

Bioinformatics analysis of stress-induced Hsp20 protein family members in ash tree genome

Esra Nurten Yer^{1,*}, Mehmet Cengiz Baloglu², Sezgin Ayan¹

¹Kastamonu University, Faculty of Forestry, Silviculture Department, Kastamonu, Turkey

²Kastamonu University, Faculty of Engineering and Architecture, Department of Genetics and Bioengineering, Kastamonu, Turkey

* Corresponding author: esranurtenyer@gmail.com

Abstract: Small heat shock proteins are characterized by a conserved sequence which acts as molecular chaperones. They can bind to denatured substrates of proteins and result in preventing unfolding and denaturation of proteins. They are associated with different abiotic stress tolerance such as regulation of salinity, drought and extreme temperature changes. *Fraxinus excelsior* L. is one of the important forest tree whose genome project was completed in 2017. Maintaining of protein conformation is extremely important for protecting plants against severe stresses. Although some of important protein family genes have been identified in ash tree genome, there is a limited study on genome-wide identification and characterizations of *Hsp20s* in both *F. excelsior* as well Oleaceae family members. Using various bioinformatics methods, we have analyzed ash genomes to find out *Hsp20* protein family members. PFAM analysis, motif analysis, gene structure analysis, phylogenetic tree construction and Gene Ontology (GO) analysis have been performed to characterize *Hsp20* proteins. PFAM accession number of *Hsp20* is PF00011.16 which reveals *Hsp20*/alpha crystalline family domain. Based on PFAM analysis, a total of 43 *FexHsp20* genes were found in ash tree genome. According to phylogenetic analysis, *FexHsp20* proteins could be classified into five main groups. A good number of the internal branches were observed with higher bootstrap values because of performing bootstrap analysis with 1000 replicates. It was also examined gene structure profile and motifs of *FexHsp20* genes to check the reliability of the phylogeny analysis. Protein sequence motifs were determined using the multiple EM for motif elicitation tool (MEME). MEME analysis identified 15 motifs according to their domain compositions of *Hsp20* proteins. All identified *FexHsp20* genes have no intron regions. The different phylogenetic tree was also drawn according to exon-intron organization of *FexHsp20* genes. Both phylogenetic trees were correlated each other. In addition, motif structures were variable among phylogenetic tree groups. Besides, we performed the GO analysis of *FexHsp20* proteins which had mainly binding activity and played significant roles in cellular process. In addition, they were found in different cellular localizations such as cell, cell-part, organelle, organelle part, membrane and lumen. Characterization of *FexHsp20* proteins could be used for improving stress tolerant forest plants for further studies. Additionally, results could provide detail information for comparative genomics studies for different plant species.

Keywords: Small heat shock protein, Genome analyses, Ash genome, Bioinformatics methods