Fourth International Cotton Genome Initiative (ICGI) meeting, Hyderabad (India), October 10-13 2004

Abstract of an oral communication proposed for the "Evolutionary and comparative genomics" session

Microsatellite allelic diversity within tetraploid *Gossypium* germplasm

Jean-Marc Lacape, Mazen Rajab, Dominique Dessauw, Bernard Hau

For cotton like for most important crops, future improvements of adaptation to adverse environment, of agronomic fitness and of quality of agricultural products, are expected to rely on the better utilization of genetic resources. Opportunities for such utilization are offered by the techniques of molecular biology through molecular breeding and genetic transformation. Within the major tetraploid cultivated species, *G. hirsutum*, two pools can be distinguished as sources of genetic variability for cotton breeders: the cultivated pool and the primitive, or exotic, pool. It is recognized that a severe bottleneck has accompanied the domestication process leading to modern commercially important cotton cultivars, resulting in a reduced genetic base. On another hand, the exotic tetraploid germplasm pool is constituted of landraces (7 geographical races belong to the *G. hirsutum* sp.), feral accessions and wild tetraploids, most being perennial and photoperiodic shrubs. This important gene reservoir, though constituted of morphologically and ecologically diverse germplasm, has received minor attention. Studies using isozymes and RFLP markers have early shown the limited diversity in the cultivated pool of *G. hirsutum* sp, thus explaining the choice made by many teams to initiate in the early 90s essentially interspecific marker-assisted breeding programs. The development of more efficient marker-systems, like AFLPs, and the availability of an increasing number of sources of microsatellites (SSRs) in the public (ca 1000) and private (min 2000) has now changed the vision. We used different sources of SSRs to assess the molecular polymorphism of a set of *G. hirsutum* cultivars and races, as well as of exotic tetraploid accessions. The 2 primary sources of SSRs belong to the "BNL" and "CIR" series, all chosen after placement of corresponding loci on our genetic map (Nguyen *et al.* TAG, 2004, 109:167-175). The plant material comprised a collection of 48 accessions, chosen in our gene bank and representing the variability of *G. hirsutum*-related landraces (7 geographical races and Moco cultigens) and of 3 other tetraploid species, *G. barbadense, G. tomentosum* and *G. darwinii*. This material was analysed for allelic diversity with 135 SSRs (236 SSRs analysed in a first run), which corresponded to 147 different loci. In total, 818 alleles were detected, varying between 2 and 17 alleles per locus, with an average of 5.6 alleles. The genetic distances calculated from the allelic data structured the genetic diversity in accordance with the known botanical, genealogical and/or geographical information, and confirmed previous molecular phylogenetic trees. The relative informativeness of SSRs assessed by the number of alleles or by the Nei statistics varied considerably. A "genotyping kit" was assembled based upon the most informative SSRs, that were re-associated in multiplexes by three (3-plexes are routinely used in our lab): 2 to 3 triplexes were assembled for each of the 13 pairs of homoeologous chromosomes of the tetraploid cotton genome. The applications of such a genome-wide genotyping kit of SSR markers in cotton are very important for phylogenetics, fingerprinting of cotton cultivars, or germplasm collection management.