Short sequence motifs enriched in 5 upstream and UTR regions of A.thaliana genes homologous to potato hyper- or hypo-expressed mrnas upon hydric stress

ID - 231

De Simone Vincenzo[1], Aiese Cigliano Riccardo[2], Parlati Aurora[2], Costa Antonello[2], Romano Maria[1], Martinelli Rosanna[1], Grillo Stefania[2]

[1]Department of Biochemistry and Medical Biotechnology, University of Napoli "Federico II", Napoli

[2]Department of Soil, Plant and Environment Science, University of Napoli "Federico II", Napoli

Motivation

Water supply is one of the most important limiting factors in crops production.

Understanding the molecular basis of plants adaptation to growth in water-restricted condition could be very useful to improve the management of food- and no-food crops.

In previous experiments based on cDNA microarray approach we have identified groups of hyper- and hypo-expressed genes in Solanum tuberosum cells grown water-deprived conditions. In order to understand whether groups of these genes are transcriptionally co-regulated via common cis-acting sequence elements, we decided to undertake the comparison of their transcriptional regulatory regions by searching the 5 upstream region for enriched short sequence motifs.

Since the potato genomic sequences of several of the selected genes are not yet available, we decided to perform our search by using the genomic sequences of A.
thaliana homologous gene counterparts, relying on the conservation across these species of transcriptional networks that regulate adaptation to hydric stress.

Methods

We have used the Fastx algorithm to selected the Arabidopsis thaliana homologous counterparts of Solanum tuberosums genes up- or down regulated by hydric stress.

Only the genes showing more than 60% of similarity and at least 50% of hit coverage with potatos genes in our list were selected for further analysis. We have then extracted the -400/-200, -200/+1 and +1/+200 (relative to the transcriptional start site) Arabidopsis thaliana genomic sequences corresponding to both upand
down-regulated potatos genes, obtaining six dataset.

Each dataset has been searched with the MEME algorithm, setting the program to display up to 10 motifs, and using three different width windows (5-10, 6-15 or 10-20 nucleotides).

Results

We have identified some specifically enriched consensus sequences for each dataset.

By comparison with the Jaspar database, we have identified among them the consensus sequences for the ABI4 plant transcription factors (in approx. 30% of the up-regulated genes dataset) and for the USF1 factor (in approx. 20% of the down-regulated genes dataset). The search for other known elements gave the following results: TATA box = 25% (dn), 41% (up); CAAT box = 0% (dn), 36% (up); GC-box = 21% (dn), 13% (up).

For some of the still unidentified consensus sequences EMSA experiments with both Arabidopsis and Solanum nuclear extracts are in progress.

Email: desimone@dbbm.unina.it