II.4 Use of Random Amplified Polymorphic DNA (RAPD) Markers to Assist Wide Hybridization in Cotton

G. Mergeai[1], I. Vroh Bi[1,2], J.P. Baudoin[1], and P. du Jardin[2]

1 Interest of RAPD Markers to Assist Interspecific Breeding Programs

In recent years, molecular marker-assisted selection has been integrated into several plant breeding programs to select traits of agronomical importance. Isozymes were initially sought for this purpose, but their use was hindered by the rather small number of loci and alleles available for analysis. The development of molecular biology has resulted in alternative DNA-based procedures for the detection of polymorphism. These include restriction fragment length polymorphism (Tanksley et al. 1989), random amplified polymorphic DNA (Williams et al. 1990), and cleaved amplified polymorphic sequences (Akkaya et al. 1992).

Although RFLPs were widely used in genetic studies concerning cotton (Brubaker and Wendel 1994; Reinisch et al. 1994) and other crop species they promise to be a cost-effective diagnostic technology for plant breeding. This technique uses a single short oligonucleotide primer (10 bp) of arbitrary DNA sequence and polymerase chain reaction (PCR)-mediated amplification of random fragments from genomic DNA. The main advantages of the RAPD technique include the ease and rapidity of analyses: it does not require preparation of genomic or cDNA clones, restriction enzyme digestion of DNA, Southern blotting, or hybridization with radioactively labeled probes (data are directly scored from agarose gels). Other advantages concern the use of universal sets of primers that are commercially available, its requirement for minimal substrate DNA and negligible primer quantity, and its lower cost. All these characteristics make RAPD analysis particularly well adapted for detecting polymorphic situations in unknown germplasm where no preexisting set of RFLP clones is available (Heun and Helentjaris 1993). Because of its simplicity, RAPD analysis can be performed in a moderately equipped laboratory, and since only a small amount of tissue is required, plants in a breeding program can be screened at the seedling stage. Moreover, because RAPD

[1] Unité de Phytotechnie des Régions intertropicales, Faculté Universitaire des Sciences agronomiques, B-5030 Gembloux, Belgium.
[2] Unité de Biologie végétale, Faculté Universitaire des Sciences agronomiques, B-5030 Gembloux, Belgium

Biotechnology in Agriculture and Forestry, Vol. 42
Cotton (ed. by Y.P.S. Bajaj)

analyses require much smaller amounts of DNA and less pure DNA than RFLP analysis, the RAPD method is particularly attractive to genetists working with species recalcitrant to nucleic acid extraction, such as cotton (Multani and Lyon 1995; Tatineni et al. 1996).

Clear and reproducible detection of polymorphims has been demonstrated using RAPD in many plants including *Dastica*, *Yucca*, and *Helianthus* (Fritsch et al. 1993), corn (Heun and Helentjaris 1993), wheat (Devos and Gale 1992; Vierling and Nguyen 1992), barley (Tinker et al. 1993), rice (Heun et al. 1994), and others (Hu and Quiros 1991).

Since its discovery, the RAPD technique has been applied in plants for the construction of genetic linkage maps (Reiter et al. 1992; Tulsieram et al. 1992; Kesseli et al. 1994; Rowland and Levi 1994), estimation of genetic relationships (He et al. 1992; Vierling and Nguyen 1992; Halward et al. 1991, 1992), tagging disease resistance traits (Michelmore et al. 1991), identification of cultivars (Hu and Quiros 1991; Quiros et al. 1992; Klein-Landhorst et al. 1991; Wilde et al. 1992), parentage determinations (Welsh et al. 1991), population genetics (Van Hausden and Bachmann 1992), analyzing interspecific hybrids and exotic germplasm utilization (Arnold et al. 1991; Baird et al. 1992; Veli-Matti Roka et al. 1992; McCoy and Echt 1993).

Given all the putative advantages of the RAPD analysis, we have chosen this technique to investigate the potential of molecular markers to monitor the introgression of diploid *Gossypium* genomes into upland cotton. This study was realized in the frame of a three-species breeding program involving the main cultivated amphidiploid, *G. hirsutum*, two wild diploid American cottons (*G. thurberi* and *G. raimondii*), and a diploid wild species endemic to Australia (*G. sturtianum*). Like all the other Australian cottons, *G. sturtianum* produces gossypol-free seeds, while gossypol glands are present in the rest of its aerial parts. Our goal is to transfer this glandless-seed and glanded-plant trait into cultivated cotton in order to transform it into a true food crop whilst preserving one of its natural defense mechanism (the gossypol glands) against insect pests. Because direct crossing strategies between *G. hirsutum* and *G. sturtianum* followed in the past by other laboratories were not successful to obtain the introgression of the researched character into upland cotton (Dilday 1986; Altman et al. 1987; Koto 1989; Rooney et al. 1991), we created two different trispecific hybrids by bridge crosses in order to confront all the chromosomes of *G. sturtianum* to the genome of *G. hirsutum* in a tetraploid structure allowing recombination (Mergeai 1992). In this context a system of random amplified polymorphic DNA (RAPD) markers was developed to facilitate the transfer of *G. sturtianum* genes into *G. hirsutum* genome by hybridization and backcrossing.

2 Materials and Methods

2.1 Plant Material

Two trispecific allotetraploids, including the Australian diploid species *G. sturtianum* ($2C_1$), have been produced using either *G. thurberi* ($2D_1$) or *G. raimondii* ($2D_5$) as bridge species (Mergeai et al. 1995). These hybrids are respectively designated by the initials TSH for G405 (*G. thurberi* × *G. sturtianum* × *G. hirsutum*: $D_1C_1D_hA_h$) and HRS for G376 (*G. hirsutum* × *G. raimondii* × *G. sturtianum*: $A_hD_hC_1D_5$). Figure 1A,B shows the crossing scheme followed to obtain these genotypes. Both trispecific hybrids were backcrossed by three *G. hirsutum* ($2[AD]_1$) glanded varieties originating from Africa (C2, Stam F and NC8) and one US glandless cultivar (LPB5). Twenty seven BC_1 plants obtained by backcrossing each trispecific hybrid to *G. hirsutum* were evaluated to test the suitability of RAPDs for detecting the amount of introgression from wild diploid species into *G. hirsutum*. Genomic DNA was extracted from all the genotypes involved in the breeding program. To obtain the BC_1 plants analyzed in this program it was necessary to treat the flowers immediately after pollination with a solution of gibberellic acid and naphthalene acetic acid (Gill and Bajaj 1987), to cultivate the seeds in vitro for 2 weeks on the medium of Stewart and Hsu (1977) and to graft most of them

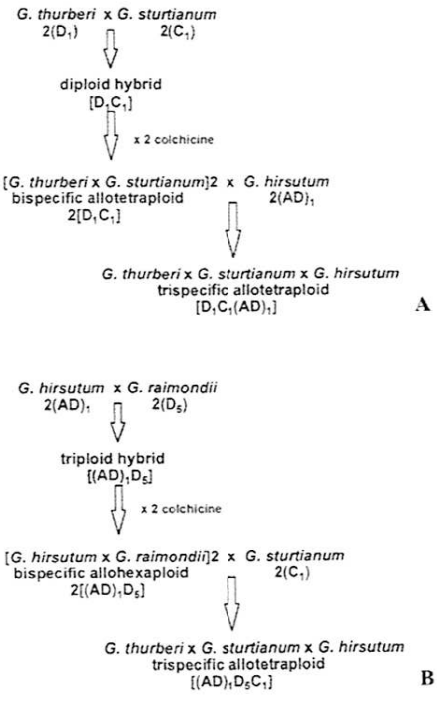

Fig. 1. **A** Creation scheme of the *G. thurberi-G. sturtianum-G. hirsutum* (TSH) tispecific hybrid (G405). **B** Creation scheme of the *G. hirsutum-G. raimondii-G. sturtianum* (HRS) trispecific hybrid (G 376)

on vigorous *G. hirsutum* seedlings. All the plants used in the experiments were grown in a greenhouse in Gembloux. Young leaves, when available, were preferably used for DNA extraction.

2.2 DNA Extraction and Amplification

The detailed protocol followed for DNA extraction has been published elsewhere (Vroh Bi et al. 1996). This protocol is adapted from the standard CTAB method of Murray and Thompson (1980), and includes incubation of the tissue homogenate with activated charcoal. This original modification proved to significantly improve the subsequent PCR amplification of the extracted DNA.

Using a Techne PHC-3 DNA thermal cycler, we optimized RAPD amplification conditions proposed by Williams et al. (1990). The main parameters studied to adapt the conditions to cotton DNA amplification are the concentration of template DNA, primers, $MgCl_2$, and dNTPs, the annealing temperature, the number of cycles, the duration of the first denaturation, different sources of polymerase, and the design of premixes to avoid variations between different reactions (Vroh Bi et al. 1997).

Seventy five decamers from Operon Technologies Inc. (Alameda, California, USA) were tested. Amplification reaction volumes were 25 µl, each containing 50 mM KCl, 10 mM Tris-HCL (pH 9), 2 mM MgCl2, 200 µM each of dATP, dCTP, dGTP, and dTTP (Pharmacia), 0.8 µM random primer, 1 U Taq Polymerase (Pharmacia), and 50 ng template DNA. Reaction mixtures were overlaid with mineral oil and exposed to the following conditions: 94 °C for 5 min, followed by 45 cycles of 1 min at 94 °C, 1 min at 34 °C, 2 min at 72 °C, ending with 7 min 30 s. at 72 °C, and then kept at 15 °C till electrophoresis. The RAPD products were resolved on 1.8% agarose gel and visualized by ethidium bromide staining.

All reactions were repeated twice, and only reproducible bands were considered.

2.3 Data Analysis

The data for RAPD analysis were scored from photographs of the ethidium bromide stained agarose gels.

Bands were considered to be the same if they occurred at exactly the same position on the electrophoresis gel. Band sizes for each genotype were determined by comparison with lambda and Phi X 174 DNA, respectively, digested by HindIII and HaeIII. For each genotype and for all the informative primers used, a RAPD bands presence (1)/absence (0) profile was recorded into a matrix. A cluster analysis based on UPGMA (unweighted pairgroup method with arithmetical mean) was then performed with systat version 5.2 software (Systat 1992). RAPDs are identified by Operon kit, primer, and fragment size (i.e., OPB-01-1500).

3 Results and Discussion

3.1 Selection of RAPD Markers Suitable for Introgression Monitoring in Cotton

The main objective of our research was to assess the suitability of the RAPD technique for monitoring the introgression of wild diploid species characters into *G. hirsutum*. For this purpose, 75 RAPD primers on four *G. hirsutum* cultivars and three diploid species implicated in the creation and backcrossing of TSH and HRS trispecific hybrids were tested. The aim of this preliminary investigation was to select primers exhibiting polymorphism between the parents of two trispecific hybrids in order to detect species-specific RAPD markers that could be used to trace the introgression of wild diploid DNA segments in the genome of *G. hirsutum* (identification of RAPD markers that are present in the trispecific hybrids and their ascendants). Among the 75 primers used, 6 (8%) primers did not amplify detectable products, 23 (30.7%) failed to amplify in one or two parents and gave clear amplification in the remaining parents, 46 (63.3%) succeeded in amplifying in all the parents. Of these 46 informative primers, 30 were tested twice to generate the results showed in this study. The sequences and the number of bands generated by the 30 primers are given in Table 1.

These primers produced 375 different RAPD bands, among which only 36 (9.6%) were shared by all the hybrids parents, the 339 (90.4%) other fragments were polymorphic.

The number of bands amplified per primer ranged from 6 (OPC-10, OPC-15 and OPC-17) to 20 (OPB-18) with a mean value of 12.5 bands per primer. These values are rather high for RAPD amplification compared to the average numbers of amplified bands recorded in other crops: (3 fragments in *Triticum turgidum* L.; Joshi and Nguyen 1993), 4.3 fragments in *Solanum tuberosum* L. (Masuelli et al. 1995), 6.7 in corn (Heun and Helentjaris 1993), and 9 in coffee (Orozco-Castillo et al. 1994).

The large size of the *Gossypium* genomes, 2246 Mbp for *G. hirsutum* (Arumunganathan and Earle 1991), the high frequency of repetitive DNA sequences (Walbot and Duke 1976), and the quality of the DNA produced by the improved extraction method are possible explanations for the relatively high numbers of amplified products obtained.

The amplified fragment sizes varied from 160 to 3200 base pairs (bp). The most reproducible types of RAPD fragments, as might be expected, were those of the greatest intensity in any one reaction. "Minor" fragments seemed to possess the greatest propensity for irreproducibility, as was also found in corn (Heun and Helentjaris 1993) and canola (Marshall et al. 1994). Species-specific RAPD fragments which were not present in both reaction repetitions were eliminated. On the average, the great majority of the amplified fragments produced by the 30 selected primers were reproducible and the interpretation of the banding pattern was straightforward. RAPD markers unique to the different parents that are also present in TSH and HRS hybrids were easily

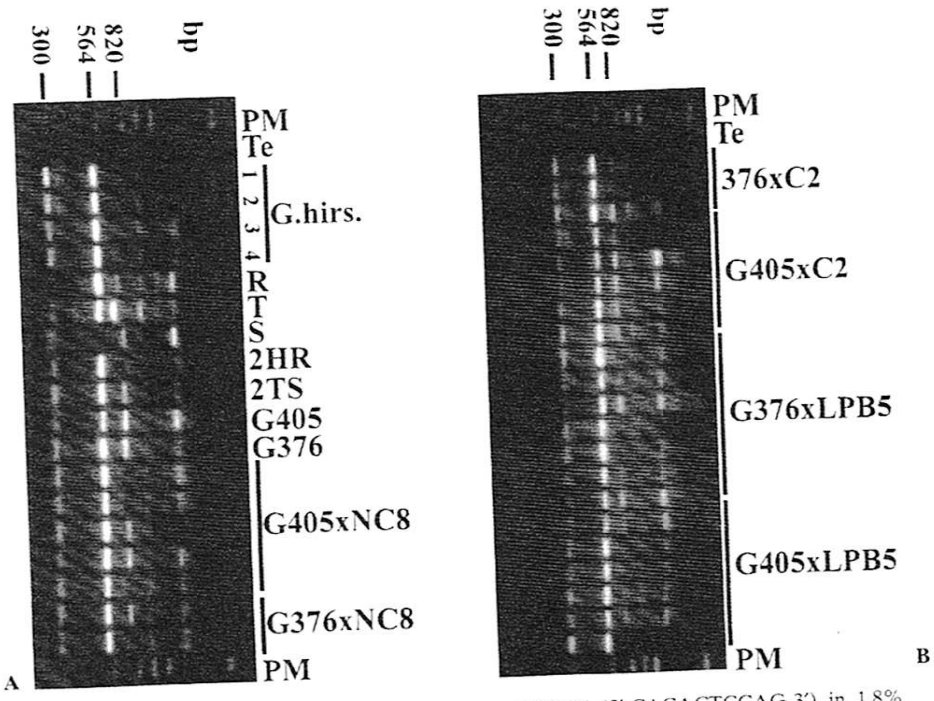

Fig. 2A,B. RAPD bands generated using primer *OPC-16* (5′-CACACTCCAG-3′) in 1.8% agarose gels. The cultivated upland cotton *G. hirsutum* = *G. hirs.* (with four cultivars: *1 Stam F*; *2 NC8*; *3 C2*; *4 LPB5*); *R G. raimondii*; *T G. thurberi*; *S G. sturtianum*; *2HR* (*G. hirsutum* × *G. raimondii*) doubled; *2TS* (*G. thurberi* × *G. sturtianum*) doubled; Trispecies hybrids *G405* = (2TS × *G. hirsutum*) and *G376* = (2HR × *G. sturtianum*). The differents groups of backcrosses 1 are indicated by their crosses (*G405* × *NC8*, *G376* × *NC8*, *G376* × *C2*, *G405* × *C2*, *G376* × *LPB5*, and *G405* × *LPB5*). Each lane contains 12 µl of RAPD products; *PM* represents the molecular size marker (DNA of lambda/Hind III and PhiX 174/Hae III); *Te* is the negative control without DNA, and molecular weights are given in base pairs (bp). The 820-bp fragment is specific to the Australian wild diploid donor *G. sturtianum*; this fragment is present in the two trispecies hybrids, G405 and G376, and clearly segregates in the backcrosses 1. The polymorphism between all the genotypes under study is easily detectable

Table 2. Transmission of RAPD diploid wild species-specific markers to trispecific hybrids

Tested genotype	No. of specific RAPD markers in the parent genotypes	No. of specific RAPD markers in the TSH hybrid (G405)	No. of specific RAPD markers found in the HRS hybrid (G376)
G. sturtianum	49	18[a]	20[a]
G. raimondii	12	0	4
G. thurberi	13	4	0

[a] The two trispecies hybrids shared 18 bands specific to *G. sturtianum*.

that some *G. sturtianum*, *G. thurberi*, and *G. raimondii* markers are not present in the trispecies hybrids. The same kind of observation was made by McCoy and Echt (1993) on *Medicago* trispecies hybrids. Actually, one can speculate that the loci for markers not present in the trispecies hybrids are on chromosomes that were not transmitted to them. Another complementary explanation is given by Heun and Helentjaris (1993), who stated that many unexpected RAPD inheritance results could be attributed to the "competition" amongst fragments where one very intense product seemed to outcompete and overwhelm other expected fragments, even those of very high yields in inbred situations. This may also explain the reappearance in the BC_1 progeny of some RAPD markers that were identified as characteristic of the parent species and vanished in the trispecific hybrids. Indeed, as the presence of some fragments can interfere with the amplification of others, the absence of these latter from some genotypes does not prove the absence of the corresponding target sequences, which may "reappear" in segregating situations if the former fragments are allowed to segregate away.

To minimize the risk of interference between fragments for amplification, we took into consideration only the products less sensitive to the competition of other fragments, that means intense specific bands that were present in the parent species and in the trispecific hybrids. A total of 22 different markers of this kind were detected for *G. sturtianum* in both trispecific hybrids against 4 markers for *G. raimondii* and *G. thurberi*.

As the two wild diploid species belonging to genome D (*G. thurberi* and *G. raimondii*) are phyletically very close to the D_h subgenome of *G. hirsutum*, it is not surprising to observe for them lower numbers of specific RAPD markers than for *G. sturtianum*, which belongs to the more remote C genome. Specific RAPD bands detected in both D genome species seem to be more sensitive to the competition of other markers for amplification.

This higher sensitivity to competition may limit the usefulness of the RAPD markers to trace introgression of genes from diploid species considered as the most closely related to *G. hirsutum* compared to introgression of genes from a more distant genome such as *G. sturtianum*.

3.2 Assessment of Introgression in Upland Cotton Using RAPD Markers

The species-specific RAPD markers detected in our investigations were used to assess the introgression of cotton wild diploid DNA segments in the BC_1 plants produced by backcrossing the TSH and HRS trispecific hybrids with four *G. hirsutum* cultivars.

The distribution of the species-specific markers observed in each BC_1 plant are given in Tables 3, 4, and 5.

Figure 2A,B shows the RAPD markers produced with primer OPC-16 by all the genotypes compared in our work. RAPD markers specific to the wild diploid parents of the trispecific hybrids that are also present in some plants of their BC_1 progeny can be easily identified. This is, in particular, the case for the band OPC 16-820, which is specific to *G. sturtianum*. This fragment is present

in both trispecific hybrids and in 17 introgressed genotypes out of the 27 BC_1 plants analyzed.

In the case of a totally randomly distribution of the chromosomes during meiosis, a 50% transfer of a RAPD wild species-specific marker observed in one of the trispecific hybrids to the BC_1 generation should be expected if this marker is present in only one chromosome and if intergenomic recombination and deletion are rare.

Table 3 shows the segregation of the *G. sturtianum*-specific markers in the backcross generation. All 22 *G. sturtianum*-specific markers selected from the analysis of the trispecific hybrids were found among the tested BC_1 plants. The number of *sturtianum*-specific markers per plant ranged from 7 to 20 (mean = 13 markers per plant).

The marker OPB-01-1500 was found in significantly less than 50% of the BC_1 plants, OPC-08-760 was significantly found in more than 50% of the BC_1. Among the rest of the RAPD markers, they were systematically observed in all the BC_1 plant and 17 fit the expected 1:1 Mendelian segregation ratio. As was noted by Masuelli et al. (1995) in the case of *Solanum*-interspecific hybrids, one can speculate that the loci for markers not present in the BC_1 plants are on chromosomes that were not transmitted from the trispecies hybrids.

On the contrary, the markers that are systematically present in the backcrossed progeny of the trispecific hybrids (OPC-04-1200, OPB-10-1078, and OPB-18-3200) should be carried on *G. sturtianum* chromosomes that are preferentially transmitted to the BC_1 plants. The average numbers of chromosome pairs in AC and DC diploid hybrids are respectively 8 and 5 (Endrizzi et al. 1985). This means that at least five chromosomes of the Australian diploid species present enough remnant homology to pair regularly with the chromosomes of the A_h and D_h subgenomes. These chromosomes with a higher pairing affinity certainly have more chance to be transmitted to the progeny of the trispecific hybrids than the others.

The presence of markers at more than one chromosome in the *G. sturtianum* genome might also explain the transfer frequency of RAPD markers significantly higher than 50%. Extensive gene duplication has been found within diploid cotton species from A and D genomes (Reinisch et al. 1994). Therefore, it is likely that some RAPD markers could be present at several unlinked loci. However, it is worthwhile to note that the number of markers that are transmitted with an abnormally high frequency is rather low and could be explained for most of them by the preferential pairing capacities of the chromosome that carries them.

With the detection of more *G. sturtianum*-specific RAPD markers it should be possible to identify synteny groups among them and assess their organization. The number of Australian species with chromosomes that are preferentially transferred to the BC_1 and the types of translocation, deletion, or insertion that occur in BC_1 should be revealed by coupling synteny groups analyses with GISH (genomic in situ hybridization) (Jacobsen et al. 1995).

All but one of the D species-specific RAPD markers fit the 1:1 segregation ratio (Tables 4, 5). No explanation can be given so far for this exception. The data recorded for the rest of the *thurberi*- and *raimondii*-specific markers

Table 3. Segregation of RAPD markers specific to *G. sturtianum* in the 27 BC_1 plants obtained from G376 and G405

Cross formulas and genotype nos. / Specific RAPD markers	G405 × NC8				G376 × NC8		G376 × C2			G405 × C2				
	5	6	9	7	3	4	11	S1	2	10	5	7	6	3
OPC04-1200	+	+	+	+	+	+	+	+	+	+	+	+	+	+
OPC08-840	−	+	+	−	−	+	+	−	−	−	−	+	+	+
OPC08-760	+	+	+	+	+	+	+	+	+	+	−	−	+	−
OPC13-1250	+	+	+	+	+	+	+	+	−	−	−	−	+	−
OPC13-800	−	−	−	+	−	−	+	+	−	+	+	+	+	−
OPC14-1300	+	+	−	−	−	−	+	−	+	−	−	+	+	+
OPC16-820	−	−	+	+	−	+	−	+	−	+	+	+	+	+
OPC18-740	+	+	−	−	−	+	+	−	−	−	+	−	+	+
OPC19-872	−	−	−	+	+	+		+	+	−	+	+	+	+
OPC19-500	+	−	+	−	−	−	−	−	−	−	+	+	−	−
OPB01-1500	+	+	−	+	−	−	+	−	+	+	−	+	−	−
OPB03-360	+	+	+	−	−	+	+	+	−	−	+	−	+	+
OPB03-180	−	−	+	+	+	−	+	+	+	+	−	−	+	+
OPB04-740	+	+	+	+	+	−	−	−	−	−	+	+	−	−
OPB10-1078	+	+	+	+	+	+	+	+	−	−	−	+	+	+
OPB18-3200	+	+	+	+	+	+	+	+	+	+	+	+	+	+
OPD03-690	+	−	+	+	+	−	+	−	−	−	−	−	+	+
OPD03-510	+	+	+	−	+	+	−	−	−	−	−	+	+	+
OPD03-271	−	+	+	−	−	−	−	+	−	−	+	−	+	−
OPD-220	+	+	+	+	+	−	+	+	−	+	+	+	−	−
OPD13-310	−	+	−	+	−	+	−	+	−	+	+	+	+	−
OPD18-690	+	−	+	+	+	+	−	+	+	−	+	+	−	−
Total of introgressed bands per backcross	15	15	16	15	12	13	14	14	9	10	13	17	17	12

[a] $\chi^2 > 3.85$ ($P < 0.05$) indicates that the frequency of transfer of the marker is significantly different from 50%.

[b] χ^2 values were not calculated for markers with presence or absence number <5.

G376 × LPB5							G405 × LPB5						segregation in backcrosses I	
17	18	6	11	10	14	5	6	8	1	10	12	4	Presence:absence	Chi-square
+	+	+	+	+	+	+	+	+	+	+	+	+	27:0	[a]
+	−	−	+	−	+	+	+	+	+	+	−	−	15:12	0.33
+	−	−	−	+	+	+	+	−	+	+	+	+	20:7	6.25[a]
+	+	+	−	+	−	−	+	−	+	+	+	+	18:9	3
+	+	+	−	−	−	+	+	−	−	+	+	+	15:12	0.33
−	+	+	+	+	−	−	−	−	+	−	+	−	13:14	0.03
+	−	+	+	−	−	+	+	+	−	−	+	+	17:10	1.06
+	−	−	+	+	+	+	+	−	−	+	+	+	16:11	0.925
−	−	+	+	+	−	−	+	−	−	+	+	+	16:11	0.925
+	−	−	−	−	−	−	+	+	−	−	+	−	10:17	1.06
−	−	+	−	−	−	−	−	−	−	−	−	−	8:19	4.48[a]
+	−	−	−	−	−	−	−	−	+	+	+	−	13:14	0.03
−	−	−	+	+	+	+	−	−	−	+	+	−	15:12	0.33
−	+	−	−	−	−	+	−	+	−	+	+	−	12:15	0.33
+	+	+	+	+	+	+	+	+	+	+	+	+	27:0	[b]
+	+	+	+	+	+	+	+	+	+	+	+	+	27:0	[b]
+	−	−	+	−	−	+	−	+	+	+	+	+	16:11	0.925
−	−	+	−	−	−	−	+	+	−	+	+	−	14:13	0.03
+	−	−	−	+	+	−	−	+	+	−	+	−	10:17	1.06
−	−	−	−	−	−	−	−	−	+	+	+	+	13:14	0.03
+	−	+	−	−	+	+	+	+	−	+	+	+	17:10	1.06
+	−	−	−	−	+	+	+	+	−	+	+	+	16:11	0.925
15	7	10	10	10	10	13	13	12	11	16	20	13		

Table 4. Segregation of RAPD bands specific to *G. ramonalii* in the G376 BC_1 plants

Cross formulas and genotype nos. / Specific RAPD markers	G376 × NC8	G376 × C2				G376 × LPB5								Segregation in backcrosses I	
	3	4	11	S1	2	17	18	6	11	10	14	5	Presence : absence	Chi-square
OPB01-1130	–	+	+	–	+	–	+	+	–	+	–	–	6:6	0
OPB05-980	–	+	+	+	+	–	–	–	+	+	–	+	7:5	0.33
OPB05-872	–	–	–	+	+	–	+	–	–	–	–	–	3:9	3
OPC11-672	–	–	+	–	+	–	+	+	–	–	–	–	5:7	0.33
Total of introgressed bands per backcross	0	2	3	2	4	4	3	2	1	2	0	1		

Table 5. Segregation of RAPD bands specific to *G. thurberi* in the G405 BC_1 plants

Cross formulas and genotype nos. / Specific RAPD markers	G405 × NC8				G405 × C2					G405 × LPB5						Segregation in backcrosses I	
	5	6	9	7	10	5	7	6	3	6	8	3	10	12	4	Presence:absence	Chi-square
OPC13-460	–	–	+	–	–	–	–	–	–	–	+	+	–	+	+	5:10	1.66
OPC18-370	–	–	–	–	+	+	+	–	–	+	+	–	–	–	–	5:10	1.66
OPC18-1078	+	+	–	–	–	+	+	–	–	+	+	+	–	–	–	7:8	0.066
OPB10-496	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	1:14[a]	12.13[a]
Total of introgressed bands per backcross	1	1	1	0	2	2	2	0	0	2	3	2	0	1	1		

[a] $\chi^2 > 3.85$ ($P < 0.05$) indicates that the frequency of transfer of the marker is significantly different from 50%.

are consistent with the pairing affinity of D chromosomes with the D_h subgenome.

3.3 Prospects of Gene Mapping with RAPD Markers in the Progeny of Trispecific Hybrids

The backcross progeny should be either deficient for entire *G. sturtianum* chromosomes or for fragments of chromosomes, depending on whether homeologous recombination took place. In either case, groups of linked markers are expected to be absent from the genome of most individuals. Linkage could thus be established by the comparative testing of a relatively small number of backcross progeny (20–100, depending on the rate of chromosome loss). If no homeologous recombination is taking place, linkage groups will correspond to *G. sturtianum* chromosomes. If homeologous recombination takes place, linkage groups as defined by a single individual may cover only part of a chromosome. However, chromosomal groups could be defined by comparing individuals containing overlapping sets of missing markers.

If the backcross 1 progeny is tested phenotypically for the presence of genes of interest, positive and negative plants could be bulked and subjected to bulk segregant analysis, as described by Michelmore et al. (1991). For dominant genes, the markers absent in the negative plants should be linked to the gene of interest. Therefore, in order to facilitate the transfer of the glanded-plant and glandless-seed genes from *G. sturtianum* to *G. hirsutum*, we plan to analyze the first backcross progeny for *G. sturtianum* RAPD markers linked to the desired character by bulk segregant analysis.

3.4 Genetic Relationships Between Introgressed Tetraploid Cotton and Their Ascendants

One of the goals of our study was to investigate the efficiency of RAPD markers in determining accurately the genetic relationships between the BC_1 and their parents (*G. hirsutum*, *G. sturtianum*, *G. raimondii*, *G. thurberi*, and the trispecies G405 and G376).

Figure 3 represents the dendrogram generated by UPGMA analysis of the crosses G405 × NC8 and G405 × C2.

Three clusters can be observed. The first cluster comprises only *G. sturtianum*, showing clearly that this Australian wild diploid species of genome C is the most remote from the other genotypes studied. The American wild diploid *G. thurberi* (genome D) clusters also apart; it is situated between *G. sturtianum* and the remaining genotypes, which form the third cluster. This last cluster is divided into four subgroups; the first consists of the two cultivars, NC8 and C2, of *G. hirsutum* (genome AD), the second is the backcross 1 G405 × C2/10 that failed to integrate the group in which it was expected to be. With this exception, all the backcrosses derived from the same cross group together, G405 × NC8 backcrosses forming the third group and backcrosses G405 × C2

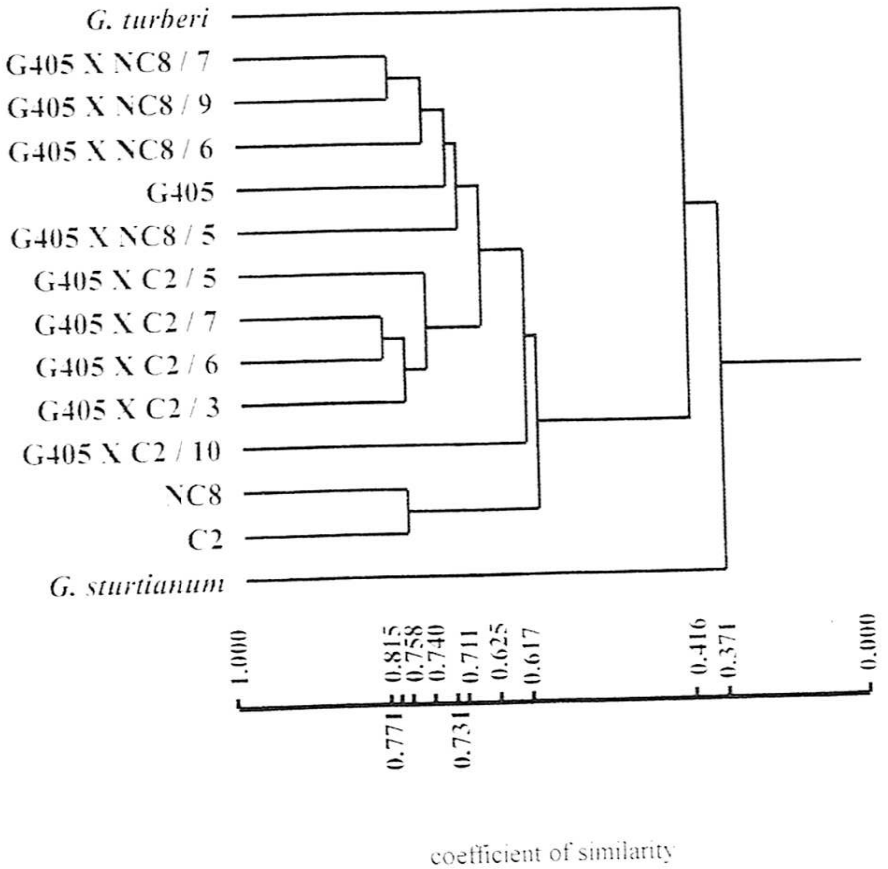

Fig. 3. Dendrogram obtained by UPGMA cluster analysis showing the genetic relationship between the backcrosses 1 and their parents used in the crosses G405 × NC8 and G405 × C2

the fourth. The placement of the trispecies G405 into the G405 × NC8 group instead of the G405 × C2 group is not surprising since G405 was itself generated by the cross (*G. thurberi* × *G. sturtianum* × *G. hirsutum* NC8); it is due to the fact that in a backcross 1 of G405 × NC8, the cultivated cotton NC8 is double-dosed whereas in a backcross 1 of G405 × C2, NC8 is only single-dosed.

In all crosses, the similarity between the wild species and the cultivated parent ranges from 31.30 to 39.20% and the similarity between the three upland cotton cultivars used varies from 78.3 to 78.70%. RAPD data also show that a large amount of genetic variation exists in the backcrosses 1, with a similarity ranging from 57.70 to 81.50% between them. This wide variation in the backcrosses may suppose that genetic events such as translocations, deletions, and homeologous recombinations frequently occur in such materials.

The composition of the trispecies hybrids G405 and G376 was confirmed by RAPD; these hybrids are more similar to the cultivated upland cotton than to the wild species parents *G. raimondi*, *G. thurberi*, and *G. sturtianum*.

These results are in close agreement with the current taxonomy of the *Gossypium* genus. They confirm the conclusions of previous research works (Multani and Lyon 1995; Tatineni et al. 1996) concerning the interest of RAPDs to determine genetic relationships within a diverse array of *Gossypium* germplasm.

4 Conclusion and Prospects

A system of random amplified polymorphic DNA (RAPD) markers has been developed to monitor the transfer of wild diploid cotton genes into *Gossypium hirsutum* by hybridization and backcrossing. It is shown here that RAPD markers, specific to wild diploid species, can be generated quickly and efficiently. As Klein-Lankhorst et al. (1991) demonstrated for potato, and McCoy and Echt (1993) for *Medicago*, the RAPD technique provides a highly effective means to fingerprint cotton species and interspecific hybrids. Moreover, our results demonstrate the ability of RAPD markers to reliably differentiate between *G. hirsutum* varieties and provide a molecular tool to assess introgression from several wild diploid species (*G. sturtianum*, *G. thurberi*, and *G. raimondii*) into *G. hirsutum*.

Given the important genetic similarities existing between the diploid species belonging to A and D cytotypes and both *G. hirsutum* subgenomes, RAPD seems to be more convenient to detect useful specific markers in species from phyletically more remote genomes (such as the C cytotype).

One obvious use of the markers we selected is to identify the BC_1 individuals introgressed for a useful character and which simultaneously show the minimum number of wild species RAPD markers.

It should also be possible to use them in a bulk segregant analysis according to the method developed by Michelmore et al. (1991) in order to identify markers associated with the glandless-seed and glanded-plant trait of *G. sturtianum*.

For homozygous parents, the gel pattern of RAPD products of interspecific hybrid should be the sum of the parental patters. However, we found a relatively high rate of exception to this prediction. The same kind of observation was made in the analysis of potato and alfalfa interspecific hybrids by Masuelli et al. (1995) and McCoy and Echt (1993). The appearance of products that are not present in the parents, or the disappearance of some parental products, could be explained by primer competition during annealing (Williams et al. 1990) or by inversion and deletion due to heterogenetic pairing in the intermediate tetraploid (for the HRS hybrid) or hexaploid (for the TSH hybrid) stages.

RAPD analysis is sensitive to minor changes in the amplification procedure because it uses short (10-mer primers). To limit this problem, Paran and

Michelmore (1993) proposed the sequencing of RAPD products and their conversions to normal PCR products by the use of longer primers. These markers constitute a type of sequence tagged site (STS, Olson et al. 1989) and have been termed sequence characterized amplified regions (SCARs).

As was observed by Masuelli et al. (1995) in the framework of an interspecific breeding program of potato, the conversion into SCARs of the RAPD primers we selected should improve the accuracy of the analysis of introgression from wild diploid species in subsequent generations. Furthermore, combination of primers in RAPD may provide additional species-specific markers.

Cluster analysis of genetic distances between the ascendants and the progeny of the trispecific hybrids shows phylogenetic relationships in agreement with the known lineage of the compared genotypes.

The clear and reliable detection of polymorphism among different upland cotton species, cultivars, and backcrosses makes us anticipate that RAPD markers will prove to be extremely useful in introgressing economic traits in cotton.

Acknowledgment. This research was supported by the Belgian Fonds de la recherche fondamentale collective.

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Altman DW, Stelly DM, Kohel RJ (1987) Introgression of the glanded-plant and glandless-seed trait from *Gossypium sturtianum* Willis into cultivated upland cotton using ovule culture. Crop Sci 27:880–884

Arnold ML, Buckner CM, Robinson (1991) Pollen-mediated introgression and hybrid speciation in Louisiana irises. Proc Natl Acad Sci USA 88:1398–1402

Arumunganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Baird E, Lojkowska E, Ehlenfeldt MK, Kelman A, Helgeson JP (1992) Molecular characterization of inter- and intra-specific somatic hybrids of potato using random amplified polymorphic DNA (RAPD) markers. Mol Gen Genet 233:469–475

Beckman JS, Soller M (1986) Restriction fragment length polymorphism and genetic improvement of agricultural species. Euphytica 35:111–124

Brubaker CL, Wendel JF (1994) Reevaluating the origin of domesticated cotton (*Gossypium hirsutum*, Malvaceae) using nuclear restriction fragment length polymorphisms (RFLPs). Am J Bot 81(10):1309–1326

Devos KM, Gale MD (1992) The use of random amplified polymorphism DNA markers in wheat. Theor Appl Genet 84:194–200

Dilday RH (1986) Development of cotton plant with glandless seeds and glanded foliage and fruiting forms. Crop Sci 26:639–641

Endrizzi JE, Turcotte EL, Kohel RJ (1985) Genetics, cytology and evolution of *Gossypium*. Adv Genet 23:271–275

Fritsch P, Hanson MA, Spore CD, Pack PE, Rieseberg L (1993) Constancy of RAPD primer amplification stength among distantly related taxa of flowering plants. Plant Mol Biol Rep 11:10–20

Gill MS, Bajaj YPS (1987) Hybridization between diploid (*Gossypium arboreum*) and tetraploid (*Gossypium hirsutum*) cotton through ovule culture. Euphytica 36:625–630
Halward T, Stalker HT, LaRue E, Kochert G (1991) Genetic variation detectable with molecular markers among unadapted germplasm resources of cultivated peanut and related wild species. Genome 34:1013–1020
Halward T, Stalker HT, LaRue E, Kochert G (1992) Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogea*). Plant Mol Biol 18:315–325
He S, Ohm H, Mackenzie S (1992) Detection of DNA sequence polymorphisms among wheat varieties. Theor Appl Genet 84:573–578
Heun M, Helentjaris T (1993) Inheritance of RAPDs in F_1 hybrids of corn. Theor Appl Genet 85:961–968
Heun M, Murphy JP, Phillip TD (1994) A comparison of RAPD and isozyme analysis for determining the genetic relationships among *Avena sterilis* L. accessions. Theor Appl Genet 87:689–696
Hu J, Quiros CF (1991) Identification of brocolli and cauliflower cultivars with RAPD markers. Plant Cell Rep 10:505–511
Jacobsen E, De Jong JH, Kamstra SA, Van Den Berg PM, Ramanna MS (1995) Geomic in situ hybridization (GISH) and RFLP analysis for the identification of alien chromosomes in the backcross progeny of potato (−) tomato fusion hybrids. Heredity 74:250–257
Joshi CP, Nguyen HT (1993) Application of the random amplified polymorphic DNA technique for the detection of polymorphism among wild and cultivated tetraploid wheats. Genome 36:602–609
Kesseli RV, Paran I, Michelmore RW (1994) Analysis of a detailed genetic linkage map of *Lectuca sativa* (lettuce) constructed from RFLP and RAPD markers. Genetics 136:1435–1446
Klein-Lankhorst RM, Vermunt A, Weide R, Liharska T, Zabel P (1991) Isolation of molecular markers for tomato (*Lycopersicon esculentum*) using random amplified polymorphic DNA (RAPD). Theor Appl Genet 83:108–114
Koto E (1989) Tentative de transfert du caractère "retard à la morphogénèse des glandes à gossypol" I. Caractéristiques des hexaploides *G. hirsutum* × *G. sturtianum* et *G. hirsutum* × *G. australe*. 1ère Conférence de la recherche cotonnière africaine, Lomé, Togo, 31 janv.–2 fév., 1989, IRCT, Montpellier, Tome I:167–173
McCoy TJ, Echt CS (1993) Potential of trispecies bridge crosses and random amplified polymorphic DNA markers for introgression of *Medicago daghestanica* and *M. pironae* germplasm into alfalfa (*M. sativa*). Genome 36:594–601
Marshall P, Marchand MC, Lisieczko Z, Landry BS (1994) A simple method to estimate the percentage of hybridity in canola (*Brassica napus*) F_1 hybrids. Theor Appl Genet 89:853–858
Masuelli RW, Tanimoto EY, Brown CR, Comai L (1995) Irregular meiosis in a somatic hybrid between *Solanum bulbocastanum* and *S. tuberosum* detected by species-specific PCR markers and cytological analysis. Theor Appl Genet 91:401–408
Mergeai G (1992) New perspectives concerning the methodology to be used for introgression of the glanded-plant and glandless-seed character in cultivated cotton (*Gossypium hirsutum* L.). Coton Fibres Trop 47:113–119
Mergeai G, Vroh Bi I, du Jardin P, Baudoin JP (1995) Introgression of glanded-plant and glandless-seed trait from *G. sturtianum* Willis into tetraploid cotton plants. Proc Beltwide Cotton Improvement Conf, San Antonio, 6–7 Jan 1995, pp 513–514
Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating population. Proc Natl Acad Sci USA 88:9828–9832
Multani DS, Lyon BR (1995) Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome 38:1005–1008
Murray M, Thomson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4325
Olson M, Hood L, Cantor C, Botstein D (1989) A common language for physical mapping of the human genome. Science 245:1434–1435
Orozco-Castillo C, Chalmers KJ, Wangh R, Powell W (1994) Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. Theor Appl Genet 87:934–940

Paran I, Michelmore W (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85:985–993

Quiros CF, Hu J, This P, Chevre AM, Delseny M (1992) Developmental and chromosomal localization of genome-specific markers by polymerase chain reaction in *Brassica*. Theor Appl Genet 82:627–632

Reinisch AJ, Dong Jian-min, Brubaker CL, Stelly DM, Wendel JF, Paterson AH (1994) A detailed RFLP map of cotton, *Gossypium hirsutum* × *Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. Genetics 138:829–847

Reiter RS, Williams JGK, Feldmann KA, Rafalski JA, Tingey SV, Sclonik PA (1992) Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Proc Natl Acad Sci USA 81:8014–8018

Rooney WL, Stelly DM, Altman DW (1991) Identification of four *Gossypium sturtianum* monosomic alien addition derivatives form a backcrossing program with *G. hirsutum*. Crop Sci 31:337–341

Rowland LJ, Levi A (1994) RAPD-based genetic linkage map of blueberry derived from a cross between diploid species (*Vaccinium darrowi* and *V. elliottii*). Theor Appl Genet 87:863–868

Stewart J McD, Hsu CL (1977) In-ovulo embryo culture and seedling development of cotton (*Gossypium hirsutum* L.). Planta 137:113–117

Tanksley SD, Young ND, Pat AH, Bonierbale MW (1989) RFLP mapping in plant breeding – new tools for an old science. Bio/Technology 7:257–264

Tatineni V, Cantrell RG, Davis DD (1996) Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Sci 36(1):186–192

Tinker NA, Fortin MG, Mather GE (1993) Random amplified polymorphic DNA and pedigree relationships in spring barley. Theor Appl Genet 85:976–984

Tulsieram LK, Glaubitz JC, Kiss G, Carlson JE (1992) Single tree linkage mapping in conifers using haploid DNA from megagametophytes. Biotechnolgy 10:686–690

Van Hausden AW, Bachmann K (1992) Genotype relationships in *Microseries elegans* (Asteraceae, Latcuceae) revealed by DNA amplification from arbitrary primers (RAPDs). Plant Syst Evol 179:221–233

Veli-Matti Rokka, Yong-Shen Xu, Jyri Kankila, Anja Kuusela, Seppo Pulli, Eija Pehu (1992) Identification of somatic hybrids of dihaploid *Solanum tuberosum* lines and *S. brevidens* by species-specific RAPD patterns and assessment of disease resistance of the hybrids. Euphytica 80:207–217

Vierling RA, Nguyen HT (1992) Use of RAPD markers to determine the genetic diversity of diploid wheat genotypes. Theor Appl Genet 84:835–838

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

Vroh Bi I, du Jardin P, Mergeai G, Baudoin J-P (1997) Optimisation et application de la RAPD (random amplified polymorphic DNA) dans un programme de sélection récurrente chez le cotonnier (*Gossypium* spp.). Biotechnol Agron Soc Environ 1:142–150

Walbot V, Duke LS (1976) Developmental biochemistry of cotton seed embryogenesis and germination. 7. Characterisation of the cotton genome. J Mol Biol 101:503–536

Welsh J, Honeycutt RJ, McClelland M, Sobral BWS (1991) Parentage determination in maize hybrids using arbitrarly primed polymerase chain reaction (AP-PCR). Theor Appl Genet 82:473–476

Wilde J, Waugh R, Powell W (1992) Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. Theor Appl Genet 83:871–877

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingery SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 22:6531–6535