Identification of motifs in the promoter region of genes related to the ABA-dependent pathway in sugarcane

Mauro de Medeiros Oliveira[1], Alan Durham[1], Glaucia Souza Mendes[1],

1 University of Sao Paulo

Abstract

In general, the promoter region consists of different regulatory elements, such as Transcription Factor Binding Sites (TFBS), which are responsible for the activation of gene transcription. TFBSs can be characterized using different experimental processes such as Chip-Seq. However, these experimental present low reproducibility for non-model organisms. For these organisms the dominant form of TFBS discovery is computational estimations using techniques such as expectation maximization (EM). The goal of this work is to characterize the PR of ABA signaling pathway (ABAsp) genes using an in silico approach. To perform our analyzes, we used expression data of the sugarcane variety RB83-5486 to select all ABAsp genes differentially expressed in drought-tolerant plants. The 22 selected genes were mapped on the sugarcane genome SP80-3280 and the regions of 2000 nucleotides usptream from the transcription start site of each gene were extracted as putative promoter regions. For these regions we tried 3 different motif-finding approaches: Gibbs Sampling (using GLAM2), position-specific score matrices for previously characterized motifs in the JASPAR plant databases, and expectation maximization (using MEME). Only the last approach resulted in consistent results. GLAM2 showed a bias for AT-rich motifs and none of the results had any TOMTOM match against the JASPAR plant databases. Analysis using PSSMs did not find any candidates with significant scores. We parametrized MEME to find motifs from 5 to 15 nucleotides and maximum of 6 different motifs. We only considered motifs with a TOMTOM match to JASPAR plant databases. In general, 50% of the sequences presented similar TFBS architectures, with the bZIP, WRKY and AP2 / ERF classes of TFBSs as the most representative. Moreover, we distinguished different architectures for up and for down-regulated genes: in up-regulated genes we found motifs associated to ARR and bHLH TFBS classes, and in down-regulated genes were found motifs associated tot he HD-Zip, MYB and NAC TFBS classes. In this scenario it is possible to infer that the drought tolerance may be due to the crossing of different signaling pathways for water stress. Since two groups of TFBS distinct from the ABAsp were identified in the promoter region, one associated with up-regulated genes, and one associated with down-regulated genes, the architecture of the promoter region may be the factor necessary to activate the drought tolerant character observed in the evaluated plants.

Funding: FAPESP and Capes