

GENOMIC-ASSISTED BREEDING OF SUGARCANE: USING MOLECULAR MARKERS FOR UNDERSTANDING THE GENETIC ARCHITECTURE OF QUANTITATIVE TRAITS AND IMPLEMENT MARKER ASSISTED SELECTION

Anete Pereira de Souza

Center of Molecular Biology and Genetic Engineering / University of Campinas (UNICAMP)

FAPESP Process 2008/52197-4 | Term: Aug 2009 to Jan 2014 | Thematic Project

Co-PIs: Roland Vencovsky, Antônio Augusto Franco Garcia

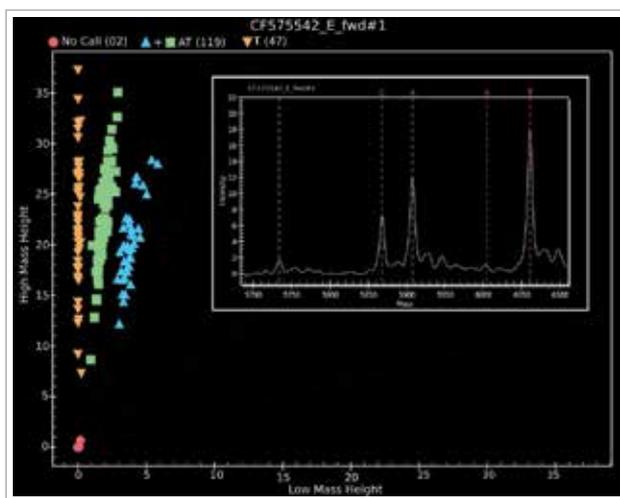


Figure 1. Example of results obtained with the SNP technology. Upper right: results for a single individual, showing the results for two loci; each peak corresponds to a nucleotide (A and C for the first, A and T for the second, that is homozygous). Left: genotyping of the segregating population, showing 3 classes

Breeding programs have been successfully over the years in the generation of new improved sugarcane varieties (*Saccharum spp.*), which are more productive and resistant to pests, diseases and abiotic stresses. These varieties are of central importance for sugar and ethanol production. However, breeding process takes about 10 to 15 years to release new varieties, mainly because of the difficulty to correctly identify good genotypes on the fields, since there is strong influence of environmental conditions. This process could be speed up with the development and the use of genetic markers, which are genomic regions that could be observed (evaluated) on each individual. By studying the segregation of those markers, it is possible to estimate the genetic distances between them, resulting in the so called genetic maps. After, linkage studies are performed in order to associate genotype (based on molecular traits) and phenotypes (traits that are evaluated on field conditions, such as sugar and fiber content). If the genomic regions are strongly linked with genes that control agronomic traits, they could be used for help the breeding process. Since most of the traits of agronomic and economic importance are

quantitative (controlled by many loci), the major goal is to identify genomic regions associated with such traits, named quantitative trait loci (QTL). The use of markers in genetic studies, including QTL mapping, has allowed important progress in the knowledge of the genomic structure, genetics and evolution of sugarcane. In this project, new markers will be developed and used for QTL mapping. A new class of very useful markers called genic molecular markers (GMMs) will be developed and used. This kind of marker (EST-SSRs and SNPs), also named Functional Markers, will be obtained from expressed sequences from the SUCEST data bank, and sequencing genes from BAC clones. BAC libraries will be constructed using DNA from parental sugarcane varieties employed in the biparental crosses used for genetic mapping. The GMMs developed will be used for: 1) QTL mapping in biparental crosses; 2) association studies using sugarcane genotypes important in breeding programs, using a panel of about 150 genotypes important for Brazilian breeding programs. Once the genomic regions are found, strategies for marker assisted selection will be developed.

SUMMARY OF RESULTS TO DATE AND PERSPECTIVES

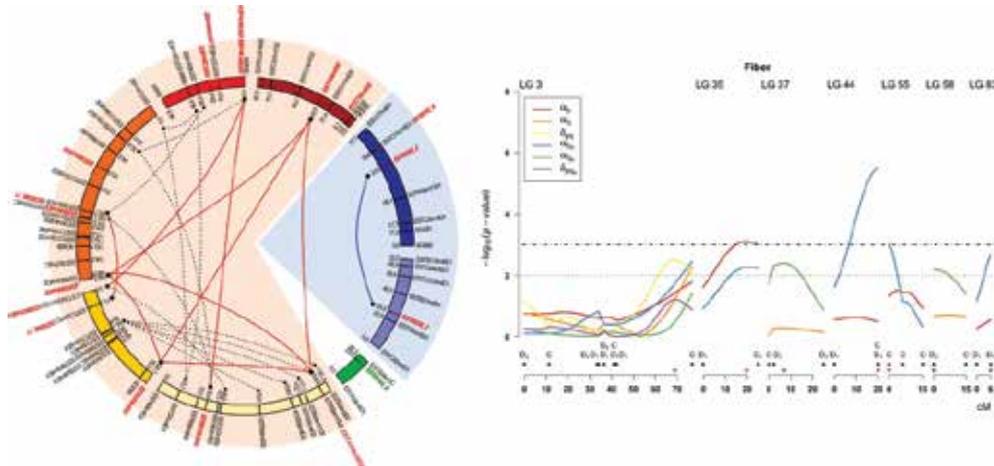


Figure 2. Example of mapping results. Left: Eight linkage groups obtained from markers with single (names in black), double (red) and triple dose (green). Lines are connecting the homology groups based on properties of the markers (EST-SSR and EST-RFLP) (dotted lines) or on loci with more than one copy on the genome (red and blue lines). Right: QTL mapping for fiber content. Each curve indicates the statistical evidence of QTL presence. Genetic effects for each parent, location and harvest were investigated

DNA from 220 individuals from the mapping population was extracted and quantified. The same was done with the genotypes from the association mapping panel. About 350 SNPs developed by international collaborators (Southern Cross University, Australia) were used to genotype the parents of the mapping population, as well as 14 random selected individuals from the progeny. This was done to check if the SNPs were polymorphic, i. e., if they were segregating on the population, providing information for the mapping studies. About one third of the SNPs are informative, a number which should be considered satisfactory, given that the SNPs were developed for Australian varieties. Also, Dr. Glaucia Souza provided information about 438 EST sequences that showed differential gene expression, and from them about 2200 SNPs were found, showing that this marker has a great potential for sugarcane studies, due to its abundance. At this time, more sequences have been evaluated with the goal of having at least 700 polymorphic SNPs.

So far, about 40 SNPs were used to genotype the whole biparental mapping population. Several challenges arose, due to the genome complexity of sugarcane, that is a polyploid species. First, the software that is provided with the SNP technology that has been used (Sequenom Inc.) was developed for diploids, and can not be directly used to interpret sugarcane data. Therefore, alternatives were investigated and a new computer program is under development. Second, the SNP technology allows the usage of markers with higher doses, i. e., with more than one copy on the polyploid genome. Statistical methods available to analyze

this type of data are not satisfactory, since they are based on unrealistic biological and statistical assumptions. Several new methods were then developed to estimate the dose of the markers, to allocate them in the genetic maps and to estimate homology groups. Third, currently none of the models used for QTL mapping can correctly deal with information from these new maps with a mixture of several doses. The common approach is based on single marker analysis of only simplex loci. Clearly this is not adequate for modern data and do not allow studies of the genetic architecture of quantitative traits in sugarcane. Several alternatives have been investigated, including multiple interval mapping and mixed models. The results are promising and will likely give a significant contribution in a near future.

Anete Pereira de Souza

Centro de Biologia Molecular e Engenharia Genética
Universidade Estadual de Campinas (Unicamp)
Cidade Universitária Zeferino Vaz – Barão Geraldo
Caixa Postal 6109
CEP 13083-970 – Campinas, SP – Brasil

+55-19-3521-1132
anete@unicamp.br