S2-1

NON-INDEPENDENCE BETWEEN MARKERS ON HOMOEOLOGOUS CHROMOSOMES IN AN INTERSPECIFIC ALLOPOLYPLOID COTTON RILS POPULATION

Viot C.(1), Jacobs J.(2), Arioli T.(3), Llewellyn D.(4), Derijcker R.(2), Claverie M.(1), Lacape JM.(1)

(1) UMR DAP, CIRAD, Montpellier, France, christopher.viot@cirad.fr
(2) Bayer BioScience N.V., Gent, Belgium, (3)Bayer Cropscience, Bioscience Research, Lubbock, USA, (4)CSIRO Plant Industry, Black Mountain, Canberra, Australia

Cotton, as the world's main natural textile fibre, is the focus of many studies for genetic improvement of fibre quality. Two allotetraploid (AtDt genome, 2n=4X=52) species dominate world production: *G. hirsutum* (*Gh*) with medium fibre quality and *G. barbadense* (*Gb*) with high fibre quality, accounting for 95% / 3% of production respectively. A RIL population originating from a *Gh-Gb* cross is the base material of the CIRAD-Bayer CropScience-CSIRO ANR research project Cotton_RILs, aiming at the genetic and genomic dissection of cotton fibre quality for introgression of high fibre quality genes of *Gb* into *Gh* germplasm. Until now, classical breeding did not succeed at satisfactorily combining traits from both cotton species, in spite of high apparent synteny conservation between their chromosomes.

The fixed heterozygosity of allopolyploids is supposed to result in gene expression changes, or expression subfonctionalization, a partitioning of the ancestral expression domains among duplicate genes. It is hypothesized that allopolyploids should be subject to more genetic and epigenetic regulatory changes than autopolyploids.

We report here on evidence of interactions between markers on pairs of homoeologous chromosomes in the *Gh-Gb* RIL population studied (140 lines genotyped with 800 markers, assembled in a saturated map of 26 linkage groups). In five of the 13 homoeologous chromosome pairs (c12-c26, c6-c25, c11-c21, c3-c17 and c13-c18), markers from homoeologs displayed unexpectedly high linkage (LOD from 9 to 14). Such linkage was not observed between chromosomes from non-homoeologous pairs. The association concerns 1-3 markers from one chromosome against stretches of 5-8 or more markers on the homoeolog. Gametic disequilibrium (GD) has been assessed pairwise between all markers. Positive GD is highly dominant between markers of homoeologous chromosomes, but negative as well as positive GD are observed between markers of non-homoeologs, with uneven distribution. In positive GD, frequencies of *Gh-Gb* allelic combinations are very low, meanwhile in negative GD, there is uniform low frequency of *Gb-Gb* allelic combinations.

Diverse genetic and epigenetic mechanisms have been characterized in polyploid plants, including unequal expression of duplicate genes, segregation distortion and restricted recombination; in cotton, lesser retention of *Gb* alleles is a common feature of advanced-generation backcross or RIL interspecific populations. Our results support hypotheses of intergenomic incompatibility, with selection during inbreeding that favoured elimination of *Gh-Gb* allelic combinations that were too conflicting regarding the expression of duplicate genes.

Keywords : Cotton, interspecific RILs, intergenomic incompatibility, gametic disequilibrium

Non-independence between markers on homoeologous chromosomes in an interspecific allopolyploid cotton RILs population

Christopher Viot[1], John Jacobs[2], Tony Arioli[3], Danny Llewellyn[4], Ruud Derijcker[2], Michel Claverie[1], Jean-Marc Lacape[1]
[1]CIRAD, UMR DAP, 34398 Montpellier, France
[2]Bayer BioScience N.V., Ghent, Belgium
[3]Bayer Cropscience, Bioscience Research, Lubbock, USA
[4]CSIRO Plant Industry, Black Mountain, Canberra, Australia
Corresponding author: christopher.viot@cirad.fr

This study reports on evidence of interactions between genes on several pairs of homoeologous chromosomes in a population of interspecific RILs from a cross between allotetraploid cotton species *Gossypium hirsutum* and *G. barbadense*. Very high LOD scores were needed to separate some groups of markers from two homoeologs into independent linkage groups and positive gametic disequilibrium at higher frequency than between non-homoeologs was observed.

Our observations can be explained by selection acting against certain *G. hirsutum-G. barbadense* allelic combinations in homoeologs during inbreeding generations following the interspecific cross. These results support hypotheses of intergenomic incompatibility, with elimination of *G. hirsutum-G. barbadense* allelic combinations that conflict in the expression of duplicate genes.

Cotton, as the world's main natural textile fibre, is the focus of many studies for genetic improvement of fibre quality. Two allotetraploid (AtDt genome, 2n=4X=52) species, *G. hirsutum* (*Gh*) with medium quality fibre, higher yield potential and broader environmental adaptability, and *G. barbadense* (*Gb*) with high quality fibre but lower yield potential and adaptability, dominate world cotton production with a 95% / 3% share respectively. Despite their common origin (Fig.1) and the high apparent synteny conservation between their chromosomes, classical breeding has not succeeded at satisfactorily combining traits from both cotton species, possibly due to genomic incompatibility resulting in particular from a partitioning of gene expression between the two subgenomes (subfunctionalization) that differs in the two species. A test for this hypothesis could be the existence of interactions between genes present on homoeologous chromosomes from interspecific *Gh-Gb* plants.

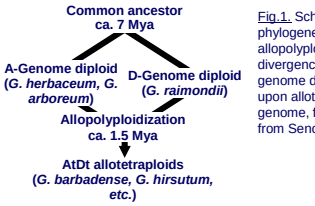

Fig.1. Schema of the phylogenetic origin of recent allopolyploid cottons : divergence of A- and D-genome diploids and reunion upon allotetraploid, AtDt genome, formation (adapted from Senchina *et al.* 2003).

Material and Methods

A RIL population originating from a *G. hirsutum* x *G. barbadense* cross in F6 to F8 inbreeding generation was genotyped (140 lines x 800 markers - 517 SSR, 281 AFLP, 2 morphological markers, assembled in a saturated map of 26 linkage groups) as part of the CIRAD-Bayer CropScience-CSIRO, ANR-funded, research project *Cotton_RILs* (Lacape *et al.* 2007, 2009).

LOD values were calculated with JoinMap 4 (Kyazma B.V., Wageningen, Netherlands). Gametic disequilibrium (GD) was measured with the standardized D' calculated as in Lewontin (1988) pairwise between all markers ; GD significance test used was the conditional 2-sided p-value (Kulinskaya 2009) calculated with LDtests R-package (Lewin 2008) ; calculation of GD parameters was made on a subset of 656 markers chosen so as to have no more than one marker in each 1 cM bin. Threshold for GD between marker pairs to be considered as informative was p<0.05 and |D'|>0.5.

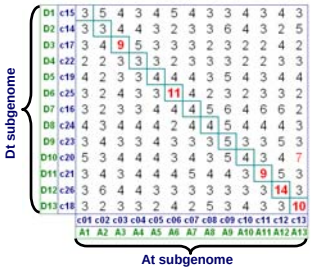

Table 1. Maximum LOD thresholds for independence of markers on pairs of chromosomes from the two sub-genomes At and Dt.

Average LOD are : 6.4±3.7 between homoeologs (data inside rectangles), and 3.4 ±1.1 for non-homoeologs.

Results and Discussion

Independence LODs: In five of the 13 homoeologous chromosome pairs (c12-c26, c6-c25, c11-c21, c3-c17 and c13-c18), some groups of markers from homoeologs necessitated the use of unexpectedly high LOD thresholds (from 9 to 14) to be separated in independent linkage groups (Tab. 1). Such linkage is not observed between chromosomes from non-homoeologous At-Dt chromosome pairs (p<0.01 for the difference), as well as from each sub-genome independently. Taking LOD>3 as a threshold, the association concerns 1-3 markers from one chromosome against stretches of 5-8 or more markers on the homoeolog ; such LOD values normally correspond to recombination frequencies lower than 0.10, indicative of synteny at close distances of 10 cM or less.

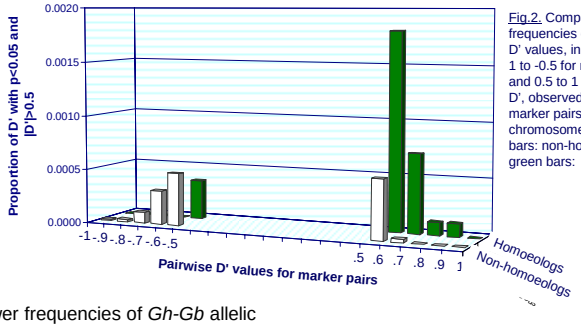

Fig.2. Comparison of frequencies of significant D' values, in the ranges -1 to -0.5 for negative D' and 0.5 to 1 for positive D', observed between marker pairs on At x Dt chromosome pairs; white bars: non-homoeologs; green bars: homoeologs.

Gametic disequilibrium (GD) : Positive GD is highly prevalent between markers on homoeologous chromosomes, while between markers of non-homoeologs, negative as well as positive GD were observed with a slight excess of negative GD (Fig. 2). Between chromosomes intra each At or Dt sub-genome, distribution of GD was very similar to that between non-homoeologs At-Dt (results not shown).

The threshold of |D'|>0.5 corresponds in the case of syntenic markers to distances ~< 15 cM between pairs of markers (results not shown).

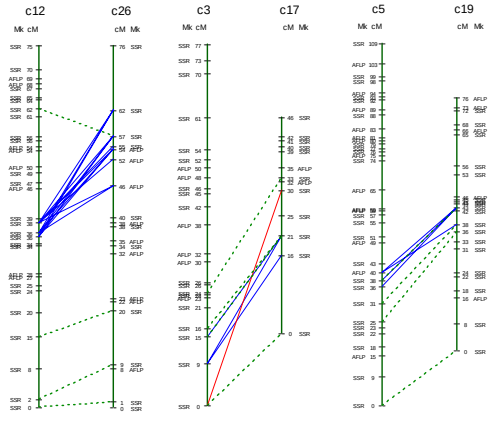

Fig.4. Gametic disequilibrium between markers on pairs of homoeologous chromosomes c3-c17, c5-c19, c12-c26. Lines join positions of markers in significant GD, blue for positive and red for negative GD. Dotted green lines join homoeolog locus pairs.

Positive GD is characterized by lower frequencies of *Gh-Gb* allelic combinations for At-Dt marker pairs, relative to global frequencies, meanwhile in negative GD, there is uniform very low frequency of *Gb-Gb* allelic combinations (Fig. 3). These altered frequencies can be considered as indicative of selection acting against recombinant (*Gh-Gb*) allelic associations for markers in positive GD, and against *Gb* parent (*Gb-Gb*) allelic associations and in favour of 'recombinant' (*Gh-Gb*) allelic associations for markers in negative GD.

The selection against *Gh-Gb* allelic associations in pairs of homoeologous chromosomes hints at incompatibility between some genes from the two species.

Some correspondence can be observed between regions with markers in positive GD and loci of homoeolog markers (Fig. 4), indicating that in some cases at least, positive GD could link homoeolog genes from the two subgenomes.

Such direct interactions between homoeolog genes can be related to the differential expression of duplicate genes, or subfunctionalization, instead of additivity, that has been characterized in cotton and other allotetraploids (Doyle *et al.* 2008). Thus, differences of adaptation of structural genes could interact with the gene network regulating the subfunctionalization. More generally, the results indicate that there is notable incompatibility between some homoeolog regions of the genomes of *G. hirsutum* and *G. barbadense*.

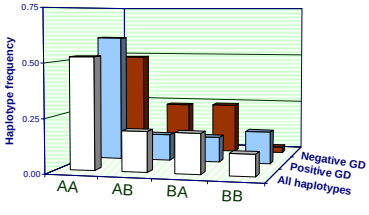

Fig.3. Distribution of haplotype frequencies for marker pairs on At-Dt chromosome pairs, compared between three categories: positive GD (D'>0.5, p<0.05), negative GD (D'<-0.5, p<0.05), all marker pairs. A=*Gh* allele, B=*Gb* allele.

This work was supported by ANR (project GPLA06027C), CIRAD and BAYER

References
Lacape, J. M., J. Jacobs, *et al.* (2007). Reconciliation of genetic and genomic approaches to cotton fiber quality improvement. WCRC4, Lubbock, USA, 10-14 Sep.2007.
Lacape, J.-M., J. Jacobs, *et al.* (2009). A new interspecific, *Gossypium hirsutum x G. barbadense*, RIL population: towards a unified consensus linkage map of tetraploid cotton. Theor. Appl. Gen.,*in press*.
Doyle, J. J., L. E. Flagel, *et al.* (2008). Evolutionary genetics of genome merger and doubling in plants. Annu. Rev. Genet. 42: 443–61.
Lewin, A. 2008. The LDtests package. http://www.r-project.org/
Kulinskaya, E. (2008, submitted). On two-sided p-values for non-symmetric distributions. http:/arxiv.org/abs/0810.2124.
Lewontin, R. C. (1988). On measures of gametic disequilibrium. Genetics 120: 849–852.
Senchina, D. S., I. Alvarez, *et al.* (2003). Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. Mol. Biol. Evol. 20(4): 633–643.

ICPHB 2009, May 17 – 20, 2009 – Saint Malo, FRANCE