Review

Genetic diversity and population structure of cotton (*Gossypium hirsutum* L. race *latifolium* H.) using microsatellite markers

Leonel Domingos Moiana[2], Pedro Soares Vidigal Filho[1]*, Maria Celeste Gonçalves-Vidigal[1], Giselly Figueiredo Lacanallo[1], Marta Zulema Galván[3], Luiz Paulo de Carvalho[4], Manuel Pedro Maleia[2], Cynthia Maritza Pacheco[1], Tiago Ribeiro[1], Hugo Zeni Neto[1] and Gislayne Kelly Coimbra[1]

[1]Universidade Estadual de Maringá, Pós-Graduação em Genética e Melhoramento, Av. Colombo, 5790, Bloco J45, CEP 87020-900, Maringá-PR-Brazil.
[2]Instituto de Investigação Agrária, Av. das FPLM 2698, C.P. 2698, Maputo– ozambique.
[3]CONICET- Lab. de Biot. EEA INTA, Ruta Nac 68 Km 172 (4403) Cerrillos, Salta, Argentina.
[4]Embrapa Algodão. Rua Osvaldo Cruz, 1143, CEP 58428-095, Campina Grande–PB-Brazil.

Cotton (*Gossypium* spp.) is the world's leading natural fiber crop and is cultivated in diverse temperate and tropical areas. In this sense, molecular markers are important tools for polymorphism identification in genetic diversity analyses. The objective of this study was to evaluate genetic diversity and population structure in tetraploid cotton (*Gossypium hirsutum* L. race *latifolium* H.) by means of microsatellite markers. 43 cultivars and inbred lines from Africa, United States and Brazil were analyzed. From a total of 33 SSRs markers, 15 markers revealed 104 polymorphic SSR alleles. Four groups were identified applying different methods (the probabilistic method, Principal Coordinates Analysis and Neighbor Joining tree). American cultivars and inbred lines were included in group I; African cultivars in group II; and Brazilian cultivars in groups II, III and IV. The F_{ST} index indicated high genetic variability among the cultivars and inbred lines studied. In general, American cultivars were the most divergent compared to African and Brazilian ones. The dissimilarity index ranged from 0.06 to 0.90 and the lowest genetic divergence was observed between TAMCOT22 and TAM96WD-69s(L). Combination of American cultivars and inbred lines with African and Brazilian cultivars is recommended for obtaining superior segregant in order to improve yield.

Key words: Dissimilarity index, *Gossypium hirsutum L.*, polymorphism, SSRs markers.

INTRODUCTION

The genus *Gossypium* includes approximately 50 species distributed worldwide, from which five are tetraploids and belong to subgenus Viz *karpas* (Cronn and Wendel, 2004). Diploid species with 26 chromosomes are divided into eight cytogenetic genomic groups denominated from A to G, and K; whereas tetraploids species with 52 chromosomes are established in group AD (Fryxell, 1984;

Wendel and Cronn, 2003; Cronn and Wendel, 2004; Lapace et al., 2007).

In cotton industrialization, 1 kg of fiber can be obtained from about 1.65 kg of seed with 21% oil and 23% proteins. This makes cotton the fifth crop for oil production and second for protein source in the world (Wallace et al., 2008; Benbouza et al., 2010).

There are 60 countries around the world that cultivate cotton in 34 million hectares. The main areas of production belong to Brazil (1,338 kg.ha-[1]), United States of America (985 Kg.ha-[1]), India (550 kg.ha-[1]), China:

---

*Corresponding author. E-mail: psvfilho@uem.br.

(1,265 kg.ha-[1]); Mexico (1,247 kg.ha-[1]); Australia (2,000 kg.ha-[1]); (Fengguo et al., 2007; Beltrão and Azevedo, 2008; Khadi et al., 2010). Seven races had been described by Hutchinson (1951), six with different stages of domestication *"marie-galante", "punctatum", "richmondi", "morrilli", "palmeri"* and "*latifolium*" and a single wild race "y*ucatanense"* (Lacape et al., 2007).

In cotton breeding programs, genetic dissimilarity analysis is one strategy for parents' selection. However, the use of morphological traits has been limited due to the restricted number of markers, and the existence of environment and plant development influence (Lukonge et al., 2007). Biochemical markers like storage proteins and isoenzymes were also used to study genetic diversity in the genus *Gossypium* (Wendel et al., 1992; Brubaker and Wendel, 1994). With the advent of molecular marker technology, it became possible to develop multiple strategies to identify polymorphism. Several studies applying different molecular markers have been carried out including amplified fragment length polymorphisms (AFLPs) (Pillay and Myers, 1999; Abdalla et al., 2001; Lukonge et al., 2007; Iqbal et al., 2001; Zhang et al., 2008; Rakshit et al., 2010), random amplification of polymorphic DNA (RAPD) (Khan et al., 2000; Zhang et al., 2008; Maleia et al., 2010a), restriction fragment length polymorphisms (RFLPs) (Brubaker and Wendel, 1994; Zhang et al., 2008) and simple sequence repeats (SSR) (Gupta and Varshney, 2000; Zhang et al., 2008; Rakshit et al., 2010). Of these markers, SSR was the most outstanding in recent years.

SSR or microsatellites are abundant across genomes, showing high levels of polymorphism. SSRs are co-dominant markers present in coding and noncoding regions distributed throughout the nuclear genome (Collard et al., 2005; Agarwal, 2008; Lin et al., 2010; Kalia et al., 2011).

There are few studies regarding cotton population structure and genetic diversity in Brazil and Africa, especially in Mozambique, using molecular markers. Some research groups have conducted genetic analyses of cultivars from these locations, but no comparison among locations has been carried out (Lacape et al., 2007; Lukonge et al., 2007; Maleia et al., 2010a). Therefore, the objective of this study was to evaluate the genetic diversity and population structure of cotton cultivars and inbred lines from Mozambique and Brazil, using microsatellite markers.

DNA extraction and quantification

Genomic DNA was extracted from fresh young leaves, from four days seedlings (Zhang and Stewart, 2000) of each cultivar and inbred line (Table 1). After collection, leaves were immediately placed in eppendorf tubes and maintained in liquid nitrogen to preserve the DNA. DNA concentration was measured by a Fluorometer QuBit[TM]

and samples were diluted with buffer TE 0.1X to a concentration of 30 ng µl[-1] and stored at -20°C.

SSR genotyping

For the genetic divergence assay, 33 SSR primers were obtained from *Invitrogen-Induslab*, but only 15 were selected since they considered the most polymorphic ones and also because they presented well defined bands (Table 2).

Amplification reactions were conducted in a total volume of 20 µL which contained: 10 mM Tris HCl, pH 8.3; 50 mM KCl; 0.1% Triton X-100; 0.2 mM de dNTP; 1 U *Taq* DNA polymerase; 4 mM of each primers; $MgCl_2$ in a concentration compatible to the primer pair; 30 ng genomic DNA; and ultra pure water. Polymorphism was verified comparing the amplified products on Applied Biosystems (AB), and the equipment of Veritti[TM] following Williams et al. (1990)'s protocol.

The BNL primers amplification was performed using initial desnaturation of 95°C by 12 min, followed by 30 cycles of 93°C by 1 min, annealing at 51 to 55°C by 2 min and extension of 72°C by 3 min. The last cycle was followed by extension of 72°C by 7 min. The NAU and JESPR primers amplifications were performed using a initial denaturation of 94°C by 5 min, followed by 35 cycles of 94°C by 30 seg, annealing at 51 to 53°C by 1 min and an extension of 72°C by 1 min. The last cycle was followed by final extension of 72°C by 8 min.

Statistical analysis

The software LAB IMAGE 1D, version 1.10, Loccus Biotecnologia[TM] was used to genotype the samples. The higher allelic frequency per locus, the number of alleles per locus, the average of alleles per locus, the average of high allelic frequencies and genetic diversity per locus were obtained using the software Powermarker 3.25 (Liu and Muse, 2005). Population structure analysis of cotton cultivars and inbred lines was conducted using the software Structure 2.3.3. (Pritchard et al., 2000). Cotton variability was also structured using GenAlEx 6.3. (Peakall and Smouse, 2006), which provided a genetic distance matrix used for Principal Coordinates analysis. In addition, genetic diversity among cultivars/inbred lines was also studied by the construction of a Neighbour Joining tree based on C.S. Chord distance (Cavalli-Sforza and Edwards, 1967).

RESULTS AND DISCUSSION

A total of 104 alleles were detected analyzing 35 cultivars and eight inbred linesof*G hirsutum* L. by means of 15 SSR primers. Most of them were polymorphic among the

Table 1. List of cultivars and inbred lines used to analyze a genetic diversity and population structure.

Number of order	Cultivar/inbred line	Geographic group	Origin	Number of order	Cultivar	Geographic group	Origin
1	Alb.sz 9314	African	Zimbabwe	22	CD 404	American	Brazil
2	Alb.FQ 902	African	Zimbabwe	23	CD408	American	Brazil
3	Alb.BC 853	African	Zimbabwe	24	CD410	American	Brazil
4	Stam 42	African	Senegal	25	ECC 2606	American	Brazil
5	CA-222	African	Ivory Cost	26	ECC 2623 M	American	Brazil
6	CA-324	African	Ivory Cost	27	ECC 2627 M	American	Brazil
7	IRMA 12-43	African	Cameron	28	ECC 2645 M	American	Brazil
8	ISA-205	African	Ivory Cost	29	Mocó	American	Brazil
9	REMU-40	African	Mozambique	30	DP 604 BG	American	Brazil
10	TAMCOT 22	American	USA-Texas	31	Sicala 40	American	Brazil
11	TAM 96WD-69s [L]	American	USA-Texas	32	FMT 705	American	Brazil
12	TAMCOT Pyramid	American	USA-Texas	33	FMT 910	American	Brazil
13	TAM 98D -102 [L]	American	USA-Texas	34	FMT 993	American	Brazil
14	TAM 96WD-18 [L]	American	USA-Texas	35	FMT 966	American	Brazil
15	TAM 94J-3 [L]	American	USA-Texas	36	FMT 701	American	Brazil
16	TAM 88G-104 [L]	American	USA-Texas	37	BRS Rubi	American	Brazil
17	TAMCOT Sphinx	American	USA-Texas	38	BRS Verde	American	Brazil
18	TAM 98D-99ne [L]	American	USA-Texas	39	BRS Safira	American	Brazil
19	TAM 94WE-37s [L]	American	USA-Texas	40	BRS Buriti	American	Brazil
20	TAM 94L-25 [L]	American	USA-Texas	41	BRS 293	American	Brazil
21	TAMCOT Luxor	American	USA-Texas	42	BRS Aroeira	American	Brazil
				43	BRS Araripe	American	Brazil

USA,United States of America; [L] Inbred lines. Numbers of order 10 to 21 are cultivars and inbred lines introduced from USA in 2006.

materials analyzed. The maximum number of alleles found was 10 and corresponded to *loci* JESPR292, BNL-1694 and JESPR152, with an average of 6.9 alleles per locus (Table 3). The higher allelic frequency values ranged from 0.326 for *locus* JESPR152 to 0.774 for *locus* BNL3816, with a mean of 0.453 (Table 3). Similar results were obtained by Liu et al. (2000) who reported an average of five alleles per locus, ranging from 2 to 11 alleles, evaluating 97 cultivars and primitive species of *Gossypium* using 62 SSR markers. The population structure analysis applying the Probabilistic method described by Pritchard et al. (2000) showed the existence of four distinguished groups.

The software Structure was run a single time for each K value ranging from K=2 to K=15. Each run was performed using the admixture model and 10,000 replicates for burn-in and 100,000 during the analysis. The K=5 analysis was of particular interest to identify four groups that consisted of American cultivars and inbred lines, African and Brazilian cultivars, BRS Brazilian cultivars and FM Brazilian cultivars (Figure 1). Structure harvester (Earl, 2010) was used to determine the optimal value for K which resulted to be K=5 (Figure 1). Maleia et al. (2010a), using RAPD markers, found a similar population structure in relation to American cultivar/inbred lines and African cultivars. The principal Coordinates Analysis also

revealed the existence of four groups (Figure 2). Populations 1 and 2 corresponded to American cultivars and African and Brazilian cultivars, respectively, while populations 3, 4 and 5 corresponded to Brazilian cultivars.

On the other hand, Lacape et al. (2007) found an average of 5.6 alleles per locus (from 2 to 17), while analyzing 47 accessions of *Gossypium* with SSR markers.

In this study, microsatellite markers exhibited PIC values that varied from 0.367 to 0.771, with a mean value of 0.646. The locus with higher diversity among cultivars and inbred lines was JESPR152 with a value of 0.797, followed by loci BNL1053 and BNL4035 with values of 0.770 and 0.753, respectively; whereas the locus with the lowest genetic diversity was BNL3816 with 0.409 (Table 3). These values are higher than those found by Liu et al. (2000), while testing a collection of 97 cultivars and primitive species of *Gossypium* through 62 SSR markers that obtained PIC values from 0.05 to 0.82 (means of 0.31). In other research, Lacape et al. (2007), evaluating 47 accesses of *Gossypium,* obtained PIC values from 0.08 to 0.89 (mean of 0.55), while, Lukonge et al. (2007), evaluating 26 cotton cultivars using AFLP markers, found values between 0.35 and 0.57 (mean of 0.47). This aspect showed a high genetic diversity among the

Table 2. List of SSR primers utilized in polymerase chain reaction (PCR) amplification and genomic DNA analysis.

Number of order	SSR marker	Motif	Nucleotides sequence (5´- 3´)	Direction
1	BNL1694	(AG)19, (TC)19	CGTTTGTTTTCGTGTAACAGG TGGTGGATTCACATCCAAAG	Forward Reverse
2	BNL3408	(GT)2AT(GT)12	ATCCAAACCATTGCACCACT GTGTACGTTGAGAAGTCATCTGC	Forward Reverse
3	BNL2495	(AG)14, (TC)14	ACCGCCATTACTGGACAAAG AATGGAATTTGAACCCATGC	Forward Reverse
4	BNL2572	(GA)23	GTCCTATTACTAAAATTGTTAATTTAGCC CGATGTTAAATCAATCAGGTCA	Forward Reverse
5	BNL1053	(AC)16	AGGGTCTGTCATGGTTGGAG CATGCATGCGTACGTGTGTA	Forward Reverse
6	BNL3649	(TC)20	GCAAAAACGAGTTGACCCAT CCTGGTTTTCAAGCCTGTTC	Forward Reverse
7	BNL2543	(AG)11	GCCGAAACTAAAACGTCCAA TCCTTACTCACTAAGCAGCCG	Forward Reverse
8	BNL2494	(AG)14, (TC)14	ACCGCCATTACTGGACAAAG AATGGAATTTGAACCCATGC	Forward Reverse
9	BNL3816	(TG)15, (TG)5TA(TG)15	GTTAGCCACGTGTTAGTTCTATG ATCGATCACTTGCTGGTTCC	Forward Reverse
10	BNL3904	(GT)11	ATGCATTAATGAGTCGATAGGC GCACAAAGAAAACAAACTGTACG	Forward Reverse
11	BNL3998	(AC)11	CGGCGAGAAGTTGAAAGATC TGTGCAAAAGTGGGTGGTAA	Forward Reverse
12	BNL4035	(TG)10	TGCATCTGCATTTGGGATTA TAGCCAACCGTTACACACCA	Forward Reverse
13	NAU864	(TG)10	GGATTAATTAGCCCCCACAT TCTTTTTCAGCTTGGGTTCT	Forward Reverse
14	JESPR292	(CTT)7	GCTTGCAATCTCCTACACC GAATATGTTTCATAGAATGGC	Forward Reverse
15	JESPR152	(GAA)50	GATGCACCAGATCCTTTTATTAG GGTACATCGGAATCACAGTG	Forward Reverse

Source: Cotton Marker Database (2010).

cultivars and inbred lines used in this research.

The American cultivars and inbred lines were placed in very distinguished groups from the African and Brazilian cultivars according to the First Principal Coordinate (34.51%). The African and some Brazilian cultivars were in separate groups according to the Second Principal Coordinate that represented 20.35% of the sample variability. Brazilian cultivars BRS and Brazilian cultivars FM formed isolated groups. The first two Principal Coordinates represented 54.86% of the total variation of the sample.

The F_{ST} value according to Wright's method was 0.236 showing a high genetic variability among the cultivars and inbred lines evaluated. Population subdivision determined

Table 3. Estimated indexes of genetic diversity per microsatellite locus evaluated.

Locus	Allele number	Haf	GD	PIC	Minimum (bp)	Maximum (bp)
BNL-3408	5	0.395	0.728	0.685	136	214
BNL-3816	4	0.743	0.409	0.367	116	223
BNL-3904	4	0.465	0.613	0.535	154	187
BNL-4035	7	0.326	0.753	0.712	135	213
BNL-3649	8	0.419	0.74	0.706	80	178
BNL-2544	5	0.512	0.671	0.632	195	297
BNL-1053	7	0.372	0.77	0.739	164	297
NAU-864	6	0.465	0.67	0.614	139	192
JESPR-292	10	0.419	0.708	0.665	161	206
BNL-1694	10	0.432	0.737	0.707	114	235
BNL-2572	8	0.432	0.707	0.666	219	291
BNL-3998	4	0.512	0.645	0.592	118	156
JESPR-152	10	0.326	0.797	0.771	110	198
BNL-2495	9	0.432	0.715	0.677	123	214
BNL-2494	7	0.512	0.664	0.623	107	269
Means	6.9	0.453	0.688	0.646		

Haf=Higher allelic frequency; GD=genetic diversity; PIC=polymorphism information content, allele size (bp) values.
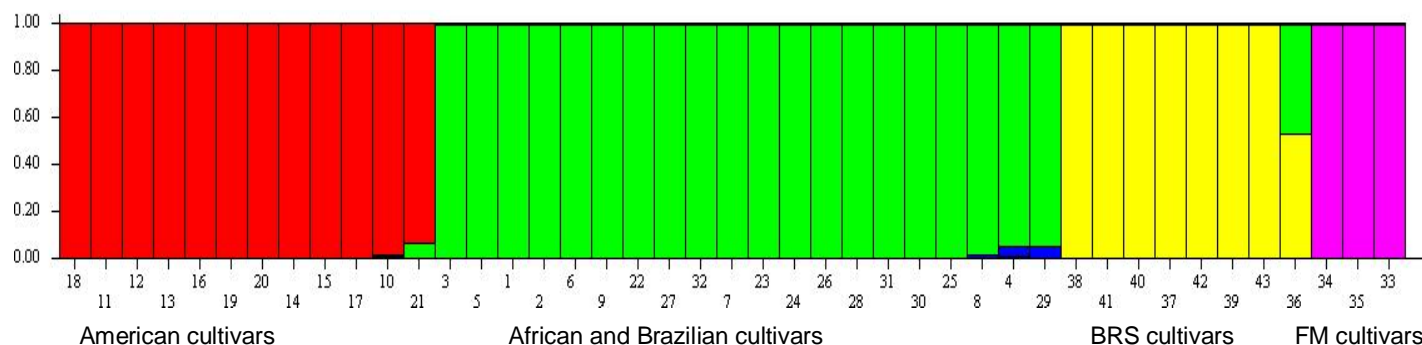
Figure 1. Population structure of cotton cultivars and inbred lines. K=5 groups obtained using software Structure 2.3.3. (Pritchard et al., 2000).

by structure analysis was coincident with the clusters observed in the Neighbor-joining tree generated based on CS Chord genetic distance (Cavalli-Sforza and Edwards, 1967) (Figure 3). Cultivars and inbred lines from USA (cluster I) were divergent from cultivars from Africa (cluster III) and Brazil (clusters II, III and IV). Most of the American cultivars and inbred lines have commun ancestors as DP 90, DP 50 and other unrealized inbred lines (Thaxton and Smith, 2005; Wallace et al., 2008). Divergence between American and African cultivars was also reported by Maleia et al. (2010a) while evaluating the genetic divergence of 21 cultivars and inbred lines, using RAPD markers.

All the groups identified applying different methods were consistent with pedigree data. African and Brazilian cultivars share ancestors in their pedigrees (cultivars DP61, Sicala and Deltapine) (Bertini et al., 2006). Cultivar Sicala originated from CSIRO cultivars selected from

DP16; this cultivar is one of the genitors of CA 324 and belongs to African groups. While, cultivars ISA 205 and Stam-42, both from the African group originated fromcultivar Deltapine (Dessauw [1]).The highest dissimilarity Cs Chord index between the materials analyzed was 0.9, and corresponded to TAMCOT 22, TAM 96WD-69s, TAMCOT Pyramid, TAM 98D-102 (L), TAM 96WD-18 (L), TAM 94J-3 (L), TAM 94J-3 (L), TAMCOT Sphinx, TAM 98D-99ne (L), TAM 94WE-37s (L), TAM 94L-25 (L), FM 993, and FM 966. These cultivars tend to promote hete-rosis effects when used in cotton breeding programs (Falconer and Mackay, 1996). Similar results, in relation to dissimilarity indexes were obtained by Bertini et al. (2006), while evaluating a sample of 53 cultivars using SSR markers.

[1]Personal information: Dessauw, Dominique.

Researcher of CIRAD. France.

Principal coordinates
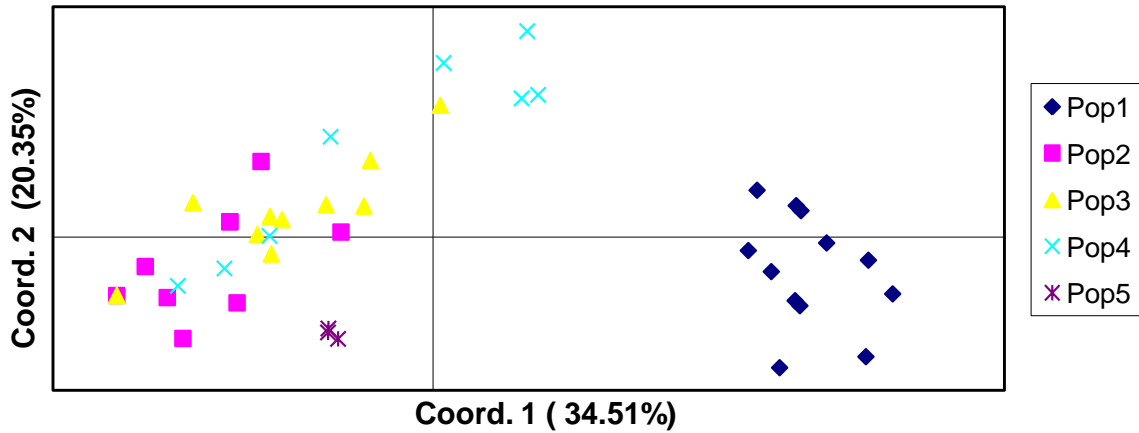
Figure 2. Population structure between cotton cultivars and inbred lines by Principal Coordinates analysis (PC1 vs PC2). Pop1= USA Cultivars and inbred lines; Pop 2=African cultivars; Pop 3= Brazilian cultivars; Pop 4= Brazilian (BRS and one FMT); Pop 5= Brazilian cultivars (FMs).
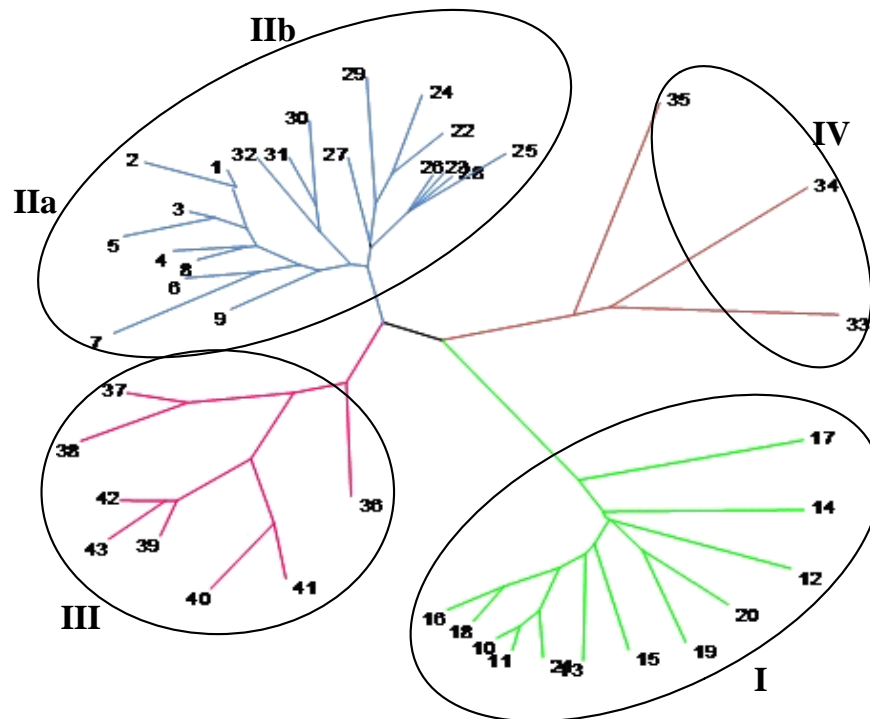
Figure 3. Distribution of 43 cotton cultivars and inbred lines according to *Neighbor-Joining* tree composing four groups. Each main branch is in agreement with K=5 formed groups by *Structure* software.

For Cotton Breeding Programs in Africa, particularly in Mozambique, hybridizations between TAMCOT 22 x Albar sz 9314, TAM 98D-99ne (L) x Albar FQ 902, TAM 96WD-69s (L) x FMT 705, TAMCOT Pyramid x CA-222, TAMCOT Pyramid x Albar BC 853, TAMCOT Sphinx x Stam-42, TAM 96WD-69s (L) x IRMA 12-43, TAMCOT Sphinx x ISA-205, TAMCOT 22 x REMU-40, TAMCOT Sphinx x CD 404, TAMCOT Pyramid x FM 993, FM 966 x

BRS 293, TAMCOT 22 x BRS Buriti, TAMCOT 22 x CA-324 would be the most promising and recommended combinations.

Cultivars Albar sz 9314, Albar FQ 902, Albar BC 853, CA-222, IRMA 12-43, ISA-205 and REMU-40 show wide adaptability and high tolerance to jassids (*Empoasca facialis*). Moreover, CA324 and Stam42 demonstrated specific adaptability to high and low quality environments and mean and low tolerance to jassids, respectively (Maleia et al., 2010b). In addition, TAMCOT Sphinx showed a high resistance to nematodes (*Rotylenchus reniformis*) (El-Zik and Thaxton, 1996) and the inbred line TAM 96WD-69s (L) is pilose, facilitating mechanic harvest. TAMCOT Pyramid is a cultivar that originated from the cross between TAMCOT Sphinx and CD3HGCBU8S-1-91 (unreleased inbred line), and shows improved characteristics as high yield, moderate pilosity, good fiber quality and high adaptability to diverse environments in Texas (Thaxton and Smith 2005; Wallace et al., 2008). TAMCOT 22 originated from crosses involving TAM 88G-104; is a cultivar with high yield and medium cycle and has high resistance to white fly (*Bemesis argentifolli*). This inbred line has high-yielding and excellent fiber properties (Thaxton and Smith, 2005; Wallace et al., 2008).

Conclusions

A great variability was observed among the 43 cotton cultivars and inbred lines (*Gossypium hirsutum*) analyzed. The cultivars from USA were more divergent in relation to cultivars from Brazil and Africa as revealed by the probabilistic method, Principal Coordinates Analysis and Neighbor-Joining tree. Also, some Brazilian cultivars (BR and FM) showed low genetic similarity with other cultivars analyzed.

For Cotton Breeding Programs in Africa, particularly in Mozambique, hybrid combinations between African and Brazilian cultivars with USA cultivars/inbred lines could be applied to acquire high productivity.

ACKNOWLEDGEMENTS

REFERENCES

Abdalla AM, Reddy EKM, Pepper AE (2001). Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. Theor. Appl. Genet. 102:222-229.

Agarwal M, Shrivastava N, Padh H (2008). Advances in molecular marker techniques and their applications in plant science. Plant Cell Rep. 27:617-631.

Beltrão NEM, Azevedo DMP (2008). Agronegócio do Algodão no Brasil. Brasília: Embrapa Informação Tecnológica p. 1309.

Benbouza H, Lacape JM, Jacquemin JM, Courtois B, Diouf FBH, Sarr D, Konan N, Baudoin JP, Mergeai G (2010). Introgression of the low gossypol seed & high-gossypol plant trait in upland cotton: analysis of [(*Gossypium hirsutum L. x G. raimondii*)[2] x *G. sturtianum*] tri-specific hybrid and selected derivatives using mapped SSRs. Mol. Breed. 25:273-286.

Bertini CHCM, Schuster I, Sediyama T, Barros EG, Moreira MA (2006). Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genet. Mol. Biol. 2: 321-329.

Brubaker CL, Wendel JF (1994). Reevaluating the origin of domesticated cotton (*Gossypium hirsutum L.*; Malvaceae) using nuclear restriction fragment length polymorphism (RFLPs). Am. J. Bot. 81:1309-1326.

Cavalli-Sforza LL, Edwards AWF (1967). Phylogenetic analysis: models and estimation procedures. Am. J. Hum. Genet. 19: 233-257.

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica 142:169-196.

Cronn R, Wendel JF (2004). Cryptic trysts, genomic mergers, and plant speciation. New Phytologist. 161: 133-142.

Earl DA (2010). Structure Harvester Vol. 3. Available in:http://users.soe.ucsc.edu/~dearl/software/struct_harvest/. Acess in: 20th Dezember.

El-Zik KM, Thaxton PM (1996). Registration of 'Tamcot Sphinx' cotton. Crop Sci. 36:1074.

Falconer DS, Mackay TFC (1996). Introduction to quantitative genetics. England: Longman Group Limited, p. 464.

Fengguo S, Xiangyun Z, Junlan L, Baosheng G, Junyi G, Yongzeng L, Shulong J, Zhinhong L, Yancai Z, Weiming L, Shezeng L, Wenlian P, Li Z, Chaojiang W, Senmao G (2007). Training course on cotton breeding and management technology. Shijiazhuan, China p. 231.

Fryxell PA (1984) Taxonomy and germplasm resources. In: Kohel RJ, Lewis CF (eds.), Cotton. Agron. Monogr. 24: 27-57.

Gupta PK, Varshney RK (2000). The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113:163-185.

Hutchinson JB (1951). Intra-specific differentiation in *Gossypium hirsutum.* Heredity 5:169-193.

Iqbal J, Reddy Ouk, El-Zik Km, Pepper AEA (2001). genetic bottleneck in the evolution under domestication of upland cotton *Gossypium hirsutum L.* examined using DNA fingerprinting. Theor. Appl. Genet. 103:547-554.

Kalia Rk, Raí Mk, Kalia S, Rohtas S, Dhawan AK (2011). Microsatellite markers: an overview of the recent progress in plants. Euphytica 177:309-334.

Khadi Bm, Santhy V, Yadav MS (2010). Cotton: an introduction. Biotechnol. Agric. For. 65:1-14.

Khan Sa, Hussain D, Askari E, Stewart JM, Malik KAZ (2000). Molecular phylogeny of *Gossypium* species by DNA fingerprinting. Theor. Appl. Genet. 101: 931-938.

Lacape JM, Dessauw D, Rajab M, Noyer JL, Hau B (2007). Microsatellite diversity in tetraploid *Gossypium* germplasm: assembling a highly informative genotyping set of cotton SSRs. Mol. Breed. 19:45-58.

Lin Z, Yuan D, Zhang X (2010). Mapped SSR markers unevenly distributed on the cotton chromosomes. Agric. China. 4:257-264.

Liu S, Cantrell RG, Mcarty JCJ, Stewart JM (2000). Simple sequence repeat-based assessment of genetic diversity in cotton race stock accessions. Crop Sci. 40:1459-1469.

Liu KJ, Muse SV (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128-2129.

Lukonge E, Herselman L, Labuschangne MT (2007). Genetic diversity of Tanzanian cotton (*Gossypium hirsutum L.*) revealed by AFLP analysis. Afr. Crop Sci. 8:773-776.

Maleia MP, Vidigal Filho PS, Vidigal MC G, Gonela A, Lacanallo GF, Moiana LD, Chamuene A, Sousa LL, Darben LM (2010a). Genetic divergence among african and american cotton (*Gossypium hirsutum L.* race *latifolium* H.) cultivars inbred lines through random

amplification of polymorphic DNA (RAPD) markers. Afr. J. Biotechnol. 9:8539-8548.

Maleia MP, Vidigal Filho PS, Kvitschal MV, Vidigal MCG (2010b). Stability and adaptability of commercial cotton cultivars (*Gossypium hirsutum L. race latifolium H.*) in Mozambique. Afr. J. Agric. Res. 5:539-550.

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

Pillay M, Myers GO (1999). Genetic diversity in cotton assessed by variation in ribosomal RNA genes and AFLP markers. Crop Sci. 39:1881-1886.

Pritchard J, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Rakshit A, Rakshit S, Singh J, Chopra SK, Balyan HS, Gupta PK, Bhat SR (2010). Association of AFLP and SSR markers with agronomic and fibre quality traits in *Gossypium hirsutum L.* J. Genet. 89:155-162.

Thaxton PM, Smith CW (2005). Registration of TAM 96WD-69s glabrous upland cotton germplasm line. Crop Sci. 45:1172-1173.

Wallace TP, Bowman D, Campbell BT, Chee P, Gutierrez OA, Kohel RJ, Mccarty J, Myers G, Percy R, Robinson F, Smith W, Stelly DM, Stewart JM, Thaxton P, Ulloa M, Weaver DB (2008). Status of the USA cotton germplasm collection and crop vulnerability. Genet. Res. Crop Evol. 56:507-532.

Wendel JF, Cronn RC (2003). Polyploidy and the evolutionary history of cotton. Adv. Agron. 78:139-186.

Wendel JF, Brubaker CL, Percival AE (1992). Genetic diversity in *Gossypium hirsutum L.* and the origin of Upland cotton. Am. J. Bot. 79:1291-1310.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990). DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18:6532-6535.

Zhang J, Stewart JM (2000). Economical and rapid method for extracting cotton genomic DNA. J. Cotton Sci. 4:193-201.

Zhang HB, Li YN, Wang BH, Chee PW (2008). Recent advances in cotton genomics. Int. J. Plant Genomics p. 20.