

Genome and Transcriptome-wide analysis of Cellulose Synthase Gene Superfamily: Insights in Soybean Seed Development

Mawra Nadeem¹, Muhammad Amjad Nawaz¹, Hafiz Mamoon Rehman¹, Sung-June Chu¹, Seung Hwan Yang¹ and Gyuhwa Chung^{1*}

¹Dept. of Biotechnology, Chonnam National University, Chonnam, 59626, KOREA

[Introduction]

The plant cellulose synthase gene superfamily, belongs to the category of type-2 glycosyltransferases, is involved in cellulose and hemicellulose biosynthesis. These enzymes are vital for maintaining cell-wall structural integrity throughout plant life. The major superfamily of cell wall-synthesizing enzymes is cellulose synthase-like (CSL) proteins, a group of type 2 glycosyltransferases found throughout the plant kingdom. This superfamily is further divided into cellulose synthase-A (CESA) and nine CSL subfamilies (CSL A/B/C/D/E/F/G/H/J).

[Materials and Methods]

We first mined the publically available data. Analysis was done to get insight into evolution, functional relationship, and transcriptional association and coordination. To further aid future molecular breeding and biotechnological applications, all the identified genes were subjected to predict the presence of SSR markers and miRNA target sites. Furthermore, publically available microarray datasets for various soybean plant developmental stages, anatomical parts as well as under different biotic and abiotic stresses were analyzed to study the expression potential of all gene families involved in CWRD. Finally, gene co-expression networks were studied to visualize transcriptional coordination within studied gene families. soybean genome to identify 78 putative CSs based on gene models in Arabidopsis. Furthermore, we confirmed their differential expression in 5 developmental stages, 68 anatomical parts, and 40 tissues/compartments of the developing soybean seed. Gene co-expression networks were analyzed to determine functional relationship between different soybean CSs. Finally, we proposed CS gene models for soybean seed development and confirmed the expression of 15 soybean CSs through RT-PCR.

[Results and Discussions]

Here, we identified 78 putative cellulose synthases (CS) in the soybean genome. Phylogenetic analysis against 40 reference Arabidopsis CS genes clustered soybean CSs into seven major groups (CESA, CSL A, B, C, D, E and G), located on 19 chromosomes (except chromosome 18). Soybean CS expansion occurred in 66 duplication events. Additionally, we identified 95 simple sequence repeat makers related to 44 CSs. We next performed digital expression analysis using publically available datasets to understand potential CS functions in soybean. We found that CSs were highly expressed during soybean seed development, a pattern confirmed with an Affymatrix soybean IVT array and validated with RNA-seq profiles. Within CS groups, CESAs had higher relative expression than CSLs. Soybean CS models were designed based on maximum average RPKM values. Gene co-expression networks were developed to explore which CSs could work together in soybean. Finally, RT-PCR analysis confirmed the expression of 15 selected CSs during all four seed developmental stages.

[Acknowledgements]

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2015R1D1A1A09060925).

*Corresponding author: Tel. +82-61-659-7302, E-mail, chung@chonnam.ac.kr