Introgression of Pigment Gland Morphogenisis Delay into Upland Cotton: Potential of DNA Markers to Monitor Parental Contribution to Progenies

I. Vroh Bi[1&3], G. Mergeai[1], J.P. Baudoin[1], J.M. Jacquemin[2], P. du Jardin[3]

[1]Unité de Phytotechnie des Régions Intertropicales, Faculté Universitaire des Sciences Agronomiques, 2, Passage des déportés, B-5030 Gembloux, Belgium

[2]Station d'Amélioration des Plantes, Centre de Recherches Agronomiques, 4, rue du Bordia, 5030 Gembloux

[3]Unité de Biologie Végétale, Faculté Universitaire des Sciences Agronomiques, 2, Passage des déportés, B-5030 Gembloux, Belgium

ABSTRACT

The delay of pigment gland morphogenisis in the seed confers to several Australian wild diploid cottons the glandless-seed/glanded-plant trait. To introgress this trait from G. sturtianum Willis (2m = 2x = 26, 2C1 genome) into G. hirsutum L. (2n = 4x = 52, 2(AD)1 genome), we used bridge crosses to synthesize two trispecies hybrids, G. hirsutum-G. raimondii Ulbrich – G. sturtianum (HRS) and G. thurberi Torado – G. sturtianum – G. hirsutum (TSH). Recurrent backcrossing of these hybrids to G. hirsutum produced progenies expressing the desired trait at different levels. The objective of this study was to assess the genomic contribution of the parental species to their progenies with random amplification polymorphic DNA (RAPD) and restriction fragment length polymorphism (RFLP) probes. The use of 30 decamer primers yielded 375 bands of which 339 were polymorphic between parents. Among 49 bands specific to the donor G. sturtianum, 20 and 18 that segregated in backcrosses, were observed in HRS and TSH, respectively. The American wild diploid species used as bridge showed 12 and 13 specific bands for G. raimondii (2n = 2x = 26, 2D5 genome) and G. thurberi (2n = 2x = 26, 2D1 genome), respectively. Genetic distances between G. hirsutum and the wild species involved in the cross were determined from RAPD data. This study allowed choice within backcross progenies those that shared the highest similarity to the cultivated cotton. The parental origin of chromosomes in the trispecis hybrids and the backcrosses were then identified using RFLP probes specific to cotton chromosomes. The results are discussed in relation to the expression of the desired trait.

Introduction

Gossypium contains about 50 diploid and tetraploid species. The diploid species (2n= 2x= 26) fall into 8 different cytotypes designated A, B, C, D, E, F, G and K (Fryxell, 1979; Endrizzi *et al.*, 1985; Gorham *et al.*, 1996). Cotton is not only the leading fiber crop, it is also the second best potential source of plant proteins, and the fifth best oil-producing plant (Textier, 1993). One of the main traits delineating the *Gossypium* genus is the presence of pigment glands throughout the plant. Pigment glands (also called gossypol glands), contain polyphenolic compounds which confer insect resistance to cotton plant. However, the presence of glands in cottonseed is undesirable because gossypol and derivatives are toxic to man and other monogastric animals. Glanded cottonseed kernel containing from 0.6 to 2.0% gossypol has limited nutritional uses (Lusas and Jividen, 1987), and completely glandless cotton (McMichael, 1954) is susceptible to several insect pests. The objective of our research is to develop an upland cotton *G. hirsutum* L. [2n= 4x = 52; 2(AD)1 genome], having glandless or low-gossypol seed for feed and food uses, and a high level of gossypol in the remaining organs to resist pests. Using the Australian wild diploid species *G. sturtianum* Willis (2C1 genome) as donor, and an American wild diploid species (*G. raimondii* Ulbr, 2D5 genome, or *G. thurberi* Tod.; 2D1 genome) as bridge, trispecies hybrids and backcross progenies were obtained through recurrent selection.

In the past few years, new strategies based on marker-assisted selection have been proposed to reduce time and effort in developing new varieties (Rafalski and Tingey, 1993 ; Young and Tanksley, 1989; Lee 1998). Molecular markers are very efficient tools to monitor alien DNA introgression during breeding programs. In the present study, the trispecies hybrids and their backcross progenies (BC) were evaluated to test the suitability of RAPDs and RFLPs for detecting introgression, and improving selection efficiency.

Materials and methods

Plant materials. Two trispecies hybrids [*G. thurberi-G. sturtianum-G. hirsutum* (TSH), and *G. hirsutum - G. raimondii - G. sturtianum* (HRS)] were involved in recurrent backcrosses with *G. hirsutum* to produce

BC_1, BC_2, BC_3 and BC_2sp (BC_2 selfed progenies), thanks to *in vitro* culture of seed embryos, application of growth regulators after manual pollinations, and grafting of unbalanced progenies on vigorous *G. hirsutum* plantlets.

DNA extraction. An important step in DNA marker-assisted selection is the efficient isolation of plant DNA. We used activated charcoal in a modified CTAB method to obtain a simple but efficient DNA extraction procedure (Vroh Bi *et al.*, 1996).

Molecular analysis. RAPD reactions and statistic analyses were performed according to Vroh Bi *et al.*, 1997 and Mergeai *et al.*, 1998. In addition, the robustness of similarity estimates was analyzed by the bootstrap technique (Felsenstein, 1985), using the genetic distance of Nei (Nei and Lei 1979). Restriction endonuclease digestion, southern blotting, labelling and hybridization of RFLP probes were performed as described in Reinisch *et al.* (1994), with slight modifications (7 µg of genomic DNA/reaction, 50 ng of PCR-amplified probe-DNA, 10 ml hybridisation mix, Kodak XOMAT AR film, one week-exposure minimum). Forty-nine RFLP markers of known map positions (Reinisch *et al.*, 1994) were tested on genomic DNA digested with six restriction enzymes (BamH1, Cfo1, EcoR1, EcoR5, Hind3, Xba1). These clones were kindly provided as M13 inserts by Professor A.H. Paterson of Texas A&M University (USA).

Results and discussion

RAPD analysis. Seventy five decamer primers were screened on parental species, trispecies hybrids and 27 BC_1 plants. Thirty primers showing consistently reproducible bands were used for further studies. The analysis of species-specific bands confirmed the tri-parental origin of both hybrids TSH and HRS. Owing to the dominant nature of RAPD markers, only markers specific to the wild parents (bands present in the wild parents and not in the cultivated cotton) could be used to detect introgression. An example of introgression and segregation of such specific RAPD bands in the trispecies hybrid HRS and its BC_1 is summarized in Table 1. The thirty primers detected 12, 13, and 49 specific bands from *G. raimondii*, *G. thurberi* and *G. sturtianum* (the donor parent), respectively. Of the 49 *G. sturtianum*-specific bands, 22 were present in the trispecies hybrids and 17 segregated in the BC_1 progenies (Mergeai *et al.*, 1998). Markers that are not transmitted to progenies at each cycle of cross are either located on chromosomes that are not transmitted during meiosis, or constitute markers undergoing recombinations that can modify the primer binding sites. Few specific markers of *G. sturtianum* were systematically present in all the BC_1 analyzed. Such bands should be located on C genome chromosomes that are preferentially transmitted to the BC_1, due to their higher pairing affinity with the A and D chromosomes of the other parents, or represent repeated DNA dispersed throughout the genome of *G. sturtianum.*

The reliability of our RAPD data to establish genetic distances within cotton germplasm was checked by studying relationship between the parental species, using UPGMA (unweighted pair group method with arithmetical mean) and Jaccard's distances (Figure 1). The pattern obtained is in agreement with the current phylogenetic classification of *Gossypium* species based on morphological and cytogenetical studies (Fryxell, 1979; Endrizzi *et al.*, 1985). Taking the most remote species, *G. sturtianum* as outgroup, the robustness of the remaining nodes was assessed on 1000 bootstrap runs by the neighbour-joining method of PHYLIP (Felsenstein 1993). Results are shown on Figure 2. All the four varieties of *G. hirsutum* are tightly clustered together (bootstrap values of 840 to 992), showing that such studies can also reliably group cotton varieties. Analysis of similarity showed that both trispecies hybrids were closer to cultivated cotton than to wild diploids. This is certainly due to the tetraploid nature of the cultivated parent that contributed twice to the hybrids composition, compared to the contribution of each wild diploid species. We used these data to generate Jaccard's coefficients of similarity between genotypes. Analysis of all crosses revealed 30.6 to 39.4% similarity between *G. hirsutum* and the wild species and 74.3 to 89.0% between the four cultivated varieties, while similarity between BC progenies and cultivated parent varied from 57.7 to 67.2%. This facilitated the choice of BC plants sharing the highest similarity with cultivated cotton.

RFLP analysis. Of the 49 RFLP markers amplified by PCR, 41 having specific amplification were labelled and used as probes. Twenty five probes distributed across chromosomes 1, 6, 10, 14, 15, 17, 20, 22, 23 and linkage groups A02, A07, D01, D03, D04, D07, U01, U07 hybridized successfully and generated 106 RFLPs of which, 54 (50.90%) were polymorphic. Analysis of introgression from the donor parent (*G. sturtianum*) showed the presence of 11 and 7 chromosomal segments in the trispecies hybrids HRS and TSH respectively. Introgression of these chromosomal segments was traced in subsequent backcross generations. In the BC_1 expressing the "glandless-seed and glanded-plant" trait, the presence of segments from chromosomes 1, 10, and linkage groups A07 and U07 of *G. sturtianum* was evidenced. This plant contains also a segment of chromosome 1 of *G. thurberi*. Introgression from both bridge species *G. raimondii* and *G. thurberi* was also identified in hybrids and backcrosses, but introgression seemed to be most common from *G. raimondii* than from *G. thurberi*. Among all the linkage groups analyzed, the chromosome 1 is the most introgressed with four markers in one or another backcross. In addition to these four markers, the introgression of wild diploid specific markers located on chromosome 15 indicated

probably introgression of chromosome 1 segments, since chromosome 15 is homoeologous to chromosome 1 (Reinisch *et al.*, 1994). One of such chromosome 1 markers was introgressed from *G. sturtianum* and *G. thurberi* respectively.

The introgression from wild species was of two types. In the most common type, RFLP alleles characteristic of cultivated and wild cottons were present, and the plants were heterozygous at the introgressed locus. In the second type, one of the *G. hirsutum* RFLP allele was replaced by the corresponding *G. sturtianum* or *G. raimondii* allele. This last type, evidenced on chromosome 1 with the probes A1204 and A1593, is consistent with reciprocal recombinations due probably to multivalent configurations observed at metaphase I (Vroh Bi *et al.*, 1998). Evidence of reciprocal recombinations also indicates that the crossing schemes developed in this study can lead to homoeologous chromosomes pairing and intergenomic exchanges.

Conclusion

The interest of molecular markers in breeding programs to tag introgression from parents has been demonstrated in many plants (Tanksley and Hewit 1988; Lee 1988). First, we have evaluated RAPD markers as tools for determining relationships between species, varieties, hybrids and backcross progenies of cotton. The results showed that RAPD can be used to differentiate cultivated genotypes of cotton, but also to estimate the genetic contribution of each parent to each member of a segregating population (i.e. backcross). Introgressed individuals whose genome composition most resembles the cultivated cotton genome can be selected for the next cross through genetic distance estimations. This could potentially accelerate the introgression of traits from genetically distant parents like those used in the present program. Second, using markers selected from the developing RFLP map of cotton (Reinisch *et al.*, 1994), we have demonstrated the ability to follow introgression of specific chromosome regions from parents to descendants through multiple generations. The fact that cotton RFLP markers initially developed from the cross *G. hirsutum* x *G. barbadense* (Reinisch *et al.*, 1994) works also in wide crosses involving other species shows that this map will prove to be useful in a wide range of cotton breeding programs assisted by DNA markers. Although the introgression of the "glandless-seed and glanded-plant" trait from Australian wild diploid cottons was previously attempted by cotton breeders (Dilday 1986; Shuijing and Biling 1993), chromosomes acting for the variable expression of this trait in introgressed genetic backgrounds remain unknown. Since the BC progenies analyzed here are segregating for both RFLP markers of known chromosomal positions and the desired trait, the present study is an important step towards the mapping of the "low-gossypol seed and high-gossypol plant" trait. Indeed, using more RFLP markers in large segregating populations could establish association between chromosomal segments or loci of *G. sturtianum* and the gland levels in different organs of the plants.

Acknowledgements

We thank Professor A.H. Paterson of Texas A&M University (USA) for providing the mapped RFLP markers. This research was supported by the convention 2.4565.95 of the Belgian 'Fonds de la recherche fondamentale collective'. The PhD scholarship of the first author was provided by the Ministry of Scientific Research of Ivory Coast. We are grateful to the International Cotton Advisory Committee (ICAC) for sponsoring the participation of the first author to the WCRC-2.

References

Dilday, R.H. (1986): Development of cotton plant with glandless seeds and glanded foliage and fruiting forms. Crop Sci. 26:639-641.

Endrizzi, J.E., E.L. Turcotte and R.J. Kohel. (1985): Genetics, cytology, and evolution of *Gossypium*. Adv. Genet. 23:271-375.

Felsenstein, J. (1985): Confidence limits on phylogenies: an approch using the bootstrap. Evolution 39:783-791.

Felsenstein, J. (1993): PHYLYP (Phylogeny Interference Package) version 3.5c. University of Seattle, Washington, USA.

Fryxell, P.A. (1979): The natural history of the cotton tribe. Texas A&M Univ. Press, College Station, TX, USA.

Gorham, J., and E.M. Young. (1996): Wild relatives of cotton and rice as sources of stress resistance traits. Proc. Meeting Tropical plants. Montpellier, 11-15 March 1996. CIRAD Montpellier France, Pp.39-52.

Lee, M. (1998): Genome projects and gene pools: New germplasm for plant breeding? Proc. Natl. Acad. Sci. USA. 95:2001-2004.

Lusas E.W. and G.M. Jividen. (1987): Glandless cottonseed: A review of the first 25 years of processing and utilization research. J. Am. Oil Chem. 64:839-854.

McMichael, S.C. (1954): Glandless boll in upland cotton and its use in the study of natural crossing. Agron. J. 46:527-528.

Mergeai G., I. Vroh Bi, J.P. Baudoin and P. du Jardin. (1998): Use of randomly amplified polymorphic DNA (RAPD) markers to assist wide hybridizaton in cotton. *In*: Biotechnology in Agriculture and Forestry, Vol. 42, pp122-139. Cotton YPS Bajaj (Ed). Springer-Verlag Berlin Heidelberg.

Nei M. and W. Lei. (1979): Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76:5269-5273.

Rafalski, J.A. and S.V. Tingey. (1993): Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends in genet. 9:275-280.

Reinisch A.J., J. Dong, C.L. Brubaker, D.M. Stelly J.F. Wendel and A.H. Patterson. (1994): A detailed RFLP map of cotton, *Gossypium hirsutum* x *Gossypium barbadense*: Chromosome organization and evolution in a disomic polyploid genome. Genetics 138:829-847.

Shuijin, Z. and L. Biling. (1993): Studies of the glandless seeds - glanded plant trait from *Gossypium bickii* into cultivated upland cotton (*G. hirsutum*). Coton Fibers Trop. 48:195-199.

Tanksley, S.D. and J. Hewitt. (1988): Use of molecular markers in breeding for soluble solids content in tomato - a re-examination. Theor. Appl. Genet. 75:811-823.

Texier, P.H. (1993): Le coton, cinquième producteur mondial d'huile alimentaire. Coton et Developpement 8:2-3.

Vroh Bi, I., J.P. Baudoin and G. Mergeai. (1998): Cytogenetics of cotton plants involved in the introgression of the "glandless-seed and glanded-plant" trait from *Gossypium sturtianum* Willis into upland cotton *Gossypium hirsutum* L. Plant Breed. 117 (in press).

Vroh Bi, I., P. du Jardin, G. Mergeai and J.P. Baudoin. (1997): Optimization and application of RAPD (random amplified polymorphic DNA) in a recurrent selection programme of cotton (*Gossypium* spp.). Biotechnol. Agron. Soc. Environ. 1:142-150.

Vroh Bi, I., L. Harvengt, A. Chandellier, G. Mergeai and P. du Jardin. (1996): Improved RAPD amplification of recalcitrant plant DNA by the use of activated charcoal during DNA extraction. Plant Breed. 115:205-206.

Young, N.D. and S.D. Tanksley. (1989): RFLP analysis of the size of chromosomal segments retained around the *Tm-2* locus of tomato during backcross breeding. Theor. Appl. Genet. 77:353-359.

Table 1. Segregation of 15 RAPD markers specific to *G. sturtianum* in 12 BC_1 derived from the trispecies hybrid *G.hirsutum-G.raimondii-G.sturtianum.*

	N° of BC1												Segregation
Specific RAPD markers	1	2	3	4	5	6	7	8	9	10	11	12	presence (+): absence (-)
OPC04-1200	+	+	+	+	+	+	+	+	+	+	+	+	12 : 0
OPC08-760	+	+	+	+	+	+	-	-	-	+	+	+	9 : 3
OPC13-1250	+	+	+	+	-	+	+	+	-	+	-	-	8 : 4
OPC16-820	-	+	-	+	-	+	-	+	+	-	-	+	6 : 6
OPC19-872	+	+	-	+	+	-	-	+	+	+	-	-	7 : 5
OPC19-500	-	-	-	-	-	+	-	-	-	-	-	-	1 : 11
OPB01-1500	-	-	+	-	+	-	-	+	-	-	-	-	3 : 9
OPB03-360	-	+	+	+	-	+	-	-	-	-	-	-	4 : 8
OPB03-180	+	-	+	+	+	-	-	-	+	+	+	+	8 : 4
OPB04-740	+	-	-	-	-	-	+	-	-	-	-	+	3 : 9
OPB10-1078	+	+	+	+	+	+	+	+	+	+	+	+	12 : 0
OPD03-690	+	-	+	-	-	+	-	-	+	-	-	+	5 : 7
OPD03-271	-	-	-	+	-	+	-	-	-	+	+	-	4 : 8
OPD-220	+	-	+	+	-	-	-	-	-	-	-	-	3 : 9
OPD13-310	-	+	-	+	-	+	-	+	-	-	+	+	6 : 6
Total of introgressed bands	9	8	9	11	6	10	4	7	6	7	6	8	

Figure 1. Dendrogram of cultivated and wild species of cotton based on UPGMA cluster analysis.

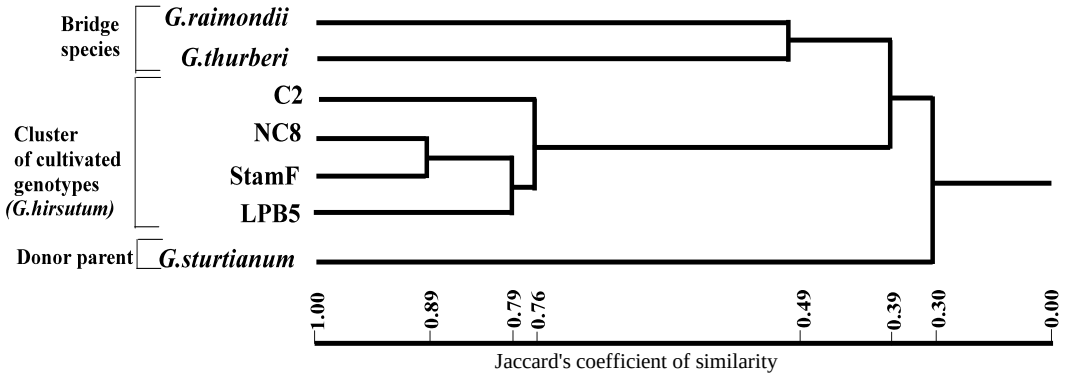

Figure 2. Neighbour-joining tree based on genetic distances of Nei. The tree is rooted with data from *G. sturtianum* chosen as outgroup taxon. Numbers adjacent to nodes indicate percentage of bootstrap replicates supporting that node.