Phylogenetic Analysis to Provide Insight on the Trihelix Gene Function in Arabidopsis thaliana

Muthukumar, Pragati (School: Commack High School)

Plant growth and productivity is greatly affected by climate change, especially in hot and dry climates. Scientists use various methods to study gene families and find genes that respond to various abiotic stresses. The Trihelix gene family is a plant-specific transcription factor family that has known roles in heat and drought stress. However, this gene family has not been functionally characterized in most plants. The purpose of this study was to use a genome-wide method to identify, classify and predict gene function in the Trihelix gene family. This was done by incorporating a phylogenetic study, looking at gene expression and protein-protein interaction. Thirty members of the Trihelix gene family were identified using Pfam ID(PF13837) and their respective MYB-like/SANT binding domain was verified using InterProScan. From the phylogenetic tree, the gene family was divided into two sub-classes and four sub-groups. A second phylogenetic tree including trihelix members of other plant species including Zea mays, Sorghum bicolor and Orzya Sativa helped to identify dicot and monocot specific members. The Trihelix gene family was further grouped by the expression at different developmental stages and with different abiotic stresses (salt, light and heat) using publicly available data. The data shows that 5 gene pairs were determined to share similar gene function and their protein similarity was further tested with a percent dot matrix. Two particular gene pairs, AT2G38250-AT5G01380 and AT3G11100-AT5G05550 may be induced by high temperatures and codes for heat stress proteins. Further experimentation is currently being done with these Trihelix mutant plant lines in order to determine biological function and understand the relationship between genotype and phenotype.

Awards Won:
Second Award of $2,000