

## COMPARISON OF WILD ACCESSIONS OF *Gossypium. barbadense* L. FROM PERU AND BRAZIL VIA MICROSATELLITE MARKERS

### COMPARAÇÃO ENTRE ACESSOS SILVESTRES DE *Gossypium barbadense* L. DO PERU E BRAZIL VIA MARCADORES MICROSSATÉLITES

Josiane Isabela da Silva RODRIGUES<sup>1</sup>; Luiz Paulo de CARVALHO<sup>2</sup>;  
Francisco José Correia FARIAS<sup>2</sup>

1. Pós-doutoranda em Melhoramento de Plantas, Embrapa Algodão; 2. Pesquisador, Embrapa Algodão. josianeisabela@gmail.com

**ABSTRACT:** Peru and Brazil are recognized as primary and secondary centers of *Gossypium barbadense* L., which is a species that stands out for producing long cotton fiber. However, since its introduction in Brazil, *G. barbadense* L. populations have reduced its occurrence and genetic variability. Thus, the objective of this study was to compare the diversity level among wild accessions of *G. barbadense* L. from Peru and Brazil, for conservation and management of its genetic resource. Fifteen accessions from Peru, 13 from Brazil, and the cultivar Pima S7 were analyzed using 29 microsatellite markers. Genetic diversity in the 29 accessions was analyzed using descriptive statistics and methods based on distance index (UPGMA, 3D projection, and principal coordinate analysis), as well as ancestry model (Structure 2.3.4). Based on these analyses, it is verified that there is similar variability level between the Brazilian and the Peruvian accessions. At least 5 Brazilian accessions come from gene pools different from those of the rest of the Brazilian accessions. Peru's accessions originate from six different gene pools. Nine out of 29 accessions present genetic admixture of at least three different gene pools. Variability of Peruvian and Brazilian accessions may prove to be crucial for cotton crop worldwide.

**KEYWORDS:** *G. barbadense* L.. Molecular markers. Genetic diversity. Cotton. Genetic resources.

## INTRODUCTION

*Gossypium* spp. is a vegetable considered to be the main source of textile fiber. This genus belongs to the Malvaceae family, originated between 5-10 million years ago, and includes more than 50 species, being 45 diploids (A-G and K genomes,  $2n = 2x = 26$ ), and 5 allotetraploids (AD genome,  $2n = 4x = 52$ ). Due to the variety of species, *Gossypium* spp. is cytogenetically subdivided, and based on meiotic compatibility, it is divided in eight genomic groups (A-G and K). Of the total species, only four were domesticated by humans to produce fiber: *G. herbaceum* L. ( $2n = 2x = 26$ ), *G. arboreum* L. ( $2n = 2x = 26$ ), *G. hirsutum* L. ( $2n = 4x = 52$ ), and *G. barbadense* L. ( $2n = 4x = 52$ ) (WANG et al., 2012).

Species with allotetraploid genome (AD genome,  $2n = 4x = 52$ ) belong to the primary gene pool. They all had monophyletic origin 1 million to 2 million years ago, from a genomic duplication after hybridization between an invasive diploid species of genome A from the Old World and other genome D diploid species of the New World. Over thousands of years, this line spread in three modern lines, and two of them formed the economically important species *G. hirsutum* L. and *G. barbadense* L. (WENDEL et al., 2009).

In one of the modern lines, *G. mustelinum* Miers ex Watt, *G. darwinii* Watt and *G. tomentosum* Nuttall ex Seemann formed wild and endemic species in Brazil, in the Galapagos Islands and in Hawaii Islands, respectively (WESTENGEN et al., 2005; CAMPBELL et al., 2010). In other two lines, *G. hirsutum* L. and *G. barbadense* L. were independently domesticated for about four thousand years ago. Over pre-Columbian diffusion years, these species achieved fairly wide geographic distribution (WENDEL et al., 1995). Wild *G. hirsutum* L. is currently predominantly distributed in Mesoamerica and the Caribbean. *G. barbadense* L. is distributed in South America and in the Caribbean (HOVAV et al., 2008). Cultivars of both species occur in various parts of the world.

*G. hirsutum* L. is the most cultivated species in the world due to the yield potential and to the wide range of adaptation (CAMPBELL et al., 2010). The second most cultivated species, *G. barbadense* L., is cultivated due to the long, strong and fine fiber of its improved cultivars in parts of Asia, Egypt, Sudan, India, China, USA and the Andes (Colombia, Ecuador, Bolivia), especially in Peru. Due to its long fiber, it is highly valued (WESTENGEN et al., 2005).

The elite germplasm *G. barbadense* L., including Egyptian and Pima cottons, among others, are originated in "Sea Island" cottons. These cottons, in turn, were developed in the islands of

Georgia's and South Carolina's coasts, USA (WENDEL et al., 2010), and may have arisen from west Peruvian Andean introductions (HUTCHINSON, 1947) and/or west Indian introductions (STEPHENS, 1976). The exact origin of the first "Sea Island" cottons is still uncertain, but it is known that "Sea Island" lines contributed to the development of the Egyptian cottons. These Egyptian cottons were then reintroduced in the US, becoming part of the gene pool of Pima cottons (PATERSON et al., 2009; WENDEL et al., 2010).

The oldest archaeological remnant of *G. barbadense* L. is 5,500 years old, and is found in Peru's center-coastal region. Their primitive characteristics led to the hypothesis that it was originated from the first domesticated forms, and the initial domestication occurred in some part of this region (HUTCHINSON et al., 1947). Thus, from archaeological evidence, Hutchinson et al. (1947) proposed the *G. barbadense* L. diversity center in northwestern South America. On the other hand, Rossen et al. (1996) and Piperno and Pearsall (1998) proposed the initial domestication center in Peru's and Ecuador's coasts.

Soon after, based on isozymes, Percy and Wendel (1990) proposed the center of origin and diversity of *G. barbadense* L. in northwestern South America and/or west of the Andes, which in turn showed congruence with the distribution of wild populations. From this center, a dispersion route to the east of South America, the Caribbean and Central America was traced. It seems to have occurred a trans-Andean initial expansion from the west of the Andes to the northeast South America, and a second expansion to Central America, the Caribbean and the Pacific. New post-Columbian dispersion of the same gene pool seems to have occurred in the west Andean to Argentina and Paraguay. Thus, new trans-Andean dispersion seems to have occurred towards the northeast of South America and, from there, to Central America, to the Caribbean and to the Pacific (WENDEL et al., 2010). In fact, there are wild types of *G. barbadense* L. in South America, in southeast Meso-America, and in the Caribbean basin (FRYXELL, 1976). Finally, Westengen et al. (2005), based on AFLP, proposed the initial domestication center of the species around the Gulf of Guayaquil, in coastal regions of northwestern Peru and southwestern Ecuador, where the species would spread to the Andes and expand to its pre-Columbian range, including southern Bolivia, eastern Brazil, northern Colombia and Venezuela, and the Caribbean and Pacific Islands.

On the coast of Peru and Ecuador, native *G. barbadense* L. has characteristics intermediate to wild and domesticated types. There are also reports on wild types in the region, but they are not very clearly defined in terms of characteristics related to wild survival (PERCY, WENDEL, 1990). Wild populations, for example, are described close to Guayas and Losno Rivers, in Ecuador, and Tumbes River, Peru (PERCIVAL, KOHEL, 1990). In turn, indigenous *G. barbadense* L. is grown in gardens and farms in northwestern South America, where it is referred to as dooryard cotton or commensals. This cotton is typically a perennial shrub that reaches up to four meters and fiber in a variety of colors (WESTENGEN et al., 2005).

In Brazil, *G. barbadense* L. occurs as dooryard cotton, and its origin is traced to the forms of cultivation by local pre-Columbian indigenous people, and there are reports of two botanical races: "brasiliense", which is domesticated in the Amazon, with seeds attached one to another, also known as kidney-seed; and "quebradinho", with typical seed type. *G. barbadense* L. from the USA and from Israel was also grown in the northeast and southeast, in the colonial period, but it was gradually replaced by herbaceous cottons (*G. hirsutum* L. r. *latifolium*) and 'mocó' cotton (*G. hirsutum* L. r. *marie-galante*), which were also introduced in Brazil (ALMEIDA et al., 2009). Thus, since its introduction in eastern Brazil, *G. barbadense* L. populations decreased, and so did the genetic variability. Therefore, the objective of this study was to compare the diversity level of *G. barbadense* L. wild accessions from Peru, which is the center of origin, and from Brazil, and to evaluate the diversity as a whole for the conservation and management of the species genetic resource.

## MATERIAL AND METHODS

Fifteen *G. barbadense* L. accessions from Peru and 13 from Brazil were compared regarding genetic diversity (Table 1). These accessions were obtained from GRIN/USA and from the germplasm bank of Embrapa/BRA, respectively. Brazilian samples were originally collected in urban and rural residences, and had been previously found by interviews with residents and farmers from various states in Brazil. Cultivar Pima S7 was obtained from the germplasm bank of Embrapa/BRA and was included in this study to constitute an external group in each analysis.

DNA extraction was carried out from the seeds by the method of McDonald et al. (1994). Twenty-nine microsatellite markers (SSR) were

analyzed from the DNA samples. The SSR analysis was performed using the same conditions described by Ulloa et al. (2013). From the molecular data, it was carried out analysis by descriptive statistics and analysis by methods based on distance index and on ancestry model. In the first case, it was calculated allele frequency, observed and expected heterozygosity under the HWE, inbreeding coefficient, polymorphism information content (PIC), total and effective number of alleles per locus, allelic ratio, and number of rare and unique alleles in each population by country of origin, excluding Pima S7. Afterwards, it was carried out AMOVA and partition of the fixation index.

For analysis based on distance index, it was estimated genetic distances between the 29 accessions by the complement of the weighted similarity index. From the distance matrix, it was carried out UPGMA, 3D projection and principal coordinates analysis. In the analysis by the UPGMA method, it was used the Mojema method and value of  $K = 1.25$ . Finally, it was carried out analysis by the Structure 2.3.4 software, using length of burning period and number of MCMC replications after burning of 10,000, and a model which takes over admixture and correlated allele frequencies. Ten independent runs were carried out for each simulated value of  $K$ , ranging from 1 to 15. For each value of  $K$ , the estimated log probability of data ( $\text{Pr}(X/K)$ ) was used to calculate  $\Delta K$ . Among the groups defined by the Structure 2.3.4 software, Nei's genetic distance (NEI, 1972) and the  $F_{ST}$  fixation index were calculated. The following softwares was used: Genes (<http://www.ufv.br/dbg/genes/>), Structure 2.3.4 (<http://pritch.bsd.uchicago.edu/structure.html>), Structure Harvest (<http://taylor0.biology.ucla.edu/structureHarvester/>), and GenAlex (<http://biology-assets.anu.edu.au/GenAlEx/Welcome.html>).

## RESULTS AND DISCUSSION

This study uses 29 molecular markers of *G. hirsutum* L. to analyze the genetic diversity of 15 wild accessions of *G. barbadense* L. from Peru, and 13 *G. barbadense* L. accessions from Brazil, and included the cultivar Pima S7, which was used to be an outgroup (Table 1). Among the 29 accessions, it was observed 129 alleles and 5 heterozygotes (<1%). The number of alleles per locus ranged from 2 to 7, and mean was 4.4. Allele frequencies ranged from 0.02 to 0.89, and the overall mean was 0.22. Ten alleles (7.8%) had frequency >0.5, and 41

(32%) had frequency <0.1. Most alleles (62%) had low frequency ( $\leq 0.25$ ), and ten alleles (7.8%) were unique to a particular accession.

In subpopulations from Peru and Brazil, it was observed 105 and 108 alleles in total, respectively. The number of alleles per locus ranged from 2 to 5 and had mean and/or effective number of 3.62 and 3.72, respectively. PIC values ranged from 0.11 to 0.73, and from 0.23 to 0.73, and the mean was 0.52 and 0.55, respectively. It was observed 1 rare allele in the subpopulation from Peru, 2 in the subpopulation from Brazil; 21 unique alleles in the subpopulation from Peru, and 24 in the subpopulation from Brazil. Among the 29 SSR, 25 had  $\text{PIC} \geq 0.5$ , and 22 had at least 4 alleles per locus. Twenty-one SSR had simultaneously  $\text{PIC} \geq 0.5$  and at least four alleles per locus. In turn, 8 SSR had  $\text{PIC} \geq 0.6$  and 5 alleles per locus, and among them, five SSR had  $\text{PIC} \geq 0.7$ . Locus is polymorphic if the  $\text{PIC} \geq 0.1$ , or if the most frequent allele's frequency is lower than 0.95, and highly polymorphic if  $\text{PIC} \geq 0.7$  (CRUZ, 2008). Thus, all the 29 SSR were polymorphic, and 5 of them were highly polymorphic (Table 2).

Peru's subpopulation mean of PIC, diversity index, heterozygosity and inbreeding coefficient were  $\overline{\text{PIC}} = 0.522$ ,  $\overline{H} = 0.5803$ ,  $\overline{H}_0 = 0.007$  and  $F = 0.987$ . Brazil's subpopulation had  $\overline{\text{PIC}} = 0.547$ ,  $\overline{H} = 0.6045$ ,  $\overline{H}_0 = 0.008$  and  $F = 0.986$  (Table 2). In turn, the variation within subpopulations (97.051%) was higher than between subpopulations (2.949%), and the proportion of variance between subpopulations, relative to the total variance, was significant by AMOVA ( $\Phi_{ST} = 0.0295$ ) ( $p < 0.05$ ). For the same subpopulations, the components  $F_{IS} = 0.9828$ ,  $F_{ST} = 0.034$  and  $F_{IT} = 0.9834$  were estimated through the partition of the fixation index. These parameters indicate high informativeness for the SSR loci in general, and high variability, both between the Peruvian accessions and between the Brazilian accessions, and the largest variation between the Brazilian accessions is attributed to sampling. This variability is greater within each subpopulation. High fixation level occurs in each subpopulation due to the reproductive system, and low fixation level occurs due to this subdivision, indicating some considerable differentiation. In turn, low heterozygosity is related to the high inbreeding level, which is indicative of a reproductive system almost exclusively by self-pollination. Richness of alleles in the two subpopulations is consistent with Peru and Brazil being the primary and secondary centers of the species, respectively. The similar

variability level is consistent with the Peruvian and Brazilian evaluated samples. Thus, there is great diversity in the genotypes from Brazil, which is comparable to that of the exemplars from Peru,

whose variation of alleles per locus is also similar to the trend observed in the genus *Gossypium* spp. (FANG et al., 2013; MOIANA et al., 2015).

**Table 1.** Origin and profile of *G. barbadense* L. accessions from Brazil and Peru.

Accession	Origin	Latitude	Longitude	Phenotypic profile <sup>a</sup>
1-Pima S7	Arizona, USA	-	-	Yield/821 kg ha <sup>-1</sup> of lint, fibre length/1.35 inch, fibre strength/ 38.3 g/tex.
2-CE04-47	Ceará, Brazil	2° 57' S	40° 21' W	Plant height/2-3 m, flower stain/weak, seeds/strong kidney, lint/colored, fiber/white, leaf/green.
3-AP04-01	Amapá, Brazil	0° 2' N	51° 3' W	Plant height/2-3 m, flower stain/strong, seeds/separate, lint/absent, fiber/white, leaf/green.
4-PE06-71	Pernambuco, Brazil	7° 55' S	34° 53' W	Plant height/2-3 m, seeds/strong kidney, lint/colored, fiber/white, leaf/green.
5-MG04-43	Minas Gerais, Brazil	17° 52' S	41° 19' W	Plant height/1-2 m, seeds/strong kidney, fiber/white, leaf/green.
6-RR04-01	Roraima, Brazil	2° 27' N	60° 55' W	Plant height/≥3 m, flower stain/strong, seeds/separate, lint/colored, fiber/white, leaf/purple.
7-AL06-14	Alagoas, Brazil	9° 9' S	35° 17' W	Plant height/2-3 m, seeds/weak kidney, lint/colored, fiber/white, leaf/green.
8-MT04-01	Mato Grosso, Brazil	15° 41' S	55° 55' W	Flower stain /absent, seeds/strong kidney, lint/absent, fiber/white, leaf/green.
9-BA04-60	Bahia, Brazil	17° 32' S	39° 44' W	Flower stain/strong, seeds/separate, lint/colored, fiber/white, leaf/green.
10-PA04-03	Pará, Brazil	1° 2' S	47° 54' W	Plant height/2-3 m, flower stain/strong, seeds/strong kidney, lint/white, fiber/white, leaf/green.
11-MA04-01	Maranhão, Brazil	3° 43' S	43° 13' W	Plant height/≥3 m, flower stain/absent, seeds/strong kidney, lint/colored, fiber/white, leaf/purple.
12-PI04-02	Piauí, Brazil	9° 6' S	44° 21' W	Plant height/1-2 m, flower stain/strong, seeds/weak kidney, lint/absent, fiber/white, leaf/green.
13-PB05-01	Paraíba, Brazil	6° 55' S	37° 0.5' W	Plant height/2-3 m, flower stain/strong, seeds/separate, lint/absent, fiber/white, leaf/green.
14-SE06-02	Sergipe, Brazil	10° 20' S	36° 48' W	Plant height/≥3 m, flower stain/strong, seeds/strong kidney, lint/absent, fiber/white, leaf/green.
15-PI528329	Peru	-	-	Plant height/0.90 m, arealometer A/391 mm <sup>2</sup> mm <sup>-3</sup> , arealometer D/9, boll length/2.6 mm, boll width/21 mm, calyx hairs/absent, E1/8,6%, fuzz/3.0, leaf lac/0.38, locules/boll/3.3, petal color/segregating, pollen color/cream, pubescence/hairy plant, reflectance/41%, seeds/boll/23, UHM/0.97, yellowness/15.6, maturity/no bolls open, lint (%) /30.1%, lint index/3.4%, mean/ 0.76 inches, micronaire/5.60, productiveness/fair production, seed index/7.9 g, T0/4.58 g force/tex, T1/28.40 g force/tex.
16-PI435247	La Union Market, Piura, Peru	5° 30' S	80° 30' W	Primitive cultivar, fiber white and light brown, streaked.

17-PI435267	Chongayape, Lambayeque, Peru	6° 30' S	79° 30' W	Primitive cultivar, flood water irrigated, plant 3 m tall with trunk 4 cm thick, fiber grayish-white.
18-PI528213	Peru	-	-	Leaf lac/0.45, petal spot/medium spot, pubescence/hairy plant.
19-PI528243	Peru	-	-	Petal spot/heavy spot, pubescence/few plant hairs.
20-PI528086	Peru	-	-	Leaf lac/0.45, pubescence/no plant hairs.
21-PI608352	Loreto, Peru	07° 52' S	74° 46' W	-
22-PI528062	Peru	-	-	Plant height/1.3 m, boll length/39 mm, boll width/28 mm, leaf lac/0.36, petal color/segregating, petal spot/heavy spot, pollen color/cream, pubescence/few plant hairs, maturity/no bolls open, productiveness/fair production.
23-PI435242	Piura, Peru	5° 30' S	80° 30' W	Flood water irrigated, plants 1-1.5 m tall, some disease and insect damage, fiber brown.
24-PI435241	Piura, Peru	5° 30' S	80° 30' W	Flood water irrigated, plants 1-1.5 m tall, some disease and insect damage, fiber medium brown.
25-PI608353	Loreto, Peru	07° 52' S	74° 46' W	-
26-PI435257	Morrope market, Lambayeque, Peru	6° 32' S	80° 0' W	Non-irrigated, fiber light to medium brown, streaked.
27-PI435250	Catacaos market, Piura, Peru	5° 30' S	80° 30' W	Primitive cultivar, fiber light, medium, and purplish brown.
28-PI435248	La Union market, Piura, Peru	5° 30' S	80° 30' W	Primitive cultivar, fiber light through dark brown, some with purple tint.
29-PI435275	Morrope market, Peru	-	-	Primitive cultivar, fiber purple and reddish-brown.

<sup>a</sup>**Plant Height**, plant height in m. **Boll length**, length from the base to the tip in mm. **Boll width**, cross section at the widest part in mm. **Leaf lac**, B/A where B = distance from the petiole to the sinus between the middle lobe and the first lateral lobe, and A = leaf length from the petiole to the tip of the middle lobe. **Petal spot**, type of petal spotting. **Arealometer A**, measure of the external surface area of the fibers of a given volume of material (mm<sup>2</sup>mm<sup>-3</sup>). **Arealometer D**, difference between the specific area determined at high pressure (AH) and standard pressure (Areo A). D is presumably a measure of flatness of the fiber ribbon. **E1**, percentage elongation at break of the center 1/8 inch of the fiber bundle measured for T1 strength on the Stelometer. **Pubescence**, hairiness of plant in general. **Fuzz**, amount of fuzz in seeds (1.0 = heavy fuzz, 18.0 = naked seed). **Maturity**, relative maturity of entries when the earliest cotton had all the bolls open. **Locules/Boll**, mean number of locules in a 10 boll sample. **Reflectance**, percentage of reflectance on a Nickerson-Hunter Colorimeter. **UHM**, length in inches of the half of the fibers by weight that contains the longer fibers. **Yellowness**, measure of increasing yellowness of the cotton, measured with a Nickerson-Hunter Colorimeter. **Lint (%)**, weight of lint ginned from a sample of seed cotton expressed as a percentage of the weight of seed cotton. **Lint index**, weight of lint from 100 seed in g. **Mean**, mean length in inches of all fibers longer than 1/4 inch. **Micronaire**, fineness of the sample taken from the ginned lint, expressed in standard curvilinear micronaire units. **Seed index**, weight of 100 seed in g. **T0**, fiber strength of a bundle of fibers measured on a Stelometer with the jaws holding the fiber bundle tightly appressed (g force tex<sup>-1</sup>). **T1**, fiber strength of a bundle of fibers measured on a Stelometer with the jaws separated by a 1/8 inch spacer (g force tex<sup>-1</sup>). **Source**: Albrana (<http://www.cnpa.embrapa.br/albrana>), GRIN (<http://www.ars-grin.gov>).

**Table 2.** Descriptive analysis of the genetic diversity of wild accessions of *G. barbadense* L. from Brazil and Peru.

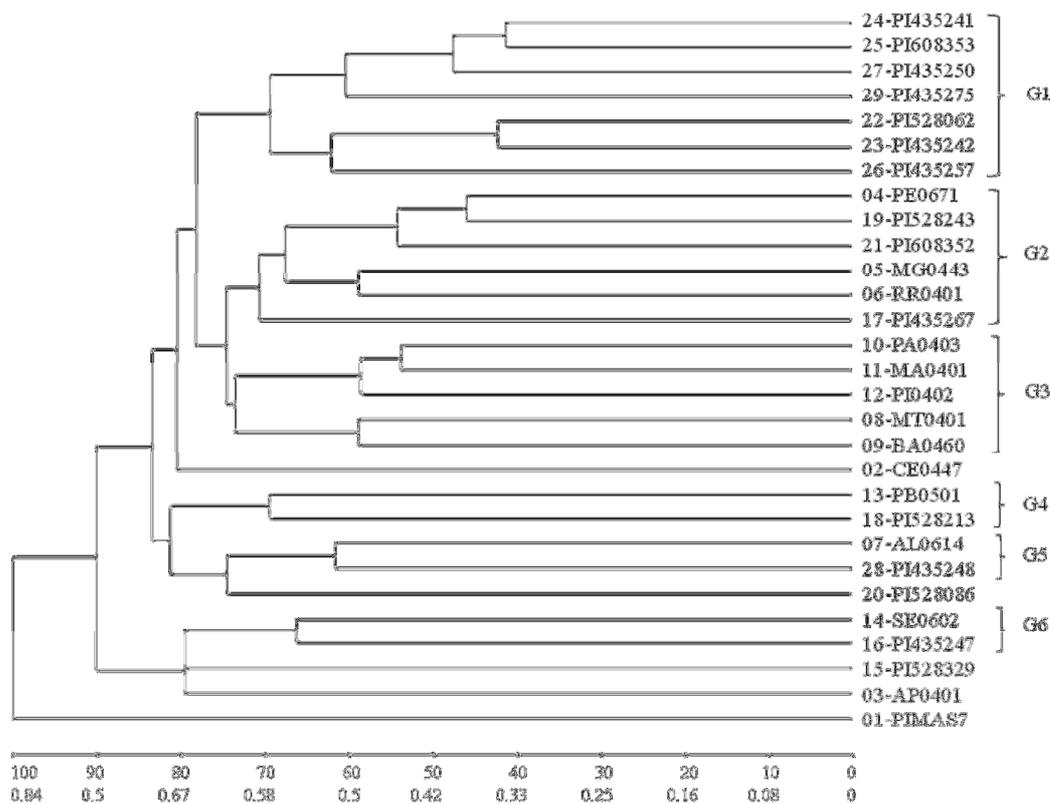
Locus/ Allele	Brazil										Peru							F <sub>ST</sub> <sup>d</sup>																	
	1	2	3	4	5	6	7	H <sup>a</sup>	PIC	N <sup>c</sup>	Maximum	1	2	3	4	5	6		7	H <sup>a</sup>	PIC <sup>b</sup>	N <sup>c</sup>	Maximu												
1-bnl0169	0.115	0.576	0.153	0.153	0			0.6065	0.56	4	0.5769	0	0.666	0.166	0	0.166			0.5	0.449	3	0.6667	-0.017												
2-bnl0840	0	0.25	0.416	0.25	0.083			0.6944	0.63	4	0.4167	0.066	0.4	0.133	0.4	0			0.6578	0.592	4	0.4	-0.005												
3-bnl1059	0	0.692	0	0.307	0			0.426	0.33	2	0.6923	0.218	0.375	0.062	0.281	0.062			0.7246	0.677	5	0.375	0.050												
4-bnl1350	0.333	0.458	0.083	0.083	0.041			0.6632	0.60	5	0.4583	0.166	0.5	0.333	0	0			0.6111	0.535	3	0.5	-0.036												
5-bnl1421	0.5	0.166	0.083	0.083	0.166			0.6806	0.64	5	0.5	0.066	0.266	0	0.2	0.466			0.6667	0.609	4	0.4667	0.193												
6-bnl1551	0.538	0.307	0.153	0				0.5917	0.51	3	0.5385	0	0.6	0.333	0.066				0.5244	0.440	3	0.6	0.074												
7-bnl1597	0.307	0.461	0.153	0.076				0.6627	0.60	4	0.4615	0.266	0.533	0.2	0				0.6044	0.535	3	0.5333	-0.027												
8-bnl1665	0.153	0.538	0.230	0.076				0.6272	0.57	4	0.5385	0.181	0.818	0	0				0.2975	0.253	2	0.8182	0.120												
9-bnl1694	0.5	0.25	0.083	0.166				0.6528	0.59	4	0.5	0	0.428	0.428	0.142				0.6122	0.529	3	0.4286	0.311												
10-bnl2495	0.846	0.153						0.2604	0.22	2	0.8462	0.875	0.125						0.2188	0.194	2	0.875	-0.017												
11-bnl2553	0.076	0.615	0.230	0.076	0			0.5562	0.50	4	0.6154	0.066	0.366	0.466	0	0.1			0.6333	0.564	4	0.4667	0.026												
12-bnl2895	0.384	0.384	0.230					0.6509	0.57	3	0.3846	0.266	0.533	0.2					0.6044	0.535	3	0.5333	-0.015												
13-bnl3034	0	0.583	0.166	0.25	0			0.5694	0.50	3	0.5833	0.312	0.25	0.25	0.125	0.062			0.7578	0.716	5	0.3125	0.064												
14-bnl3008	0	0.333	0.25	0.166	0.25			0.7361	0.68	4	0.3333	0.187	0.437	0.312	0	0.062			0.6719	0.611	4	0.4375	-0.034												
15-bnl3649	0.153	0.307	0.230	0.230	0.076			0.7692	0.73	5	0.3077	0.187	0.437	0.125	0.125	0.125			0.7266	0.690	5	0.4375	-0.028												
16-bnl3902	0	0.833	0.083	0.083				0.2917	0.27	3	0.8333	0.062	0.937	0	0				0.1172	0.110	2	0.9375	0.058												
17-bnl4053	0	0.454	0.181	0.272	0	0.091		0.6777	0.62	4	0.4545	0.214	0.285	0.071	0.285	0.142	0		0.7653	0.726	5	0.2857	-0.021												
18-bnl4094	0.333	0	0.083	0.5	0.083	0	0	0.625	0.55	4	0.5	0.333	0.266	0	0	0.2	0.066	0.133	0.7556	0.715	5	0.3333	0.313												
19-cir0085	0.416	0.333	0.25					0.6528	0.57	3	0.4167	0.214	0.5	0.285					0.6224	0.551	3	0.5	0.001												
20-cir0109	0.153	0.538	0.307					0.5917	0.51	3	0.5385	0.125	0.625	0.25					0.5313	0.468	3	0.625	-0.034												
21-cir0148	0.076	0.384	0.384	0.153				0.6746	0.61	4	0.3846	0.071	0.428	0.5	0				0.5612	0.465	3	0.5	-0.017												
22-cir0036	0.25	0.333	0.25	0.083	0.083			0.75	0.70	5	0.3333	0.333	0.2	0.133	0.266	0.066			0.7556	0.715	5	0.3333	-0.027												
23-cir0246	0.076	0.230	0.615	0	0.076			0.5562	0.50	4	0.6154	0	0.375	0.562	0.062	0			0.5391	0.446	3	0.5625	-0.037												
24-cir1414	0.333	0.25	0.333	0.083				0.7083	0.65	4	0.3333	0.25	0.625	0.062	0.062				0.5391	0.483	4	0.625	0.085												
25-bnl2895	0.076	0.538	0.384					0.5562	0.46	3	0.5385	0.066	0.266	0.666					0.48	0.412	3	0.6667	0.072												
26-bnl3103	0.076	0.692	0.153	0.076				0.4852	0.45	4	0.6923	0.187	0.687	0.125	0				0.4766	0.427	3	0.6875	-0.035												
27-bnl3359	0.458	0.333	0	0.166	0.041			0.6493	0.58	4	0.4583	0.062	0.343	0.375	0.187	0.031			0.7012	0.646	5	0.375	0.171												
28-bnl3479	0.166	0	0.583	0.083	0.166			0.5972	0.55	4	0.5833	0	0.066	0.333	0.533	0.066			0.5956	0.525	4	0.5333	0.011												
29-bnl3590	0.076	0	0.461	0.461	0			0.568	0.47	3	0.4615	0	0.062	0.312	0.562	0.062			0.5781	0.509	4	0.5625	-0.018												
Nt <sup>e</sup>	108					$\bar{H} = 0.6045$					$\bar{PIC} = 0.5472$					105					$\bar{H} = 0.5803$					$\bar{PIC} = 0.52$					$\bar{Fst} = 0.034$				
Pa <sup>f</sup>	0.84															0.81																			
Ne <sup>g</sup>	3.72															3.62																			
Nr <sup>h</sup> (%)	1.85															0.95																			

<sup>a</sup>H, expected heterozygosity ( $H=1-\text{Sum}(P^2)$ ). <sup>b</sup>PIC, polymorphism information content. <sup>c</sup>N, number of allele observed per locus, ranging from 1 to 7. <sup>d</sup>F<sub>ST</sub>, fixation index due to the subdivision of the population, which leads to its genetic differentiation. <sup>e</sup>Nt, total number of alleles per locus. <sup>f</sup>Pa, proportion of alleles in each population (Pa = number of alleles in the population/total number of alleles in the species). <sup>g</sup>Ne, effective number of alleles per locus (Ne = total number of alleles of polymorphic loci/total number of polymorphic loci). <sup>h</sup>Nr, number of rare alleles with frequency lower than 0.05.

Genetic distance values between the 29 accessions obtained by the complement of the weight similarity index, from molecular markers, ranged between 0.59 and 0.98, and the mean was 0.66. Higher distance values involved cultivar Pima S7. Lower values involved pairs of Peru accessions, and then pairs between Peruvian and Brazilian accessions. Dissimilarity mean, in relation to the total accessions, was higher for Pima S7, followed by PI528329; and lower for PI528243, followed by MG04-43. This range of distance index is also close to those obtained with SSR within *G. barbadense* L. (WANG et al., 2011; ABDELLATIF et al., 2012).

In the projection of distance matrix by UPGMA, Pima S7 is the most external accession, and thus the most genetically distant in relation to most accessions. From top to bottom, 7 accessions from Peru (24-PI435241, 25-PI608353, 27-

PI435250, 29-PI435275, 22-PI528062, 23-PI435242, 26-PI435257) form G1 group. Below it, 3 accessions from Peru and 3 accessions from Brazil form the G2 group, and 2 accessions from Brazil form a subgroup (5-MG0403, 6-RR0401). Five Brazilian accessions form G3 group (10-PA0403, 11-MA0401, 12-PI0402, 8-MT0401, 9-BA0460). Subsequently, three other groups present 1 accession from Peru and 1 from Brazil, each (G4 to G6). Then, by carrying out a cut at 74%, one between the distance values indicated by the Mojema method, using value of  $k = 1.25$ , six groups are defined, G1 to G6. In this projection, the values of cophenetic correlation, distortion and stress were equal to 0.72, 1.2% and 10.9%, respectively, indicating good adjustment of the original and graphic values (Figure 1).



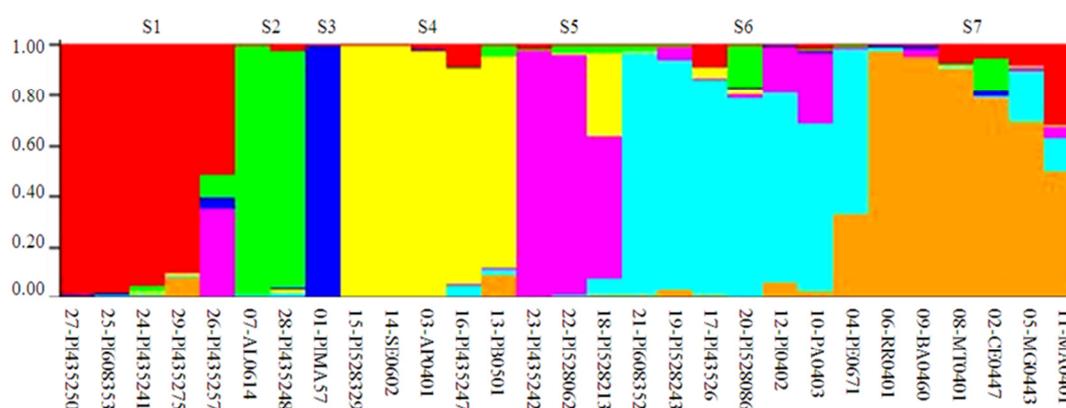
**Figure 1.** Grouping of 29 *G. barbadense* L. accessions from Brazil and Peru by UPGMA.

By the UPGMA method projection, both Peruvian and Brazilian accessions show high allelic variability. In this projection, despite the fairly large distance, some Brazilian accessions (4-PE0671, 13-PB0501, 7-AL0614, 14-SE0602) are closer to different Peruvian accessions, indicating origin and/or ancestry related to different stocks and/or gene pools from Peru. Other accessions from Brazil (3-AP0401, 2-CE0447) have great divergence over all the other Brazilian and Peruvian accessions,

indicating origin and/or ancestry more related to other gene pools from Brazil or Peru, or another not studied region. In turn, 2-CE0447, 13-PB0501, 7-AL0614, 14-SE0602, and 3-AP0401 may have originated from distinct gene pools in relation to the other Brazilian studied accessions, while the studied Peruvian accessions may have originated from six different gene pools, based on the melting levels at 74% of the highest estimated divergence.

In the analysis by the Structure 2.3.4 software, the distribution of the values of  $\Delta k$  had peak at  $\Delta k = 7$ , indicating the subdivision of the 29 accessions into seven groups (which equals to six groups, if the group formed by Pima S7, which is the external group in the analysis, is not taken into account). By assigning each accession to their respective group, using a membership probability threshold  $\geq 0.5$ , Pima S7 constitutes the blue group (S3); five accessions from Peru, the red group (S1); three accessions from Peru, the pink group (S5); and six accessions from Brazil, the orange group (S7).

Peruvian and Brazilian accessions form the groups green (S2), yellow (S4) and light blue (S6) (Figure 2). The ancestry inferred in the groups, from S1 to S7, ranged from 0.51-0.99, 0.94-0.98, 0.99-0.99, 0.84-0.99, 0.57-0.97, 0.65-0.96, and 0.51-0.98, and had mean of 0.87, 0.96, 0.99, 0.93, 0.83, 0.79, and 0.80, respectively. The groups which obtained the highest mean inferred ancestry were S3 (blue), followed by S2 (green). S6 (light blue), followed by S7 (orange) had the lowest mean inferred ancestry. Mean inferred ancestry is, in turn, a measure of the group's consistency.



**Figure 2.** Grouping of 29 *G. barbadense* L. accessions from Brazil and Peru by the Structure 2.3.4 software, for  $k = 7$  value, from ancestry models, taking over genetic admixture and correlated allele frequencies.

Comparing the results of the Structure 2.3.4 software with the UPGMA projection, G1 brings together the accessions of S1 (red) and S5 (pink), except 18-PI528213, from S5; G2 and G3 overlap groups S6 (light blue) and S7 (orange); G5 brings together the accessions of S2 (green); G6 overlaps S4 (yellow); and Pima S7 is not associated with any other accession in both analysis, and there was good overlap in the groups by the both methods. Thus, G1 presents ancestry related to the red and pink groups; G2 presents ancestry related to the light blue and orange groups; G3 presents ancestry related to the light blue, orange and pink groups; G4 presents ancestry related to the yellow group (but in the case of 18-PI528213, it also presents ancestry related to pink group); G5 presents ancestry related to the green group; and G6, as well as 15-PI528329 as 3-AP0401, presents ancestry related to the yellow group. On the other hand, 2-CE0447 and 20-528086 present ancestry related to the orange and light blue groups, respectively.

Since Structure 2.3.4 software does not take over hierarchical structure, which is a reasonable assumption in higher taxonomic levels, but not necessarily true within a single species, it was investigated the relationship between pairs of groups

of Structure 2.3.4. The Nei's genetic distance between pairs of groups of Structure 2.3.4, excluding Pima S7, ranged from 0.38 to 0.87, with mean of 0.59. On the other hand,  $F_{ST}$  index ranged from -0.01 to 0.26, with mean of 0.08. By the respective parameters, S1/S2 and S1/S4 pairs of groups presented divergence of allele frequency and genetic differentiation. However, individually, S4/S7 and S2/S5 presented greater genetic distance and genetic differentiation degree, respectively. Among the groups, nine accessions with ancestry  $\leq 80\%$  can be characterized by genetic admixture (2,20,12,5,10,4,18,26,11). Among them, most are identified by mixture between groups S6 (light blue), S7 (orange) and S5 (pink).

By the 3D projection, it is also observed outstanding overlap between the relation of accessions in this projection of the distance matrix and G1 to G6 groups, in the first analysis, by the UPGMA method, and S1 (red), S2 (green), S3 (blue), S4 (yellow) and S7 (orange) groups, in the second analysis, by the Structure 2.3.4 software. The values of correlation, distortion and stress in 3D projection were, respectively, 0.67, 26.7% and 36.3%, which are acceptable values. By the

principal coordinates analysis, it was also observed similar patterns of genetic relationship, despite being small the variation explained by the two main axes (<20%). In turn, the correspondence in relationships between the 29 accessions in these analytical methods is indication of consistency in genetic structure found.

In the only previous study on *G. barbadense* L. from Brazil (states of Amapá and Pará), 12 microsatellite markers produced 2.53 alleles per locus, and in total, eight unique alleles for these states were produced (approximately 20% of the total alleles). Variation within (64%) was higher than between collections of the two states (36%). However, variation between collections was also significant ( $p < 0.05$ ) (ALMEIDA et al., 2009). Comparing the studies, diversity between Brazilian accessions, in the present study, is higher than that between the accessions from the states of Pará and Amapá, of the previous study, based on the number of alleles per locus and of unique alleles, as it is expected. Both studies found high proportion of unique alleles and lack of heterozygotes between *G. barbadense* L. The high proportion of unique alleles characterizes isolation and low gene flow, and the lack of heterozygotes characterizes the reproductive system based on self-pollination. Once cotton is also pollinated by insects, especially bees, the lack of heterozygotes is also explained by isolation, since in the collection site, plants were generally found alone or in groups of two or three plants, and isolated from neighboring properties.

After the previously cited study, another one, perhaps more related to this research, which included Brazilian genotypes, evaluated the diversity of *G. barbadense* L. accessions in countries of South America and Central America, along the probable route of pre-Columbian expansion, using AFLP (WESTENGEN et al., 2005). Diversity estimated within *G. barbadense* L. in this study reflected indeed the geographic patterns, in which the accessions with primitive characteristics of the Peruvian coast exhibited greater diversity. The far north of Peru's accessions and southwestern Ecuador grouped basally to the accessions of the east of Andes, and these accessions of the east of Andes were grouped together with accessions from Bolivia, Brazil, Colombia, Venezuela and the Caribbean and Pacific Islands. Thus, northwestern Peru and/or southwestern Ecuador were pointed out to be the original center of *G. barbadense* L., from where it would have spread across the Andes. This hypothesis is, in turn, supported by another study based on isozymes, including specimens from South

America, Central America and the Caribbean (PERCY, WENDEL, 1990).

There are at least five other studies regarding *G. barbadense* L. diversity outside South America. In the first, Hussein et al. (2007) evaluated materials from Egypt with RAPD, SSR, EST and AFLP, which reported similarity between 13.8 and 99.3%. Boopathi et al. (2008) evaluated accessions from India with SSR, and found 1.72 alleles per locus, and PIC ranging from 0.11 to 0.73, and mean of 0.39. Wu et al. (2010) evaluated accessions from the former USSR and from Egypt, China and the USA, and observed 1.62 alleles per locus SSR, and similarity index between 0.66 to 0.94, and mean of 0.81. Wang et al. (2011) also evaluated cultivars from China, and introductions from other countries, with SSR, and reported PIC between 0.03 and 0.86, and mean of 0.32. Finally, Abdellatif et al. (2012) evaluate varieties and hybrids from Egypt with ISSR, EST and SSR, and reported PIC values between 0.8 and 0.91, and mean of 0.86 for SSR. Thus, based on these studies, there is still reasonable diversity in *G. barbadense* L. in other regions of the world.

Since there is diversity not only in Peru, but also in Brazil, as shown in this study, the maintenance of this genetic resource should be prioritized. It is necessary to know the variability throughout the country, and the possible germplasm conservation alternatives. In Brazil, *G. barbadense* L. practically does not occur in a natural environment, but as dooryard plants. Thus, ex-situ conservation, in germplasm bank, appears to be the most viable alternative, as well as the safest long-term preservation. Since diversity is not adequately represented in germplasm banks, further efforts should be made to indeed increase Brazilian collections.

## CONCLUSIONS

There is comparable variability level in specimens from Brazil and Peru, according SSR loci.

Accessions from Brazil show fairly large genetic distance from the most similar accessions from Peru, indicating different origin for some accessions, and closer ancestry to other not studied gene pools for other accessions.

At least five accessions from Brazil come from different gene pools from the rest of the Brazilian accessions.

Peru's accessions may have originated from six different gene pools.

Nine accessions have genetic admixture of at least three gene pools.

In Brazil, due to the diversity and to the reduction of local populations, preservation of genetic resources should be intensified, which can be fundamental for the future improvement of

upland cotton. In turn, variability of Peru may prove to be crucial for cotton crop worldwide.

#### ACKNOWLEDGMENTS

We thank Richard Percy for providing us cotton seeds from GRIN.

**RESUMO:** O Peru é reconhecido como o centro primário e o Brasil como o centro secundário de *G. barbadense* L., espécie que se destaca pela fibra longa do algodão. Mas, desde a sua introdução no Brasil, as populações de *G. barbadense* L. vêm se reduzindo e também sua variabilidade genética. Em razão disso, o objetivo deste trabalho foi comparar o nível de diversidade de acessos de *G. barbadense* L. do Peru e Brasil, para fins de conservação e manejo de seu recurso genético. Quinze acessos do Peru e treze do Brasil, além da cultivar Pima S7, foram analisados com 29 marcadores microssatélites. A diversidade genética foi analisada por meio de estatísticas descritivas e por métodos baseados em índice de distância (UPGMA, projeção 3D e análise de coordenadas principais) e modelo de ancestralidade (Structure 2.3.4). Por estas análises, existe nível de variabilidade similar entre os acessos do Brasil e Peru. Ao menos cinco acessos do Brasil vêm de pools gênicos distintos daqueles dos outros acessos brasileiros. Os acessos do Peru têm origem traçada a seis pools gênicos distintos. Nove entre os 29 acessos possuem mistura genética de ao menos três pools gênicos diferentes. A variabilidade do Peru e Brasil pode vir a ser crucial para a cultura do algodão em todo o mundo.

**PALAVRAS-CHAVE:** *G. barbadense* L.. Marcadores moleculares. Diversidade genética. Algodão. Recursos genéticos.

#### REFERENCES

- ABDELLATIF, K. F.; KHIDR, Y. A.; EL-MANSY, Y. M.; EL-LAWENDEY, M. M.; SOLIMAN, Y. A. Molecular diversity of Egyptian cotton (*Gossypium barbadense* L.) and its relation to varietal development. **Journal of Crop Science and Biotechnology**, Seoul, v. 15, p. 93-99, 2012. <http://dx.doi.org/10.1007/s12892-011-0120-5>
- ALMEIDA, V. C. D.; HOFFMANN, L. V.; YOKOMIZO, G. K. I.; COSTA, J. N. D.; GIBAND, M.; BARROSO, P. A. V. In situ and genetic characterization of *Gossypium barbadense* populations from the states of Pará and Amapá, Brazil. **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, p. 719-725, 2009. <http://dx.doi.org/10.1590/S0100-204X2009000700011>
- BOOPATHI, N. M.; GOPIKRISHNAN, A.; SELVAM, N. J.; RAVIKESAVAN, R.; IYANAR, K.; MUTHURAMAN, S.; SARAVANAN, N. Genetic diversity assessment of *G. barbadense* accessions to widen cotton (*Gossypium* spp.) gene pool for improved fibre quality. **Journal of Cotton Research and Development**, Haryana, v. 22, p. 135-138, 2008.
- CAMPBELL, B. T.; SAHA, S.; PERCY, R.; FRELICHOWSKI, J.; JENKINS, J. N.; PARK, W.; MAYEE, C. D.; GOTMARE, V.; DESSAUW, D.; GIBAND, M.; DU, X.; JIA, Y.; CONSTABLE, G.; DILLON, S.; ABDURAKHMONOV, I. Y.; ABDUKARIMOV, A.; RIZAEVA, S. M.; ABDULLAEV, A.; BARROSO, P. A. V.; PÁDUA, J. G.; HOFFMANN, L.; PODOLNAYA, L. Status of the global cotton germplasm resources. **Crop Science**, Madison, v. 50, p. 1161-1179, 2010. <http://dx.doi.org/10.2135/cropsci2009.09.0551>
- CRUZ, C. D. **Programa Genes: diversidade genética**. Viçosa: Editora UFV, 2008. 278 p.
- FANG, D. D.; HINZE, L. L.; PERCY, R. G.; LI, P.; DENG, D.; THYSSEN, G. A microsatellite-based genome-wide analysis of genetic diversity and linkage disequilibrium in Upland cotton (*Gossypium hirsutum* L.) cultivars from major cotton-growing countries. **Euphytica**, Wageningen, v. 191, p. 391-401, 2013. <http://dx.doi.org/10.1007/s10681-013-0886-2>

- FRYXELL, P. A. **Germplasm utilization: *Gossypium*, a case history**. New Orleans: USDA-ARS, 1976. 16 p.
- HOVAV, R.; CHAUDHARY, B.; UDALL, J. A.; FLAGEL, L.; WENDEL, J. F. Parallel domestication, convergent evolution and duplicated gene recruitment in allopolyploid cotton. **Genetics**, Bethesda, v. 179, p. 1725-1733, 2008. <http://dx.doi.org/10.1534/genetics.108.089656>
- HUSSEIN, E. H.; OSMAN, M. H.; HUSSEIN, M. H.; ADAWY, S. S. Molecular characterization of cotton genotypes using PCR-based markers. **Journal of Applied Science Research**, Maan, v. 3, p. 1156-1169, 2007.
- HUTCHINSON, J. B.; SILOW, R. A.; STEPHENS, S. G. **The evolution of *Gossypium* and the differentiation of the cultivated cottons**. London: Oxford University Press, 1947. 160 p.
- MCDONALD, M. B.; ELLIOT, L. J.; SWEENERY, P. M. DNA extraction from dry seeds for RAPD analyses in varietal identification studies. **Seed Science and Technology**, Bassersdorf, v. 22, p. 171-176, 1994.
- MOIANA, L. D.; GONCALVES-VIDIGAL, M. C.; CARVALHO, L. P. Genetic diversity and population structure of upland cotton Brazilian cultivars (*Gossypium hirsutum* L. race 'latifolium' H.) using SSR markers. **Australian Journal of Crop Science**, Winnipeg, v. 9, p. 143-152, 2015.
- NEI, M. Genetic distance between populations. **The American Naturalist**, Chicago, v. 106, p. 283-292, 1972. <http://dx.doi.org/10.1086/282771>
- PATERSON, A. H. **Genetics and genomics of cotton**. New York: Springer Science, 2009. 515 p. <http://dx.doi.org/10.1007/978-0-387-70810-2>
- PERCIVAL, A. E.; KOHEL, R. J. Distribution, collection, and evaluation of *Gossypium*. **Advances in Agronomy**, Delaware, v. 44, p. 225-256, 1990. [http://dx.doi.org/10.1016/S0065-2113\(08\)60823-8](http://dx.doi.org/10.1016/S0065-2113(08)60823-8)
- PERCY, R. G.; WENDEL, J. F. Allozyme evidence for the origin and diversification of *Gossypium barbadense* L. **Theoretical and Applied Genetics**, Berlin, v. 7, p. 529-542, 1990. <http://dx.doi.org/10.1007/bf00226164>
- PIPERNO, D. R.; PEARSALL, D. M. **The origins of agriculture in the lowland neotropics**. San Diego: Academic Press, 1998. 178 p.
- ROSSEN, J.; DILLEHAY, T. D.; UGENT, D. Ancient cultigens or modern intrusions? Evaluating plant remains in an Andean case study. **Journal of Archaeological Science**, London, v. 23, p. 391-407, 1996. <http://dx.doi.org/10.1006/jasc.1996.0035>
- STEPHENS, S. G. The origin of Sea Island cotton. **Agricultural History**, Winter Park, v. 50, p. 391-399, 1976.
- ULLOA, M.; HUTMACHER, R. B.; ROBERTS, P. A.; WRIGHT, S. D.; NICHOLS, R. L.; DAVIS, R. M. Inheritance and QTL mapping of Fusarium wilt race 4 resistance in cotton. **Theoretical and Applied Genetics**, Berlin, v. 126, p. 1405-1418, 2013. <http://dx.doi.org/10.1007/s00122-013-2061-5>
- WANG, P.; ZHU, Y.; SONG, X.; CAO, Z.; DING, Y.; LIU, B.; ZHANG, T. Inheritance of long staple fiber quality traits of *Gossypium barbadense* in *G. hirsutum* background using CSILs. **Theoretical and Applied Genetics**, Berlin, v. 124, p. 1415-1428, 2012. <http://dx.doi.org/10.1007/s00122-012-1797-7>
- WANG, X. Q.; FENG, C. H.; LIN, Z. X.; ZHANG, X. L. Genetic diversity of sea-island cotton (*Gossypium barbadense*) revealed by mapped SSRs. **Genetics and Molecular Research**, Ribeirão Preto, v. 10, p. 3620-3631, 2011. <http://dx.doi.org/10.4238/2011.December.8.5>
- WENDEL, J. F. Cotton. In: SIMMONDS, N.; SMARTT, J. (Ed.). **Evolution of crop plants**. London: Longman, 1995. p. 358-366.

Comparison of wild accessions...

RODRIGUES, J. I. S.; CARVALHO, L. P.; FARIAS, F. J. C

WENDEL, J. F.; BRUBAKER, C.; ALVAREZ, I.; CRONN, R.; STEWART, J. M. Evolution and natural history of the cotton genus. In: PATERSON, A. H. (Ed.). **Genetics and genomics of cotton**. New York: Springer Science, 2009. p. 3-22. [http://dx.doi.org/10.1007/978-0-387-70810-2\\_1](http://dx.doi.org/10.1007/978-0-387-70810-2_1)

WENDEL, J. F.; BRUBAKER, C. L.; SEELANAN, T. The origin and evolution of *Gossypium*. In: STEWART, J. M.; OOSTERHUIS, D.; HEITHOLT, J.; MAUNEY, J. R. (Ed.). **Physiology of cotton**. Netherlands: Springer, 2010. p. 1-18. [http://dx.doi.org/10.1007/978-90-481-3195-2\\_1](http://dx.doi.org/10.1007/978-90-481-3195-2_1)

WESTENGEN, O. T.; HUAMAN, Z.; HEUN, M. Genetic diversity and geographic pattern in early South American cotton domestication. **Theoretical and Applied Genetics**, Berlin, v. 110, p. 392-402, 2005. <http://dx.doi.org/10.1007/s00122-004-1850-2>

WU, D. P.; FANG, X. X.; MA, M. N.; CHEN, J. H.; ZHU, S. J. Genetic relationship and diversity of the germplasm in *Gossypium barbadense* L. from four different countries using SSR markers. **Cotton Science**, Baton Rouge, v. 22, p. 104-109, 2010.