

# Estimation of linkage disequilibrium on a Brazilian collection of sugarcane (*Saccharum* spp.) genotypes from AFLP, SSR and EST-SSR data

Rosa, JRBF<sup>1</sup>; Silva, RR<sup>2</sup>; Pastina, MM<sup>3</sup>; Barreto, FZ<sup>4</sup>; Balsalobre, TWA<sup>5</sup>; Souza, AP<sup>5</sup>; Carneiro, MS<sup>4</sup>; Garcia, AAF<sup>1</sup>

<sup>1</sup>Escola Superior de Agricultura Luiz de Queiroz, USP, Piracicaba, SP; <sup>2</sup>University of Wisconsin, Madison, Estados Unidos;

<sup>3</sup>Embrapa Milho e Sorgo, Sete Lagoas, MG; <sup>4</sup>Universidade Federal de São Carlos, UFSCar, Araras, SP; <sup>5</sup>Universidade Estadual de Campinas, UNICAMP, Campinas, SP

jrbachega@usp.br

**Keywords:** *Saccharum* spp., polyploids, dominant markers, non-random associations, mapping

Linkage disequilibrium (LD) is one of the most important population parameters in genetic studies and has been extremely useful for association mapping in plants in the last years. However, despite its important and increasing application, few papers have been reported for LD analysis in sugarcane (*Saccharum* spp.), which is a highly polyploid species that represents a great importance for sugar and alcohol production in Brazil. In this sense, aiming the association mapping of important traits, we assessed the LD patterns in a subset of 80 sugarcane genotypes belonging to the Brazilian Panel of Sugarcane Varieties. We scored as presence or absence a total of 2,519 polymorphic markers, of which 1,474 and 1,121 were obtained from 100 SSRs (86 EST-SSRs and 14 genomic SSRs) and 21 AFLPs primers, respectively. Fisher exact test was used as a LD measure and performed for each possible pairwise combination involving all the polymorphic markers. In order to avoid the false positives detection, i.e. incorrectly detected significant associations, we used both Bonferroni and FDR (False Discovery Rate) as correction procedures. A total of 3,317,460 Fisher exact tests were carried out between the 2,519 polymorphic markers, of which 142,232 (4.48%) showed to be significant (or in LD) using a threshold of 0.05 (5%). However, since this threshold should not be appropriate due to the multiple tests, we initially corrected it using Bonferroni procedure, which accounted for 277 (0.09%) associations in LD. Nevertheless, because of Bonferroni is a very conservative procedure, FDR was also established to consider a less conservative threshold in the context of this work. Thus, we found 3,908 (0.12%) associations in LD using FDR procedure, which indicated a significant increasing in comparison to Bonferroni. These results suggest that there would be potential for sugarcane association mapping in Brazil, since many linked markers should be present between detected non-random associations. However, as many of the associations that were in LD could also involve unlinked markers, additional strategies should be established in order to control them before mapping. Furthermore, despite the fact that the LD detected here was based on many polymorphic markers from the sugarcane genome, it was not inferred considering the high and complex ploidy levels of this species. Therefore, we believe that recent advances involving SNP genotype calling in sugarcane, which provides allele dosage and different ploidy levels from its genome, will certainly be helpful to our understanding respect to genome-wide LD in order to refine the association mapping of desirable traits in Brazil. Financial Support: CNPq and FAPESP.