



Determining genetic diversity in cotton genotypes to improve variability¹

Melissa Cristina de Carvalho Miranda^{2*}, Daniel Bonifácio Oliveira Cardoso³, Thatiane de Sousa Paiva³, Francisco José Correia Farias⁴, Larissa Barbosa de Sousa³

10.1590/0034-737X202067060006

ABSTRACT

A better understanding of genetic diversity in cotton cultivars is needed to make the most of the genetic resources and guarantee continuing improvements in the breeding programs. Our objective was to evaluate the genetic diversity among cotton genotypes from the top breeding programs in Brazil and recommend hybrid combinations that would increase variability. The field experiments we conducted on three cotton genotypes from each of the following Brazilian breeding programs: EMBRAPA, MONSANTO DELTAPINE, FIBERMAX, IMA, TMG and UFU. Several phenotypic evaluations were carried out at V5, B1, full flower and full maturity. Moderate genetic divergence was observed among the genotypes from these breeding programs. Hybridization of the BRS 433 FL B2RF and FM 980 GLT genotypes would likely produce segregating populations with greater genetic variability, yield potential, lint yield and desirable fiber quality. Maturity, micronaire and fiber length were the strongest contributors to genetic divergence.

Keywords: *Gossypium hirsutum* L.; genetic vulnerability; quantitative characters.

INTRODUCTION

Cotton is one of the most recent and most promising agricultural sectors in Brazil. The 2019/2020 harvest was the fourth largest in the world (Conab, 2020). The success of cotton cultivation in Brazil is largely due to investments in research, production technology and domestic breeding programs.

The continued development of cultivars is dependent on making the most of available genetic resources. Because of numerous species and varieties, genetic variety within the *Gossypium* genus is vast (Vidal Neto & Freire, 2013). However, to produce genotypes with favorable agronomic and technological traits, breeders frequently choose successful cultivars as parent plants, which narrows the genetic base and leads to high levels of homogeneity across large areas of cropland (Mehboob-ur-rahman, 2012; Borém & Miranda, 2013).

This homogeneity poses risks such as greater vulnerability to biotic and abiotic stress and restricted variability in quantitative characters (Cruz *et al.*, 2011; Borém & Miranda, 2013). Thus, breeding programs need a broad genetic base to build diverse populations with satisfactory agronomic performance that can be used to obtain genetic gains and to identify hybrids with greater heterotropy (Cruz *et al.*, 2011; Ludke *et al.*, 2017).

Most research on genetic divergence in cotton focusses only on quantitative characters, such as productivity, yield and fiber quality, while disregarding the underlying phenotypic characters. However, it is important to integrate all possible variables to avoid segmentation in inferences about genetic divergence between genotypes and provide a better basis for evaluating strategies that improve variability and lead to better decisions in breeding programs (Carvalho *et al.*, 2017).

Submitted on September 25th, 2019 and accepted on September 4th, 2020.

¹ This work is part of the Master's Dissertation of the first author.

² Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz", Departamento de Genética e Melhoramento de Plantas, Piracicaba, São Paulo, Brazil. melissamiranda94@gmail.com

³ Universidade Federal de Uberlândia, Instituto de Ciências Agrárias, Uberlândia, Minas Gerais, Brazil. danieludia13@hotmail.com; thatianepaiva31@hotmail.com; larissa@ufu.br

⁴ Embrapa Algodão, Centro Nacional de Pesquisa de Algodão, Campina Grande, Paraíba, Brazil. francisco.farias@embrapa.br

*Corresponding author: melissamiranda94@gmail.com

Therefore, the present study evaluates the genetic diversity of cotton genotypes used in the top breeding programs in Brazil and recommends the best hybrid combinations for increasing variability.

MATERIAL AND METHODS

The experiment was carried out in the field using the 2017/18 harvest in Uberlândia, Minas Gerais Brazil (18° 52' S and 48° 20' W, elevation 805 meters).

The three most-cultivated cotton genotypes (*Gossypium hirsutum* L.) were selected from each of the five top breeding programs in Brazil and the breeding program at the Universidade Federal de Uberlândia – UFU (PROMALG) (Table 1).

The trial was set up using a randomized completely block design (RBD) with 18 treatments and three repetitions. Each experimental plot consisted of four rows of cotton, spaced one meter apart and four meters long. Only the two centermost rows from each plot, less half a meter from each end of the rows, were evaluated.

Agronomic aspects of five randomly sampled plants from each plot were evaluated using the cotton scale (*Escala do Algodão* in Portuguese) proposed by Marur & Ruano (2001). These evaluations were carried out at (V5), appearance of the first boll (B1), full flowering (FF) and full maturity (MAT) (Table 2).

Yields of seed cotton, cotton seed plus lint and lint were determined after evaluating the plants at full maturity. In addition, standard samples (SS) from each lint plot were

evaluated using High Volume Instruments (HVI) to determine indicators of fiber quality such as upper-half mean fiber length (UHM), the fiber length uniformity index (LU), short fiber index (SFI), fiber strength (STR), fiber elongation (ELG), micronaire index (MIC) and fiber maturity (MAT).

The resulting data from the 18 genotypes and 28 variables were evaluated by univariate and multivariate analyses using the GENES software package (Cruz, 2013). Univariate analysis was performed first and, depending on significance, the Scott-Knott test was then used to determine genotype groups ($p < 0.05$ and $p < 0.01$).

Several genotypic parameters were estimated, such as the genotypic determination coefficient (h^2) and the ratio of CVg to CVe (CVr). Phenotypic (rf) and genetic (rg) correlations were determined for characters with the highest CVr and h^2 values.

Genetic dissimilarity was calculated using the generalized Mahalanobis distance ($D^2_{ii'}$) (Cruz *et al.*, 2011). Then genotype clusters were determined using Torcher optimization and UPGMA clustering and displayed as dendrogram.

Dendrogram clusters were chosen using the method proposed by Mojena (1977) and adapted by Milligan & Cooper (1985). The cophenetic correlation coefficient (Mantel, 1967) was calculated to check the quality of UPGMA clustering. Finally, the relative contribution of each character to genetic divergence was quantified using the criterion proposed by Singh (1981).

Table 1: Characterization of the cotton genotypes (*Gossypium hirsutum* L.) used in the study

Genotype	Breeder	Cycle	Genetic Transformation Event
BRS 368RF	EMBRAPA	Medium to early	MON88913
BRS 372	EMBRAPA	Medium to late	Absent
BRS 433 FL B2RF	EMBRAPA	Medium to late	MON15985 and MON88913
DP Delta Opal	MONSANTO DELTAPINE	Medium to late	Absent
DP 1228 B2RF	MONSANTO DELTAPINE	Late	MON15985 and MON88913
DP 1552 B2RF	MONSANTO DELTAPINE	Medium	MON15985 and MON88913
FM 975 WS	FIBERMAX	Medium to late	281-24-236 and 3006-210-23
FM 980 GLT	FIBERMAX	Late	GHB 614, T304-40 and GHB119
FM 982 GL	FIBERMAX	Medium to late	GHB614 and LLCotton25
IMA 2106 GL	IMA	Medium to early	GHB614 and LLCotton25
IMA 5675 B2RF	IMA	Medium to early	MON15985 and MON88913
IMA 8405 GLT	IMA	Late	GHB 614, T304-40 and GHB119
TMG 45 B2RF	TMG	Medium to early	MON15985 and MON88913
TMG 47 B2RF	TMG	Late	MON15985 and MON88913
TMG 82 WS	TMG	Medium to late	281-24-236 and 3006-210-23
UFUJP-H	UFU	Late	Absent
UFUJP-P	UFU	Late	Absent
UFUJP-B	UFU	Late	Absent

*MON88913: genetically modified cotton tolerant to glyphosate; MON15985: genetically modified cotton resistant to Lepidoptera; 281-24-236: genetically modified cotton resistant to insects; 3006-210-23: genetically modified cotton tolerant to the herbicide glufosinate ammonium (Widestrike cotton); GHB614: genetically modified cotton tolerant to the herbicide glyphosate (GlyTol cotton); LLCotton25: genetically modified cotton tolerant to the herbicide glufosinate ammonium (LibertyLink cotton); T304-40 x GHB119: genetically modified cotton tolerant to the herbicide glufosinate ammonium and resistant to Lepidoptera (TwinLink Cotton).

RESULTS AND DISCUSSION

Analysis of the phenotypic characters of the cotton genotypes

Base populations can be studied to select parent plants that would provide promising crosses, genetic diversity and satisfactory agroeconomic performance (Cruz *et al.*, 2011). Thus, the mean genotype values of the characters were grouped in order to evaluate the potential of each cultivar.

Lint yield ranged from 35.67% to 45.84% (Table 3). UFUJP-P, UFUJP-H and BRS 433 FL B2RF had the lowest yields while FM 980 GLT, TMG 47 B2RF and IMA 8405 GLT had the highest values, which were greater than the 40% target used in breeding programs (Vidal Neto & Freire, 2013). Carvalho *et al.* (2003) found lint yields from 26.4% to 41.7% for genotypes in the EPAMIG germplasm bank, whereas Santos *et al.* (2017) found lint yields from 41.37% to 45.56% for cultivars from various breeding programs.

Table 2: Field evaluations at each developmental stage of the cotton plant according to the cotton scale (*Escala do Algodão*) proposed by Marur and Ruano (2001)

V5*	FF	B1	MAT
Plant height	Plant height	Unopened boll diameter	Plant height
Stem Diameter	Stem Diameter	Number of reproductive structures	Stem Diameter
Leaf Area	Leaf Area		Number of reproductive structures
	Height of the first reproductive branch		Mean boll weight
	Number of nodes		Number of nodes
	Number of vegetative branches		Seed cotton yield (kg ha ⁻¹)
	Number of reproductive structures		Seed plus lint yield (kg ha ⁻¹)
			Yield (%)

*V5: plants with 5 fully developed leaves, FF: plants in full flower, B1: plants with first boll; MAT: plants at full maturity.

Table 3: Mean of lint percentage, seed cotton yield, lint yield, and seed yield of cotton genotypes

Genotype	LP	SCY	LY	SY
BRS 368RF	39.23c	3988.07a	1535.03c	2391.81b
BRS 372	40.14c	3694.78a	1486.72c	2169.86b
BRS 433FL B2RF	35.87d	3103.32a	1149.74c	1981.82b
DP Delta Opal	38.41c	3348.02a	1290.61c	2034.37b
DP 1228 B2RF	39.52c	3354.68a	1326.43c	1952.58b
DP 1552 B2RF	41.67b	4702.32a	1960.80b	2712.77b
FM 975 WS	40.12c	3847.97a	1542.83c	2295.93b
FM 980 GLT	45.84a	3075.26a	1414.23c	1634.75b
FM 982 GL	39.46c	4051.40a	1620.85c	2333.89b
IMA 2106 GL	41.33b	3651.61a	1509.81c	2115.33b
IMA 5675 B2RF	39.90c	3705.07a	1487.60c	2216.89b
IMA 8405 GLT	44.10a	3575.72a	1566.28c	2021.25b
TMG 45 B2RF	41.07b	6834.01a	2763.37a	4090.61a
TMG 47 B2RF	44.46a	4023.33a	1810.73c	2137.07b
TMG 82 WS	38.85c	4688.51a	1804.00c	2874.73b
UFUJP-H	35.74d	3353.12a	1501.64c	2186.27b
UFUJP-P	35.67d	3759.09a	1427.54c	2576.18b
UFUJP-B	36.63d	2783.48a	1009.65c	1796.35b
Mean	39.89	3688.57	1543.94	2306.80
F trat	17.55*	2.24*	3.18*	2.61*
CV (%)	3.04	26.78	23.78	25.10
CVg/Cve	2.35	0.64	0.85	0.73
h ² (%)	94.30	55.50	68.58	61.63

¹Means followed by same lowercase letters within a column belong to different groups (Scott-Knott test, $p < 0.05$). LP: lint percentage (%); SCY: seed cotton yield (kg ha⁻¹); LY: lint yield (kg ha⁻¹); SY: seed yield (kg ha⁻¹). * Significant at 5%, F test; CV: coefficient of variation; CVg/Cve: ratio between the genetic and environmental coefficients of variation; h²: coefficient of genotypic determination.

The genotypes did not cluster by cotton seed yield, which may be due to high CV values. However, two and three groups were found for seed and lint yield, respectively, with TMG 45 B2RF producing the highest values (Table 3). DP 1552 B2RF, FM 982 GLT, TMG 45 B2RF, TMG 47 B2RF and TMG 82 WS produced higher yields than the averages for the 2017/18 crop in Minas Gerais, Brazil (3,966 kg ha⁻¹ of seed cotton, 1,586 kg ha⁻¹ of lint and 2,380 kg ha⁻¹ of seed) (Conab, 2018).

Globally, cotton bolls average 61.5% seeds and 38.5% fiber and in general, larger seeds indicate lower fiber quantities (Beltrão, 2001). This assertion was also determined by Fang *et al.* (2017) who showed that cotton plants contain two gene loci that control the pleiotropic and inverse association of fiber percentage with seed index and seed size. Most of the seed cotton yield from the PROMALG UFU genotypes was due to superior seed yield. Specifically, the percentages of seed weight to total weight for UFUJP-H, UFUJP-P and UFUJP-B were 65.20%, 68.53% and 64.54%, respectively, which were higher than those of the other genotypes.

Intrinsic fiber quality is cultivar specific and although influenced by environmental conditions, is mainly controlled by genetic factors (Freire, 2015). Understanding associations among characters is essential in choosing breeding strategies. Therefore, this paper analyzes correlations among characters and discusses mean indicators of fiber quality, where CV_g/CV_e and h² values were the highest.

Genotypic correlations were greater than or equal to phenotypic correlations in 91.67% of the pairs of characters, suggesting that genetic factors contributed more to these correlations than did environmental factors (Table 4). Phenotypic correlations have greater practical value since selections are generally based on phenotype (Cruz *et al.*, 2012). The strongest and most significant phenotypic correlations were between the short fiber index and length uniformity (-0.92), maturity and micronaire (0.88), and resistance and length (0.76).

Four groups were formed according to fiber quality, with mean fiber lengths (UHM) varying from 27.47 to 32.08 (Table 5). Only TMG 45 B2RF had medium-length fibers, while the rest were classified as having long fibers, which is the goal of breeding programs. BRS 433 FL B2RF had the longest fibers (32.08 mm), which were greater than those proposed by Embrapa (2002) and classified as extra-long according to Cotton Incorporated (2018).

BRS 433 FL B2RF produced the greatest average fiber length but lower lint yield. Similar results were found by Carvalho *et al.* (2015), who found a negative correlation between yield and fiber length in extra-long fiber cultivars, but many breeders break this linkage (Kennedy, 2018). However, given that only BRS 433 FL B2RF had extra-long fibers (Embrapa, 2002), no correlation could be shown between these characteristics. Santos *et al.* (2017) state that some genotypes may have both higher average fiber yields and fiber lengths, as was observed in FM 980 GLT, IMA 8405 GLT and TMG 47 B2RF.

Table 4: Phenotypic and genotypic correlations among seed cotton yield, lint percentage, fiber length, fiber length uniformity index, short fiber index, micronaire index, fiber strength, fiber elongation and fiber maturity of 18 cotton genotypes

Character	Correlations	LP	UHM	LU	SFI	MIC	STR	ELG	MAT
SCY	rf	0.19 ^{ns}	-0.55*	-0.04 ^{ns}	0.15 ^{ns}	0.17 ^{ns}	-0.43 ^{ns}	0.33 ^{ns}	-0.04 ^{ns}
	fg	0.27 ^{ns}	-0.80 ⁺⁺	-0.14 ^{ns}	0.24 ^{ns}	0.17 ^{ns}	-0.63 ⁺	0.46 ^{ns}	-0.14 ^{ns}
RP	rf		-0.15 ^{ns}	-0.17 ^{ns}	0.39 ^{ns}	0.24 ^{ns}	0.05 ^{ns}	-0.27 ^{ns}	0.34 ^{**}
	fg		-0.18 ^{ns}	-0.22 ^{ns}	0.44 ⁺	0.24 ^{ns}	0.05 ^{ns}	-0.29 ^{ns}	0.35 ⁺⁺
UHM	rf			0.43 ^{ns}	-0.58*	0.28 ^{ns}	0.76 ^{**}	-0.29 ^{ns}	0.45 ^{ns}
	fg			0.41 ^{ns}	-0.56 ⁺	0.33 ^{ns}	0.77 ⁺⁺	0.40 ^{ns}	0.50 ⁺
LU	rf				-0.92 ^{**}	0.42 ^{ns}	0.63 ^{**}	0.38 ^{ns}	0.26 ^{ns}
	fg				-0.95 ⁺⁺	0.49 ^{ns}	0.70 ⁺⁺	0.09 ^{ns}	0.29 ^{ns}
SFI	rf					-0.40 ^{ns}	-0.63 ^{**}	-0.40 ^{ns}	-0.22 ^{ns}
	fg					-0.43 ^{ns}	-0.65 ⁺	-0.42 ^{ns}	-0.22 ^{ns}
MIC	rf						0.53*	0.11 ^{ns}	0.88 ^{**}
	fg						0.63 ⁺	0.12 ^{ns}	0.86 ⁺⁺
STR	rf							-0.22 ^{ns}	0.66 ^{**}
	fg							-0.22 ^{ns}	0.74 ⁺⁺
ELG	rf								-0.37 ^{ns}
	fg								-0.40 ^{ns}

** *: Significant according to the t test at 0.01 and 0.05; ++, + significant according to the bootstrap test with 10000 simulations at 0.01 and 0.05. Estimates of phenotypic (rf) and genotypic (rg) correlations among characters: seed cotton yield (SCY); lint percentage (LP); fiber length (UHM); fiber length uniformity index (LU); short fiber index (SFI); micronaire index (MIC); fiber strength (STR); fiber elongation (ELG); fiber maturity (MAT).

UHM is correlated with resistance (STR). In other words, longest fibers more strength and fineness. The same was found in the present study and evidenced by a strongly positive correlation ($r_f = 0.76$). The correlation data also showed that STR is directly related to the length uniformity index ($r_f = 0.63$) and micronaire ($r_f = 0.53$), and inversely associated with the short fiber index ($r_f = -0.63$). These correlations are important for cotton breeding since they show that the selection of only one of these characters, preferably the one with the highest coefficient of genotypic determination (UHM), favors the simultaneous selection of the other characters.

The average STR values of these genotypes were ideal (greater than 27 gf tex⁻¹) and could be separated into four groups. The BRS 433 RF B2RF genotype alone classified as very strong (34.92 gf tex⁻¹), while the others were classified as highly resistant (greater than 28.09 gf tex⁻¹).

A lower short fiber index (SFI) is desirable since shorter fibers perform better in processing and result in better quality yarns (Santos *et al.*, 2017; Cordão Sobrinho *et al.*, 2015). BRS 433 FL B2RF had the lowest mean SFI (5.72%) of the group and was classified as very low (Embrapa 2002). Negative correlations between SFI and UHM ($r_f =$

-0.58), SFI and STR ($r_f = -0.63$), SFI and LU ($r_f = -0.92$) showed that genotype BRS 433 FL B2RF had the highest average resistance, length uniformity, and elongation (very high), which suggest that the fibers from this cultivar would suffer limited damage during processing. The same characteristics were found in genotypes studied by Santos *et al.* (2017).

The genotypes could be separated into four groups according to fiber elongation. DP 1552 and BRS 368, DP Delta Opal, TMG 45, UFUJP-H, UFUJP-P and UFUJP-H produced the highest averages (greater than 7.6%) and were classified as having very high elongation. The mean elongation values of all genotypes, except TMG 47 B2RF (medium elongation, 6.74%), showed that the fibers should withstand the high accelerations of processing without rupturing (Belot, 2018).

Three groups were formed among the genotypes regarding uniformity of length (LU). The fiber of the group composed of BRS 433 FL B2RF (85.00%) and DP 1552 B2RF (84.92%) had higher LU values and was classified as uniform, which is better for spinning (Lana *et al.*, 2014). The other groups had medium to uniform LU values that were close to the 83% target of breeding programs (Cunha Neto *et al.*, 2015).

Table 5: Mean of fiber length, fiber length uniformity index, short fiber index, micronaire index, fiber strength, fiber elongation, and fiber maturity of cotton genotypes

Genotype	UHM	LU	SFI	MIC	STR	ELG	MAT
BRS 368RF	28.35c	82.67c	8.78d	3.20b	30.04d	8.01b	82.32b
BRS 372	29.67b	82.94c	8.28c	3.76a	33.02b	7.45c	84.21a
BRS 433FL B2RF	32.08a	85.00a	5.72 ^a	3.79a	34.92a	7.65c	84.20 a
DP Delta Opal	29.16b	82.64c	8.05c	3.68a	31.55c	7.91b	83.61 a
DP 1228 B2RF	28.46c	83.32b	8.00c	3.83a	31.67c	7.51c	84.26 a
DP 1552 B2RF	29.13b	84.92a	6.74b	4.15a	31.31c	9.03a	83.95 a
FM 975 WS	29.15b	82.77c	8.71 ^a	3.46a	31.44c	7.42c	83.42 a
FM 980 GLT	29.59b	83.84b	7.66c	3.36b	32.71b	7.78c	82.97 b
FM 982 GL	29.58b	82.80c	8.23c	3.71a	31.26c	7.40c	84.05 a
IMA 2106 GL	30.13b	82.14c	8.82d	3.55a	31.30c	7.25c	83.74 a
IMA 5675 B2RF	28.36c	82.08c	9.41d	3.05b	28.09d	7.69c	82.10b
IMA 8405 GLT	29.07b	81.24c	9.65d	3.75a	29.75d	7.38c	84.09a
TMG 45 B2RF	27.47d	82.27c	8.64d	3.57a	28.12d	8.46b	82.83b
TMG 47 B2RF	29.87b	82.17c	8.90d	3.57a	30.61d	6.74d	84.11 a
TMG 82 WS	28.30c	83.18c	8.26c	3.28b	29.77d	7.38c	82.94 b
UFUJP-H	29.56b	83.55b	7.37b	3.21b	29.95d	8.12b	82.27b
UFUJP-P	29.25b	82.65c	7.92c	3.41b	28.94d	8.09b	82.74b
UFUJP-B	29.12b	81.51c	8.60d	3.12b	29.32d	7.99b	82.11b
Mean	29.24	82.87	8.21	3.52	30.76	7.74	83.33
F trat	12.83*	3.75*	8.92*	4.04*	6.37*	11.50*	4.88*
CV (%)	1.61	1.08	6.59	7.05	3.89	3.39	0.75
CVg/Cve	1.99	0.96	1.63	1.01	1.34	1.87	1.14
h ² (%)	92.21	73.3	88.8	75.28	84.3	91.31	79.53

*Means followed by distinct lowercase letters within a column belong to distinct groups, Scott-Knott (0.05). UHM: fiber length (mm); LU: fiber length uniformity index (%); SFI: short fiber index (%); MIC: micronaire index; STR: fiber strength (gf tex⁻¹); ELG: fiber elongation (%); MAT: fiber maturity (%). * F test (5%); CV: coefficient of variation; CVg/CVe: ratio between the genetic and environmental coefficients of variation; and h²: coefficient of genotypic determination.

The micronaire index of the genotypes varied from 3.05 to 4.15. Except for DP 1552 B2RF (4.15), all genotypes had fine fibers that fell within the optimal range defined by breeding programs (3.6 to 4.2 mm) (Freire, 2015).

Although micronaire is important, it should not be considered in isolation given its correlation with MAT (rf = 0.88). Thus, fibers with low MIC and high MAT values are the most desirable since they produce fine, strong yarns and fabrics (Cunha Neto *et al.*, 2015). In addition to the micronaire index, lint yield (rf = 0.34) and fiber strength (rf = 0.66) were also significantly correlated with maturity.

Although two groups were formed regarding fiber maturity, all genotypes were classified as below average according to Embrapa (2002). However according to Santana *et al.* (2008), maturity values greater than 80%, as was the case for all genotypes in this study, indicate that the fibers are capable of maximizing dye uptake and retention (Belot, 2018).

Genetic divergence

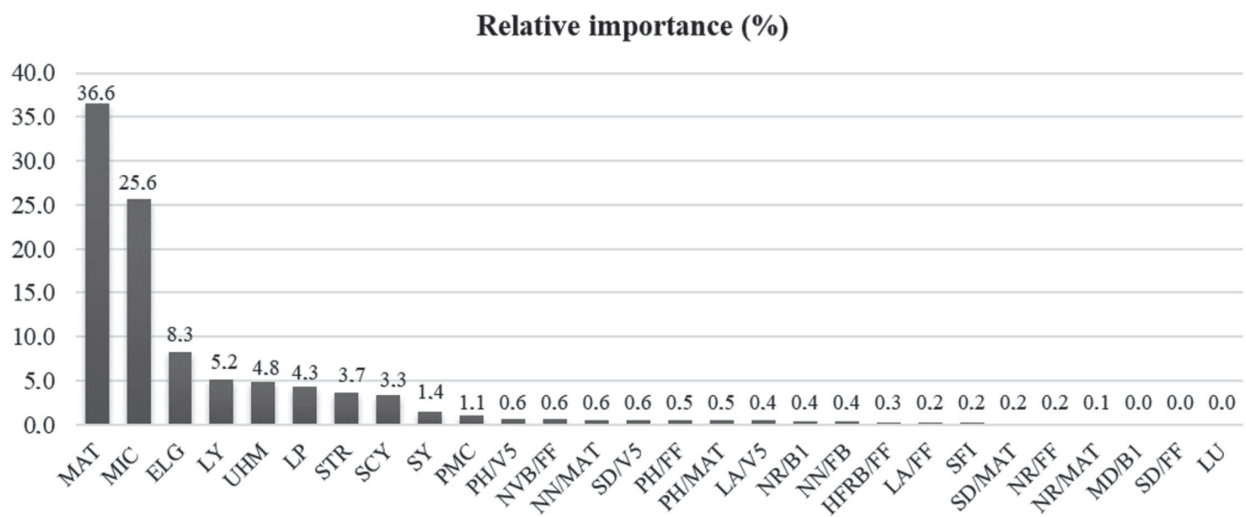
Fiber maturity provided the highest relative contribution (36.57%) to genetic divergence (Singh, 1981), followed by micronaire (25.61%) and fiber elongation (8.31%) (Figure 1). In contrast, Cunha Neto *et al.* (2015) evaluated divergence among genotypes from the same breeding programs with white and colored fibers and found that the technological characteristics of the fiber contributed least to genetic diversity, while fiber yield and percentage contributed the most.

LU, SD / FF, SD / MAT, MD / B1, NR / MAT, NR / FF, SFI, LA / FF contributed little to the detection of genetic diversity. The lower genotype determination coefficients and lower CVg / CVe ratios explain the reduced importance of most of these characters, but not of SFI, MD / B1 and LU, which may have resulted from variations already represented by other characters (Cruz *et al.*, 2011).

The relative importance of character data showed that the lint quality characters were more important than the agronomic variables (Figure 1). This may have occurred because these cultivars came from breeding programs and have already reached advanced levels for yield and production but are still evolving in quality. Fang *et al.* (2017) performed genomic analysis and detected more gene loci associated with lint yield than with fiber quality. This suggests that breeding for higher lint yield has been the main emphasis of cotton breeding over time.

The dendrogram based on UPGMA clustering and the Mahalanobis generalized distance (Figure 2) shows that the genotypes are separated into six clusters. Torcher clustering also separated the 18 genotypes into six distinct clusters (Table 6).

Combining the UPGMA and Tocher methods guarantees good estimates of genetic divergence (Gilio *et al.*, 2017). These methods appear to be in partial agreement since they both produced the same number of clusters; however, the constituents of these clusters do differ somewhat due to the different ways of calculating genetic dissimilarity and of defining proximity between an indivi-



*PH/V5: plant height at V5; SD/V5: stem diameter at V5; LA/V5: leaf area at V5; PH/FF: plant height at full flower; SD/FF: stem diameter at full flower; LA/FF: leaf area at full flower; NN/FF: number of nodes at full bloom; NVB/FF: number of vegetative branches at full flower; HFRB/FF: height of the first reproductive branch at full bloom; NR/FF: number of reproductive structures at full bloom; NR/B1: number of reproductive structures at onset of first boll; MD/B1: mean diameter of an unopened boll at onset of first boll; PH/MAT: plant height at full maturity; SD/MAT: stem diameter at full maturity; NN/MAT: number of nodes at full maturity; NR/MAT: number of reproductive structures at full maturity; PMC: mean boll weight; SCY: seed cotton yield; LY: lint yield; SY: seed yield; LP: lint percentage; UHM: fiber length; LU: length uniformity; SFI: short fiber index; MIC: micronaire index; STR: fiber strength; ELG: fiber elongation; MAT: fiber maturity.

Figure 1: Estimates of the relative importance of characters affecting genetic divergence in cotton genotypes.

dual and an existing group or between any two groups (Buttow *et al.*, 2010; Cruz *et al.*, 2011).

All clustering methods distributed the Embrapa, TMG and Fibermax cultivars into three clusters, except the Tocher method which distributed the Fibermax genotypes into two groups. The IMA cultivars separated into two clusters regardless of method; however, the composition of these clusters did vary by method. The consistent distribution of these cultivars suggests similarities within the breeding programs.

Regardless of clustering method, the PROMALG UFU genotypes were always found within a single cluster, while only one of the Monsanto DeltaPine cultivars (DP 1552 B2RF) was found in a separate cluster, indicating genetic similarity within these breeding programs.

Although the PROMALG UFU genotypes were arranged in a cluster with other cultivars by the Tocher optimization method, they were isolated by the UPGMA method. This differentiation occurred because the UFU genotypes were bred by crossing *Gossypium hirsutum* L. with *Gossypium barbadense* L. in order to incorporate the

excellent fiber quality of *G. barbadense* L. in *G. hirsutum* L. cultivars.

UPGMA clustering and the Tocher optimization method grouped DP 1552 B2RF and BRS 433 B2RF separately from the other genotype, while BRS 433 B2RF was grouped separately by all clustering methods because of its dissimilarity to the other genotypes. The distance between these cultivars and the rest may be due to superior fiber quality (Table 4), especially BRS 433 B2RF, and lint yield (DP 1552 B2RF) (Table 3).

In general, the largest clusters consisted of genotypes from the greatest number of different breeding programs. Bertini *et al.* (2006) also identified clusters of cultivars from various breeding programs, suggesting that different programs share similar germplasm. In addition, Amalraj (1982), Singh & Gill (1984) found that no relationship between genetic and geographic diversity in cotton, since varieties from the same geographical origin could be found in different groups. They attributed this phenomenon to the selection and adaptation of populations. So, breeders should be depending on genetics rather than geographical distribution.

Dissimilarity within the clustering methods demonstrates divergence among these genotypes. However, at least one cultivar was similar across all the breeding programs. These cultivars should not be used in future crosses to maintain genetic variability and ensure selection gains. Genetically related parents tend to share many genes or alleles and produce crosses with low levels of allelic heterozygosity and consequently low levels of vigor (Cruz, 2012).

Proximity between genotypes suggests that similar germplasm sources with the same alleles have been used to breed Brazilian cotton cultivars, which have also been influenced by varieties from the United States and Australia. According to Gutiérrez *et al.* (2002), these varieties have a narrow genetic base that has been selected from existing cultivars, while underutilizing wild germplasm (Penna, 2005; Iqbal *et al.*, 2001; Bertini *et al.*, 2006). Bertini *et al.* (2006) found that many cotton cultivars are descended from a few parents and states that new alleles need to be introduced into cotton breeding.

The mean Mahalanobis distances (Table 7) were highest in BRS 433 B2RF and FM 980 GLT, demonstrating

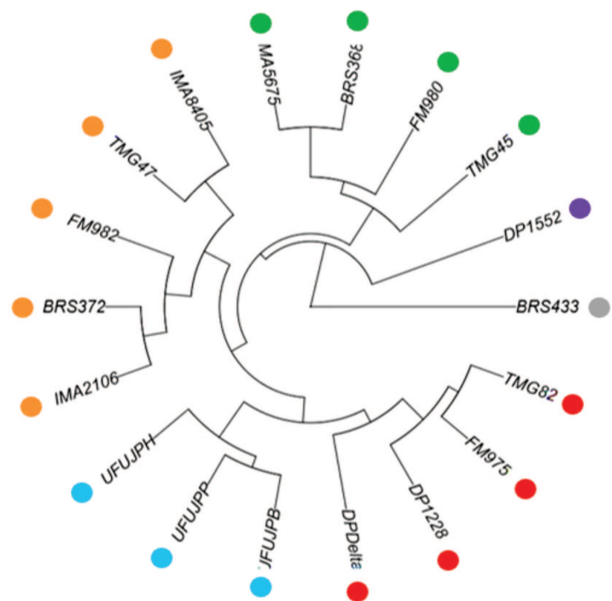


Figure 2: Dendrogram of genetic divergence among 18 genotypes of cotton, provided by the UPGMA clustering method and based on the generalized distance of Mahalanobis (D^2). Colored circles represent distinct clusters as defined by Mojena (1977). Cophenetic correlation coefficient (r): 0.60.

Table 6: Clustering of 18 cotton genotypes by the Tocher method and the generalized distance of Mahalanobis

Cluster	Nº of genotypes	Genotypes
1	6	BRS 368RF, IMA 5675 B2RF, FM 975 WS, IMA 8405 GLT, FM 980 GLT, TMG 45 B2RF
2	6	UFUJP- P, UFUJP- B, UFUJP- H, BRS 372, FM 982 GL, IMA 2106 GL
3	3	DP 1228 B2RF, TMG 82 WS, DP Delta Opal
4	1	DP 1552 B2RF
5	1	TMG 47 B2RF
6	1	BRS 433 FL B2RF

Table 7: Genetic divergence among 18 cotton genotypes based on 28 characters and the generalized Mahalanobis distance

Genotype	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
2	329.2																
3	862.2	370.2															
4	202.9	271.9	453.4														
5	323.4	206.9	504.8	310.9													
6	323.5	269.7	599.9	208.0	525.1												
7	146.5	112.0	471.9	149.9	115.3	263.6											
8	234.8	376.5	1121.3	505.4	431.1	570.4	252.9										
9	413.3	150.5	286.9	332.8	146.1	508.9	133.4	443.8									
10	428.3	95.7	328.1	322.6	362.2	285.2	153.5	426.9	167.6								
11	68.6	296.4	825.0	284.8	310.1	386.1	118.8	170.8	364.9	314,0							
12	255,1	167,7	684,9	330,3	328,9	360,1	105,4	158,4	227,7	125,6	138,4						
13	260,1	266,6	836,4	468,0	256,0	453,0	198,3	205,6	269,8	350,4	244,4	219,8					
14	420,7	344,3	757,3	436,5	474,7	502,0	214,1	257,2	311,0	198,2	276,0	110,7	375,8				
15	193,9	249,6	540,6	144,2	130,7	360,4	92,6	432,5	237,6	354,4	237,7	315,2	269,5	412,1			
16	465,7	193,5	249,3	324,6	296,3	394,2	296,2	762,1	278,4	315,9	501,2	528,2	450,6	764,5	273,7		
17	487,0	260,4	241,8	220,4	351,3	306,2	268,6	844,6	289,0	311,5	552,1	522,1	522,6	627,5	201,0	121,2	
18	296,1	155,6	234,4	174,6	193,3	294,4	120,4	529,9	142,4	175,5	285,8	285,5	323,9	391,9	151,2	125,0	91,7

1: BRS 368RF; 2: BRS 372; 3: BRS 433FLB2RF; 4: DP Delta Opal; 5: DP 1228 B2RF; 6: DP 1552 B2RF; 7: FM 975 WS; 8: FM 980 GLT; 9: FM 982 GL; 10: IMA 2106 GL; 11: IMA 5675 B2RF; 12: IMA 8405 GLT; 13: TMG 45 B2RF; 14: TMG 47 B2RF; 15: TMG 82 WS; 16: UFUJP-H; 17: UFUJP-P; 18: UFUJP-B.

that these genotypes were the most divergent among those tested.

Direct comparison between BRS 433 B2RF and FM 980 GLT shows that the fiber quality of BRS 433 B2RF was greater than the other genotypes in this study and greater than Brazilian standards. Morello *et al.* (2017) also found that the fiber quality of BRS 433 FL B2RF was among the best of the cultivars grown in Brazil.

Hybridization between clusters is more efficient than within cluster to produce better progenies. This due to genetic dissimilarity is lower within clusters than between clusters.

Conversely, FM 980 GLT showed better production results (MD/B1, NN/MAT), especially lint yield, but lower fiber quality. According to Zeng *et al.* (2018), there may be a negative relationship between fiber yield and fiber quality.

Understanding the distance between genotypes is important for choosing the parents used in breeding since hybridization choices should be based on the magnitude of dissimilarities (Santos *et al.*, 2017). Thus, in order to obtain genetic gains and to make the most of genetic variability among this set of cultivars, hybridizations between BRS 433 B2RF and FM 980 GLT could yield segregate populations with higher productive potential, lint yield and fiber quality.

BRS 433 B2RF was isolated from the other clusters, was generally quite divergent from the other genotypes and was least distant from UFUJP-B (234.41), UFUJP-P (241.80) and UFUJP-H (249.31). This probably resulted from crossing of germplasm with greater genetic variability, which provided genetic gains and generated recombinants. This in turn allowed the selection of the best genetic combinations and explains the excellent fiber quality of BRS 433 B2RF.

Crosses with FM 980 GLT should improve the genetics of the breeding program at UFU since this cultivar is genetically distant and has strong agronomic characteristics, good fiber quality and high lint yield that could improve the proportion of fiber and seed in the cotton bolls of PROMALG's genetic material.

Genetic diversity among the cotton genotypes in this study ranged from 68.64 to 1121.27 (Table 7). This amplitude echoes that of Araújo *et al.* (2014) who found distances of 23.01 to 1172.28 while evaluating the fiber quality and productivity of 11 cultivars resulting only from Embrapa's breeding program. These results indicate the limited genetic base of current cultivars.

Furthermore, the relationship between the highest and lowest value of D^2 in the present study was 16.34. This suggests moderate genetic divergence according to Nardino *et al.* (2017) and Paixão *et al.* (2008) who classified values of 33.6 and 39.0, respectively, as wide variability

among corn genotypes. These findings show that cotton breeding programs in Brazil need to make better use of genetic resources.

CONCLUSIONS

Moderate genetic divergence was found among genotypes from the main cotton breeding programs in Brazil.

To make the most of the genetic variability among these cultivars, hybridizations between BRS 433 B2RF and FM 980 GLT should generate segregate populations with higher productive potential, lint yield and fiber quality. Crosses between the genotypes developed by PROMALG and FM 980 GLT could increase fiber yield and quality of UFUJP-B, UFUJP-H and UFUJP-P.

Fiber quality, maturity, micronaire and fiber elongation contributed more to detecting genetic divergence than did the other characters evaluated in this study.

The fiber quality traits had higher heritability more than yield components. Therefore, selection for fiber traits is facilitated and can be done in the early generation. While for yield traits, the selection should be in an early generation or direct in late generations.

CONFLICT OF INTERESTS

There is no conflict of interests in carrying the research and publishing the manuscript.

REFERENCES

- Amalraj SFA (1982) Genetic divergence in *Gossypium hirsutum* L. *Genetic Agriculture*, 36:23-30.
- Araújo LF, Almeida WS de, Bertini CHCM, Vidal Neto FC & Bleicher E (2014) The use of different clustering methods in the evaluation of genetic diversity in upland cotton. *Revista Ciência Agronômica*, 45:312-318.
- Belot J (2018) A indústria têxtil e a qualidade de fibra. Primavera do Leste, IMAMT. 191p. (Boletim técnico).
- Beltrão NEM (2001) Componentes da produção na cotonicultura: uma visão integrada. In: Congresso Brasileiro de Algodão, Campo Grande. Anais, EMBRAPA-CNPA. p.605- 608.
- Bertini CHCM, Schuster I, Sediya T, Barros EG & Moreira MA (2006) Characterization and genetic diversity analysis of cotton cultivars using microsatellites. *Genetics and Molecular Biology*, 29:321-329.
- Borém A & Miranda GV (2013) Melhoramento de plantas. 6th ed. Viçosa, Editora UFV. 523p.
- Buttow MV, Barbieri RL, Neitzke RS, Heiden GE & Carvalho FIF (2010) Diversidade genética entre acessos de pimentas e pimentões da Embrapa Clima Temperado. *Ciência Rural*, 40:1264-1269.
- Carvalho LP, Farias FJC & Rodrigues JIS (2015) Selection for increased fiber length in cotton progenies from Acala and Non-Acala types. *Crop Science*, 55:1-7.
- Carvalho LP, Farias FJC, Rodrigues JIS & Suassuna ND (2017) Genetic diversity among exotic cotton accessions as for qualitative and quantitative traits. *Genetics and Molecular Research*, 16:1-10.

- Carvalho LP, Lanza MA, Fallieri J & Santos JW (2003) Análise da diversidade genética entre acessos de banco ativo de germoplasma de algodão. *Pesquisa Agropecuária Brasileira*, 38:1149-1155.
- Companhia nacional de abastecimento (2018) Acompanhamento da safra brasileira grãos safra 2017/18. Brasília, CONAB. 148p. (Boletim técnico, 12).
- Companhia nacional de abastecimento (2020) Acompanhamento da safra brasileira grãos safra 2019/20. Brasília, CONAB. 73p. (Boletim técnico, 10).
- Cordão Sobrinho FP, Guerra HOC, Araujo WP, Pereira JR, Zonta JH & Bezerra JRC (2015) Fiber quality of upland cotton under different irrigation depths. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 19:1057-1063.
- Cotton incorporated (2018) Ratings of fiber properties. Available at: <https://www.cottoninc.com/cotton-production/quality/us-cotton-fiber-chart/ratings-of-fiber-properties/>. Accessed on: August 24th, 2018.
- Cruz CD (2013) GENES: a software package for analysis in experimental statistics & quantitative genetics. *Acta Scientiarum Agronomy*, 35:271-276.
- Cruz CD, Ferreira FM & Personi LA (2011) *Biometria aplicada ao estudo da diversidade genética*. 1st ed. Viçosa, Suprema. 620p.
- Cruz CD, Regazzi AJ & Carneiro PCS (2012) *Modelos biométricos aplicados ao melhoramento genético*. Viçosa, UFV. 668p.
- Cunha Neto J, Bertini CHCM & Silva APM (2015) Divergência genética entre genitores de algodoeiro de fibras brancas e coloridas. *Revista Brasileira de Ciências Agrárias*, 10:492-498.
- EMBRAPA (2002) Resultados de ensaio HVI e suas interpretações (ASTM D-4605). Campinas Grande, EMBRAPA. 13p.
- Fang L, Wang Q, Hu Y, Jia Y, Chen J, Liu B & Mei G (2017) Genomic analyses in cotton identify signatures of selection and loci associated with fiber quality and yield traits. *Nature Genetics*, 49:1089-1098.
- Freire EC (2015) Fatores que afetam a qualidade das fibras In: Freire EC (Ed.) *Algodão no cerrado do Brasil*. Brasília, Positiva. p.653-750.
- Gilio TAS, de Araújo DV, Krause W, Rosa HHR & Ascari JP (2017) Divergência genética em genótipos de algodão em condições de safra e safrinha. *Revista Caatinga*, 30: 377-390.
- Gutiérrez AO, Basu S, Saha S, Jenkins JN, Shoemaker CL, Cheatham CL & Mccarty JC (2002) Genetic distance among selected cotton genotypes and relationship with F2 performance. *Crop Science*, 42:1841-1847.
- Iqbal MJ, Reddy OUK, El-zik KM & Pepper AE (2001) A genetic bottleneck in the evolution under domestication of upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. *Theoretical and Applied Genetics*, 103:547- 554.
- Kennedy HD (2018) *Selection and Response of Yield and Fiber Traits in Upland Cotton*. Tese de Doutorado. Texas A&M University, College Station. 94p.
- Lana V, Ribas PV & Chanselme J (2014) A classificação do algodão. Primavera do Leste, IMAMT. 369p. (Boletim técnico).
- Ludke WH, Andrade ACB, Volpato L, Almeida DP, Oliveira ICM, Paiva JT, Silva MJ, Del Conte MV, Silva TC, Almeida VC & Pinto VB (2017) *Desafios biométricos no melhoramento genético*. 1st ed. Viçosa, GenMelhor. 166p.
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27:209-220.
- Marur CJ & Ruano O (2017) Escala do Algodão. Available at: http://www.iapar.br/arquivos/File/zip_pdf/EscaladoAlgodao.pdf. Accessed on: March 31th, 2017.
- Mehboob-ur-rahman Shaheen T, Tabbasam N, Atif Iqbal ML, Ashraf M, Zafar Y & Paterson AH (2012) Cotton genetic resources. A review. *Agronomy for Sustainable development*, 32:419-432.
- Milligan GW & Cooper MC (1985) An Examination of Procedures for Determining the Number of Clusters in a Data Set. *Psychometrika*, 50:59-179.
- Mojena R (1977) Hierarchical grouping methods and stopping rules: an evaluation. *The Computer Journal*, 20:359-363.
- Morello CL, Suassuna ND, Pedrosa MB, Silva Filho JL, Barroso PAV, Suassuna TMF, Perina FJ, Sofiatti V, Farias FJC & Magalhaes FOC (2017) BRS 433 FL B2RF: Cultivar de algodoeiro de fibra longa com resistência a Lepidópteros e tolerância ao herbicida Glifosato. In: Congresso Brasileiro de Algodão, Maceió. Anais, Associação Brasileira dos Produtores de Algodão.
- Nardino M, Baretta D, Carvalho IR, Follmann DN, Ferrari M, Pelegrin AJ, Szareski VJ, Konflanz VA & Souza VQ (2017) Genetic divergence among corn (*Zea mays* L.) genotypes in distinct environments. *Revista de Ciências Agrárias*, 40:164-174.
- Paixão SL, Cavalcante M, Ferreira PV, Madalena JASE & Pereira RG (2008) Divergência genética e avaliação de populações de milho em diferentes ambientes no estado de Alagoas. *Revista Caatinga*, 21:191-195.
- Penna JCV (2005) Melhoramento do algodão. In: Borém A (Ed.) *Melhoramento de espécies cultivadas*. Viçosa, UFV. p. 15-53.
- Santana JCF, Wanderley MJR, Beltrão NEM, Azevedo DMP, Leão AB & Vieira DJ (2008) Características da fibra e do fio do algodão: análise e interpretação dos resultados. In: Beltrão NEM & Azevedo DMP (Eds.) *Embrapa: O agronegócio do algodão no Brasil*. Brasília, Embrapa Informação Tecnológica. p.1099-1120.
- Santos IG, Teodoro PE, Farias FC, Farias FJC, Carvalho LP, Rodrigues JLS & Cruz CD (2017) Genetic diversity among cotton cultivars in two environments in the State of Mato Grosso. *Genetics & Molecular Research*, 16:1-8.
- Singh D (1981) The relative importance of characters affecting genetic divergence. *Indian Journal of Genetics and Plant Breeding*, 41:237-245.
- Singh TH & Gill SS (1984) Genetic diversity in upland cotton under different environments. *Indian Journal and Plant Breeding*, 44:506-513.
- Vidal Neto FC & Freire EC (2013) Melhoramento genético do algodoeiro. In: Neto FCV & Cavalcanti JJV (Eds.) *Melhoramento genético de plantas no Nordeste*. Brasília, Embrapa. p.49-83.
- Zeng L, Stetina SR, Erpelding JE, Bechere E, Turley RB & Scheffler J (2018) History and Current Research in the USDA-ARS Cotton Breeding Program at Stoneville, MS. *Journal of Cotton Science*, 22:24-35.