Journal of Plant Physiology 215 (2017) 163–175

Contents lists available at ScienceDirect

Journal of Plant Physiology

journal homepage: www.elsevier.com/locate/jplph

Genome and transcriptome-wide analyses of cellulose synthase gene superfamily in soybean

Muhammad Amjad Nawaz[a], Hafiz Mamoon Rehman[a], Faheem Shehzad Baloch[b], Babar Ijaz[c], Muhammad Amjad Ali[d], Iqrar Ahmad Khan[e], Jeong Dong Lee[f], Gyuhwa Chung[a,*], Seung Hwan Yang[a,*]

[a] Department of Biotechnology, Chonnam National University, Yeosu, Chonnam, 59626, Republic of Korea
[b] Department of Field Crops, Abant Izzet Baysal University, Bolu 14280, Turkey
[c] Department of Plant Genetics and Breeding, China Agricultural University, Beijing 100193, China
[d] Department of Plant Pathology, University of Agriculture, Faisalabad 38040, Pakistan
[e] Institute of Horticultural Sciences, University of Agriculture, Faisalabad 38040, Pakistan
[f] Division of Plant Biosciences, Kyungpook National University, Daegu 702-701, Republic of Korea

ARTICLE INFO

Keywords:
Cellulose
Hemicellulose
Cellulose synthase
Cell wall
Soybean
Seed development

ABSTRACT

The plant cellulose synthase gene superfamily belongs to the category of type-2 glycosyltransferases, and is involved in cellulose and hemicellulose biosynthesis. These enzymes are vital for maintaining cell-wall structural integrity throughout plant life. 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.

1. Introduction

Rich in active compounds and carbohydrates, soybean (*Glycine max* (L) Merr.) is an important legume predominantly cultivated for vegetable oil and protein that is used in human consumption and animal feed (Chung and Singh 2008; Nawaz et al., 2017). Soybean is also a model legume for comparative genomics due to the availability of its whole genome sequence, its substantial germplasm, and its widespread commercial applications (e.g., in processed soy foods and pharmaceuticals) (Rehman et al., 2016).

In plants, the cell wall is the main structural and functional organelle controlling shape, overall growth, and expansion, while serving as a barrier between inner and outer cellular transport (Richmond and Somerville, 2000; Persson et al., 2007). Cell-wall biosynthesis is a complex process controlled by specific enzymes (e.g., cellulose synthases, xyloglucan xylosyltransferases, galactomannan galtransferases, xyloglucan fucosyltransferases, xyloglucan galactosyltransferases, homogalacturonan α-1,4-galacturonosyltransferases, laccases and glucan synthases). Cellulose synthesized in the plasma membrane is the main load-bearing and organizing structure of the cell wall (Rai et al., 2016; Maleki et al., 2016). Unlike cellulose, other polysaccharides of the cell wall matrix, such as pectins and hemicelluloses, are synthesized in the Golgi (Suzuki et al., 2006). Post-synthesis, these various molecules are cross-linked, embedded into the cell wall, or both. Many databases have classified cell wall-related enzymes into different groups according to relevant biological processes (Lombard et al., 2014).

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) (Richmond and Somerville 2000; Yin et al.,

* Corresponding authors.
E-mail addresses: chung@chonnam.ac.kr (G. Chung), ymichigan@chonnam.ac.kr (S.H. Yang).

http://dx.doi.org/10.1016/j.jplph.2017.04.009
Received 9 February 2017; Received in revised form 12 April 2017; Accepted 14 April 2017
Available online 29 April 2017
0176-1617/ © 2017 Elsevier GmbH. All rights reserved.

2009, 2014). In *Arabidopsis*, at least six CSL gene families are present (A, B, C, D, E, and G) beside CESA, while CSL F, H, and J are presumed to be monocot-related (Genes et al., 2006). The Csl superfamily is distantly related to bacterial cellulose synthases (CSs). Complex cell wall meshwork is formed by cellulose (synthesized by CESAs) and hemicelluloses (synthesized by CSLs) as well as non-cellulosic polysaccharides to create a complex cell-wall meshwork. Major conserved active sites in these cell wall-related integral proteins are aspartyl D,D,D residues followed by the QxxRW motif. CESA members harbor an additional CQIC motif that is absent in CSL members (Kaur et al., 2016). Apart from having many structural similarities in conserved residues, domains, and transmembrane domains (TMDs), the subfamilies differ in protein length, topology, TMD number, and cellular localization.

Cellulose microfibrils are core components of cell walls that typically comprise up to 36 hydrogen-bonded chains with 500–14000 β-1,4-linked glucan residues (Pear et al., 1996). The synthesis of these interlinked glucan chains are coordinated, and they form overlapping layers around cells that greatly increase tensile strength, rigidity, as well as resistance to osmotic pressure and chemical attack (Li et al., 2017; Scheible et al., 2001). Cellulose synthesis is controlled by CESAs, first identified in cotton based on sequence similarity with bacterial CS (Pear et al., 1996). Individual CESA hexamers form a large synthetic complex in the plasma membrane. *Arabidopsis* is currently known to have 10 CESAs and 30 CSLs (Richmond and Somerville 2000). In terms of the other CSL members, CSLA catalyzes mannan and glucomannan synthesis (Goubet et al., 2009; Liepman et al., 2005), CSLC catalyzes xyloglucan synthesis (Cocuron et al., 2007), CSLDs might act as mannan synthases (Verhertbruggen et al., 2011). However exact polymer products of CLSDs is still under debate (Yin et al., 2014), while CSL F, H, and J catalyzes (1,3) (1,4)-β-glucan synthesis (mixed linkage glucan) (Burton et al., 2006). The function of other CSLs currently remains unknown (Burton et al., 2004; Genes et al., 2006; Doblin et al., 2009).

It is believed that cell wall biosynthesis is a complex process in which polysaccharides and glycoproteins are synthesized as oligomeric structures in Golgi and are then transported to periphery to take part in cell wall assembly (Geisler et al., 2008; Hansen et al., 2014). The exceptional polymers of cellulose are synthesized at the plasma membrane by CESAs complex made up of three different types of CESA proteins. Primary cell wall (PCW) and secondary cell wall (SCW) proteins forms microfibrils, which co-align with microtubules through POM2 and/or KORRIGAN mediating interaction (McFarlane et al., 2014; Bringmann et al., 2012). Transcriptional coordination analysis of cell wall related gene families in Arabidopsis revealed that many of the proteins tend to be co-expressed i.e. xylan and lignin synthesis is co-expressed with PCW and SCW proteins, and members of chintinase family show associated functions with CESAs.

Functional characterization of this superfamily in plants revealed a role in growth and development (Gall et al., 2015; Redekar et al., 2015), including elongation, root hair growth, and seed coat development (Testa et al., 2010). CSLs are also implicated in resistance to abiotic stressors, such as osmotic stress (Rai et al., 2016; Zhu et al., 2011), and negative environmental stimuli (cold, heat, and salt) (Guerriero et al., 2014). Hamann et al. (2004) studied CSL global expression in *Arabidopsis* and found that CESA genes are expressed at higher levels than CSL genes. Maize CESA genes are highly expressed in seedlings, roots, leaf, stalk, shoot, kernel, tassel, endosperm, and embryo (Dhugga et al., 2004). Finally, transcriptome profiling revealed differential expression of CESA during seed development (Severin et al., 2010).

With whole-genome-sequence availability, microarrays, mRNA profiling and available co-expression network platforms, we can now further validate the role of CSs in soybean seed development (Trans and Roberts, 2016; Mutwil et al., 2009). In this study, we conducted dry lab experiments supported with bioinformatics techniques and reverse transcriptase PCR (RT-PCR). We first mined the publically available soybean genome to identify 78 putative CSs based on gene models in *Arabidopsis*. We next subjected these putative genes to phylogenetic analysis, gene structure analysis, and physical chromosomal mapping, as well as duplication and synteny analysis. The identified genes were also examined for simple sequence repeats (SSRs) to aid molecular breeding and biotechnological applications. 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.

2. Material and methods

2.1. Data retrieval and identification of putative CS genes

A BLASTp 2.2.28+ search of the Phytozome 11.0 database (https://phytozome.jgi.doe.gov/pz/portal.html; Goodstein et al., 2012), using *Arabidopsis* CESA (10) and CSL (30) genes as query terms, was performed to identify putative CS genomic, CDS, and peptide sequences. Seventy-eight genes were obtained with an expectation-value (*e*-value) cutoff was $1e^{-10}$. To validate the membership of all CSL genes, three potential cellulose synthase domains were scanned and confirmed in EMBL InterProScan (https://www.ebi.ac.uk/interpro/): nucleotide-diphospho-sugar transferase (*Nucleotide-diphossugar_trans*, IPR029044), RING-type zinc-finger (*CES_Znf_RING*; IPR027934), and cellulose synthase (*Cellulose_synt*, IPR005150). Gene ID, functional annotation, peptide size, and chromosomal locations were obtained from Phytozome. Isoelectric point (pI) and protein molecular weight (kDa) were calculated in ExPASy (http://web.expasy.org/compute_pi/; Gasteiger et al., 2005). The TMDs for each putative peptide were predicted with TMHMM server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/).

2.2. Phylogenetic analysis of soybean CSs

An alignment of 118 CS peptides was generated using ClustalW in MEGA 7.0 (Stecher et al., 2016). All conserved residue sites were used to construct evolutionary relationships with the neighbor-joining (NJ) method (500 replicates). All positions containing gaps and missing data were excluded to yield a final dataset of 109 positions. Evolutionary analyses were conducted in MEGA 7.0. The final phylogenetic tree was visualized and edited in iTOL (http://itol.embl.de/; Letunic and Bork, 2016). A manual scan for active residues D,D,D and conserved motif QxxRW in the aligned sequences (Carroll and Specht, 2011) further confirmed the genes' membership in the CS superfamily. Conserved motifs were found via scanning peptide sequences through MEME suite 4.11.2 (http://alternate.memesuite.org/; Timothy et al., 2009).

2.3. Gene structure analysis and physical mapping

An intron-exon map of CS was constructed in accordance with previously published methods (Le et al., 2016; Rehman et al., 2016; Rai et al., 2016). Based on coding sequence alignment with genomic sequences, intron phase, length, and number, as well as exon length and number, were obtained to study the gene structure in Gene Structure Display Server 3GSDS; (http://gsds.cbi.pku.edu.cn/; Hu et al., 2015). The exon and intron boundaries were determined with the Analyze feature in PhytoMine (https://phytozome.jgi.doe.gov/phytomine/). The physical chromosomal locations of all CSs were retrieved from the soybean genome assembly Glyma.Wm82.a2 at soybase.org (https://soybase.org/) and then visualized in Map Chart 2.2 (Voorrips, 2002), with the output drawn using MapInspect v 1.0 (http://www.plantbreeding.wur.nl/uk/software_mapinspect.html).

Download English Version:

https://daneshyari.com/en/article/5518010

Download Persian Version:

https://daneshyari.com/article/5518010

Daneshyari.com