

PLANT CO-EXPRESSION ANNOTATION RESOURCE 2.0: A WEB TOOL FOR THE ASSOCIATION OF PROTEINS OF UNKNOWN FUNCTION TO ABIOTIC STRESSES IN PLANTS

Maurício de Alvarenga Mudadu, Adhemar Zerlotini Neto, Marcos José Andrade
Viana

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Abstract

The development of genetically modified (GM) crops includes the discovery of candidate genes through bioinformatics analysis, using genomic data, gene expression, among others. Proteins of unknown function (PUFs) are interesting targets for pipelines of GM crops due to the novelty associated and also to avoid copyright protections. One method to infer the possible function of PUFs is to relate them to factors of interest, such as abiotic stresses, using orthology and coexpression networks, applying the guilt by association approach. The objective of this work is the discovery of PUFs involved in responses to abiotic stresses in plants. For this, we analyzed and processed the genomic data of 67 plant species, including 5 important species tolerant to abiotic stresses. Diamond and InterproScan were used to discover PUFs in all of these species. RNA-seq data related to abiotic stress was downloaded from NCBI / GEO and used as inputs to the LSTrAP software to build coexpression networks and clusters whose members are most likely related to the molecular mechanisms associated with abiotic stress in plants. Ortholog groups were created with Orthofinder using all proteins as input. So, we searched for PUFs associated with these groups of ortholog and simultaneously with some gene coexpression cluster related to abiotic stress. With that, we stored in a database provided by the ax software (<https://github.com/lmb-embrapa/machado>) 2, 136, 336 genes and 2, 714, 161 mRNA, together with their translated proteins. We recovered 78, 416 PUFs with Diamond and Interproscan analyzes, created 91, 172 groups of orthologists and 1, 975 coexpression clusters. We developed a protocol to search for PUF annotations and retrieve PUFs that belong to an ortholog group that also contains proteins whose mRNA belongs to a coexpression cluster related to abiotic stresses. After running this protocol, we recovered 4, 673 PUFs. We conducted a literature search on the proteins that belong to the orthologs groups, for all the PUFs that belong to the species Pearl millet, *Populus simonii*, *Oropethium thomaeum* e *Boea hygrometrica*, all known to be tolerant to abiotic stress (517 PUFs). We found studies related to abiotic stresses, on average, for 67.5% of PUFs. A webserver <https://www.machado.cnptia.embrapa.br/plantannot2> is freely available and provides indexed queries of PUFs possibly associated with abiotic stresses. We hope that Plantannot2 will be useful for researchers trying to obtain genes related to responses to abiotic stresses for the production of new GM crops tolerant to the risks of climate change.

Funding:

Link to Video: