5TH INTERNATIONAL CONFERENCE OF THE BRAZILIAN ASSOCIATION FOR BIOINFORMATICS AND COMPUTATIONAL BIOLOGY

x-meeting
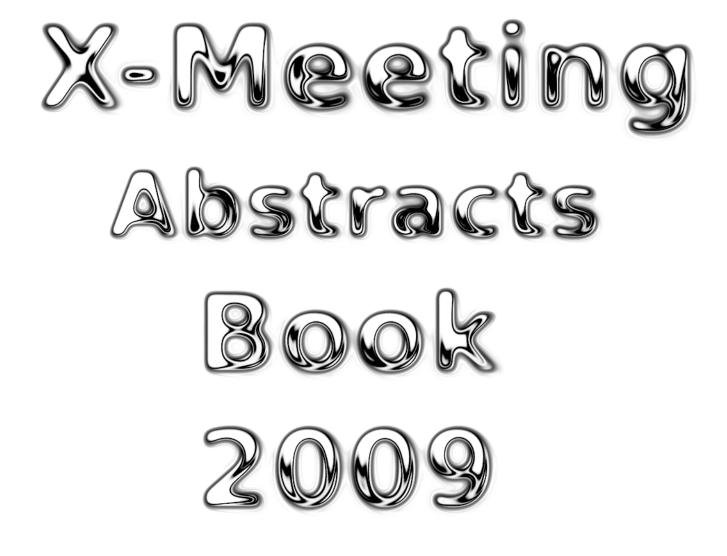
October 18 to 22, 2009

Hotel do Frade - Angra dos Reis - Rio de Janeiro - Brazil

X-Meeting Abstracts Book 2009

Angra dos Reis
Rio de Janeiro
Brazil

Topic: **Genomics, Evolution and Phylogeny**

BIOINFORMATICS ANALYSIS APPLIED TO GENOSOJA PROJECT

LC Nascimento[1], GGL Costa[1], RO Vidal[1], L Meyer[1], E Binneck[2], EA Kido[3], F Rodrigues[2], FR Kulcheski[4], AL Nepomuceno[2], RV Abdelnoor[2], GAG Pereira[1], MF Carazzolle[1]

[1]*Laboratório de Genômica e Expressão - Departamento de Genética, Evolução e Bioagentes, Instituto de Biologia, Universidade Estadual de Campinas, Campinas - SP, Brazil*
[2]*Embrapa Soja, Londrina, PR, Brazil*
[3]*Universidade Federal de Pernambuco, Recife, PE, Brazil*
[4]*Departamento de Plantas de Lavoura, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, RS, Brazil*

Soybean is the legume of most economic importance in the international market, with world production of almost two hundred and thirty million tons in the 2007/2008 harvest. The Brazil appears as the largest exporter of the product in the world, with about twenty-five percent of the world production. In 2007 the brazilian government started the GENOSOJA project with the main objective of discover new treats to improve the plant production process, emphasizing in stresses that affect the national production like the occurrence of droughts, pests attacks and the Asian rust disease. This work is inserted in the GENOSOJA scope and aims to generate bioinformatics tools to integrate the public soybean data like ESTS and genomic sequences, with all data generated during the project like SuperSAGE tags, MicroRNAs and subtractive libraries sequences analysis. These sequences were generated using high throughput sequencing technologies and for each analysis a specific pipeline have to be implemented. Actually, we have processed and performed a hybrid assembly of 1,276,813 ESTs available at NCBI (sequenced by sanger and pyrosequencing tecnologies) resulting in 30,809 contigs and 29,938 singlets, supersage analysis of 2,334,864 tags sequenced by Ilumina/Solexa from several libraries resulting in 54,052 unique tags, mapped into soybean genome and assembly of 11,783,331 short-sequences sequenced by Ilumina/Solexa generated by subtracted library experiments from leaf and root tissues submitted to drought stress resulting in 5,084 genes. All these results are making available to the final user in a web interface that allows searches by keywords, statistics comparisons, visualization of automatic annotation (AutoFACT pipeline) and the gene ontology classification. Supported by: CNPq