism of 88.8 %. The overall gene frequency analysis reveals a maximum of 1 and minimum of 0.2000. Observed number of alleles (1.3333 ± 0.4830), Effective number of alleles (1.2504 ± 0.3906), Gene diversity (0.1390 ± 0.2076) and Shannon's Information index (0.20160 ± 0.2970) was found in the Genic variation statistics for all loci. The total number of polymorphic loci was found 7 and the percentage of polymorphic loci was 33.33 and the PIC value was found in between 0.980 – 0.466. Phylogenetic trees among 10 tobacco varieties were constructed and 65% genetic similarity was found in the cluster diagram with 2 major groups. This first sub group was divided in two groups where in one group, GT-5 100% was similar with A-2. In second group, MRGTH-1 & GT-9 was 100% similar with ABD-101. Second major group was divided into two sub groups where ABD-118 was similar with A-119 & GT-4. This two tobacco varieties A-119 & GT-4 have 100% similarity. First major group was divided into four subgroups and the second major group again divided into two subgroups at 76.5% genetic similarity which contained ABD-118 and A-119 & GT-4 with 85% genetic similarity. Now, in the first major group the first subgroup was divided at 72% genetic similarity into two another subgroup first subgroups had GT-5 and A-2. The second subgroup was divided into two as MRGTH-1, ABD-101 & ABT-10 and another subgroup had GT-7. GT-9 showed maximum genetic similarity (1.0000) and A-119 showed minimum genetic similarity (0.6667). ABD-118 and GT-4 varieties showed maximum genetic distance of 0.2719 while GT-9 showed no genetic distance (0.0000). Based on study, the large range of similarity values for related cultivars using microsatellites provided greater confidence for the assessment of genetic diversity and relationships. This information (genetic variability at molecular level) will certainly be helpful to identify and develop genetically unique germplasm that compliments the existing cultivars. Further practical approaches are required to unveil the differences among these closely related species of the tobacco varieties studied in the present investigation.

Keywords: *genetic diversity, Nicotiana tabacum, SSR marker, genetic diversity, cluster analysis*

Cite This Article: Mukund Chandra Thakur, Sunil Agrawal, Mitesh Patel, Arif Khan, and Hiren V. Doshi, "Genetic Variability of *Nicotiana tabacum* (Linn.) Using SSR Marker." *World Journal of Agricultural Research* 1, no. 6 (2013): 124-129. doi: 10.12691/wjar-1-6-6.

1. Introduction

Tobacco is an agricultural product processed from the leaves of plants in the genus *Nicotiana*. Tobacco plants are also used in plant bioengineering, and some of the 60 species are grown as ornamentals. It is a very important economic crop worldwide. It is most commonly used as a recreational drug, and is a valuable cash crop for countries such as Cuba, China and the United States. *N. rustica*, a mild-flavored, fast-burning species, was the tobacco originally raised in Virginia, but it is now grown chiefly in Turkey, India, and Russia. Approximately 4916 million kg

of tobacco, *Nicotiana tabacum* L. (2n = 48), were produced worldwide in 2007 [1], making it one of the most economically important nonfood cultigens. There are more than 70 species of tobacco are used. Tobacco breeding programs in the United States have made significant advancements over the last 50 yr, especially for burley and flue-cured tobacco, by crossing existing cultivars and elite breeding lines within specific market classes [2]. Tobacco lines that have been derived from crosses between lines from different market classes have tended to exhibit undesirable quality & characteristics, and thus such breeding strategies have typically been avoided [3,4]. The U.S. Regional Minimum Standards Program, which introduced stringent industry requirements for numerous quality attributes of cured tobacco leaf [5], has only reinforced the practice of advanced cycle pedigree breeding. Murphy et al. [6] found a considerably large value for the average coefficient of parentage, 0.41, among historically important U.S. flue-cured cultivars dating back the early 1930s. In general, exotic germplasm has not been utilized for the development of improved tobacco cultivars other than for the incorporation of a few relatively simply inherited disease resistance traits [7]. Such *N. tabacum* germplasm may harbor, however, desirable allelic variation of possible value for modification of cured leaf chemistry, agronomic characteristics, or even the development of novel tobacco market types. *N. tabacum* likely evolved on the eastern slopes of the Andes Mountains in Bolivia or northern Argentina due to a chance hybridization event(s) between *N. sylvestris* and a member of section *Tomentosae* followed by spontaneous chromosome doubling [8].

Although a tremendous amount of phenotypic variability exists within the species as a whole [9], little is known about the DNA variation underlying this diversity. A relatively low level of genetic diversity has been revealed by molecular marker systems such as amplified fragment length polymorphisms (AFLPs; [10] and randomly amplified polymorphic DNA (RAPD; [11]). Nonetheless, some studies have been able to link markers to loci impacting disease resistance [12]. The development of microsatellite, or simple sequence repeat (SSR), markers for tobacco has offered increased possibilities for characterizing and utilizing genetic variation in *N. tabacum* [13]. Simple sequence repeat markers have become increasingly popular for molecular marker studies because they are highly polymorphic, co-dominant, and highly reproducible [14,15]. They have been used extensively to study genetic diversity within the germplasm pools of other cultivated crops [16,17,18,19].

Therefore, competitive PCR was performed using a pair of simple sequence repeat (SSR) primers. The copy number of the exogenous gene in transgenic tobacco can be determined by using software according to the intensities of endogenous and exogenous bands in ethidium bromide-stained gels.

2. Materials & Methods

2.1. Collection of Seed Samples

Seeds of 10 different varieties of tobacco were obtained from the germplasm collection centre of "Agricultural Tobacco Breeding Research Centre, Anand Agricultural University, Anand". The varieties with desirable agronomic characteristics, such as large leaf size, high leaf yield, low nicotine content, or resistance to various diseases or insects [20], the optimal characterize to evaluate the different varieties. Those varieties were Gujarat Tobacco-9 (GT-9), Mosaic Resistance Gujarat Tobacco Hybrid – 1 (MRGTH-1), Gujarat Tobacco-7 (GT-7), Gujarat Tobacco-4 (GT-4), Anand-2 (A-2), Anand BT-10 (ABT-10), Anand BD-101 (ABD-101), Anand BD-118 (ABD-118) and Anand-119 (A-119).

2.2. Surface Sterilization of Tobacco Seeds

All operations for surface sterilization of seeds were performed at room temperature. Rice seeds obtained from the germplasm centre were gently washed with sterile distilled water and treated with 70 % ethanol for 10 min. Subsequently, the seeds were immediately washed with sterile distilled water (3 x 10 min.) and incubated in 0.1 % mercuric chloride solution for 4 min. Following this treatment, the seeds were washed repeatedly in excess amounts of sterile distilled water for 5 – 6h on a shaker before seeding them in petridishes containing tryptone glucose yeast extract agar medium and incubated in the dark at 30°C to test the possible contamination. Using this method, more than 98 % of the seeds germinated, and 95 % of these seedlings were found to be contamination-free. The seedlings devoid of any contamination were used for DNA isolation.

For each variety, leaves from a pool of 6 plants were combined for DNA extraction. DNA Quantification was carried out by agarose gel analysis using uncut λ phage DNA (50 ng/ µl) and Nanodrop spectrophotometer Amplification of Extracted Genomic DNA by Polymerase Chain Reaction (PCR). DNA amplification is a very simple method for in vitro amplification of specific nucleic acids using Taq DNA polymerase and minimum two oligonucleotides specific to the DNA to be amplified. The concentration of isolated genomic DNA was determined by the formula N (µg/ml) = 70260/280 – 40A260/280 [21].

2.3. PCR Amplification of 16S rDNA

16S rDNA of two isolates were amplified by PCR (NyxTechnique, Master cycler gradient; Model No. – AG 22331, Hanburg, Germany) using the SSR primer (Table 1) which were purchased from SIGMA , U.K.

2.4. PCR condition and Amplification

A master mix was prepared where the constituents common to all the reactions were combined in one tube, multiplying the volume for one reaction for total number of samples. Later, the appropriate amount of master mix was dispensed to each tube and template DNA was added separately in each tube. The genomic DNA was amplified using SSR primers, which also depicts the details like primer sequence.

PCR was performed in 25 µl volumes containing approximately 25–50 ng of template DNA, 1x PCR buffer (50 mM KCl, 10 mM Tris–HCl pH 8.3), 1.5 mM $MgCl_2$, 0.15 mM of each primer, 0.25 mM dNTPs and 0.3U of Taq polymerase (New England Biolab, Fermentas),

Thermocycling started with a denaturation step for 5 min at 94°C followed by 45 cycles of 1 min at 94°C, 45 sec at 55°C, 1 min at 72°C and stopped after a final extension step of 72°C for 7 min. The amplified products were analyzed electrophoretically using 2% agarose gel (prepared 1x TBE). PCR amplified product (8 µl) and 8 µl gel loading dye. The separated bands were visualized under UV Transilluminator and photographed using Alpha Innotech Gel Documentation system.

Table 1. List of SSR primers

Sr. No	Primers	Primer type	Primer Sequence
1	SSR1	Forward(5'→3')	CCTGTAACATCTGAGTGCC
		Reverse(5'→3')	CCCATAAGTCCAACATGC
2	SSR2	Forward(5'→3')	TCAAAGTCTTTTACATTGACG
		Reverse(5'→3')	GGCAATGTTGTGGCTTAC
3	SSR3	Forward(5'→3')	GCCGAATTAAACCAACCAAA
		Reverse(5'→3')	ACCGGATTGCTTAATTGTCG
4	SSR4	Forward(5'→3')	TGTCTCGTGAAGCATGAA
		Reverse(5'→3')	GGAAATGGAGGATCTCGT
5	SSR5	Forward(5'→3')	CTTCTTCCTAAGCCGAGGGT
		Reverse(5'→3')	TTGATGATAGAACGCAACTCG
6	SSR6	Forward(5'→3')	AACCATACGCCTTCAGATCG
		Reverse(5'→3')	TGGTTTGAGTAAAGAAATGTTGTGA
7	SSR7	Forward(5'→3')	AGGAGGCGAAGAAAGAGGAG
		Reverse(5'→3')	CCCATGAATTCGTAACAGCA
8	SSR8	Forward(5'→3')	CTGCTCCATCATTGCTCAAA
		Reverse(5'→3')	GCAACATATCCGAACCTCCA
9	SSR9	Forward(5'→3')	AAGCTGCCCTAGCTCAATCA
		Reverse(5'→ 3')	AACATCACCATTTCCACAAGTTT
10	SSR10	Forward(5'→3')	TTCCAGTGAACTCTGTGGTGA
		Reverse(5'→3')	CAATATGGCAGCATCTGTAGG

2.5. Data Scoring and Data Analysis

Clear and distinct bands amplified by SSR primers study were identified by the software to find number of allelic frequency, standard deviation, expected heterozygosity and observed heterozygosity. Each amplified product was scored across all genotypes for its presence or absence. The scores 1 & 0 indicates the presence & absence of bands, respectively. The data were entered to in binary matrix and subsequently analyzed using NTSYSpc version 2.02 [22] PopGene32 and MEGA4 software. Coefficients of similarity were calculated by using simple matching coefficient (GDSM), and Jaccard's coefficient (GDJ) by SIMQUAL function. The matrix of similarity was clustered using UPGMA algorithm under Sequential Agglomerative hieraachical Nesting (SAHN) module of the NTSYSpc. Relationships between the *Nicotina tabacum* cultivars were portrayed graphically as dendrograms. Grouping of species were also evaluated by principle coordinate analysis (PCA) as reported by Thomas et al. [23]. PCA was performed by extracting Eigen value and Eigen vectors from a correlation matrix which was generated using a standardized data matrix 2-D and 3-D plots were constructed to evaluate the groupings of *N. tabacum*.

The cophenetic correlation analysis was carried out using the COPH function of NTSYS-pc. In this method the dendrogram and similarity matrix were correlated to find the goodness of fit of the dendrogram constructed based on the similarity coefficients. The comparison between the SSR based on similarity coefficient matrices were tested based on cophenetic correlation analysis and mantel matrix correspondence test. The mantel matrix correspondence test was also carried out.

Amplified products from microsatellite analyses were scored qualitatively for the presence and absence of each marker allele genotype combination. Each SSR band amplified by a given primer was treated as a unit character. Data were entered into a binary matrix as discrete variables, that is, 1 for presence and 0 for absence of the character. Gene diversity, also referred to as expected heterozygosity, of each marker locus was calculated by PIC value. The PIC value of a marker was calculated according to the following formula:

$$\underset{i}{\mathrm{PIC}} = 1 - \sum_{j=1}^{n} P^2_{ij}$$

Where Pij is the frequency of jth allele for the ith marker [24], and summed over n alleles. Average number of alleles, average PIC value and average genetic similarity were computed on the basis of different pear accessions/cultivars and microsatellite classes. Pair-wise comparisons of the ecotypes based on the proportion of unique and shared amplification alleles were used to measure the genetic similarity by Dice coefficients using Simqual sub-program in a similarity routine of NTSYS-pc software package version 2.2 [22]. Estimates of genetic similarity (F) were calculated between all pairs of the genotypes according to Nei and Li [25] based on the following formula:

$$\textbf{Similarity}\,(\mathbf{F}) = \mathbf{2Nab} / (\mathbf{Na} + \mathbf{Nb})$$

Where Na = the total number of fragments detected in individual 'a', Nb = the total number of fragments shown by individual 'b' and Nab = the number of fragments shared by individuals 'a' and 'b'.

The resultant similarity matrix data were employed to construct a dendrogram using Sequential Agglomerative Hierarchical Nesting (SAHN) based on unweighted pair-group method with an arithmetic average (UPGMA) to infer genetic relationships and phylogeny among the accessions.

3. Results and Discussion

A Total ten varieties of tobacco seeds Gujarat Tobacco-9 (GT-9), Mosaic Resistance Gujarat Tobacco Hybrid – 1 (MRGTH-1), Gujarat Tobacco-7 (GT-7), Gujarat Tobacco-4 (GT-4), Anand-2 (A-2), Anand BT-10 (ABT-10), Anand BD-101 (ABD-101), Anand BD-118 (ABD-118) and Anand-119 (A-119) were selected and grown on simple agar medium. After 5 to 7 days incubation in dark condition and result thus obtained is given in Figure 1a & Figure 1b.

3.1. Simple Sequence Repeats (SSR)

Total 10 pairs of SSR primers were screened for the amplification of the isolated DNA for all ten different varieties of tobacco.

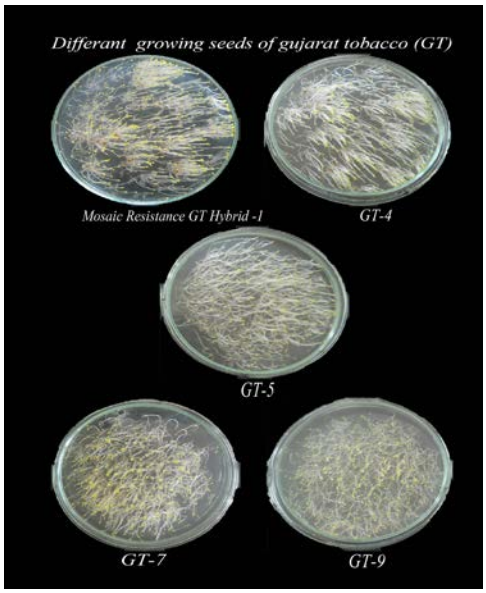

Figure 1.a. Growing seeds of *N. tabacum* (5 Varieties)

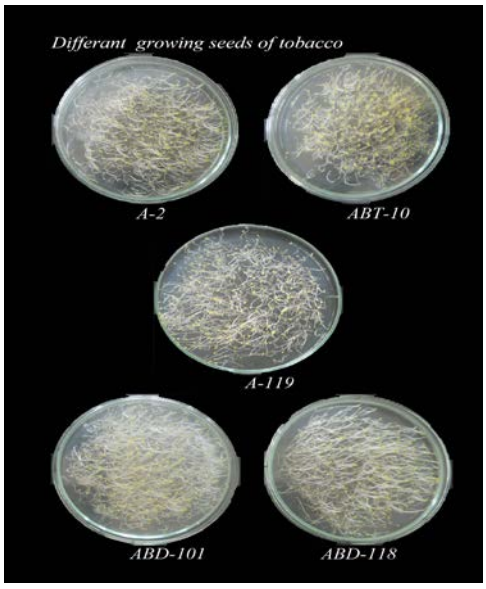

Figure 1.b. Growing seeds of *N. tabacum* (5 Varieties)

SSR primers tested in present investigation produced fragments of different length. The minimum (150bp) sized fragment was amplified by primer SSR4, whereas maximum 540bp and 600 bp sized fragment was amplified by primer SSR5 & SSR8 respectively. The highest (100%) polymorphism was exhibited by primer SSR1, SSR2, SSR3, SSR5, SSR6, SSR7, SSR9, and SSR10, while the lowest polymorphism (82.30%) was observed with primer SSR4 and SSR 8 (95.30%). The maximum scorable bands (34) were generated by primer SSR5, SSR9, & SSR10, whereas the minimum scorable bands (5) generated by primer SSR2. In the SSR analysis, 10 primers were used for polymorphism screening. A total of 18 bands in the size range of 200 to 600 bp were produced by examining across genotypes with 10 SSR primers, with an average polymorphism of 88.8 %. The PIC value was found in between 0.980 – 0.466

3.1.1. Clustering Analysis

The SSR bands were scored as 1 for present or 0 for absent across the genotypes and only those that were well defined and consistently repeatable in two independent amplifications were included in the final analysis. All clear and intense bands were scored for the construction of the data matrix.

The data were scored in an Excel sheet and was converted manually in a text format for PopGene software. POPGEN version 32 was used to calculate all genetic parameters. Variety-specific SSR markers were scored on the basis of gene frequency. The genetic relationship between the populations was determined by calculating genetic distances (GD) and genetic identities (GI) for all possible population pairs as per Nei's coefficient [26]. Cluster analysis was performed using the unweighted pair group method with arithmetic averages (UPGMA). Dendrograms were constructed using the UPGMA algorithms in the MEGA 4.0 software [27]. The result of the POPGEN software analysis is given in Figure 2. Dendogram is based on Nei's [28] gemetic distance method: = UPGMA and modified from NEIGHBOR procedure of PHYLIP Version 3.5 (Dendogram of SSR).

Phylogenetic trees among 10 plant varieties were constructed based on Nei's coefficient [26] standard genetic distances using UPGMA (Unweighted Pair Group Method with Arithmetic Mean) method for SSR analysis at 65% genetic similarity the cluster diagram indicated 2 major groups (dendogram). This first sub group divided in two groups where in one group, GT-5 100% was similar with A-2. In second group, MRGTH-1 & GT-9 was 100% similar with ABD-101. Second major group was divided into two sub groups where ABD-118 was similar with A-119 & GT-4. This two tobacco varieties A-119 & GT-4 have 100% similarity. First major group was divided into four subgroups and the second major group again divided into two subgroups at 76.5% genetic similarity which contained ABD-118 and A-119 & GT-4 with 85% genetic similarity.

Now, in the first major group the first subgroup was divided at 72% genetic similarity into two another subgroup first subgroups had GT-5 and A-2. The second subgroup was divided into two as MRGTH-1, ABD-101 & ABT-10 and another subgroup had GT-7.

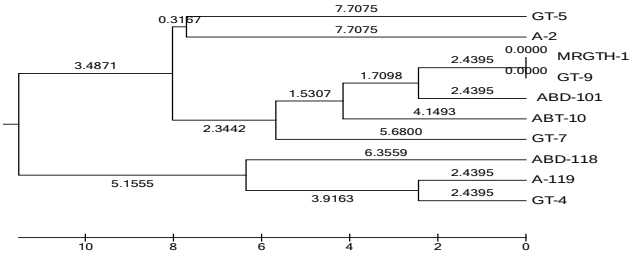

Figure 2. Dendogram based on UPGMA method showing phylogenetic relationship among ten Nicotiana spp. revealed by SSR primers.

3.2. Similarity Analysis

The genetic similarities and genetic distances were calculated for all possible population pairs using SSR data (Table 2) as per Nei's coefficient [26]. This Nei's coefficient is also called Jaccard's coefficient. GT-9 showed maximum genetic similarity (1.0000) and A-119 showed minimum genetic similarity (0.6667). ABD-118 and GT-4 varities showed maximum genetic distance of 0.2719 while GT-9 showed no genetic distance (0.0000).

Table 2. Nei's original measures of genetic similarity and genetic distance for SSR analysis.

Pop ID	GT-5	A-2	MRGTH-1	GT-9	GT-7	ABT-10	ABD-101	ABD-118	A-119	GT-4
GT-5	*****	0.8571	0.9048	0.9048	0.8095	0.8095	0.8571	0.7619	0.8095	0.7619
A-2	0.1542	*****	0.8571	0.8571	0.8571	0.8571	0.8095	0.8095	0.6667	0.7143
MRGTH-1	0.1001	0.1542	*****	1.0000	0.9048	0.9048	0.9524	0.7619	0.8095	0.7619
GT-9	0.1001	0.1542	0.0000	*****	0.9048	0.9048	0.9524	0.7619	0.8095	0.7619
GT-7	0.2113	0.1542	0.1001	0.1001	*****	0.9048	0.8571	0.8571	0.8095	0.8571
ABT-10	0.2113	0.1542	0.1001	0.1001	0.1001	*****	0.9524	0.8571	0.8095	0.8571
ABD-101	0.1542	0.2113	0.0488	0.0488	0.1542	0.0488	*****	0.8095	0.8571	0.8095
ABD-118	0.2719	0.2113	0.2719	0.2719	0.1542	0.1542	0.2113	*****	0.8571	0.9048
A-119	0.2113	0.4055	0.2113	0.2113	0.2113	0.2113	0.1542	0.1542	*****	0.9524
GT-4	0.2719	0.3365	0.2719	0.2719	0.1542	0.1542	0.2113	0.1001	0.0488	*****

3.3. MXPLOT (2D biplots) and 3-MOD (3D plot)

In this study, the larger range of similarity values for cultivars revealed by micro satellite markers provides greater confidence for the assessments of genetic diversity and relationships, which can be used in future breeding programs. Principle component analysis was also done to visualize genetic relationships among the elite breeding lines (Figure 3 & Figure 4). The results were similar to UPGMA results.

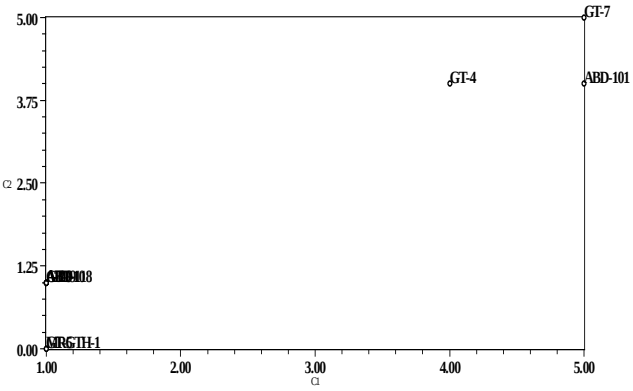

Figure 3. Principle Component Analysis grouping of 10 *Nicotiana* spp based on pooled SSR markers.

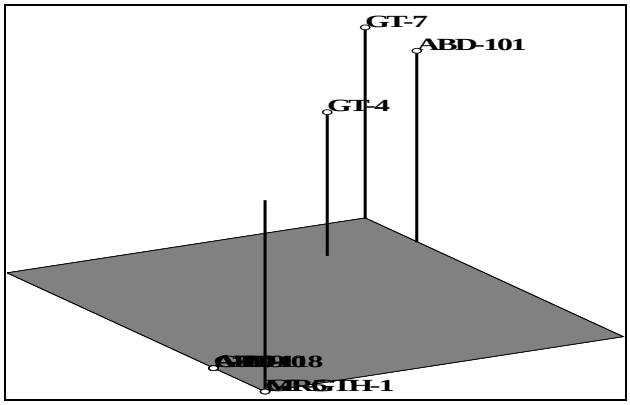

Figure 4. 3D plot showing phylogenetic relationship among 10 Nicotiana spp based on pooled SSR markers.

The reproducibility of microsatellite amplification, allele sizes produced in the current study for the cultivars of *N. tabacum* were in accordance with the study of Bindler et al. [13].

This study revealed that a moderate proportion of *N. tabacum* microsatellite markers can be found among the closely relate diploid relatives to know polymorphism in those species. This polymorphism could have attributed due to the transferability of the microsatellite markers within closely or distantly related species. The microsatellite markers were transferable at a lower rate to more distantly related *Nicotiana* species (such as Gujarat Tobaco – 4 and Gujarat Tobaco-5). Other researchers have also reported on the transferability of microsatellite markers within a genus [5,29,30,31,32,33,34].

RAPD, AFLPs, isozymes, sequence information, plant morphology and cytological observations have been used in past efforts to delineate evolutionary relationships in the genus Nicotiana [9,35,36,37,38,39]. Other authors have also predicted the amount of morphological, biochemical and DNA similarity between *N. tabacum* cultivars and suggested possible inadvertent introgression of the germplasm of this genus during development [35,40,41].

Murad et al. [42] analyzed repetitive lineage of *N. tomentosiformis* contributed to the T gemone of modern-day *N. tabacum*. In our study, average Jaccard similarity coefficients were calculated between the 10 *N. tabacum* cultivars and each of them had highest average Jaccard similarity coefficient (0.6667 to 1.0000). Our results of the Jaccard similarity coefficients are in agreement with the results obtained by Moon et al. [43].

Based on study the large range of similarity values for related cultivars using microsatellites provides greater confidence for the assessment of genetic diversity and relationships. The practical approach developed in the study is useful in DNA fingerprinting also.

Acknowledgement

I am grateful to Gujarat Council of Science & Technology (GUJCOST), Gandhinagar for funding this research work. I would like to extend my heartfelt gratitude towards Charutar Vidya Mandal (CVM) and Dr. Pradip S. Patel, Director, Ashok & Rita Patel Institute of Biotechnology and Research in Allied Sciences (ARIBAS), New Vallabh Vidyanagar, Gujarat, for providing the resources and requirements readily for the completion of this project.

References

[1] Universal Leaf Tobacco Company, Inc. supply & demand report. Available at http://www.universalcorp.com/Reports/Report Frameset.asp?ID=31201&Menu=Tob (verified 13 Oct.2008). Universal Leaf Tobacco Co., Richmond, VA, 2008.

[2] Wernsman, E.A., Rufty, R.C., Tobacco. In: Fehr W (ed). Principles of cultivar development, Macmillan Publ. New York, 1987, 2: 669-698.

[3] Garner, W.W., Allard, H.A. and Clayton, E.E., Superior germ plasm in tobacco. In 1936 yearbook of agriculture. USDA, Washington, DC., 1936, p. 785-830

[4] Clayton, E.E., The genetics and breeding progress in tobacco during the last 50 years. Agron. J., 1958, 50: 352-356.

[5] Bowman, D.T., History of the Regional Minimum Standards Program for the release of fl ue-cured tobacco varieties in the United States. Tob. Sci., 1996, 40: 99-110.

[6] Murphy, J.P., Cox, T.S., Rufty, R.C. and Rodgers, D.M., A representation of the pedigree relationships among fl ue-cured tobacco cultivars. Tob. Sci., 1987, 31: 70-75.

[7] Valleau, W.D.,Breeding tobacco for disease resistance. Econ. Bot., 1952, 6: 69-102.

[8] Lewis, R.S. and Nicholson, J.S., Aspects of the evolution of Nicotiana tabacum L.,and the status of the United States Nicotiana Germplasm Collection. Genet Res Crop Evol., 2007, 54:727-740.

[9] Goodspeed, T.H. Chromosome number and morphology in Nicotana. VII. Karyotypes of fifty-five species in relation to a taxonomic revision of the genus. Univ. Calif. Pub. Bot., 1945, 18: 345-370.

[10] Ren, N. and Timko, M.P., AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild Nicotiana species. Genome, 2001, 4: 559-571.

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

[12] Nishi, T., Tajima, T., Noguchi, S., Ajisaka, H. and Nedishi, H., Identification of DNA markers of tobacco linked to bacterial wilt resistance. Theor. Appl. Genet., 2003, 106: 765-770.

[13] Bindler, G., Vander, H.R., Gunduz, I., Plieske, J., Ganal, M., Rossi, L., Gadani, F. and Donini, P., A microsatellite marker based linkage map of tobacco. Theor. Appl. Genet., 2007, 114: 341-349.

[14] Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., Van der Wiel, C., Bredemeijer, G., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Meastri, E., Malcevschi, A., Marmiroli, N., Aert, R., Volckaert, G., Rueda, J., Linacero, R., Vazquez, A., and Karp, A., Reproducibility testing of RAPD, AFLP, and SSR markers in plants by a network of European laboratories. Mol. Breed., 1997, 3: 381-390.

[15] Milbourne, D., Meyer, R., Bradshaw, J.E., Baird, E., Bonar, N., Provan, J., Powell, W. and Waugh, R., Comparison of PCR-based marker system for the analysis of genetic relationships in cultivated potato. Mol. Breed., 1997, 3: 1277-136.

[16] Huang, X.Q., Borner, A., Roder, M.S. and Ganal, M.W., Assessing genetic diversity of wheat (Triticum aestivum L.) germplasm using microsatellite markers. Theor. Appl. Genet., 2002, 105: 699-707.

[17] Abe, J., Xu, D.H., Suzuki, Y. and Kanazawa, A., Soybean germplasm pools in Asia revealed by nuclear SSRs. Theor. Appl. Genet., 2003, 106: 445-453.

[18] Bertin, P., Gregoire, D., Massart, S. and De Froidmont, D., High level of genetic diversity among spelt germplasm revealed by microsatellite markers. Genome, 2004, 47: 1043-1052.

[19] Jayamani, P., Negrao, S., Martins, M., Macas, B. and Oliveira, M.M., Genetic relatedness of Portuguese rice accessions from diverse origins as assessed by microsatellite markers. Crop Sci., 2007, 47: 879-886.

[20] Rei, Y.H., Xu, M.L. and Huang, X.Y., Flue-cured tobacco. In Rei YH (eds): Tobacco collection in the Yunnan province. Science and Technology publishing company of Yunnan, Yunnan., 1997, pp. 38-42.

[21] Kalb, V.F. and Bernlohr, R.W., A new spectrophotometric assay for protein in cell extracts. Analyt. Biochem., 1977, 82: 362-371.

[22] Rohlf, F., NTSYS-Pc. Numerical taxonomy and multivariate analysis system version 2.02e. Exeter Software. New York, 1997.

[23] Thomas, G., Mohapatra, T., Rao, A. and Sharma, R.P., Distinguishing Indian commercial wheat varieties using RAPD based DNA fingerprints. Indian J Biotech., 2006, 5: 200-206.

[24] Nei, M., Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci., 1973, 70: 3321-3323.

[25] Nie, M. and Li, W.H., Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA, 1979, 76: 5269-5273.

[26] Nei, M., Genetic distance between populations. Am. Nat., 1972, 106: 283-292.

[27] Tamura, K., Dudley, J., Nei, M. and Kumar, S., MEGA 4 Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol., 2007, 24: 1596-1599.

[28] Nei, M., Molecular Evolutionary Genetics. Columbia Univ., New York, 1978.

[29] Van Deynze, A.E., Sorrels, M.E., Park, W.D., Ayres, N.M., Fu, H., Cartinhour, S.W., Paul, E., McCouch, S.R., Anchor probes for comparative mapping of grass genera. Theor. Appl. Genet., 1998, 97: 356-369.

[30] Tikhonov, A.P., SanMiguel, P.J., Nakajima, Y., Gorenstein, N.M., Bennetzen, J.L. and Avramova, Z., Colinearity and its exceptions in orthologous adh regions of maize and sorghum. Proc. Natl. Acad. Sci., U.S.A., 1999, 96: 74409-7414.

[31] Alvarez, A.E., Van der Wiel, C.C.M., Smulders, M.J.M. and Vosman, B., Use of microsatellites to evaluate genetic diversity and species relationships in the genus Lycopersicom. Theor. Appl. Genet., 2001, 103: 1283-1292.

[32] Dirlewanger, E., Cosson, P., Tavaud, M., Aranzana, M.J., Poizat, C., Zanetto, A., Arús P, Laigret, R., Development of microsatellite markers in peach [Prunus persica (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (Prunus avium L.). Theor. Appl. Genet., 2002, 105: 127-138.

[33] Eujayl, I., Sorrels, M.E., Baum, M., Wolters, P. and Powell, W., Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor. Appl. Genet., 2002, 104: 399-407.

[34] Kuleung, C., Baenziger, P.S. and Dweikat, I., Transferability of SSR markers among wheat, rye and triticale. Theor. Appl. Genet., 2004, 108: 11447-1150.

[35] Sheen, S.J., Isozymic evidence bearing on the origin of Nicotiana tabacum L. Evolution., 1972, 26: 143-154.

[36] Bogani, P., Liop, P., Intrieri, C. and Buiatti, M., physiological and molecular analysis of genus Nicotiana. Mol. Phylogenet. Evol., 1997, 7: 62-70.

[37] Aoki, S. and Ito, M., Molecular phylogeny of Nicotiana (Solanaceae) base on the nucleotide sequence of the matK gene. Plant Boil., 2000, 2: 316-324.

[38] Knapp, S., Chase, M.W. and Clarkson, J.J., Nomenclatural changes and a new section classification in Nicotiana (Solanaceae). Taxon., 2004, 53: 73-82.

[39] Bindler, G., Jorg, P., Nicolas, B., Irfan, G., Nikolai, I., Van der Hoeven, R., Martin, G. and Donini, P., A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. Theor. Appl. Genet., 2011, Jul ;123 (2):219-30

[40] Burk,, L.G., Partial self-fertility in a theoretical amphiploid progenitor of N. tabacum. J Hered., 1973, 64: 348-350.

[41] Lim, K.Y., Souckova-Skalicka, K., Sarasan, V., Clarkson, J.J., Chase, M.W., Kovarik, A. and Leitch, A.R., A genetic appraisal of a new synthetic *Nicotiana tabacum* (Solanaceae) and Kostoff synthetic tobacco. Am. J. Bot., 2006, 93: 875-883.

[42] Murad, L., Lim, K.Y., Christopodulou, V., Matyasek, R., Lichtenstein, C.P., Kovarik, A. and Leitch, A.R., The origin of tobacco's T genome is traced to a particular lineage within *Nicotiana tomentosiformis* (Solanaceae). Am. J. Bot., 2002, 89: 921-928.

[43] Moon, H.S., Nicholson, J.S. and Lewis, R.S., Use of transferable Nicotiana tabacum L. microsatellite markers for investigating genetic diversity in the genus Nicotiana. Genome, 2008, 51: 547-559.