DINUCLEOTIDE BLOCKS WITH EIGHT REPEATS ARE FEW IN THE HUGE GENOME OF THE GUARANA PLANT

Paula Cristina da Silva ANGELO[1*], Ana Yamaguishi CIAMPI[2], Gilvan Ferreira SILVA[1], Jorge Ivan Rebelo PORTO[3], Spartaco ASTOLFI-FILHO[4], André Luiz ATROCH[1]

RAPD had already been used for diversity analysis of guarana plants [*Paullinia cupana* (Kunth) var. *sorbilis* (Mart.) Ducke] included in Embrapa Western Amazon Guarana Germplasm Bank and another 27 clonal cultivars used for the breeding program. However, SSR or microsatellite markers can be more useful for gene flow analyses and the determination of self pollination rates in open pollination trials among high productive genotypes because these last markers are codominant. While genomic libraries were being enriched for microsatellites, the guarana plant that is part of the Sapindaceae family and native from the Amazon Forest, was confirmed as an allopolyploid with 210 chromosomes. Anyway, a search for microsatellites was performed in genomic libraries enriched using biotinilated probes $(CA)_{12}$, $(CT)_{12}$, $(CA)_{12}$+$(CT)_{12}$ e $(TC)_{14}$ and paramagnetic beads to capture Sau3AI and MseI fragments holding dinucleotide repeat blocks. In addition, a data bank of ESTs from guarana fruits with seeds was screened for repeat blocks using the TROLL software (Staden package). A huge number - more than twice the number in *A. thaliana* ESTs that has a genome 6,000 folds smaller - of very short blocks was found and there was no significant difference between genomic libraries and the EST data bank. This result is not in agreement with the hypothesis that microsatellites are rarer in big genomes because they would be preferentially localized in transcriptionally active regions, which are not so affected by the action and amplification of transposons and better represented in smaller genomes. But we considered the polyploidization of the ancestor genomes that followed an interspecific pollination 1,000 to 2,000 years ago to originate the cultivated *sorbilis* variety as the cause of the multiplication of microsatellites already present in those two genomes. As a consequence there is a myriad of short blocks of repeats (< 5) in the present day cultivars. Intriguingly, the frequency of blocks with eight or more dinucleotide repeats was very low, what is different from most plants that present nine repeats of dinucleotides per block in average. This last result could be explained in part by the inefficiency of the biotinilated probes to hybridize to blocks holding eight or more repeats in a genome dominated by very short blocks of four or five repeats. Some of the blocks identified were used as targets for primers to evaluate their utility to distinguish guarana genotypes.

Key words: *Paullinia cupana* var. *sorbilis*, Amazon Forest, Guaraná, SSR.

Financial support: FAPEAM (projeto 924/03) e Embrapa (projeto no. 02.02.4.03.00.00).

[1] Pesquisadores Doutores, Embrapa Amazônia Ocidental, Rod. AM 010, km 29, s/no, CP 319. CEP: 69010-970, Manaus - AM, Brasil.
[2] Pesquisadora Doutora, Embrapa Recursos Genéticos e Biotecnologia, Parque Estação Biológica - PqEB - Av. W5 Norte (final). CP 02372. CEP: 70770-917, Brasília - DF, Brasil.
[3] Pesquisador Doutor, INPA/NAPPA-Santarém, Rua 24 de outubro, 3389 - Salé. CEP: 68040-010. Santarém - PA, Brasil.
[4] Professor Doutor, Programa PPGBIOTEC, Centro de Apoio Multidisciplinar - CAM, UFAM, Av. General Rodrigo Octávio, 6200, Coroado I. CEP: 69077-000, Manaus - AM, Brasil.
*paula.angelo@embrapa.br - autora para correspondência