GENETIC DIVERSITY IN TOBACCO GENOTYPES EVALUATED UNDER ALL INDIA NETWORK PROJECT

K. Baghyalakshmi[1],K . Sarala[*1], K. Prabhakara Rao[1], D. Damodar Reddy[2], C.Chandrasekhara Rao[3], P. Soni[1], P. Vinay[1], G. Kiran[1] and K. Santinandivelu[1]

[1]Division of Crop Improvement, ICAR-Central Tobacco Research Institute, Rajahmundry-533105,
[2]Director, ICAR-Central Tobacco Research Institute, Rajahmundry-533105
[3]Division of Crop Chemistry & Soil Science, ICAR-Central Tobacco Research Institute, Rajahmundry-533105

Abstract:

Tobacco is high valued commercial crop with a steady increase in demand, which needs to be addressed by improvement in yield and quality for meeting national and international requirements. ICAR- Central Tobacco Research Institute and All India Network Project on Tobacco (AINPT) Centres are breeding various tobacco types for higher yield and better quality. The superior genotypes developed are evaluated in AINPT co-ordinated trials at 4 main, 7 sub and 4 voluntary centres. The significantly superior ones are tested in pre-release bulk trials. Diversity in the lines entered in the AINPT trials is essential for releasing varieties with higher genetic potential and reducing the genetic uniformity in tobacco. A study was conducted to investigate the diversity present among entries that were evaluated under AINPT co-ordinated trials from 2000 to 2015. A total of 100 AINPT lines were used, among which 87 were FCV, 7 Bidi, 4 Rustica and 2 chewing lines. Principal component analysis (PCA) was done to analyze the pattern of data matrix for determining the selection criteria and identification of morphological characters highly responsible for diversity. Based on PCA, 17 characters were selected and analyzed in DarWIN to draw the dendrogram of tobacco genotypes. The entries of FCV, Bidi, Rustica, Chewing fell on four separate clusters. Genetic distance between clusters varies from 0.04 to 2.67. Few Bidi and chewing entries were found to interlay with FCV entries as traits like plant height, plant type, floral morphology, etc., were similar between them. The results also showed that FCV38 has the highest distance from other genotypes followed by ArBD126 with a value of 2.67 and 1.69, respectively. Lower dissimilarity values were observed among genotypes indicating narrow genetic base of breeding programmes. Tobacco breeders, in general, are confined to use less divergent genotypes in order to minimize the disturbance at the genome level. This trend if continued may result in less yield gains in subsequent cultivars and increase vulnerability for biotic and abiotic stress. As bulk of tobacco genetic resources remain unrepresented in modern tobacco cultivars, genetic base need to be broadened through pre-breeding to create variability for further improvement, without quality penalty.

Introduction:

The genus *Nicotiana* is one of the five large genera of family Solanaceae. The genus is represented by about 70 recognized species, out of which 49 are native to North and South America, 20 to Australia and one species has its origin in the African continent (Lewis and Nicholson, 2007). *Nicotiana tabacum* L. and *N. rustica* L. are the two important cultivated species widely grown and consumed in various forms in India. India is the second largest producer of tobacco in the world with an annual production of about 800 million kg of cured leaves. Tobacco is one of the important high value commercial crops grown in an area of 0.433 million hectares over 15 states in India. During 2015-16, tobacco made a significant contribution of Rs. 29,376 crore to Indian economy in terms of excise revenue (Rs. 23,318 crore) and export earnings (Rs. 6058 crore). Major tobacco growing states are AP, Karnataka, Gujarat, UP, Tamil Nadu, Bihar and WB. A unique feature of tobacco production in India is that different styles of Flue-cured Virginia (FCV) and non-FCV tobacco are cultivated under varying agro-ecological situations spread all over the country. The types include Flue-cured Virginia, Burley and HDBRG, Oriental, Bidi, Cheroot/Natu/Pikka, Lanka, Chewing, cigar wrapper and filler, Hookah and snuff.

To meet the steady increase in demand for tobacco and national and international requirements, there is a need to improve the tobacco quality and yield in different types of tobacco. Varieties play a significant role in increasing the productivity and quality of any crop. The diversity present within the gernplasm is routinely being tested for its utility in breeding programmes. The assessment is done using various techniques such as morphological, biochemical and molecular markers. Morphological markers are usually based on phenotypical traits such as plant type, flower & seed color, shape, growth habits, and other visual observations which do not require expensive technology. Based on the available diversity, parents are selected in the breeding programmes for developing high yielding and stress tolerant varieties/hybrids. In tobacco, so far, 94 high yielding varieties of different types were evolved and released for cultivation in different regions of India, significantly influencing the economy of the farmers. ICAR- Central Tobacco Research Institute and All India Network Project on Tobacco (AINPT) Centres are continued to breed various tobacco types for higher yield and better quality. The superior tobacco genotypes developed by the scientists are evaluated in AINPT trials conducted at 4 main centres, 7 sub centres and 4 voluntary centres for identifying superior entries. The lines found superior among the tested entries under Initial Varietal trail (IVT), Advance varietal trial (AVT) and bulk trials are identified for commercial cultivation in AINPT workshops. Presence of higher diversity in the entries entered in the trial ensures yield increments in released varieties along with reducing the genetic vulnerability of tobacco crop. Hence, present study was undertaken to visualize the diversity present within genotypes evaluated under AINPT.

Materials and Methods:

This present study was conducted at black soil farm Katheru of ICAR-CTRI, Rajahmundry during 2017-18. A total of 100 AINPT lines including 87 FCV, 7 *Bidi*, 4 *N. rustica* and 2 chewing lines, collected from 2000 to 2015, maintained in the tobacco genebank at ICAR-CTRI were utilized in the study (Table 1).

Table 1: List of AINPT entries studied

56-3	V 4230	FCJ14	FCR 33	FCJ 32
A-3	HV 2000-2	FCJ15	FCR 34	FCJ 33
A-13	HV 2000-6	FCK2	FCR 35	FCJ 34
CY 139	PCT-09-1	FCK3	FCR 36	FCR 41
CY 149	PCT-09-2	FCK4	FCR 37	FCR 42
FCH 196	FCS-1	LR 75	FCR 38	FCR 43
FCH 197	FCS-2	LR 76	FCR 39	FCR 44
JS 77	FCR14	LR 77	FCR 40	FCR 45
KST 27	FCR15	ArBD 32	FCS 3	FCR 46

*Corresponding author:ksarala@rediffmail.com

KST 28	FCR16	ArBD 33	FCJ 16	FCR 47
LV 2	FCR17	ArR 46	FCJ 17	FCR 48
LV 10	FCR18	FCR 24	FCJ 18	FCR 49
N 98	FCR19	FCR 25	FCJ 19	FCR 50
RT 13	FCR20	FCR 26	FCJ 20	FCK 5
SBS 1	FCR21	FCR 27	FCJ 21	FCK 6
SL 15	FCR22	FCR 28	FCJ 22	ArBD 9
SL 17	FCR23	FCR 29	FCJ 23	NYBD 56
SL 21	FCJ11	FCR 30	FCJ 24	NYBD 59
SL 24	FCJ12	FCR 31	FCJ 30	ABD 152
V 4219	FCJ13	FCR 32	FCJ 31	ArBD 126

Preparatory cultivations such as deep ploughing in summer and 2-3 ploughing between July to September were carried out to make the field free of weeds. The nursery bed was raised during September and the seedlings were transplanted to the main field in first fortnight of November 2017. The spacing adopted was 70 x 60 cm and ten plants per entry were maintained. Recommended crop production and protection practices were followed to raise a healthy crop. Morphological observations were recorded in three plants after confirming the uniformity within the row. About 27 morphological observations were recorded in 106 entries to see the diversity available within the entries. The list of the characters along with the scoring are listed below (Table 2).

Table 2. Morphological observations recorded on the AINPT entries and their notes

Character	Note
Plant shape (PS)	Conical (1), Cylindrical (2), Elliptical (3)
Mean Stem Height (MSH)	Very Short (1), Short (3), Medium (5), Tall (7),Very Tall (9)
Plant Habit (PH)	Open (1), Erect(3), SemiErect(5), Squattering(7), Bouquet (9)
Plant Internodal Length (PI)	< 4 cm (1),<6 cm (2),>6cm (3)
Plant Number of Leaves (PN)	Very Few (1),Few (3),Medium Many(7),Very Many (9),
Leaf Type (LT)	Sessile (1), Petiolate (2)
Leaf Angle of Insertion (LAI)	Acute(1),Moderately Acute(2),Right Angle(3), Moderately Right(4)
Leaf length (LL)	Very Short (1),Short (3),Medium (5), Long (7),Very Long (9)
Leaf Width (LW)	Very Narrow (1),Narrow (3),Medium (5), Broad (7)
Leaf Midrib (LM)	Thin (1),Medium (2),Thick (3)
Leaf Veins-thickness (LVT)	Thin(3),Medium(5),Thick (7)
Leaf Veins& Midrib angle(LVMA)	Acute / Very Acute(1), Moderately Acute (2),Right Angle (3)
Leaf Blade Shape (LBS)	Lanceolate (1),Narrow Elliptic (2), Broad Elliptic (3),Ovate (4),Obovate (5),Cordate (6),Rounded (7)
Leaf Tip Shape (LTS)	Obtuse (1),Slightly Pointed (3), Medium Pointed (5),Strongly Pointed (7), Very Strongly Pointed (9)
Leaf Blistering of Blade (puckering) (LBB)	Absent or Very Weak (1),Weak (3), Medium (5),Strong (7),Very Strong (9)
Leaf Undulations of Margin (LUM)	Absent or Very Weak (1),Weak (3),Medium (5), Strong (7)
Leaf Development of Auricles (LDA)	Absent or Very Weak (1),Weak (3),Medium (5),Strong (7), Very Strong (9)
Leaf Color of Blade (LCB)	Yellow Green (1),White Green (2),Light Green (3),Medium Green (4),Dark Green (5)
Time of Flowering (TF) (50% of plants with at least one corolla open)	Very Early (1),Early (3),Medium (5),Late (7), Very Late (9)
Inflorescence Shape (IS)	Spherical (1),Flattened Spherical (2), Inverted Conical (3),Double Conical (4)
Inflorescence Compactness (IC)	Very Loose (1),Loose / Slightly loose(3), Medium (5),Slightly Dense / Dence (7), Very Dence (9)
Flower Length /Size (FL)	Short (3),Medium (5),Long (7)
Flower Expression -Tips of Corolla (FETC)	Absent (1),Weak (3), Medium (5),Strong (7),Very Strong (9)
Flower Colour of Corolla (FCC)	White(1), Light Pink \sh white / whitish pink (2), Medium Pink (3), Dark Pink (4),Red (5), Variegated (6),Pinkish Yellow(7),Yellow(8)
Flower Length of Pistil Relative to Stamens (FLPRS)	Shorter (1),Equal Length (2),Longer (3)
Fruit Form (FF)	Rounded (1),Intermediate (2),Ovate (3),Conical (4), Elongated (5)
Testa Colour (TC)	Light Brown (1),Dark Brown (2),Cream (3),Yellow (4)

Statistical analysis: All the morphological observations were first converted into scoring pattern to resemble qualitative characters. Principal component analysis (PCA) analysis was performed using XLStat-2018 to analyze the pattern of data matrix for determining the selection criteria and identification of morphological characters highly responsible for diversity. Those PCs with Eigen values greater than one were selected (Jeffers, 1967) for further analysis. Morphological characters that were not invariant or highly correlated to another character were excluded from further analysis.

Cluster analysis: The diversity prevailing among the AINPT entries was computed using Computer Software Program–DARwin (Perrier and Jacquemond-Collet 2006). Dissimilarly matrix for morphological observation was constructed using Rogers-Tanimoto coefficient of associations to find out genetic relationships. These data were subjected to unweighted pair groups method with arithmetic mean (UPGMA) analysis to generate dendrogram using DARwin 5.0 and dissimilarity was estimated based on the respective morphological scoring.

Results and Discussion:

Initially 27 morphological observations were recorded to know the diversity existing within 100 AINPT. Among them majority of the plants were erect and conical in shape with medium height. Leaves of the most of the entries are narrow elliptical, sessile and inserted in moderately acute angle. The midrib was found to be thin in majority of the entries. The leaves were strongly pointed in many genotypes with medium blistering. Most of the entries were early flowering with spherical medium compact inflorescence. Flower colour of the entries varied from white, pink to variegated type. The fruits were oval in shape and testa colour of the seeds was light brown in most of the entries.

The contribution of these characters to the diversity within genotypes was studied for clustering the genotypes. Principle component analysis (PCA) was done for selecting the characters that actually contribute to the diversity. Morphological observation were found to be explained by five principal components accounting for 56 % of variability with PCA 1 and PCA 2 capturing majority of the variability (Table 3). PCA1 accumulated 21 % of morphological variation and was loaded on Plant shape (PS), Mean Stem Height (MSH), Plant Number of Leaves (PN), Leaf Type (LT), Leaf Angle of Insertion (LAI), Leaf length (LL), Leaf width (LW), Leaf Tip Shape (LTS), Leaf Development of Auricles (LDA), Flower Length (FL), Fruit Form (FF) and Testa Colour (TC) and PCA2 on Leaf Midrib (LM), Leaf Veins-thickness (LVT), Leaf Blade Shape (LBS), Flower Expression of Tips of Corolla (FETC), Flower Colour of Corolla (FCC). The above characters showed high variability and are expected to provide high level of gene transfer if used in breeding programs (Gana, 2006; Aliyu *et al.*, 2000).

Table 3: Eigen value ("Load") of the correlation matrix and its contribution to total variation of AINPT entries.

Characters	PCA1	PCA 2	PCA 3	PCA 4	PCA 5
Plant					
Shape (PS)	**0.295**	0.044	0.067	0.055	0.004
Mean Stem Height (MSH)	**0.496**	0.031	0.049	0.015	0.001
Height (PH)	0.259	0.006	**0.298**	0.007	0.015
Internodal length(PI)	0.034	0.108	0.075	**0.180**	0.125
Number of Leaves (PN)	**0.553**	0.092	0.036	0.000	0.016
Leaf					
Type (LT)	**0.649**	0.065	0.090	0.002	0.001
Angle of Insertion (LAI)	**0.270**	0.134	0.089	0.008	0.001
Length (LL)	**0.555**	0.042	0.020	0.003	0.020
Width (LW)	**0.293**	0.085	0.001	0.073	0.033
Midrib (LM)	0.045	**0.534**	0.009	0.100	0.020
Veins-thickness (LVT)	0.089	**0.412**	0.038	0.170	0.026
Veins& Midrib angle(LVMA)	0.013	0.017	0.012	0.152	**0.212**
Blade Shape (LBS)	0.000	**0.612**	0.040	0.014	0.024
Leaf Tip Shape (LTS)	**0.367**	0.064	0.002	0.015	0.000
Blistering of Blade (LBB)	0.094	0.003	0.014	0.001	**0.487**
Undulations of Margin (LUM)	0.064	0.026	0.073	0.000	**0.292**
Development of Auricles (LDA)	**0.204**	0.047	0.082	0.132	0.000
Color of Blade (LCB)	0.176	0.008	0.176	**0.223**	0.007
Time to 50% flowering (TFF)	0.050	0.115	**0.266**	0.163	0.006
Inflorescence Shape (IS)	0.001	0.104	**0.143**	0.022	0.005
Inflorescence Compactness (IC)	0.010	0.060	0.071	**0.192**	0.049
Flower					
Length (FL)	**0.268**	0.004	0.029	0.093	0.016
Expression of Tips of Corolla (FETC)	0.001	**0.200**	0.006	0.004	0.014
Colour of Corolla (FCC)	0.069	**0.102**	0.094	0.006	0.060
Length of Pistil Relative to Stamens (FLPRS)	0.034	0.001	0.003	**0.060**	0.024
Fruit Form (FF)	**0.293**	0.004	0.182	0.093	0.002
Testa Colour (TC)	**0.549**	0.014	0.055	0.003	0.011
Eigenvalue	5.731	2.935	2.020	1.787	1.469

Variability (%)	21.227	10.871	7.481	6.619	5.443
Cumulative %	21.227	32.098	39.579	46.198	51.640

**Values in bold correspond for each variable to the factor for which the squared cosine is the largest*

Significant amount of difference in various observations helps in grouping the genotypes into different clusters. Hence, based on PCA1 & 2, 17 characters were selected for further analysis as the remaining variables had weak or no discriminatory power. Based on PCA1 and PCA2 the genotypes were clustered into four groups (Fig 1). The entry LR 75, LR 76, LR 77, ArBD 32 and ArBD 126 were far from the origin showing that they were distinct from other entries. The characters namely Leaf Type (LT), Plant Number of Leaves (PN), Testa Colour (TC) similarly documented for higher variability of the entries. The variables found within the same group may share some underlying biological relationship, and these associations can be useful for generating hypothesis for better understanding of knowledge on the complex traits (Maji and Shaibu, 2012).

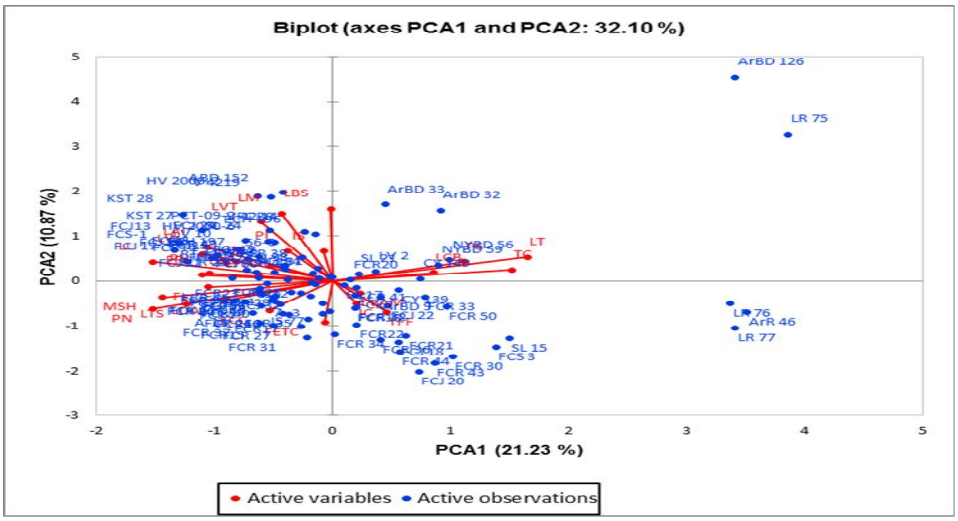

Figure1. Principle component analysis (PCA) of AINPT entries along with 27 morphological observation; plot of individual accessions on first two PCA axes. PCA data shown in Table 1.

Further in order to know the dissimilarity index and relationship between the genotypes, data was analyzed in DarWin 5.0. Clustering through unweighted pair groups produced grouping that defined four distinct clusters in 0 to 0.5 distance (Fig. 2). Majority of the FCV entries were in a single group except FCR 38 and SL 15. Similarly all the *rustica* entries formed separate cluster. The entries namely V 4230 and FCJ14 had the lowest genetic distance of about 0.04. All the FCV entries were lying with the genetic distance of about 0.25 except for FCR 38. The reason for low diversity within the cluster is that tobacco being quality oriented crop; the breeders are confined to use few genotypes in breeding programme. *Bidi* and chewing entries were found to interlay with FCV entries showing the possibility of less morphological difference between these types. Few genotypes were found to group under same cluster, this could be possible explained by the morphological similarity observed between the types of tobacco. The conclusion on diversity is made in this study based on morphological characters only and may be verified with molecular diversity for further confirmation.

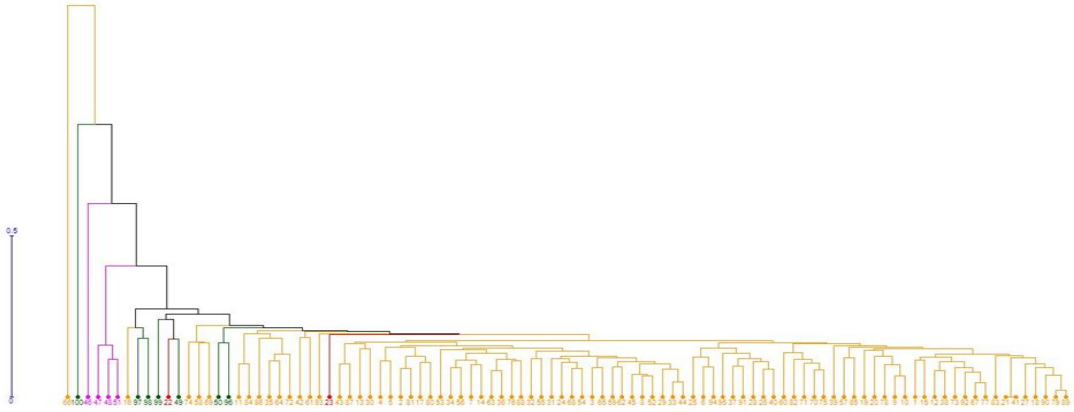

Figure 2. Clustering pattern of the 100 AINPT entries using DarWIN. Each line is named with a prefix number.

The computed tree distance shows that FCV 38 has the highest distance from other genotypes followed by ArBD 126 with a value of 2.67 and 1.69, respectively. The shortest distance is noted in between the genotype KST 27 and KST 28. It was reported by Lewis & Nicholson (2007) that genetic variability within *N. tabacum* L. was likely affected by several genetic bottlenecks, hence clustering based on morphology will lead to many sub-clusters. As reported by them the first diversity limiting step was probably due to the result of the

pathway of speciation for *N. tabacum*. Since all the cultivated forms of *Nicotiana* come under same species except *Rustica* (Goodspeed 1954; Chase *et al.* 2003; Knapp *et al.* 2004) morphological classification has its own limitations. Earlier reports also states that a high degree of genetic relatedness exists among modern varieties in the different tobacco types (Garner *et al.* 1936; Murphy *et al.* 1987; Bindler *et al.* 2005). It remains to be seen, however, whether or not the theoretically relatively low levels of genetic variability within tobacco germplasm pools (Garner *et al.* 1936; Murphy *et al.* 1987) will become an obstacle to continued cultivar improvement. The variability available in wild *Nicotiana* sp. may be utilized through pre-breeding to create variability for further broadening the genetic base of tobacco breeding programmes.

Conclusion:

All the AINPT entries studied were diversified into four clusters within a distance of 0.04 to 2.63 with majority of FCV, bidi, *rustica* and chewing entries falling into distinct clusters. The intra cluster distance was very low indicating the lack of diversity within the genotypes in a cluster. This may lead to genetic vulnerability and lower marginal yield gains in released varieties. Therefore, breeders need to overlook their breeding programme and try to bring in more diversity into the elite cultivar as much as possible. The purpose is just not to create variability for specific traits, but to broaden the genetic diversity per se so as to broader genetic base and reduce vulnerability of crop to variations in the environment. Hence, pre-breeding activities to broaden the genetic base of the breeding materials is essential in order to attain stable yields gains.

Acknowledgement:

The authors are thankful to the entire breeders who contributed their lines to AINPT, the technician of the Division of Crop Improvement, Central Tobacco Research Institute, Rajahmundry and M. Sc. Botany Students of Government College (Autonomous), Rajahmundry for assisting in executing and recording observations in the project.

References:

Aliyu, B., Akoroda, M.O., Padulosi, S. 2000. Variation within Vigna reticulata Hooke FII Nig. J. Gene. 1-8.

Bindler, G., Martin, F., Renaud, L., Kaelin, E., Donini, P., Gandani, F. and Rossi, L. 2005. Microsatellite markers for tobacco genetic fingerprinting and variety identification. Plant and Animal Genomes XIII Conference, San Diego, CA, P136 (Abstract) (abstract available at http://www.intl-pag.org/13/abstracts/PAG13_P136.html)

Chase, M.W., Knapp, S., Cox, A.V., Clarkson, J.J., Butsko, Y., Joseph, J., Savolainen, V., Parokonny, A.S. 2003. Molecular systematics and the origin of hybrid taxa in Nicotiana (Solanaceae). Ann Bot 92:107–127

Gana, A.S. 2006. Variability studies of the response of rice varieties to biotic and abiotic stresses. Unpublished Ph.D Thesis, University of Ilorin.

Garner, W.W., Allard, H., Clayton, E.E. 1936. Superior germplasm in tobacco. In: 1936 Yearbook of agriculture. USDA, Washington, DC, pp 785–830

Goodspeed, T.H. 1954. The genus Nicotiana. Chronica Botanica, Waltham Massachusetts

Jeffers, J.N.R. 1967. Two case studies in the application of principal component analysis. Appl. Stat., 16: 225-236.

Knapp, S., Chase, M.W., Clarkson, J.J. 2004. Nomenclatural changes and a new section classification in Nicotiana (Solanaceae). Taxon 52:73–82

Lewis R. S and Nicholson J. S. 2007. Aspects of the evolution of Nicotiana tabacum L. and the status of the United States Nicotiana Germplasm Collection. Genet Resour Crop Evol 54:727–740

Maji, A. T. and Shaibu, A. A. 2012. Application of principal component analysis for rice germplasm characterization and evaluation. J. Plant Breed. Crop Sci. Vol. 4(6), 87-93.

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

Perrier, X and Jacquemond-Collet, J. P. 2006. DARwin software. http://darwin.cirad.fr/darwin/