A genome-wide analysis of the *cellulose synthase-like* (*Csl*) gene family in maize

Y. LI[1], X. CHENG[2], Y. FU[1], Q. WU[1], Y. GUO[1], J. PENG[1], W. ZHANG[3]*, and B. HE[1]*

College of Life Sciences, Jiangxi Science and Technology Normal University, Nanchang 330013, P.R. China[1]
College of Life Sciences, Sichuan University, Chengdu 610064, P.R. China[2]
College of Life Sciences and Technology, Honghe University, Mengzi 661100, P.R. China[3]

Abstract

Cell walls play an important role in the structure and morphology of plants as well as in responses to various biotic and abiotic stresses. Although the comprehensive analysis of genes involved in cellulose synthase has been performed in model plants, such as *Arabidopsis thaliana* and rice, information regarding *cellulose synthase-like* (*Csl*) genes in maize is limited. In this study, a total of 56 members of *Csl* gene family were identified in maize genome and classified into six subfamilies. Analysis of gene structure and conserved motif indicated functional similarities among the ZmCsl proteins within the same subfamily. Additionally, the 56 *ZmCsl* genes were dispersed on 10 chromosomes. The expression patterns of *ZmCsl* genes in different tissues using the transcriptome data revealed that most of *ZmCsl* genes had a relatively high expression in root and tassel tissues. Moreover, the expression profiles of *ZmCsl* genes under drought and re-watering indicated that the expression of *ZmCsl* genes were mainly responsive to early stage of drought stress. The protein-protein interaction network of ZmCsl proposed some potentially interacting proteins. The data presented a comprehensive survey of *Csl* gene family in maize. The detailed description of maize *Csl* genes will be beneficial to understand their structural, functional, and evolutionary features and provide an important foundation for studying the roles of *ZmCsl* genes in response to biotic and abiotic stresses.

Additional key words: drought, *ZmCls* gene structure and function, protein-protein interaction, rehydration, *Zea mays*.

Introduction

The plant cell walls are essential for defining the shape of cells as well as providing structural integrity to plant tissues and organs. As the primary barrier, plant cell wall serves essential functions in resisting high cellular pressure potential and in protection of plants against environmental impacts (Kaur *et al.* 2017). Plant cell wall mainly consists of multiple polysaccharide fractions (Kaur *et al.* 2016). Among the various classes of polysaccharides, the major load-bearing component is cellulose. Cellulose is synthesized by cellulose synthase like (Csl) proteins, which belong to the superfamily of glycosyl transferase 2 (GT2). Due to the importance of the plant cell wall, the synthesis of cellulose and *Csl* genes have become the focus of research interest.

The *Csl* genes are found throughout the plant kingdom. However, a variable number of *Csl* genes ranging from 30 to 50 have been reported in different plant species. The members of *Csl* gene family are classified into nine subfamilies: *CslA*, *CslB*, *CslC*, *CslD*, *CslE*, *CslF*, *CslG*, *CslH*, and *CslJ*. Recently, a number of plant Csl genes have been identified, including rice (Hazen *et al.* 2002), poplar (Djerbi *et al.* 2005, Suzuki *et al.* 2006), and the moss *Physcomitrella patens* (Roberts and Bushoven 2007). The *CslJ* is a new Csl family discovered in cereals (Yin *et al.* 2009) whereas *CslB* and *CslG* families are generally lacked in cereals (Yin *et al.* 2014). Among the nine subfamilies, *CslA*, *CslC*, and *CslD* are conserved in all land plants whereas *CslF* and *CslH* families are restricted to grasses (Schreiber *et al.* 2014).

Several members of the *Csl* subfamilies have been reported to be involved in the biosynthesis of different cell wall polysaccharides. For example, *CslA* and *CslC* genes, encoding 1,4-β-glucoside synthase, have been identified and they are involved in the biosynthesis of mannan

---

and xylan, respectively (Cocuron *et al.* 2007, Liepman *et al.* 2005). Vega-Sánchez *et al.* (2012) have confirmed that rice *CslF6* gene mediates MLG (mixed-linkage glucan) biosynthesis after over-expression in *Nicotiana benthamiana*, while Schreiber *et al.* (2014) reported that the *CslF* gene family in barley may be involved in (1,3;1,4)-β-glucan synthesis. From all Csl families, *CslD*, encoding 1,4-β-linked glucan synthase, has the oldest intron-exon structure and it is found in all sequenced terrestrial plant genomes, including *Physcomitrella* and *Selaginella*, indicating that it has a highly conserved function in the entire plant community (Verhertbruggen *et al.* 2011, Douchkov *et al.* 2016). However, CslD proteins have also been suggested to synthesize xylan, homogalacturonan, or mannan polysaccharides. Therefore, the true nature of Csl biochemical activity remains an open question.

Maize is one of the three major food crops with an ancient cultivation history, and has great potential for the development of high nutrition food as well as biofuel (Wright *et al.* 2005). The synthesis of cellulose is of vital importance to protect maize against environmental impacts and provide raw materials for industrial applications. However, there are few reports on the related genes and mechanisms of cellulose synthesis in maize. Therefore, the aim of this study was to identify the *Csl* family members in the whole maize genome through large-scale data mining, to provide a basis for the cloning of *Csl* gene family in maize, and to analyze its evolution and structure. Furthermore, the transcriptomic expression analysis of *ZmCsl* genes and the results of protein-protein interaction prediction can lay a good foundation for further study on the function and the mechanism of cellulose synthesis.

Materials and methods

Determination the *Csl* family members in maize genome: Because naming rice *Csl* genes is well documented, it is used as a reference. So, all the sequences of rice *Csl* genes (*OsCsl*) were downloaded from the *Rice Genome Annotation Project* (http://rice.plantbiology.msu.edu/index.shtml). The whole genome sequence of maize was downloaded from *NCBI*. To identify the *Csl* gene family members in maize, the genome sequence of maize was aligned with all rice *Csl* genes using the *NCBI-BLAST-2.7.1+* software. The E-value was set to $1e^{-5}$. Then, the identity and cover region (more than 50 %) were used as filter criteria to eliminate improper *Csl* genes. These genes were further filtered through domain analysis to ensure that the selected sequences were non-redundant sequences to ultimately determine the candidate *Csl* gene family members. All candidate members should contain one of the two *Pfam* domain models (PF00535 and PF03552).

Multiple sequence alignment and the construction of phylogenetic tree: Multiple sequence alignment of rice *Csl* genes and *ZmCsl* gene family members was performed by the *Muscle* program with the default parameters.

MEGA (*v. 5.0.1.102*) was applied to construct a maximum likelihood tree (Tamura *et al.* 2011). The phylogenetic tree was evaluated by the bootstrap method, and the repeat value was set to 1 000. The *Figtree* program (*v. 1.4.3*) was used to visualize it.

Analysis of conserved motif and gene structure: To identify the conserved motifs of each *ZmCsl* gene, deduced ZmCsl protein sequences were subjected to *MEME v. 4.12.0*. The parameters were set to zero or one of a contributing motif site per sequence, the number of motif was chosen to be five; the motif width was set 6 to 50 (Wu *et al.* 2016). The results are visualized by the *TBtools* software. To further study the characteristics of *ZmCsl* genes, we used the *GSDS2.0* software to analyze the structure and distribution of introns and exons for each gene. The corresponding CDS sequences and genomic sequences were found in the *NCBI* database according to the *GenBank* accession number. To determine the chromosomal locations of all *Csl* genes in maize, the information of locus coordinates was obtained from *NCBI*. The location and relative distance of *Csl* gene were displayed by the *MapChart* (*v. 3.2*) software (Voorrips *et al.* 2002).

Expression of *ZmCsl* during drought stress and in different tissues: The genome-wide transcriptome data of maize under drought were obtained from *NCBI SRA* databases from accession SRR6665368 to SRR6665379. Seeds of *Zea mays* L. inbred line B73 were used as experimental material and soaked with water at 28 °C for 36 h. After germination, the seedlings were cultured under an irradiance of 600 µmol m^{-2} s^{-1}, a 16-h photoperiod, day/night temperatures of 24/18 °C, and an air humidity of 60 % for approximately 12 d. When the seedlings had thre fully expanded leaves, 20 % (m/v) polyethylene glycol (PEG) 6000 was added to the nutrient solution and after 60 h or 96 h of drought stress and after recovery for 3 d (denoted as T60, T96, and TR3d, respectively) leaves were harvested and immediately frozen in liquid nitrogen. Control seedlings were grown under the same conditions but without PEG treatment (denoted as CK60, CK96h, and CKR3d respectively).

The transcriptome data of different maize tissues have been submitted to *NCBI SRA* databases under accession No. SRP067440 (Wang *et al.* 2016). Roots were collected from 14-d-old seedlings; tassels and ears were collected during vegetative stages 7 and 8, pollen during the first reproductive stage, and embryo and endosperm from seeds 20 d after pollination. Tissues were immediately frozen in liquid nitrogen. The gene expressions were estimated by the *TopHat/Cufflinks* pipeline described in previous reports with fragments per kilobase of exon per million fragments mapped FPKM values. The heat maps for expression profiles were generated with the *OmicShare* tools (http://www.omicshare.com/tools/Home/Index/index.html).

Construction of protein-protein interaction network: Protein-protein interaction (PPI) data were obtained from the online database of *STRING*, which is an open
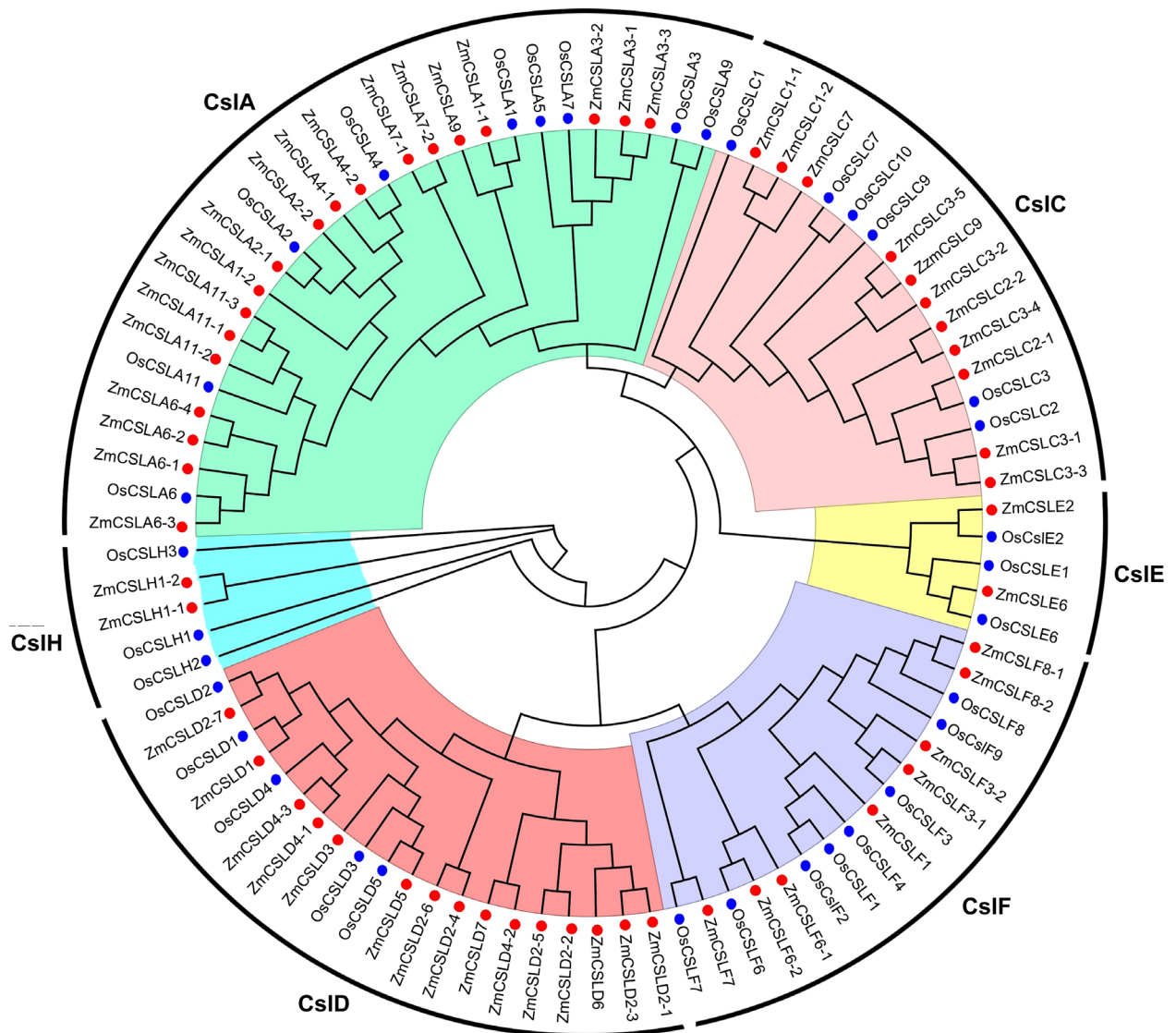
source software for predicting and visualizing complex networks. These interactions were derived from literature of experimental validation including physical interactions and enzymatic reactions found in signal transduction pathways. The PPI data were pre-processed including removing redundancy and self-loops (Zhou *et al.* 2016). Targets with a high confidence score > 0.7 were selected to construct the PPI networks (Zhu *et al.* 2016). The PPI networks were visualized in *Cytoscape* with the nodes representing proteins/genes and the edges representing interactions between any two proteins/genes.

Results

A total of 59 *Csl* genes were identified from the genome of maize using the *BLAST* programs with the query sequences

of rice *Csl* genes. Subsequently, database searches for maize, using conserved *Pfam* PF00535 and PF03552, resulted in the identification of 56 *ZmCsl* genes. The identified genes were named following the nomenclature of rice, which shares synteny with maize (Table 1 Suppl.). The length of protein ranged from 136 to 1514 aminoacids. All ZmCsl analyzed in this study possessd the structures typical of the Csl family, including a glycosyl transferase domain (PF00535) or cellulose synthase domain (PF03552). Except the structures mentioned above, ZmCslD1, ZmCslD2-7, ZmCslD3, and ZmCslD4-3 had a RING/Ubox like zinc-binding domain (PF14570) whereas ZmCslD2-1, ZmCslD2-3, ZmCslD2-4, and ZmCslD2-6 had a zinc-binding RING-finger domain (PF14569).

To investigate the phylogenetic relationship of the *Csl* gene family in maize and rice, the amino acid sequences of 56 Csl proteins from maize and 34 from rice were

Fig. 1. A phylogenetic relationship of rice and maize cellulose synthase-like (Csl). The maximum likelihood tree was constructed by the *MEGA* program with 1 000 bootstrap sampling using full-length sequences of 56 maize and 34 rice Csl proteins.

used to construct a phylogenetic tree. According it, maize Csl were classified into the same categories as rice Csl, which include CslA, CslC, CslD, CslE, CslF, and CslH subfamilies (Fig. 1). The CslB and CslG subfamilies, which lacked in cereals generally, missed also in maize. In addition, the CslJ subfamily, which was a new Csl family discovered in cereals, missed in maize as well because the sequences were unavailable. There were 19 members from maize and 9 from rice in the CslA subfamily, which was the largest subfamily. In addition, there were 11 and 13 members from maize classified into CslC and CslD subfamilies, only 2 members in CslH and CslE subfamilies, and 10 ZmCsl grouped in the CslF subfamily. In rice, 6 and 5 OsCsl members were classified into CslC and CslD subfamilies whereas 10 OsCsl grouped in the CslF subfamily. In the six clades, *Csl* genes had their clear orthologs in the genome of *O. sativa*, which suggests that these genes might be conserved for some specific functions in the two species. Furthermore, many *Csl* genes were grouped together suggesting that these homologous genes may have derived from multiple duplications after the speciation of maize during the evolution.

Analysis of exon-intron structure in a gene family is beneficial to obtain a further insight into their evolutionary trajectory. We used the *GSDS* software to compare the full-length cDNA with its corresponding genome sequence to detect the exon-intron structure of *ZmCsl* genes. The results indicate that the exon-intron structure of *ZmCs*l genes was conserved to a certain degree within subfamilies, which supportS the evolutionary relationships among members of each clade. However, as shown in Fig. 2, the number of introns in *ZmCsl* genes was highly variable among different clades, ranging from 1 to 13 introns. For example, *ZmCslF* genes contained only 1 or 2 introns whereas the majority of *ZmCsl* genes in other subfamilies contained more introns. The large variation in structures of *ZmCsl* genes suggests that the maize genome has changed significantly during its long evolutionary history.

To obtain insights into the diversity of motif compositions in ZmCsl proteins, putative motifs were predicted using the *MEME* program. A total of 10 conserved motifs were identified. The relative location of these motifs within the protein is represented in Fig. 3. The identified consensus sequence for the motifs is shown in Table 1. As expected, members that had similar motif compositions could be clustered into one class suggesting functional similarities among the ZmCsl proteins within the same subfamily. The ZmCslA and ZmCslC members, which have a closer phylogenetic relationship, contained motif 1, motif 2, motif 3, motif 4, motif 5, motif 9, and motif 10 whereas motif 6, motif 7, and motif 8 were found among other ZmCsl members. The distribution of motifs in different subfamilies implies sources of functional differentiation in *ZmCsl* genes in the evolutionary processes. Additionally, the motif distribution further confirms the accuracy of the phylogenetic relationship of *ZmCsl* genes.

The chromosomal distribution of the *ZmCsl* genes was visualized by the *MapChart* program (Fig. 4). The 56 *ZmCsl* genes were dispersed on 10 chromosomes. Chromosomes 1 and 7 harbored the most (10 of 56) *ZmCsl* genes whereas 3 *ZmCsl* genes were found on chromosome 8. In addition, there were 6 *ZmCsl* genes on chromosomes 5 and 9.

To investigate the expression patterns of *ZmCsl* genes in different tissues, we utilized publicly available transcriptome data to survey the transcription of the *ZmCsl* genes in the ear, embryo, endosperm, pollen, root, and tassel. From the results, we observed that 20 out of *ZmCsl* genes had relatively high expressions in root tissue, mainly including *ZmCslD* and *ZmCslF* genes (Fig. 5). Moreover, there were 17 *ZmCsl* genes that had very high transcript abundances in tassel, mainly *ZmCslD*, *ZmCslA*, and *ZmCslC* genes. Besides, six genes (*ZmCslA3-1*, *ZmCslA3-2*, *ZmCslA3-3*, *ZmCslC1-1*, *ZmCslD4-1*, and *ZmCslD4-3*) and two genes (*ZmCslD5* and *ZmCslD9*) were highly expressed in ear and pollen, respectively.

In order to elucidate the roles of *ZmCsl* genes in response to drought stress induced by PEG addition, we analyzed their expression profiles under drought and recovery. Obviously, more *ZmCsl* genes were up-regulated after drought treatment for 60 h (Fig. 6). For example, *ZmCslF6-1*, *ZmCslD1*, *ZmCslC2-1* and *ZmCslC2-2* displayed maximum expressions after stopping water supply for 60 h. Some genes were lightly up-regulated such as *ZmCslD2-2* and *ZmCslD4-2*. However, when recovering water supply, only *ZmCslD6, ZmCslC2-1*, and *ZmCslC2-2* exhibited high transcript accumulations. The

Table 1. The *MEME* motif sequences and lengths of maize cellulose synthase-like (ZmCsl) proteins.

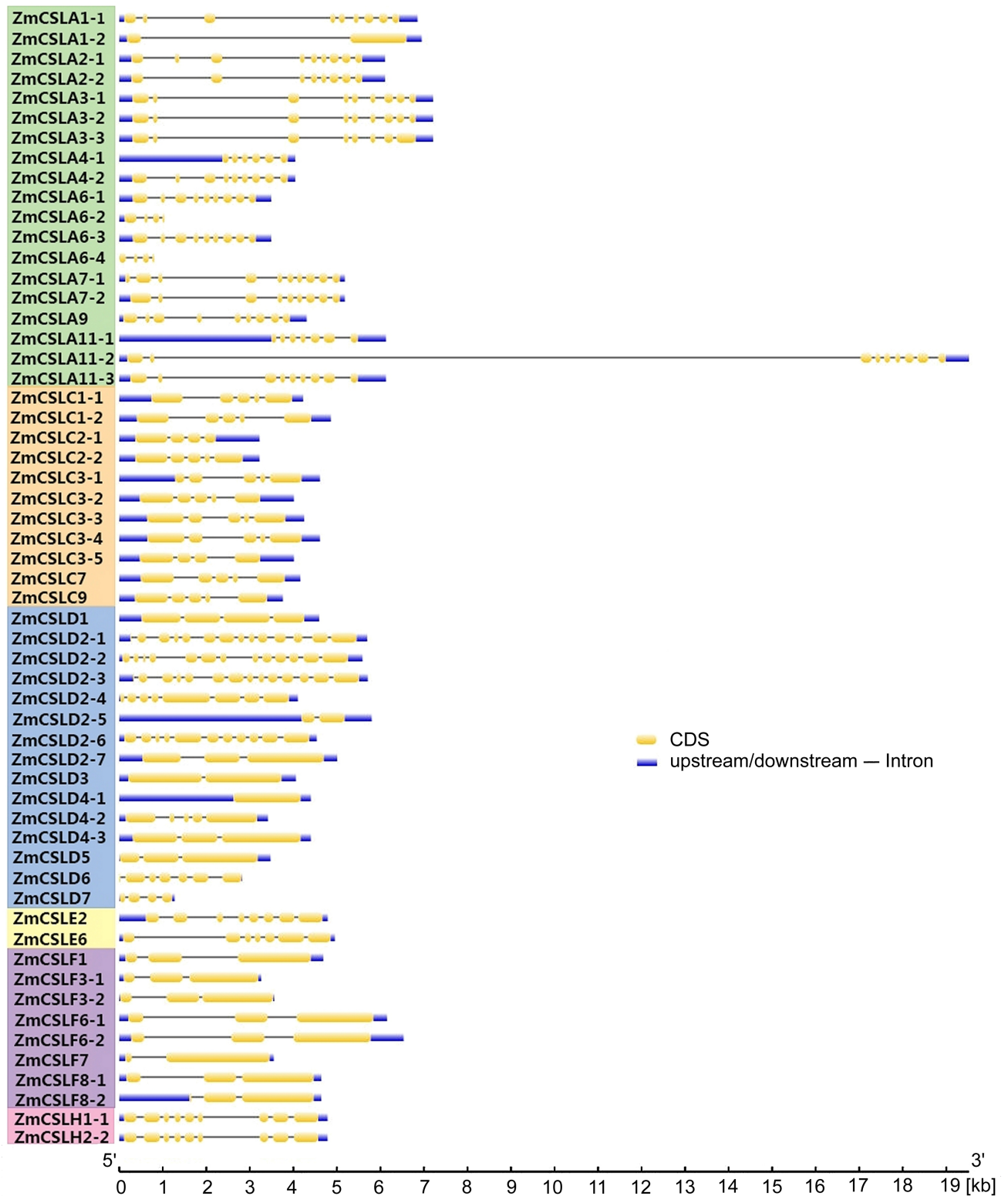| Motif	Conserved amino acid sequences	E-values	Sites	Width
1	KGWKFIYVGDVKVKSELPSTYKAYRKQQHRWSCGPANLFRKMFPEIJKSK	5.1e-953	27	50
2	LNFFGFNGTAGVWRISAJEESGGWKDRTTVEDMDJAVRAHL	3.0e-965	30	41
3	IFDADFQPEPDFLKRTVPFLVHNPEJALVQARWSFVNKDENLLTRJQEMN	3.8e-1044	29	50
4	SGYYPMVLVQIPMYNEREVYKLSIGAACGLDWPRDRFLVQVLDDSTDPVI	3.8e-868	25	50
5	SFHFVIFWILFENVMSVHRFKAAVSGLLZLG	4.2e-475	27	31
6	AYVQFPQRFDGIDPTDRYANHNRVFFDGNMRGLDGJQGPVYVGTGCVFRR	1.6e-861	26	50
7	PMLVYVSREKRPGYDHHKKAGAMNALVRVSAVLSNAPFILNLDCDHYVNN	7.2e-824	23	50
8	IYGSVTEDVVTGFRMHNRGWRSVYCSPKRDAFRGTAPINLTDRLHQVLRW	1.8e-794	24	50
9	KWAQKGVNIKYEHRVNRKGYKAGNLKSGMECDYVKDCEFVA	3.7e-650	27	41
10	VSLWKKFNLJYLFFFVRKVVAPFYTFTLYCVIJPLSVFVPE	1.6e-573	27	41

Fig. 2. The exon-intron structure of *cellulose synthase-like* (*Csl*) genes in maize. The *yellow boxes* indicate the exons and the *single lines* indicate introns. Untranslated regions are displayed by *thick purple lines* at both ends. Gene models were drawn to the *scale* as indicated at the bottom.

Fig. 3. Conserved motifs of the *cellulose synthase-like* (*Csl*) gene family in maize. All conserved motifs of the ZmCsl proteins were identified by the *MEME* program. Protein sequences are indicated by *thick gray lines*, and the conserved motifs are represented by different *colored boxes*. The length (the number of amino acids) of the protein and motif can be estimated using the *scale bar* at the bottom.

results reveal that the expressions of *ZmCsl* genes were responsive to early drought stress.

To illustrate the molecular mechanisms of ZmCsl proteins, PPI network was constructed using the data from the *STRING* database (Fig. 7). Four ZmCsl proteins ZmCslD2-4, ZmCslA4-1, ZmCslF6-1, and ZmCslF6-2

1

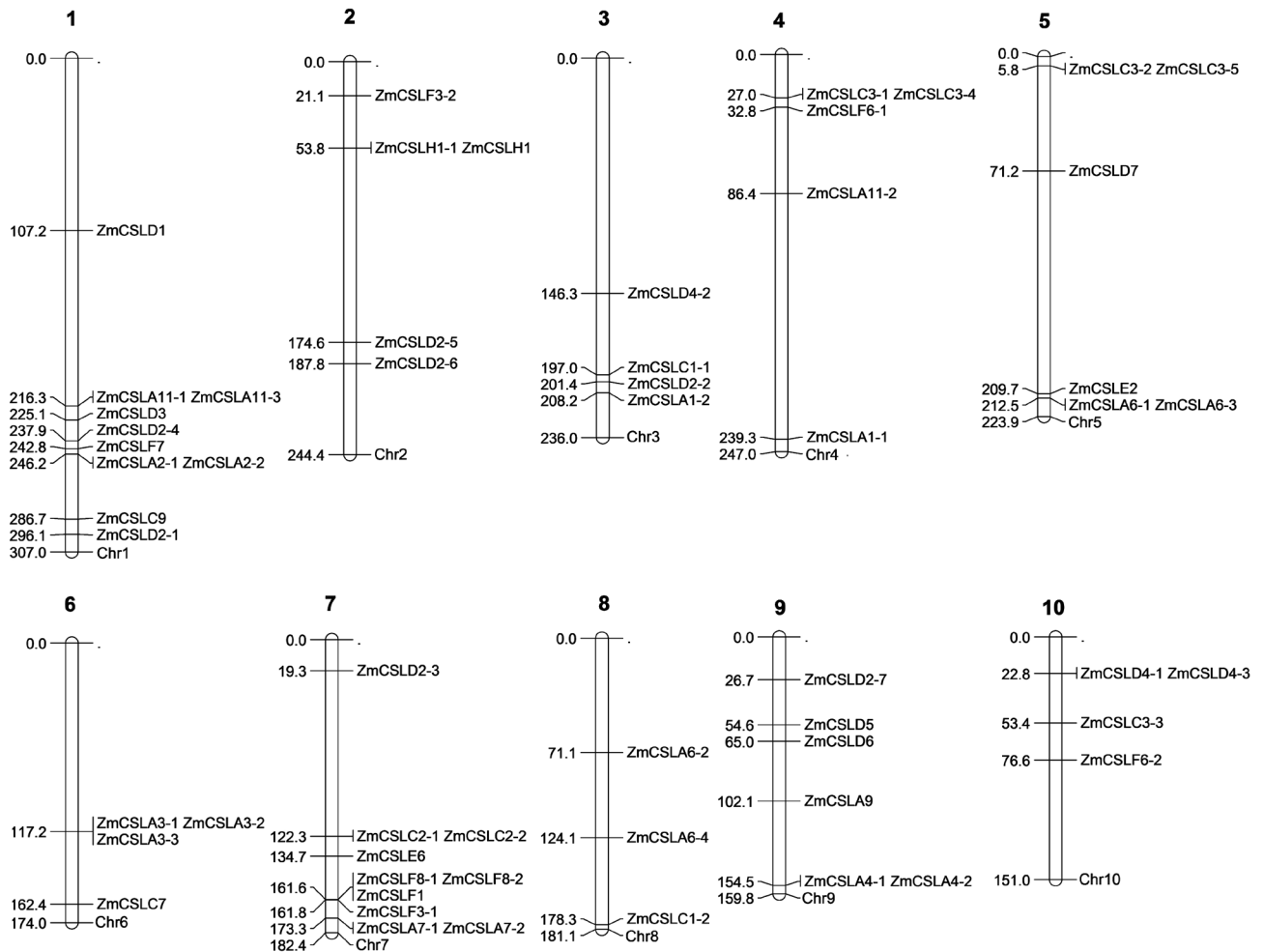|	
0.0	
107.2	ZmCSLD1
216.3	ZmCSLA11-1 ZmCSLA11-3
225.1	ZmCSLD3
237.9	ZmCSLD2-4
242.8	ZmCSLF7
246.2	ZmCSLA2-1 ZmCSLA2-2
286.7	ZmCSLC9
296.1	ZmCSLD2-1
307.0	Chr1

2

0.0	
21.1	ZmCSLF3-2
53.8	ZmCSLH1-1 ZmCSLH1
174.6	ZmCSLD2-5
187.8	ZmCSLD2-6
244.4	Chr2

3

0.0	
146.3	ZmCSLD4-2
197.0	ZmCSLC1-1
201.4	ZmCSLD2-2
208.2	ZmCSLA1-2
236.0	Chr3

4

0.0	
27.0	ZmCSLC3-1 ZmCSLC3-4
32.8	ZmCSLF6-1
86.4	ZmCSLA11-2
239.3	ZmCSLA1-1
247.0	Chr4

5

0.0	
5.8	ZmCSLC3-2 ZmCSLC3-5
71.2	ZmCSLD7
209.7	ZmCSLE2
212.5	ZmCSLA6-1 ZmCSLA6-3
223.9	Chr5

6

0.0	
117.2	ZmCSLA3-1 ZmCSLA3-2 ZmCSLA3-3
162.4	ZmCSLC7
174.0	Chr6

7

0.0	
19.3	ZmCSLD2-3
122.3	ZmCSLC2-1 ZmCSLC2-2
134.7	ZmCSLE6
161.6	ZmCSLF8-1 ZmCSLF8-2
	ZmCSLF1
161.8	ZmCSLF3-1
173.3	ZmCSLA7-1 ZmCSLA7-2
182.4	Chr7

8

0.0	
71.1	ZmCSLA6-2
124.1	ZmCSLA6-4
178.3	ZmCSLC1-2
181.1	Chr8

9

0.0	
26.7	ZmCSLD2-7
54.6	ZmCSLD5
65.0	ZmCSLD6
102.1	ZmCSLA9
154.5	ZmCSLA4-1 ZmCSLA4-2
159.8	Chr9

10

0.0	
22.8	ZmCSLD4-1 ZmCSLD4-3
53.4	ZmCSLC3-3
76.6	ZmCSLF6-2
151.0	Chr10

Fig. 4. The chromosomal locations of *cellulose synthase-like* (*Csl*) genes in the maize genome. Chromosomes 1 - 10 are depicted as *bars*. The *ZmCsl* genes are indicated by *black lines*.

were included in the network. Furthermore, ZmCslD2-4, GRMZM2G467943_P01, and GRMZM2G025016_P01 were located in the more important positions of the network suggesting those proteins play critical roles in the maintaining the whole protein interactions in the network. However, the function of GRMZM2G467943_P01 and GRMZM2G025016_P01 were unknown. The results in this study were beneficial to identify more important proteins and biological modules that interacted with ZmCsl proteins. Additionally, some proteins, such as Zm.21995 and GRMZM5G821988_P01 (eukaryotic translation initiation factors), interacted with ZmCslD2-4 as well. The detailed information of the proteins in the PPI network is listed in Table 2 Suppl.

Discussion

In recent years, a genome-wide analysis of gene family has become an efficient approach to understand gene structure, function, and evolution. The *Csl* genes are reported to be involved in cellular components and various

biological processes in many plant species. However, the comprehensive analysis of *Csl* gene family has been limited to *Arabidopsis thaliana*, rice, and wheat. In this study, we conducted a comprehensive analysis of the *ZmCsl* gene family, including identification of members, phylogenetic relationships, chromosomal organization, and expression profiles in different tissues and under drought stress conditions. The results will assist in understanding the roles of these *ZmCsl* genes and their potential molecular mechanisms in response to drought stress.

A total of 56 putative *Csl* genes were identified in the maize genome, the number of which was distinct from *A. thaliana* (36) and rice (34), a dicotyledonous and monocotyledonous plant, respectively (Schwerdt *et al.* 2015, Kaur *et al.* 2017). Comparison of the maize genome (2 500 Mb) with *A. thaliana* (125 Mb) and rice (466 Mb) reveals that maize has about 25-times and 5-times larger genome size than *A. thaliana* and rice, respectively (Inititiative 2000, Sasaki *et al.* 2002, Shukla *et al.* 2009). However, the number of *Csl* genes in these species is not proportional with the genome size. The reason may be that the maize genome has undergone significant gene losses
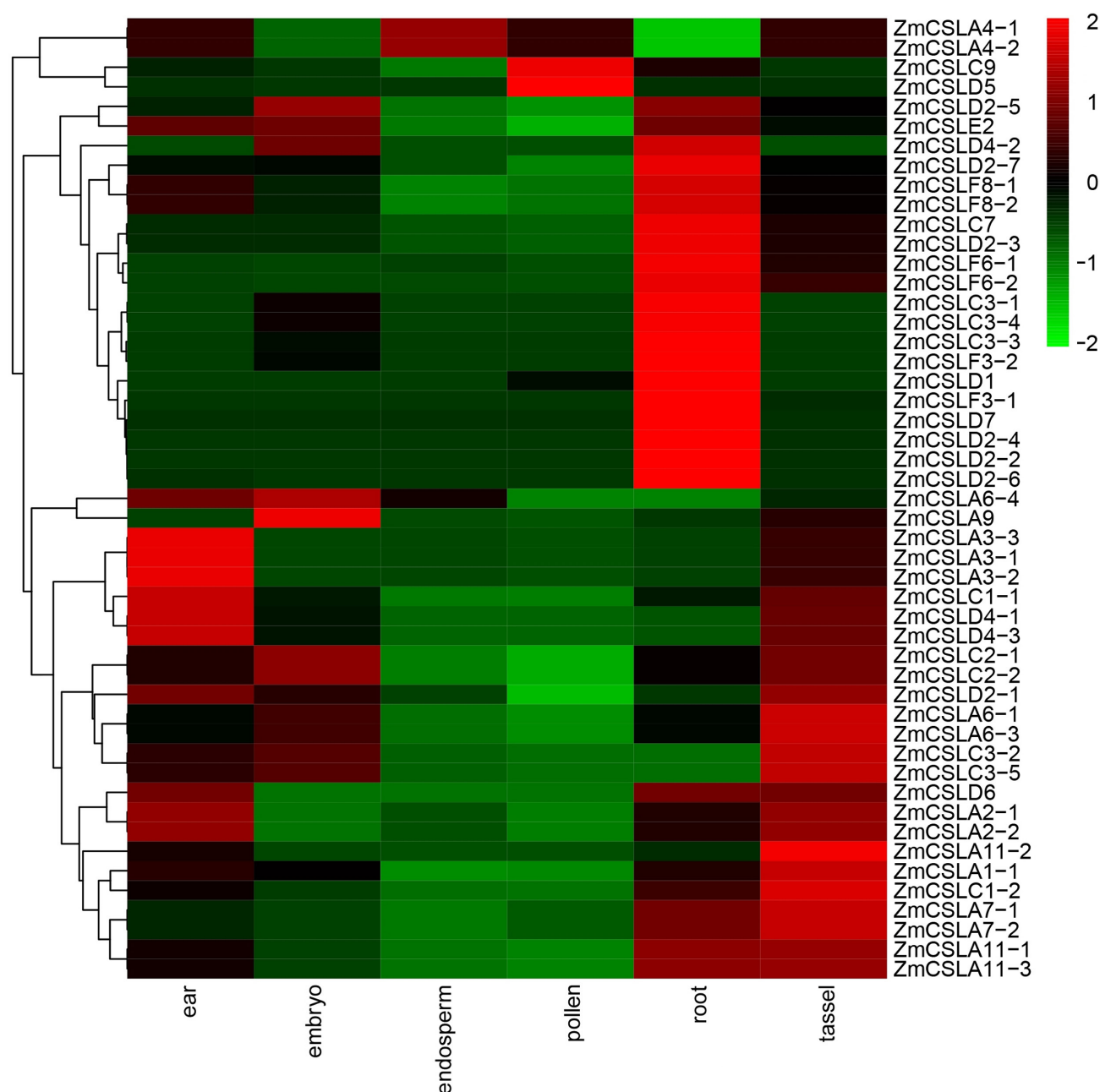
Fig. 5. The expression profiles of maize *cellulose synthase-like* (*ZmCsl*) genes across different tissues. The *ZmCsl* genes and color scale are shown at the *right* of the heat map. The genes with fragments per kilobase of exons per million fragments mapped (FPKM) equal to 0 were not used in this heat map.

since the duplication event. The results from Lai *et al.* (2004) further confirmed this speculation showing that only a small proportion of genes were conserved as duplicate factors in orthologous intervals of maize, when comparing the two homoeologous regions of the maize genome and the single homoeologous regions of the sorghum and rice genomes (Lai *et al.* 2004, Messing *et al.* 2004). Besides, phylogenetic analysis divided the members of the *ZmCsl* gene family into six distinct groups (*CslA*, *CslC*, *CslD*, *CslE*, *CslF*, and *CslH*), which are largely similar to those found in rice whereas *Csl* genes in *A. thaliana* are divided

into other six groups (*CslA*, *CslB*, *CslC*, *CslD*, *CslE*, and *CslG*). The results suggest that *Csl* genes may have a potential functional diversity between dicot and monocot plants (Penning *et al.* 2009).

The role of *Csl* genes have been reported mainly in *A. thaliana* and rice. For example, by destroying *CsLD5* expression in *A. thaliana*, Bernal *et al.* (2010) found that the growth of stem and root is significantly reduced in *CslD5* knockout plants, and the activities of xylan synthase and homogalacturonan synthase are also reduced indicating a possible role of this gene and other *CslD* genes in cell wall
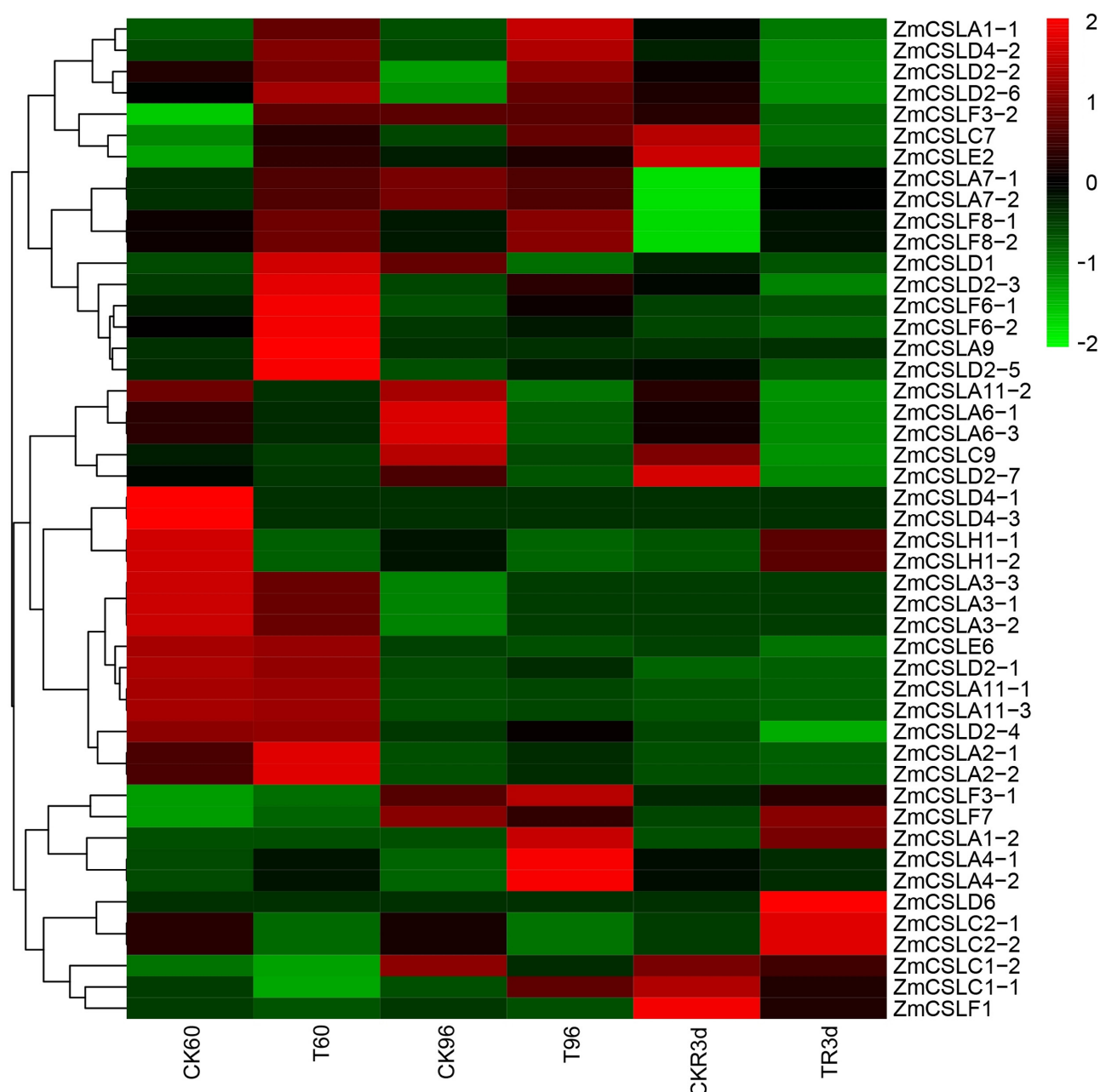
Fig. 6. The expression profiles of maize *cellulose synthase-like* (*ZmCsl*) genes in response to drought stress treatment. The color scale is shown at the *right* of the heat map. A higher expression for each gene is presented in *red*; otherwise, *green* was used. T60, T96, and TR3d indicate samples treated with PEG for 60 or 96 h and then recovered for 3 d. Control seedlings were grown under the same conditions but without PEG treatment (denoted as CK60, CK96h, and CKR3d, respectively).

biosynthesis. Goubet *et al.* (2010) observed glucomannan deficiency in a *CslA* mutant of *A. thaliana* indicating that the *CslA* family codes for glucomannan synthetase. In rice, there are continuous researches on the *Csl* gene family since the *Csl* gene family in rice has been identified (Hazen *et al.* 2002). For example, Wang *et al.* (2010) found that *Csl* genes in rice play important roles in the potentially functional complement for cell wall synthesis using transcriptional profiling and co-expression analyses. However, limited researches of *Csl* genes focused on

maize with the exception of *ZmCslD1*, which is integral to normal cell division, expansion, leaf development, and controls organ size in maize (Hunter 2010, Li *et al.* 2018).

In this study, a systematic analysis of *ZmCsl* genes was performed including the identification of members, phylogenetic analysis, and gene structure analysis. Furthermore, we explored the expression evidence for all putative *ZmCsl* genes in different tissues (ear, embryo, endosperm, pollen, root, and tassel) using the publicly available transcriptome data. The *ZmCslD* and *ZmCslF*
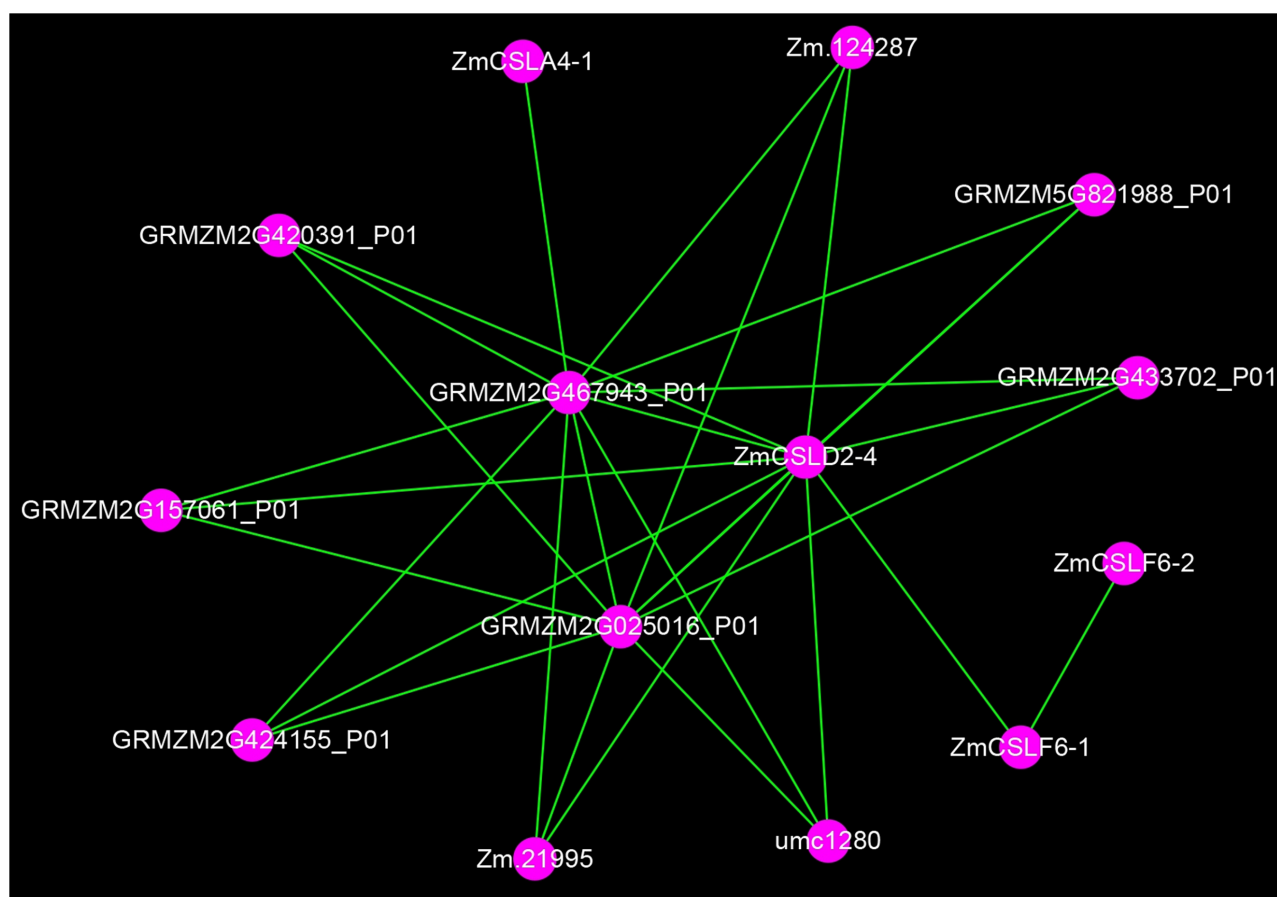
Fig. 7. A protein-protein interactions (PPI) network of maize *cellulose synthase-like* (Z*mCsl*). The *nodes* represent proteins and the *edges* represent the corresponding PPI. The confidence score was required to be greater than 0.7.

genes exhibited relatively high expressions in root tissue consistent with a previous research indicating that they could play a role in the development of tplant root (Wang *et al.* 2001, Hansey *et al.* 2010). Moreover, *Csl* genes have been shown to play crucial roles in responses to various biotic and abiotic stresses. In *A. thaliana*, *AtCslD5* is not essential for normal growth and development but play a critical role in osmotic stress tolerance, which may be likely involved in the regulation of reactive oxygen species under stress (Zhu *et al.* 2010). In another study of *A. thaliana*, a cellulose synthase gene was reported to enhance drought and osmotic stress tolerance through regulating cellulose synthesis (Chen *et al.* 2010). In *Hordeum vulgare*, *HvCslD2* represents an important defensive reaction both for nonhost resistance and for quantitative host resistance against non-adapted wheat and host-adapted barley powdery mildew pathogens, respectively, through mediating cell wall changes in the epidermal layer (Douchkov *et al.* 2016). In the present study, a significant number of *ZmCsl* genes was up-regulated after drought treatment for 60 h suggesting that these *Csl* genes may play positive roles in response to early drought stress. However, the specific mechanism of action needs to be further demonstrated and experimentally confirmed.

Conclusions

This study provides a systematic and comprehensive analysis of the *Csl* gene family in maize. In total, 56 members of the *Csl* gene family were identified in the maize genome and were classified into six subfamilies: *CslA*, *CslC*, *CslD*, *CslE*, *CslF*, and *CslH*. The expression profiles of *ZmCsl* genes in different tissues were performed to reveal that most *ZmCsl* genes had relatively high expressions in root and tassel tissues. Moreover, the expression patterns of *ZmCsl* genes under drought and re-watering indicate that the expressions of *ZmCsl* genes were mainly responsive to early stage of drought stress. The detailed description of maize *Csl* genes will be beneficial to understand their structural, functional, and evolutionary features. This information will provide an important foundation for studying the roles of these *ZmCsl* genes in response to biotic and abiotic stresses.

The sequence data are available in the *NCBI Sequence Read Archive* under accession numbers of SRP067440 for various maize tissues and SRR6665368 to SRR6665379 for maize under drought and recovery.

References

Bernal, A.J., Jensen, J.K., Harholt, J., Sørensen, S., Moller, I., Blaukopf, C., Bo, J., Lotto, R.D., Pauly, M., Scheller, H.V.: Disruption of ATCSLD5 results in reduced growth, reduced xylan and homogalacturonan synthase activity and altered xylan occurrence in *Arabidopsis*. - Plant J. **52**: 791-802, 2010.

Chen, Z., Hong, X., Zhang, H., Wang, Y., Li, X., Gong, Z.: Disruption of the cellulose synthase gene, *AtCesA8/IRX1*, enhances drought and osmotic stress tolerance in *Arabidopsis*. - Plant J. **43**: 273-283, 2010.

Cocuron, J. C., Lerouxel, O., Drakakaki, G., Alonso, A.P., Liepman, A.H., Keegstra, K., Raikhel, N., Wilkerson, C.G.: A gene from the cellulose synthase-like C family encodes a beta-1,4 glucan synthase. - Proc. nat. Acad. Sci. USA **104**: 8550-8555, 2007.

Djerbi, S., Lindskog, M., Arvestad, L., Sterky, F., Teeri, T.T.: The genome sequence of black cottonwood (*Populus trichocarpa*) reveals 18 conserved cellulose synthase (*CesA*) genes. - Planta **221**: 739-746, 2005.

Douchkov, D., Lueck, S., Hensel, G., Kumlehn, J., Rajaraman, J., Johrde, A., Doblin, M.S., Beahan, C.T., Kopischke, M., Fuchs, R.: The barley (*Hordeum vulgare*) cellulose synthase-like D2 gene (*HvCslD2*) mediates penetration resistance to host-adapted and nonhost isolates of the powdery mildew fungus. - New Phytol. **212**: 421-433, 2016.

Goubet, F., Barton, C.J., Mortimer, J.C., Yu, X., Zhang, Z., Miles, G. ., Richens, J., Liepman, A.H., Seffen, K., Dupree, P.: Cell wall glucomannan in *Arabidopsis* is synthesised by CSLA glycosyltransferases, and influences the progression of embryogenesis. - Plant J. **60**: 527-538, 2010.

Hansey, C.N., Lorenz, A.J., Leon, N.D.: Cell wall composition and ruminant digestibility of various maize tissues across development. - Bioenergy Res. **3**: 28-37, 2010.

Hazen, S.P., Scottcraig, J.S., Walton, J.D.: Cellulose synthase-like genes of rice. - Plant Physiol. **128**: 336-340, 2002.

Hunter, C., Iii, T.: The Mu transportation of *Zea mays* and their use in determining gene function: cellulose synthase-like D genes in plant and cell development. - Dissertations & Theses, Gradworks, Floroda 1016.

Inititiative, T.A.G.: Analysis of the genome sequence of *Arabidopsis thaliana*. - Nature **408**: 796-815, 2000.

Kaur, S., Dhugga, K.S., Beech, R., Singh, J.: Genome-wide analysis of the cellulose synthase-like (*Csl*) gene family in bread wheat (*Triticum aestivum* L.). - BMC Plant Biol. **17**: 193, 2017.

Kaur, S., Dhugga, K.S., Gill, K., Singh, J.: Novel structural and functional motifs in cellulose synthase (*CesA*) genes of bread wheat (*Triticum aestivum* L.). - Plos ONE **11**: e0147046, 2016.

Lai, J., Ma, J., Swigonová, Z., Ramakrishna, W., Linton, E., Llaca, V., Tanyolac, B., Park, Y.J., Jeong O.Y., Bennetzen, J.L.: Gene loss and movement in the maize genome. - Genome Res. **14**: 1924-1931, 2004.

Li, W., Yang, Z., Yao, J., Li, J., Song, W., Yang, X.: Cellulose synthase-like D1 controls organ size in maize. - BMC Plant Biol. **18**: 239, 2018.

Liepman, A.H., Wilkerson, C.G., Keegstra, K., Kende, H.J.: Expression of cellulose synthase-like (*Csl*) genes in insect cells reveals that *CslA* family members encode mannan synthases. - Proc. nat. Acad. Sci. USA **102**: 2221-2226, 2005.

Messing, J., Bharti, A K., Karlowski, W.M., Gundlach, H., Kim, H.R., Yu, Y., Wei, F., Fuks, G., Soderlund, C.A., Mayer, K.F.: Sequence composition and genome organization of maize. - Proc. nat. Acad. Sci. USA **101**: 14349-14354, 2004.

Penning, B.W., Hunter, C.T., Reuben, T., Eveland, A.L., Dugard, C.K., Olek, A.T., Wilfred, V., Koch, K.E., Mccarty, D. R.,

Davis, M.F.: Genetic resources for maize cell wall biology. - Plant Physiol. **151**: 1703-1728, 2009.

Roberts, A.W., Bushoven, J.T.: The cellulose synthase (*CESA*) gene superfamily of the moss *Physcomitrella patens*. - Plant mol. Biol. **63**: 207-219, 2007.

Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Niimura, Y., Cheng Z., Nagamura, Y.: The genome sequence and structure of rice chromosome 1. - Nature **420**: 312-316, 2002.

Schreiber, M., Wright, F., Mackenzie, K., Hedley, P.E., Schwerdt, J. G., Little, A, Burton, R.A., Fincher, G.B., Marshall, D., Waugh, R., Halpin, C.: The barley genome sequence assembly reveals three additional members of the *CslF* (1,3;1,4)-β-glucan synthase gene family. - Plos ONE **9**: e90888, 2014.

Schwerdt, J.G , Mackenzie, K , Wright, F., Oehme, D., Wagner, D., Harvey, A.J., Shirley, N.J., Burton, R.A., Schreiber, M., Halpin, C., Zimmer, J., Marshall, D.F., Waugh,R., Fincher, G.: Evolutionary dynamics of the cellulose synthase gene superfamily in grasses. - Plant Physiol. **168**: 968-983, 2015.

Shukla, V.K., Doyon, Y., Miller, J.C., Dekelver, R.C., Moehle, E.A., Worden, S.E., Mitchell, J.C., Arnold, N.L., Gopalan, S., Meng, X.: Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. - Nature **459**: 437-441, 2009.

Suzuki, S., Li, L., Sun, Y.H., Chiang, V.L.: The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus trichocarpa*. - Plant Physiol. **142**: 1233-1245, 2006.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S.: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. - Mol. Biol. Evol. **28**: 2731, 2011.

Vega-Sánchez, M.E., Verhertbruggen, Y., Christensen, U., Chen, X., Sharma, V., Varanasi, P., Jobling, S.A., Talbot, M., White, R.G., Joo, M.: Loss of cellulose synthase-like F6 function affects mixed-linkage glucan deposition, cell wall mechanical properties, and defense responses in vegetative tissues of rice. - Plant Physiol. **159**: 56-69, 2012.

Verhertbruggen, Y., Yin, L., Oikawa, A., Scheller, H.V.: Mannan synthase activity in the CSLD family. - Plant Signal. Behav. **6**: 1620-1623, 2011.

Voorrips, R.E., MapChart: software for the graphical presentation of linkage maps and QTLs. - J. Heredity **93**: 77-78, 2002.

Wang, B., Tseng, E., Regulski, M., Clark, T.A., Hon, T., Jiao, Y., Lu, Z., Olson, A., Stein, J.C., Ware, D.: Unveiling the complexity of the maize transcriptome by single-molecule long-read sequencing. - Nat. Commun. **7**: 11708, 2016.

Wang, L., Guo, K., Li, Y., Tu, Y., Hu, H., Wang, B., Cui, X., Peng, L.: Expression profiling and integrative analysis of the *CESA/ CSL* superfamily in rice. - BMC Plant Biol. **10**: 282, 2010.

Wang, X., Cnops, G., Vanderhaeghen, R., Block, S.D.: *AtCSLD3*, a cellulose synthase-like gene important for root hair growth in *Arabidopsis*. - Plant Physiol. **126**: 575-586, 2001.

Wright, S.I., Bi, I.V., Schroeder, S.G., Yamasaki, M., Doebley, J.F., Mcmullen, M.D., Gaut, B.S.: The effects of artificial selection on the maize genome. - Science **308**: 1310-1314, 2005.

Wu, Z., Cheng, J., Cui, J., Xu, X., Liang, G., Luo, X., Chen, X., Tang, X., Hu, K., Cheng, Q.: Genome-wide identification and expression profile of Dof transcription factor gene family in Pepper (*Capsicum annuum* L.). - Front. Plant Sci. **7**: 574, 2016.

Yin, Y., Huang, J., Xu, Y.: The cellulose synthase superfamily in fully sequenced plants and algae. - BMC Plant Biol. **9**: 99,

2009.

Yin, Y., Johns, M.A., Cao, H., Rupani, M.: A survey of plant and algal genomes and transcriptomes reveals new insights into the evolution and function of the cellulose synthase superfamily. - BMC Genomics **15**: 260, 2014.

Zhou, C., Yin, Y., Dam, P., Xu, Y.: Identification of novel proteins involved in plant cell-wall synthesis based on protein-protein interaction data. - J. Proteome Res. **9**: 5025-5037, 2010.

Zhu, J., Lee, B.H., Dellinger, M., Cui, X., Zhang, C., Wu, S., Nothnagel, E.A., Zhu, J.K.A.: Cellulose synthase-like protein is required for osmotic stress tolerance in *Arabidopsis*. - Plant J. **63**: 128-140, 2010.

Zhu, G., Wu, A., Xu, X., Xiao, P., Lu, L., Liu, J., Cao, Y., Chen, L., Wu, J., Zhao, X.: PPIM: A protein-protein interaction database for maize. - Plant Physiol. **170**: 618-626, 2016.