Bridging Classical and Molecular Cytogenetics of *Gossypium*

N'Guessan Olivier Konan, Jean-Pierre Baudoin, Angélique D'Hont, and Guy Mergeai

Abstract Cotton (*Gossypium* spp.) is the leading natural fiber crop in the world. The genus *Gossypium* comprises a broad genetic base that has been and continues to be actively studied by cotton genetists and breeders. Cytogenetic tools play an important role in cotton genome research and cotton breeding. Based essentially on the observation of chromosome morphologies and the analysis of chromosome pairing, classical cytogenetics has contributed greatly to understanding cotton history, taxonomy and phylogeny, and has been a great help in cotton breeding programs designed to transfer desired genes from alien species into cultivated varieties. With the advent of molecular cytogenetics in the 1980s, the field of cytogenetics has been revolutionized. Beside an increase in the speed, sensitivity and specificity of conventional cytogenetic techniques, molecular cytogenetics offers opportunities to perform a variety of tasks not achievable by classical methods. These tasks include analysis of the distribution of repeated sequences along the genome, assignment of repetitive and single copy DNA sequences to positions on chromosomes, determination of the relationship between specific chromosomes and linkage groups, determination of the relationships between physical and genetic distances, differentiation of the genomes involved in hybrids, detection of alien DNA in introgressed lines, and others. We summarize the achievements of classical and molecular cytogenetic investigations in *Gossypium* and underline the relevance of bridging these approaches in *Gossypium* genetic studies and exploitation.

1 Introduction

Cytogenetics is a branch of genetics concerning every aspect of chromosomes that can be observed at the microscopic level. The first cytogenetic investigation in *Gossypium* started during the second half of the 19th century but it is since the

N.O. Konan (✉)
Gembloux Agricultural University (FUSAGx), Unité de Phytotechnie tropicale et d'Horticulture, 2 passage des Déportés, B-5030 Gembloux, Belgium
e-mail: konan.n@fsagx.ac.be; nguessanolivier@yahoo.fr

A.H. Paterson (ed.), *Genetics and Genomics of Cotton*,
Plant Genetics and Genomics: Crops and Models 3,
DOI 10.1007/978-0-387-70810-2_11,

1920s (Denham 1924) that this science has become a powerful tool. Cytogenetic studies of cotton were first limited to examining chromosomes under an optical microscope, with staining techniques that enhance identification. Classical cytogenetics essentially allowed karyotyping, chromosome counting, detection of certain chromosomal abnormalities and analysis of chromosome pairing for assessment of genomic affinities. This conventional cytogenetics showed the existence of the diploid and polyploid species in *Gossypium* and greatly participated in the determination of the different cotton genome groups. Although this technique became very useful for *Gossypium* genome studies and cotton breeding, it was inadequate for detection of specific chromosomes or cryptic rearrangements. The remarkable development of molecular biology in the 1980s revolutionized the field of cytogenetics, with the advent of new methods which overcome many limitations of classical ones. This modern approach allows the precise physical localization of genes or DNA sequences on cytological preparations, enabling enormous progress in genome studies and breeding. Chromosome painting, study of chromosome exchanges and gene rearrangements, analysis of genome structure, development of cytomolecular maps, and monitoring of the transfer of agronomic traits are some of the interesting applications enabled by molecular cytogenetics. In spite of its success, modern techniques have not superseded classical cytogenetics; but have become an important complement that has bridged the gap between conventional cytogenetics and molecular genetic studies.

In this paper we examine the achievements obtained using classical and molecular cytogenetics to analyze and exploit *Gossypium*, emphasizing the interest of bridging these two complementary approaches.

2 Achievements of Classical Cytogenetics in Cotton

2.1 Principles of Classical Cytogenetics

Classical cytogenetics is a discipline of genetic that studies mitotic (from root tips or young leaves) or meiotic (from flower buds) chromosomes under an optical microscope using staining or banding techniques for their identification. Generally, metaphase chromosomes are studied because they are the most condensed and the most visible. The stains usually used to visualize chromosomes are ferric haematoxylin, crystal violet, acetic orceine, schiff regent, and carmine acetic acid; they allow uniform staining of chromosomes. Classical staining techniques are usually used for chromosome counting, studies of chromosome morphology, detection of chromosomal modification and analysis of chromosome pairing. In banding, chromosomes are stained with Giemsa after various treatments (enzymatic, denaturing agent, heat...) that allow staining of chromosomes with a succession of light and dark bands specific to each chromosomal pairs. The most important banding techniques are

G (Giemsa), R (reverse), BrdU (5-Bromodéoxyuridine), C (centromeric) and T (telomeric) banding. Banding techniques are useful for identifying whole chromosomes accurately (chromosomes present in interspecific hybrids, addition or substitution lines, aneuploid stocks) and for detecting chromosomal rearrangements; but this technique is laborious to apply and was much more rarely used in cotton classical cytogenetic investigations than the staining techniques.

2.2 *Contributions of Classical Cytogenetics to Taxonomic and Phylogenetic Studies of* Gossypium

2.2.1 Classical Cytogenetics in the Identification of *Gossypium* Genomes and the Understanding of *Gossypium* Phylogeny

Classical cytogenetic studies greatly contributed to current knowledge of *Gossypium* phylogeny. In the late 1920s and early 1930s, classical cytogenetic analyses of species and genome types of *Gossypium* were performed by a number of investigators (Endrizzi, Turcotte and Kohel 1985). It was Nikolajeva (1923) and Denham (1924), who the first demonstrated that there were exactly two chromosome numbers (26 and 52 chromosomes) in the genus *Gossypium*. Based on this observation, Denham (1924) divided the species of *Gossypium* into diploids (2n = 26) and tetraploids (2n = 52). The cytogenetic studies of Denham represent a decisive step in cotton systematics because they opened the door to a long series of works that contributed greatly to the classification of the genus *Gossypium*. Several subsequent workers reported chromosome numbers of 26 and 52 for many additional *Gossypium* species including wild and cultivated forms occurring in America, Asia, Africa, and Australia (Youngman and Pande 1927; Harland 1928; Banerji 1929; Baranov 1930; Kearney 1930; Longley 1933; Skovsted 1933, 1934a,b, 1935b; Webber 1934a,b, 1935, 1939; Wouters 1948). Currently the genus *Gossypium* includes 45 diploid and 5 tetraploid species (Wendel and Cronn 2003).

Meiotic studies of the diploid species and their hybrids allowed classification of cotton species into different genome groups. Beasley (1940, 1942), based on chromosomes pairing at meotic metaphase 1 in F1 interspecific hybrids, was the first to propose a genomic classification of the diploid species. Some species exhibited highly regular pairing of chromosomes in their hybrids, whereas other species showed highly irregular chromosome pairing. Beasley used the frequency of paired chromosomes to measure the affinities between species and to distinguish the different genomes. He defined five genome groups designed by the capital letters A, B, C, D, and E. In 1950, Brown and Menzel perfected the system of symbols proposed by Beasley. Similar genomes are designated by the same capital letter and closely related genomes are distinguished by a numerical subscript after each letter of that class. Eight diploid and one tetraploid genome group(s) are now recognized in *Gossypium* (Table 1). The A to E genomes were assigned by Beasley (1940), F and G genomes were assigned

Table 1 Localization and species composition of the *Gossypium* genome groups

Genomes	Localization	Species
A	Africa and Asia	*G. herbaceum* (A_1), *G. arboreum* (A_2)
B	Africa	*G. anomalum* (B_1), *G. triphyllum* (B_2), *G. barbosanum* (B_3), *G. capitis-viridis* (B_4)
C	Australia	*G. sturtianum* (C_1), *G. robinsonii* (C_2), *G. nandewarense*
D	America	*G. thurberi* (D_1), *G. armourianum* (D_{2-1}), *G. harknessii* (D_{2-2}), *G. davidsonii* (D_{3-d}), *G. klotzschianum* (D_{3-k}), *G. aridum* (D_4), *G. raimondii* (D_5), *G. gossypioïdes* (D_6), *G. lobatum* (D_7), *G. trilobum* (D_8), *G.laxum* (D_9), *G. turneri* (D_{10}), *G. schwendimanii* (D_{11})
E	Arabia and Africa	*G. stocksii* (E_1), *G. somalense* (E_2), G. *areysianum* (E_3) , *G. incanum* (E_4) , *G. benadirense*, *G. bricchettii*, *G. vollesenii*
F	Africa	*G. longicalyx* (F_1)
G	Australia	*G. bickii* (G_1), *G. australe*, *G. nelsonii*
K	NW Australia	*G. costulatum*, *G. cunninghamii*, *G. enthyle*, *G. exgiuum*, *G. nobile*, *G. pilosum*, *G. populifolium*, *G. pulchellum*, *G. rotundifolium*, *G. sp.novum*
AD	America	*G. hirsutum* $(AD)_1$, *G. barbadense* $(AD)_2$, G. *tomentosum* $(AD)_3$, *G. mustelinum* $(AD)_4$, *G. darwinii* $(AD)_5$

The lack of numerical subscript after the letter designating the genome of a species indicates that detailed cytogenetic studies have not been carried out yet for it and that its genomic classification is provisional.

respectively by Phillips and Strickland (1966) and Edwards and Mirza (1979). Table 1 shows the different genome groups in *Gossypium*, their location and the species they contain.

Moreover, classical cytogenetic studies revealed cotton chromosomes to vary widely in size (Stephens 1947, Katterman and Ergle 1970): i) the C genome has very large chromosomes; ii) E and F genomes have large chromosomes that are slightly larger than those of the A and B genomes; iii) the B genome has large chromosomes, some of which are slightly larger than those of the A genome; iv) the A genome has moderately large chromosomes; v) the G genome has moderately large chromosomes but smaller than those of the A genome; vi) the D genome has the smallest chromosomes. Cytogenetic studies showed that chromosome morphology is similar among closely related species, and related species form hybrids with normal meiotic pairing; whereas, hybrids from relatively distant species show meiotic abnormalities. It has been possible to assess the genome relationships and level of divergence based on the frequency of unpaired or univalent chromosomes in intra and intergenomic hybrids of the diploid species (Table 2). No cytogenetic studies were carried out so far on intergenomic diploid hybrids involving the F-genome species *G. longicalyx*. The relationships of this species with the other *Gossypium* genomes can however be inferred from the work of Philips and Strickland (1966) who compared the meitotic pairing in triploids and hexaploids resulting from crosses between the

Table 2 Average univalent frequency in intergenomic hybrids of diploid *Gossypium*

Intergenomic hybrid	Univalents per cell
A x B	2.82
A x C	8.50
A x D	13.98
A x E	17.13
A x G	16.00
B x C	11.17
B x D	18.19
B x E	22.35
C x D	13.10
C x E	24.68
D x E	25.15
D x F	21.60
G x C	3.84

Source : Endrizzi et al. (1985), N'Dungo et al. (1988b)

amphidiploids and diploid species belonging to all the different genomes (except K-genome). They put in evidence that the pairing affinities existing between the F-genome and the A_h subgenome chromosomes were higher than the ones observed between A_h and B.

Classical cytogenetics contributed to the understanding of the origin of the diploid species of cotton. Indeed, studies based on secondary associations that are visible during meiotic metaphase in diploid cotton (Davie 1933; Skovsted 1933; Abraham 1940; Brown and Menzel 1952), karyotype analysis by Edwards et al. (1979), and the application of BrdU-Hoechst–Giemsa chromosome banding techniques to diploid *Gossypium* (Muravenko, Fedotov, Punina, Fedorova, Grif and Zelenin 1998) suggested a paleopolyploid origin of the diploid *Gossypium*. The diploid species may have come from a single ancestral taxon, that was polyploid with the juxtaposition of 6 + 7 chromosomes.

2.2.2 Classical Cytogenetics in the Study of Tetraploid *Gossypium*

Classical cytogenetics helped to answer some questions raised by the discovery of tetraploid cottons. At first, classical cytogenetic studies demonstrated that the tetraploid cottons are true allotetraploids that contain two resident genomes, an A-genome from Africa or Asia, and a D-genome similar to those found in the American diploids. Beal (1928) had noted that the chromosomes of the allotetraploid species varied in size at metaphase I, and Davie (1933) had recorded a variation in somatic chromosome length of *G. hirsutum* ranging between 0.8 and 1.8 μm. Skovsted (1934a, b, 1935a) made the important observation that the tetraploid cottons had 13 large chromosomes with a

mean length of 2.26 to 2.36 μm and 13 small chromosomes with a mean length of 1.25–1.45 μm. Based on the meiotic behavior of a number of interspecific hybrids, and on the comparative size of the chromosomes of species in the genus, Skovsted concluded the tetraploid cottons were amphidiploids that originated by doubling of the nonhomologous chromosomes of two species with n = 13, one of which was similar to the large A chromosomes and the other of which was similar to the small D chromosomes (Endrizzi 1985). Webber (1939) and Beasley (1940, 1942) confirmed the hypothesis provided by Skovsted (1934b). Thus, by the early 1940s, it was firmly established, by classical cytogenetic studies, that the tetraploid cottons were allotetraploids that originated from combining A and D genomes.

The main contributions of classical cytogenetic studies to the understanding of *Gossypium* evolution are presented in Fig. 1.

2.2.3 Classical Cytogenetics in the Study of Translocations in *Gossypium*

Meiotic studies of cotton hybrids by classical cytogenetics revealed the presence of multivalent chromosome configurations at metaphase I. Beasley (1942) referred to the multivalents as translocations. Brown et al. (1950), Gerstel (1953), and Menzel and Brown (1954) observed also that the A genome of diploid species and the A_h subgenome of the allotetraploid differ from each other by chromosomal interchanges. The discovery of translocation in

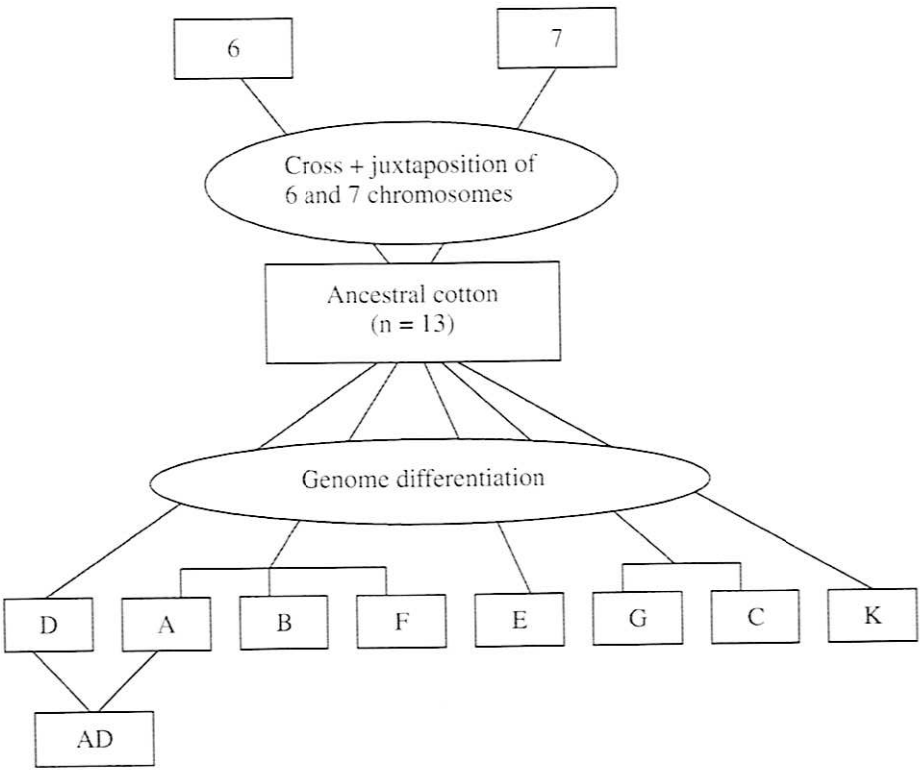

Fig. 1 Evolutionary history of *Gossypium* deduced from classical cytogenetic studies

Gossypium thanks to classical cytogenetic studies, allowed the identification of individual chromosomes. Consequently many translocations were induced in cotton by irradiation (Brown 1950; Menzel et al. 1954) and a set of 62 homologous translocation lines was developed (Brown 1980). Based on these translocation lines, 25 of the 26 chromosomes of *G. hirsutum* have been distinguished and numbered (Brown 1980). The A subgenome chromosomes were numbered A1-A13 (H1-H13) and the D subgenome chromosomes D1-D13 (H14-H26). Chromosome 26 has not been involved in a known translocation and it was identified through a process of elimination (Endrizzi et al. 1985). The translocation stocks constitute the only complete set of cytogenetic markers for the *G. hirsutum* genome (Wang, Song, Han, Guo, Yu, Sun, Pan, Kohel and Zhang 2006). The cytological procedure for assigning chromosomes in translocations to their respective genomes was described by Menzel (1955) and Endrizzi et al (1985). The identity of the chromosomes in each of the 62 translocations was determined in most cases by intercrossing the homozygous tranlocations and examing chromosomes pairing to determine whether the same or different chromosomes were involved (Brown 1980; Endrizzi et al. 1985). In some cases, stocks that were monosomic for known chromosomes were used to identify chromosomes in the translocations and vice versa (Endrizzi et al. 1985).

The translocation breakpoints with known locations can serve as reference markers in positioning genes and linkage groups on their chromosome (Menzel, Richmond and Dougherty 1985). Menzel and Brown (1978a,b) and Brown, Menzel, Hasenkampf and Naqi (1981) determined the arm location of breakpoints in the translocated chromosomes of many translocations. Menzel et al. (1985) assigned 115 translocation breakpoints to their chromosome arms in 58 translocations that involve exchanges between only two chromosomes. Using chiasma frequencies in specific chromosome regions, they estimated the recombination map length of each chromosome and the distance of each breakpoint from its centromere. Using these data they constructed the first cotton genome map with reference points on all the chromosomes, except chromosome 26. Thus, using classical cytogenetics, it has been possible to construct a recombination map of the cotton genome based on chiasma frequencies in chromosome regions defined by the breakpoints in 58 reciprocal chromosome translocations.

2.3 Classical Cytogenetics in Cotton Breeding

Once taxonomic and phyletic relationship among *Gossypium* species became better understood, classical cytogenetic studies were mostly used in breeding programs. All cotton genomes are important genetic resources that can contribute valuable genes for fibre quality, resistance to diseases and insect pests, tolerance to abiotic stress, and cytoplasmic genes conditioning male sterility along with nuclear restorer genes (Ndungo, Demol and Maréchal 1988c; Stewart 1995). In interspecific breeding programmes, classical cytogenetics is

an important discipline because it allows i) the understanding of genomic affinities in order to facilitate the planning of effective hybridization programs designed to transfer desired genes from alien species into cultivars; ii) the rational exploitation of intraspecific, interspecific, or induced variability; and iii) the explanation and resolution of some problems met by breeders such as instability or sterility.

In cotton breeding, classical cytogenetic studies focus mainly on meiotic analysis. Observations concern mainly the chromosomal configurations present at metaphase I. The goal is to establish a mean karyological formula from the analysis of a representative sample of pollen mother cells. The operation consists of the listing, for each mother cell, of the number of the different types of configurations observed (univalents, bivalents, trivalents, quadrivalents, etc.) and the total number of chismata. The average of each configuration is calculated. The deviation compared to the normal meiotic behavior (Fig. 2) can reveal the stability level of a hybrid, the type and the level of homology between the parental species (allowing the assessment of the possibilities of recombination) and certain structural rearrangements (deletions, duplications, translocations). A precise monitoring of alien chromosome segment introgression in interspectific hybrid progenies cannot however be obtained with classical cytogenetic methods.

A synthesis of the meiotic chromosome configurations observed in trispecific allotetraploids, bispecific triploids, and bispecific allohexaploids is presented in Tables 3, 4 and 5, respectively. The main lessons that can be drawn from the application of cytogenetic investigations in the framework of cotton interspecific breeding programs are summarized below.

The immediate expression of almost complete homoeologous pairing in the raw amphidiploid hybrids obtained by chromosome doubling of diploid or triploid bispecific hybrids and the lack of prevention of synapsis of homoeologous chromosomes during prophase in cotton haploids (Endrizzi et al. 1985)

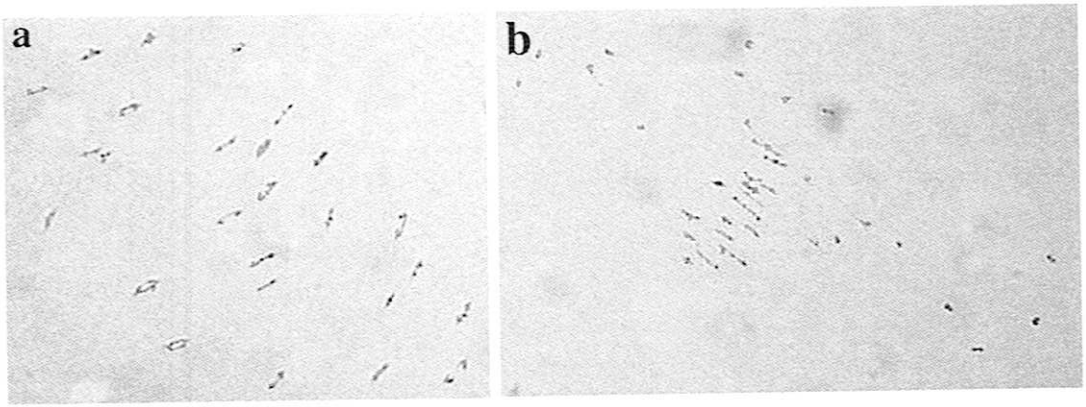

Fig. 2 Meiotic configuration at metaphase I showing regular pairing with 26 bivalent in *G. hirsutum* (a) and irregular pairing with univalent and multivalent in a trispecific hybrid [(*G. hirsutum* x *G. thurberi*)^2 x *G. longicalyx*] (b)

Table 3 Meiotic chromosome configurations in trispecific hybrids of *Gossypium* obtained with classical cytogenetic analysis

Hybrid combinations	Chr. No.	Chromosome configuration								References
		I	II	III	IV	V	VI	VIII	X	
[(*G. arboreum* x *G. thurberi*)^2 x *G. hirsutum*]	52	0.88	20.93	0.62	1.01	0.14	0.27	0.02	0.02	Brown and Menzel (1950)
[(*G. hirsutum* x *G. arboreum*)^2 x *G. harknessii*]	52	1.37	22.36	0.31	1.09	0.01	0.08	0.00	0.00	Brown et al. (1950)
[(*G. hirsutum* x *G. herbaceum*)^2 x *G. harknessi*]	52	1.84	22.14	1.14	0.54	0.00	0.05	0.00	0.00	Brown et al. (1950)
[(*G. thurberi* x *G. anomalum*)^2 x *G. hirsutum*]	52	11.45	17.13	1.22	0.50	0.04	0.03	0.00	0.00	Louant & Maréchal (1975)
[(*G. hirsutum* x *G. anomalum*)^2 x *G. harknesii*]	52	17.72	16.24	0.46	0.10	0.00	0.00	0.00	0.00	Louant et al. (1975)
[(*G. hirsutum* x *G. raimondii*)^2 x *G. sturtianum*]	52	13.64	17.04	0.85	0.35	0.00	0.07	0.00	0.00	Vroh bi, Hau, Baudoin and Mergeai (1999)
[(*G. thurberi* x *G. sturtianum*)^2 x *G. hirsutum*]	52	14.55	15.68	0.91	0.35	0.00	0.25	0.00	0.00	Vroh bi et al. (1999)
[(*G. arboreum* x *G. bickii*)^2 x *G. hirsutum*]	52	41.04	4.54	0.57	0.04	0.00	0.00	0.00	0.00	Shuijin & Biling (1993)
[(*G. hirsutum* x *G. thurberi*)^2 x *G. longicalyx*]	52	14.13	15.10	1.03	0.9	0.03	0.13	0.00	0.00	Konan, D'Hont, Baudoin and Mergeai (2007)

Table 4 Mean meiotic chromosome configurations in bispecific triploid hybrids of *Gossypium* obtained with classical cytogenetic analysis

Triploid	Chr. No.	Chromosome configuration						References
		I	II	III	IV	V	VI	
AD x A	39	13.00	11.33	0.00	0.33	0.00	0.33	Gerstel (1953); Baranov (1930); Skovsted (1934a)
AD x B	39	25.05	6.34	0.33	0.06	0.00	0.00	Poisson (1970)
AD x C	39	26.32	5.41	0.55	0.04	0.00	0.00	Skovsted (1937); Maréchal (1974)
AD x D	39	14.36	11.72	0.37	0.03	0.00	0.00	Skovsted (1937); Iyengar (1944); Endrizzi (1957); Boza and Madoo (1941); Kammacher (1960) Menzel and Brown (1954)

Table 5 Mean meiotic chromosome configurations in bispecific hexaploid hybrids of *Gossypium* obtained with classical cytogenetic analysis

Hexaploid	Chr. No.	Chromosome configuration								References
		I	II	III	IV	V	VI	VII	VIII	
AD x A	78	2.13	34.63	1.38	0.63	0.00	0.00	0.00	0.00	Iyengar (1944)
AD x B	78	2.19	36.17	0.38	0.50	0.02	0.03	0.00	0.01	Poisson (1970); Iyengar (1944), Brown and Menzel (1952);
AD x C	78	1.68	36.54	0.32	0.57	0.00	0.00	0.00	0.00	Brown and Menzel (1952)
AD x D	78	1.47	30.07	0.82	3.38	0.02	0.03	0.00	0.00	Brown and Menzel (1952); Iyengar (1944);
AD x E	78	1.26	38.07	0.09	0.08	0.00	0.00	0.00	0.00	Brown and Menzel (1952); Maréchal (1972); Schwendiman, Koto and Hau (1980)
AD x F	78	1.47	35.28	0.28	1.22	0.03	0.02	0.00	0.00	Phillips and Strickland (1966); Schwendiman et al. (1980)

suggest that if a gene functioning like *Ph1* in wheat for regulating bivalent pairing (Riley and Chapman 1958; Sears and Okamoto 1958; Riley, Chapman and Kimber 1960) probably exist in the allotetraploid cotton species, in accordance with the postulate made by Kimber (1961), its impact on the prevention of intergenomic pairing is much lower in these species than in wheat. Consequently, the introgression of alien chromosome segments in upland cotton in the framework of interspectific breeding programs should be easier in *Gossypium* than in *Triticum*.

All trispecific hybrids are obtained through the development of an intermediary hybrid, which can be either allotetraploid or allohexaploid. In these crossing schemes, besides the choice of the bridge species, the breeder cannot do much to control the level and the nature of genetic material exchanges between the different genomes that are combined in the tri-species structure. Cytogenetic analyses carried out in programs involving tri-species hybrids reveal the importance of choosing a diploid species belonging to genome D as a bridge species when creating such materials. Indeed, because chromosomes prefer autosyndetic pairing at meiosis, recombination is low for chromosomes with low homology. Genome D chromosomes being much smaller than the chromosomes of the other diploid genomes, their pairing affinities with the latter are very low. It means that if a species other than a D-genome species is used as bridge to create the trispecific hybrid, most of the chromosomes of subgenome D_h from *G. hirsutum* will remain unassociated at metaphase I and it will be almost impossible to obtain fertile progeny by backcrossing the trispecific hybrid to *G. hirsutum*. This is illustrated by the Metaphase I chromosome configurations observed by Shuijing and Biling (1993) in the ABH [(*G. arboreum* x *G. bickii*)^{2}x *G. hirsutum*, $A_hA_2D_hG_1$] hybrid and what was observed in the TSH [(*G. thurberi* x *G. sturtianum*)^2 x *G. hirsutum*, $A_hC_1D_hD_5$] and HRS [(*G. hirsutum* x *G. raimondii*)^2 x *G. sturtianum*, $A_hC_1D_hD_5$] trispecific hybrids (Mergeai, Baudoin and Vroh Bi. 1997). A rather high pairing frequency was observed in TSH and HRS hybrids (Table 3) which gave rise to fertile progeny while the number of bivalents and multivalents was very low in the ABH hybrid (2n=4x=52=41.01 I + 4.54 II + 0.57 III + 0.41 IV) from which no viable seeds could be produced. The application of growth regulators (50 $mg.l^{-;1}$ naphtoxy-acetic acid + 100 $mg.l^{-;1}$ gibberellic acid) to avoid capsule shedding after pollination and the *in vitro* rescue of mature embryos (Vroh Bi, Baudoin, Hau, and Mergeai 1999) allowed the exploitation of tri-species hybrids in which the donor species did not belong to the diploid genomes that are genetically close to the A_h and D_h subgenomes of *G. hirsutum*. This is notably the case for *G. sturtianum* (Genome C), *G. longicalyx* (Genome F) and *G. areysianum* (Genome E), which were included in fertile trispecific hybrids. Due to unfavorable linkages that exist between agronomic and fiber quality traits in the trispecific hybrids involving *G. thurberi* as bridge species (Demol 1966; Miller and Rawlings 1967; Meredith and Bridge 1971), it is recommended to use *G. raimondii* for this purpose.

The tri-specific pathway using a D-Genome species as bridge is interesting because in such allotetraploid combinations the A_h chromosomes have no autosyndetic partners and theoretically should pair with the chromosomes of the donor species. However, the successful use of trispecific synthetic tetraploids requires generally a large effort to produce fertile progeny and to eliminate the undesirable genetic material contributed by the diploid donor and bridge species (Vroh Bi et al. 1999). Although the frequency of homologous recombination between the donor species chromosomes and the A_h- or D_h-Genome chromosomes may be lower in bi-specific than tri-specific derivatives, the bi-specific pathway theoretically offers the possibility of generating more progeny in the same amount of time and thus to capture more homologous recombination events. Moreover, in direct exploitation of bi-specific hybrids through backcrossing the hexaploids to *G. hirsutum*, recombinant chromosomes are far more likely to be incorporated into fertile plants. This last method has also the advantage of allowing some control of the intensity of genetic exchanges in interspecific hybrids at the hexaploid and monosomic addition stages (Louant, Maréchal and Baudoin 1977). In such fertile interspecific structures, it is possible to accumulate recombination events over generations. Genetic material transfer from the wild diploid species to the cultivated amphidiploid is enhanced through spontaneous production, during the successive hexaploid generations, of intergenomic exchanges, which improve the affinities between the genomes. For each interspecific combination, there should exist an ideal number of hexaploid generations that would allow the production of the optimal level of intergenomic exchanges adapted to the genetic nature of the diploid species relative to *G. hirsutum*. This trend increases the chance of capturing a trait in the subsequent backcrosses to *G. hirsutum* of the allohexaploids. The interchanges occurring during successive allohexaploid generations lead to the production of pentaploid types differing according to the extent of recombinations that occurred in their interspecific parent material. For the genomes that are closely related to the amphidiploid subgenomes (A, D), the number of allohexaploid generations is limited to one or two because the very high level of genetic recombinations that occur every generation leads quickly to a complete sterility of the hybrid. The optimal number of hexaploid generations for F, B, C and G-genome species still needs to be determined. It should be inversely proportional to the multivalent frequencies observed at metaphase I in each hybrid combination (Table 5). The multiplication of hexaploid generations in bispecific hybrids involving E-genome species is not useful because in such hybrids the very low intergenomic pairing frequencies occurring at the hexaploid stage remain unchanged over generations. Monosomic addition plants are other fertile structures on which the breeder can exert some control regarding the introgression into the tetraploid cotton genome of characters from the supernumerary chromosome brought by the diploid donor species. Beside monosomic and disomic addition plants, the selfing of a monosomic addition line gives rise to euploid materials which can be introgressed by chromosomal fragments of the donor species. The extent of this introgression

depends on pairing affinities existing between the alien supernumerary chromosome and the genome of the recipient species. Once an agronomic trait of interest is identified in a monosomic addition line, plants carrying the donor diploid species supernumerary chromosome of this line can be used to conduct chromosome specific introgression by selfing them until incorporation of the desired trait into *G. hirsutum* (Hau 1981).

In order to isolate a large number of monosomic addition plants when following the bispecific introgression pathway, it is recommended, provided its pollen fertility is sufficiently high, to use the pentaploid as male parent in the backcross to *G. hirsutum* (Ahoton, Lacape, Baudoin and Mergeai 2003). In this type of cross, the progeny obtained contain a rather high proportion of monosomic addition plants (about 10 %) plus a large majority of euploid materials and almost no other aneuploid genotypes. When one uses the pentaploid as female parent in the backcross to *G. hirsutum*, most of the progeny are autosterile plants carrying several alien chromosomes. This trend is observed in the progeny of various crosses carried out between *G. hirsutum* and pentaploid hybrids involving diploid species of C, G, E and F-genome. It indirectly confirms the better tolerance of female gametes for multiple alien chromosome addition in their nucleus, and the better competitiveness of male gametes carrying only one additional alien chromosome compared to those carrying several alien chromosomes. In the progeny of the pentaploid and the monosomic addition lines, each alien chromosome addition is characterized by a particular transmission rate, which is chromosome specific. These variations can be explained by differences between the alien addition chromosome and its homology with the *G. hirsutum* genome. It can also be explained by various factors acting on the viability of aneuploid male and female gametes, on aneuploid zygote development, on aneuploid seed germination, and on the survival of plants carrying an alien supernumerary chromosome.

3 Achievements of Molecular Cytogenetics in Cotton

3.1 Principles of Molecular Cytogenetics

The coupling of molecular technologies with cytogenetics gave rise to molecular cytogenetics, largely replacing chemical stains with molecular probes and *in situ* hybridization (ISH) techniques, notably FISH (fluorescent *in situ* hybridization). FISH is based on the property of a sequence of DNA (probe) to hybridize to a complementary DNA (target). The technique involves labelling molecular probes with fluorochromes and hybridization of these fluorescently labelled probes to unique DNA sequences *in situ*. Probe detection is accomplished by ultraviolet-light excitement of fluorochromes which are directly attached to probe DNA, producing fluorescent signals that are inspected using a filter-equipped epifluorescence microscope and computer software. The different

types of probes commonly used are gene-specific probes, repetitive sequence probes, DNA clone probes (bacterial artificial chromosomes, BAC-FISH) and total genomic DNA probes (genomic *in situ* hybridization, i.e. GISH). FISH probes can range in size from over 100 kb, to less than 1 kb.

Fluorescent *in situ* hybridization is a powerful tool for genome analysis and genetic manipulation. By this sensitive method, molecular cytogenetics allows the exploration of genetic material, rendering possible the physical visualization of genes, DNA sequences, or specific chromosomes under microscope on cytological preparations. Consequently, molecular cytogenetics is a means to perform a variety of tasks such as assigning repetitive and single copy DNA sequences to positions on chromosomes, assigning molecular marker linkage groups to specific chromosomes and chromosome arms, detecting alien chromosomes in hybrids, detecting alien DNA in introgressed lines, and detecting and numbering sites of transgene inserts.

3.2 Physical Mapping of Repeated Sequences in Gossypium

Repeated sequences could be used as chromosome markers, cytological landmarks having value in chromosome identification (Mukai, Friebe and Gill 1992; Rayburn and Gill 1986; Tsujimoto, Mukai, Akagawa, Nagaki, Fujigaki, Yamamoto and Sasakuma 1997). Repeated sequences can be divided into two types based on their organization and distribution pattern. Tandem arrays are localized as clusters along the chromosomes, and dispersed repeats are interspersed with unrelated repeats and low-copy DNA over much of the genome. Among the most thoroughly investigated repetitive DNA sequences in plant species are the ribosomal RNA genes (rDNA), including 18S-26S and 5S rDNA tandem repeats (Schmidt et al. 1994; Cabrera et al. 1995; Castilho and Heslop-Harrison 1995). The rRNA genes are organized in tandem arrays within the nucleolar organizer regions (NORs). The ribosomal RNA genes can be used as probes for physical mapping in higher plants because they are arranged in tandem arrays clustered at a few sites. Visualization of these rRNA genes by FISH can provide a number of chromosomal markers to elucidate chromosome evolution and species interrelationships, including evolution of polyploid species.

In cotton, Bergey, Stelly, Price, and McKnight (1989) detected enzymatically three major 18S-26S rDNA sites (NORs) in *G. hirsutum* by meiotic ISH. Price, Stelly, McKnight, Scheuring, Raska, Michaelson and Bergey (1990) used a biotin-labeled cloned fragment of 18S-28S ribosomal DNA from soybean to hybridize DNA in meiotic chromosomes of *G. hirsutum* by *in situ* hybridization. Analysis of *in situ* hybridization to metaphase I meiocytes from two translocation heterozygotes and monosomics involving chromosome 9 indicated that a cluster of ribosomal RNA cistrons is on chromosome arm 9L, arguably the first

molecular marker mapping in cotton. Three years later, Crane, Price, Stelly and Czeschin (1993) discovered other rDNA sites. In 1996, Hanson, Islam-Faridi, Percival, Crane, Ji, McKnight, Stelly and Price used high-sensitivity mitotic FISH on somatic chromosomes to reveal at least six additional 18S-26S rDNA loci, but none of these latter loci have been mapped. Ji, De Donato, Cranel, Raska, Islam-Faridi, McKnight, Price and Stelly (1999) detected four more minor sites by meiotic FISH. It was Ji et al. (1999) who mapped and integrated all the known rDNA loci into the translocation breakpoint map of Menzel et al. (1985) by using meiotic FISH to quadrivalents of translocation heterozygotes. The detected 18S-26S rDNA loci were mapped to the right arms of chromosomes 8, 9, 15, 17, 19, 20, and 23 and the left arms of chromosomes 5, 11, 12, and 14. Using the rDNA loci as common reference points, Ji et al (1999) detected several erroneous arm assignments in the previously published map of heterozygous translocation breakpoints. Therefore, the use of meiotic FISH constitutes an important method for accurate physical mapping. The multiplicity of rDNA sites and consistent detection of many of them makes them useful in molecular cytogenetic studies of *G. hirsutum*. First, the three major NOR sites, four intermediate 18S-26S rDNA loci, and two 5S rDNA loci provide multiple, consistently detectable markers for karyological studies. Crane et al. (1993) indicated that the three major sites and the intermediate site on chromosome 7 could serve to define arm locations and thus facilitate translocation-based mapping of new loci in 18 of the 26 cotton chromosomes. Analogously, the other localized intermediate rDNA sites on chromosomes 5, 17, and 19 allow the definition of arm locations of new loci in all other chromosomes, except chromosomes 22 and 26. Chromosome 22 is associated with only one translocation, in which a minor 18S-26S site marks arm 20R. Chromosome 26 is not involved in any of the known translocations, so its arms are cytogenetically distinguishable only by monotelodisomy. The distribution of rDNA genes on chromosomal arms suggests that rRNA genes in short arms may be more likely to organize nucleoli. The map of rDNA locations will facilitate site-specific analysis of rRNA gene function and rDNA evolution, using methods for identification of rDNA-containing bacterial artificial chromosomes (Woo 1996).

In cotton, chromosome translocations are the primary sources for chromosome identification. In 1997, Ji, Raska, McKnight, Islam-Faridi, Crane, Zwick, Hanson, Price and Stelly used meiotic FISH to analyze a new monosome of *G. hirsutum*. Painting with A2-genome DNA revealed the monosome's D-subgenome origin. DAPI-PI staining showed that the monosome carries a major NOR, delimiting it to the major NOR-bearing chromosomes of the D-subgenome, i.e., 16 or 23. Dual-color FISH with 5S and 18S-28S rDNAs indicated that the monosome contains separate major clusters of each of these two tandemly repeated rDNA elements, thus delimiting the monosome to chromosome 23, for which the Cotton Cytogenetic Collection previously lacked any sort of deficiency.

3.3 *Identification of Structural Abnormalities in* Gossypium

Chromosome structural abnormalities, such as duplication-deficiencies (dp-dfs), can be used, as with other hemizygous aneuploids, to assign genes or molecular markers to their respective chromosomes. Well established translocation stocks exist in cotton (*G. hirsutum*). The cotton cytogenetic collection includes a set of 58 simple reciprocal translocations in upland cotton (*G. hirsutum*), which constitutes the primary resource for chromosome manipulation and identification. Breakpoints of those translocations collectively affect 25 of the 26 chromosomes, and most have been mapped relative to one another and to their respective centromeres and telomeres, based on interstitial and distal chiasma frequencies in translocation heterozygotes (Menzel and Richmond 1985; Menzel and Dougherty 1987; Menzel, Richmond and Dougherty 1987; Menzel et al. 1985). However, conventional means of dp-df identification are often difficult to apply. Ji, Raska, De Donato, Islam-Faridi, Price, and Stelly (1999) demonstrated that by means of meiotic fluorescent *in situ* hybridization, the identification of the four types of dp-dfs is greatly facilitated, proving that fluorescence *in situ* hybridization (FISH) overcomes the limitations of classical cytogenetics and can be used to identify the structural abnormalities. That is important since dp-dfs can be used to assign genes or molecular markers to their respective chromosomes, and thus can facilitate the integration of physical and recombination maps.

3.4 *Physical Mapping of Agronomically Important Genes*

In crops, the physical mapping of genes, especially those of high agronomic value, is a valuable step in research programs dealing with the improvement of cultivars through hybridization and gene manipulation. Physical and molecular cytogenetic maps of cultivated plants have great practical and research value. They open up the prospect of producing chromosome-specific DNA libraries from parts of chromosomes containing valuable genes. Another key aim is to integrate the physical and (or) cytological map of a crop plant with its genetic linkage map. Genes and markers are then assigned to linkage groups, each of which corresponds to a single chromosome (Jiang and Gill 1994).

Agronomically important genes are mostly unique or low copy sequences. The information about the exact physical location of agronomically important genes is useful in breeding programs as well as in understanding the organization of genomes. In plants, *in situ* hybridization techniques have been used mainly for mapping repetitive DNA sequences and multicopy gene families. Mapping of low or single-copy sequences has proven difficult in plants compared to humans. On human chromosomes, single copy sequences as small as 1 kb can be routinely detected by the standard FISH technique (Richards, Vogt, Muleris, Malfoy and Dutrillaux 1994), but in plants it is difficult to locate 10 kb

sequences (Guzzo, Campagnari and Levi 2000) even when amplification of FISH signals is applied to enhance high resolution FISH efficiency. As an alternative, large genomic clones such as lambda phages, cosmids, and bacterial artificial chromosomes (BAC), can be successfully used (Mukai 2005): the large amount of repeated sequences contained in them will facilitate homologous hybridization and signal detection.

Genome physical mapping is a centerpiece of structural, functional and comparative genomics research. Access to a genome with large fragments of DNA is currently essential for genome analysis. Cloning of large DNA fragments is possible with bacterial artificial chromosomes (BAC). Artificial chromosomes are laboratory constructs that contain DNA sequences and are used to introduce and control new DNA in a cell, to study how chromosomes function, and to map genes in genomes. This technique has been developed by Shizuya et al. (1992). BACs have been very useful in the development of genome libraries. Specific clones isolated from BAC libraries can be used for FISH mapping to generate high-density cytomolecular maps (Lichter et al. 1990). The use of genomic DNA cloned in large-insert vector BACs as probes in FISH experiments is called BAC-FISH. Hybridization of cloned BACs on metaphasic chromosomes allows their assignation to a particular chromosome, which is useful for physical mapping and chromosome identification. In cotton this technique has been used to assign unassigned linkage groups to specific chromosomes. Indeed, despite significant progress to construct genetic maps in the tetraploid cotton *G. hirsutum* (Reinisch, Dong, Brubaker, Stelly, Wendel and Paterson 1994; Rong, Abbey, Bowers, Brubaker, Chang, Chee, Delmonte, Ding, Garza, Marler, Park, Pierce, Rainey, Rastogi, Schulze, Trolinder, Wendel, Wilkins, Williams-Coplin, Wing, Wright, Zhao, Zhu and Paterson 2004; Nguyen, Giband, Brottier, Risterucci and Lacape 2004; Han, Wang, Song, Guo, Gou, Li, Chen and Zhang 2006), there remained six linkage groups that were not associated with specific chromosomes, which was a hindrance for integrated genetic map construction. These linkage groups were assigned to the A subgenome (A01, A02, and A03) or the D subgenome (D02, D03, and D08) by analysis of marker loci in two progenitor diploid species (*G. herbaceum* and *G. raimondii*) of the allotetraploid cotton (Reinisch, Dong, Brubaker, Stelly, Wendel and Paterson 1994). BAC-FISH have been used by Wang et al. (2006) to resolve this problem. Specific BAC clones constructed in *G. hirsutum* acc. TM-1 for these six linkage groups were identified by screening the BAC library using linkage group-specific simple-sequence repeats markers. These BAC clones were hybridized to ten translocation heterozygotes of *G. hirsutum* as BAC-fluorescence *in situ* hybridization probes. The results obtained allowed Wang et al. (2006) to assign the six unassigned linkage groups A01, A02, A03, D02, D03, and D08 to chromosomes 13, 8, 11, 21, 24, and 19, respectively, establishing the 13 homeologous chromosome pairs.

FISH techniques offer new potential not only for more reliable chromosome identification, but also regarding the integration of genetic and physical maps, for ordering molecular markers and measuring physical genome distances, and for structural and functional chromosome analyses. Molecular cytogenetic

markers are very important for plant genome analysis and genetic manipulation. To develop molecular cytogenetic markers in cotton, Zhang, Dong, Decanini, Lee, Ren, Yan, Kohel, Yu, Zhang and Stelly (2002) selected 8 BAC clones for FISH to somatic cells. Six of these 8 selected BAC clones yielded unambiguous signals on the distal region of A or D subgenomes.

3.5 Total Genomic DNA as Probe in FISH : GISH

Genomic *in situ* hybridization (GISH), a modification of FISH, allows chromosomes from different parents in hybrid plants to be painted in different colours. The technique uses total genomic DNA from one species as the labelled probe in hybridization experiments to chromosomal DNA. The addition of an excess amount of unlabelled DNA (blocking DNA) from the parent not used as a probe substantially increases the specificity of probing and enables more closely related species to be distinguished. The blocking DNA hybridizes to sequences in common between the block and the species used as labelled probe, so that mainly genome specific sequences remain exposed as sites for probe hybridization. In cells viewed under a fluorescence microscope, target DNA hybridized to the probe fluoresces differentially to the non-hybridized unrelated DNA sequences. Simultaneous multiple target-sequence detection (multicolor Genomic *in situ* hybridization) is now possible. GISH is a powerful tool that permits characterization of genomes and chromosomes in allopolyploid species. It can thus be used to differentiate the chromosomes of the different

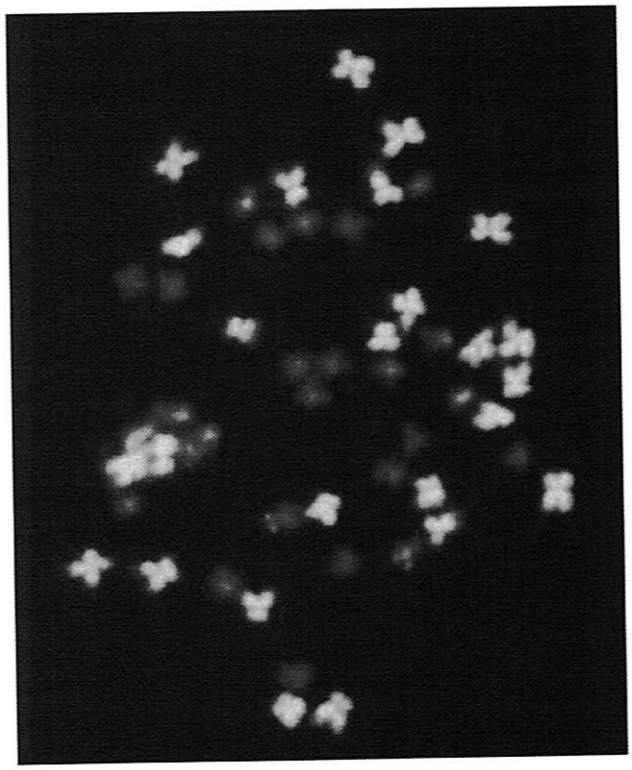

Fig. 3 Use of GISH to differentiate the A subgenome (in green) and the D sugenome (in blue) of *G. hirsutum*. Total DNA of *G. herbaceum* (A1 genome) labeled with digoxenin was used as probe; and total DNA of *G. thurberi* was used as blocking DNA (*See* Color Insert)