



Article

Genome-Wide Identification and Characterization of Copper Chaperone for Superoxide Dismutase (CCS) Gene Family in Response to Abiotic Stress in Soybean

Shuang Jiao ¹, Rui Feng ¹, Yu He ¹, Fengming Cao ¹, Yue Zhao ¹, Jingwen Zhou ¹, Hong Zhai ² and Xi Bai ^{1,*}

¹ College of Life Science, Northeast Agricultural University, Harbin 150030, China

² Key Laboratory of Soybean Molecular Design Breeding, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin 150081, China

* Correspondence: baixi@neau.edu.cn

Abstract: Copper Chaperone For Superoxide Dismutase (CCS) genes encode copper chaperone for Superoxide dismutase (SOD) and dramatically affect the activity of SOD through regulating copper delivery from target to SOD. SOD is the effective component of the antioxidant defense system in plant cells to reduce oxidative damage by eliminating Reactive oxygen species (ROS), which are produced during abiotic stress. CCS might play an important role in abiotic stress to eliminate the damage caused by ROS, however, little is known about CCS in soybean in abiotic stress regulation. In this study, 31 *GmCCS* gene family members were identified from soybean genome. These genes were classified into 4 subfamilies in the phylogenetic tree. Characteristics of 31 *GmCCS* genes including gene structure, chromosomal location, collinearity, conserved domain, protein motif, cis-elements, and tissue expression profiling were systematically analyzed. RT-qPCR was used to analyze the expression of 31 *GmCCS* under abiotic stress, and the results showed that 5 *GmCCS* genes (*GmCCS5*, *GmCCS7*, *GmCCS8*, *GmCCS11* and *GmCCS24*) were significantly induced by some kind of abiotic stress. The functions of these *GmCCS* genes in abiotic stress were tested using yeast expression system and soybean hairy roots. The results showed that *GmCCS7/GmCCS24* participated in drought stress regulation. Soybean hairy roots expressing *GmCCS7/GmCCS24* showed improved drought stress tolerance, with increased SOD and other antioxidant enzyme activities. The results of this study provide reference value in-depth study CCS gene family, and important gene resources for the genetic improvement of soybean drought stress tolerance.

Keywords: *Glycine max*; *GmCCS*s; ROS; abiotic stress



Citation: Jiao, S.; Feng, R.; He, Y.; Cao, F.; Zhao, Y.; Zhou, J.; Zhai, H.; Bai, X. Genome-Wide Identification and Characterization of Copper Chaperone for Superoxide Dismutase (CCS) Gene Family in Response to Abiotic Stress in Soybean. *Int. J. Mol. Sci.* **2023**, *24*, 5154. <https://doi.org/10.3390/ijms24065154>

Academic Editor: Hunseung Kang

Received: 8 February 2023

Revised: 24 February 2023

Accepted: 28 February 2023

Published: 8 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Abiotic stress, such as drought, high salinity, low and high temperature, heavy metals can seriously affect the normal growth and development of plants, especially the yield of crops. Drought stress can reduce the productivity of soybean by interfering with physiological and biochemical processes such as photosynthesis, translocation, respiration, and growth stimulants [1,2]. During a prolonged drought stress condition, reaction oxygen species (ROS) accumulate excessively and cause oxidative damage. Reaction oxygen species (ROS) are produced by redox reactions in plants, including aerobic respiration and photosynthesis, and modulate some important pathways, including apoptosis, energy metabolism, the inflammatory response, and stress responses. High levels of ROS such as O₂ (singlet oxygen), O₂⁻ (superoxide radical), OH (hydroxyl radical) and H₂O₂ (hydrogen peroxide) are generated during drought stress. To overcome drought-induced damage, plants activate their defense system by enhancing superoxide dismutase (SOD) activity to scavenge superfluous ROS.

CCS (copper chaperone for copper-zinc superoxide dismutase) proteins work as a copper chaperone and this factor has been characterized as a specific chaperone for SOD

(superoxide dismutase). CCS proteins were identified in the human first time which transferred, interfere and balanced copper to SOD and plan a strategy for reducing or preventing amyotrophic lateral sclerosis (ALS) [3,4]. SODs are metalloenzymes and the active site can be classified into Cu, Zn-SOD, Mn-SOD, or Fe-SOD depending on the metal. The CCS would recruit the copper ions to the copper-requiring enzyme Cu-zinc (Zn) superoxide dismutase (SOD1) whereas increasing the portion of CCS remarkably extends the survival of the transgenic mice administered with CuII (astm) [5]. A CCS-null mutant of *Drosophila* had characterized the requirement for CCS in activating SOD1, and was founded that the resemblant phenotypically of shortened adulthood [6]. The chloroplastic copper chaperone for the cuprozinc superoxide dismutase *GmCCS* gene was demonstrated to exist and cloned in 2014 in soybean [7]. A human liver fluke, *Clonorchis sinensis*, CsSOD1 can be successfully activated by the full-length CsCCS, while deletion of the N-terminal domain of CsCCS did not function [8]. CCS selectively bound to and facilitated Cu transfer to MEK1 in turn modulates the canonical mitogen-activated protein kinase pathway consisting of the RAF-MEK-ERK signaling cascade [9,10].

Nevertheless, the biological functions and molecular mechanisms of the *GmCCS* have not been revealed in soybean. The *GmCCS* members were identified in this study, and their location in the genome was determined. The motif, gene structure, conserved domains, and transcription factor prediction were also analyzed. The expression patterns of these genes in different tissues and under different stress conditions were analyzed. A series of abiotic stress models in yeast systems and hairy roots of soybean were tested. The mechanism of how this gene regulates oxidase activity during the biological stress response is gradually revealed. This study provides insights into the *GmCCS* gene family in *Glycine (G.) max* and a theoretical basis for future research on *GmCCS*s.

2. Results

2.1. Genome-Wide Identification of the *GmCCS*s Family of *Glycine (G.) Max* Subsection

To identify putative *GmCCS* genes, sequences from 3 *Arabidopsis* CCS proteins were used as queries in BLASTp search against *Glycine (G.) max* with relatively complete genome annotations. 31 *GmCCS* family members were identified and named *GmCCS1-GmCCS31* in soybean. The CDS length of *GmCCS*s ranged from 246 bp to 960 bp, the protein length was between 82 aa and 320 aa and PIs of 4.65–9.41. Nearly 70% of *GmCCS*s PI is less than 7, which is inferred to be an acidic protein (Table S1). Subcellular location prediction showed that 17 *GmCCS* members were located in the nucleus, 11 *GmCCS* members were located in the golgi apparatus and 5 *GmCCS* members were located in the cell membrane. Among them, 3 *GmCCS* members (*GmCCS1*, *GmCCS11*, *GmCCS30*) were located in the golgi apparatus and nucleus. *GmCCS13* was located in the cell membrane and chloroplast.

Chromosomal location indicated that 31 *GmCCS* genes excited on all ten chromosomes (Chrs). Thirty-one putative *GmCCS* were located on Chr19, followed by *Chr11*(5), *Chr03*(4), *Chr12*(4), *Chr05*(3), *Chr04*(2), *Chr10*(2), *Chr16*(2), *Chr20*(2), *Chr13*(1) (Figure 1A).

The orthologous relationship between *GmCCS* genes in *Glycine (G.) max*, *Arabidopsis* and *Glycine (G.) soja* genomes was assessed to clarify the evolutionary differences between CCS genes. CCSs were collinear with 47 and 15 genes in *Glycine (G.) soja* and *Arabidopsis*, respectively. The results showed that the *Glycine (G.) max* was more closely related to legume crops than other plants (Figure 1B). Gene duplication events are common in all species and generate new functional genes and drive evolution, 17 pairs of *GmCCS*s in *Glycine (G.) max* had collinearity among all 20 chromosomes (Figure 1C).

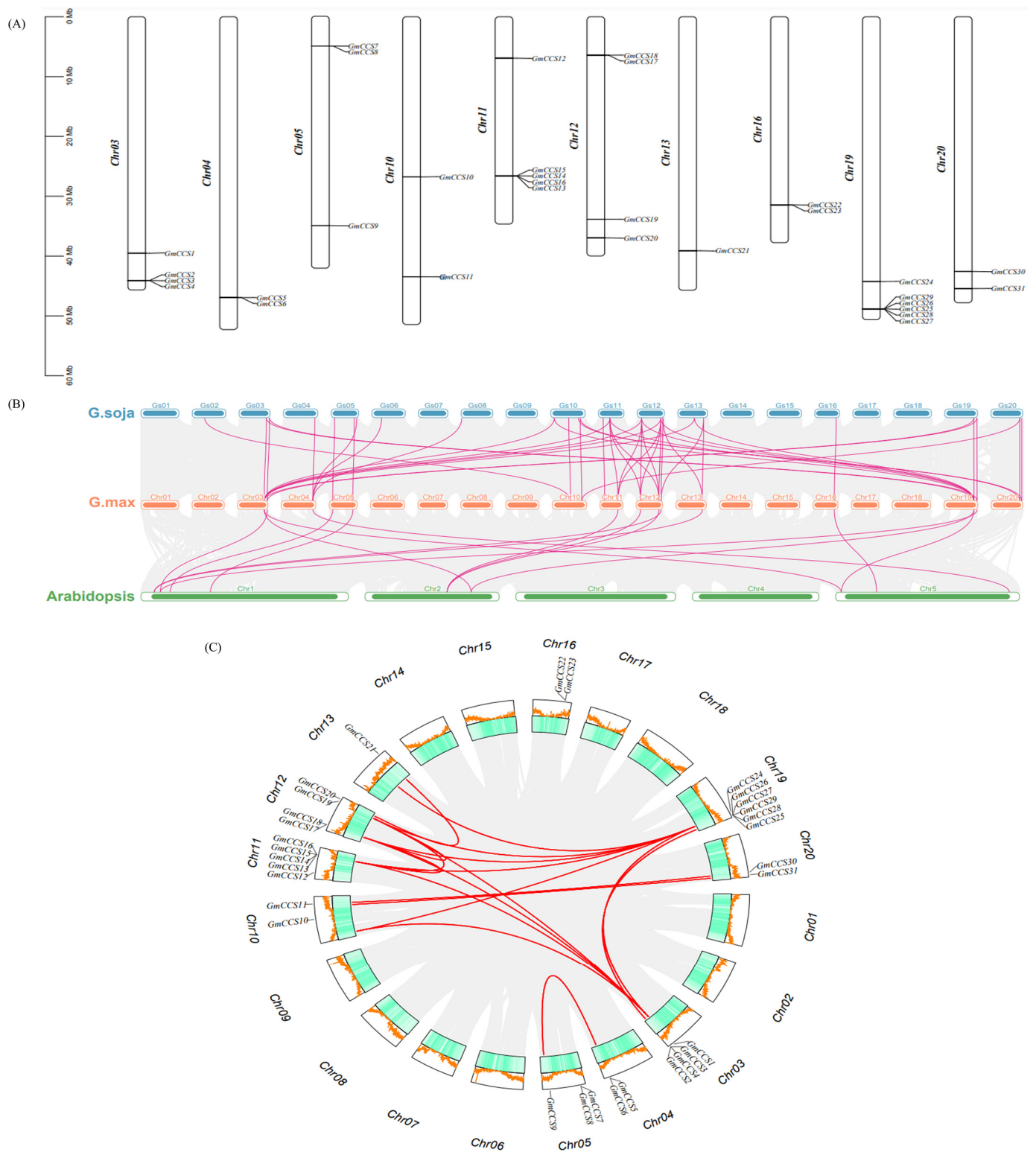


Figure 1. Chromosomal analysis of GmCCS members. **(A)** Chromosomal location of GmCCSs gene family in *Glycine (G.) max*. **(B)** Collinear analysis of GmCCSs with *Arabidopsis* and *Glycine (G.) soja*. Green, orange and blue represents *Arabidopsis*, *Glycine (G.) max* and *Glycine (G.) soja*, respectively. The bright lines indicate the connection with GmCCSs. **(C)** Chromosomal locations of GmCCS genes on 20 chromosomes of *Glycine (G.) max* and synteny analysis of interchromosomal relationships of GmCCS genes. The green blocks indicate parts of the radish chromosomes. The red lines indicate segment duplicated GmCCS gene pairs.

The maximum likelihood method was used to perform the phylogenetic analysis of GmCCS members of soybean (*Glycine (G.) max*), *Arabidopsis thaliana* and *Saccharomyces cerevisiae*. The model for MEGA prediction was used to analyze the phylogenetic relationship of GmCCS family. The GmCCS family could be classified into 4 subgroups (Groups I, II, III, IV) based on the phylogenetic tree analysis (Figure 2). GmCCSs have the same branch as different AtCCSs and ScCCS, which indicates that compared with other species, *Glycine (G.) max* is closed related to *Arabidopsis thaliana* and *Saccharomyces cerevisiae*. The yeast ScCCS1 is a multifunctional chaperone promoting all levels of SOD1 maturation [11].

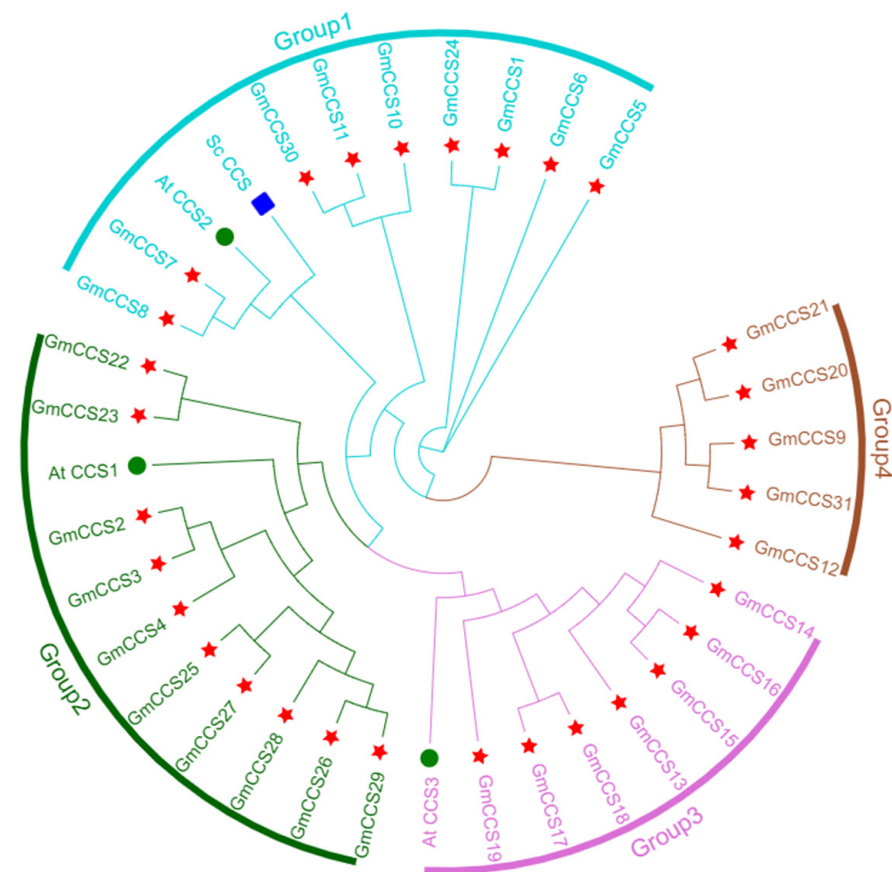


Figure 2. Phylogenetic tree of CCS proteins in *Glycine (G.) max*, *Arabidopsis thaliana* and *Saccharomyces cerevisiae*. Four subgroups were classified with different colors, and their names are marked in the corresponding positions. The circle, five-pointed star, and square text indicate *Arabidopsis thaliana*, *Glycine (G.) max* and *Saccharomyces cerevisiae*, respectively.

2.2. Gene Structure and Conserved Domain Analysis

Genome sequence and annotation files of AtCCSs, ScCCS and GmCCSs were downloaded and queried by TBtools. The exon-intron organization of 31 GmCCS genes was examined to obtain information on the structure, diversity, and evolution of GmCCS families in soybean, and the number of exons in GmCCS family members were from 3–7 (Figure 3B). Closely related genes have similar exon-intron arrangements with similar protein structures.

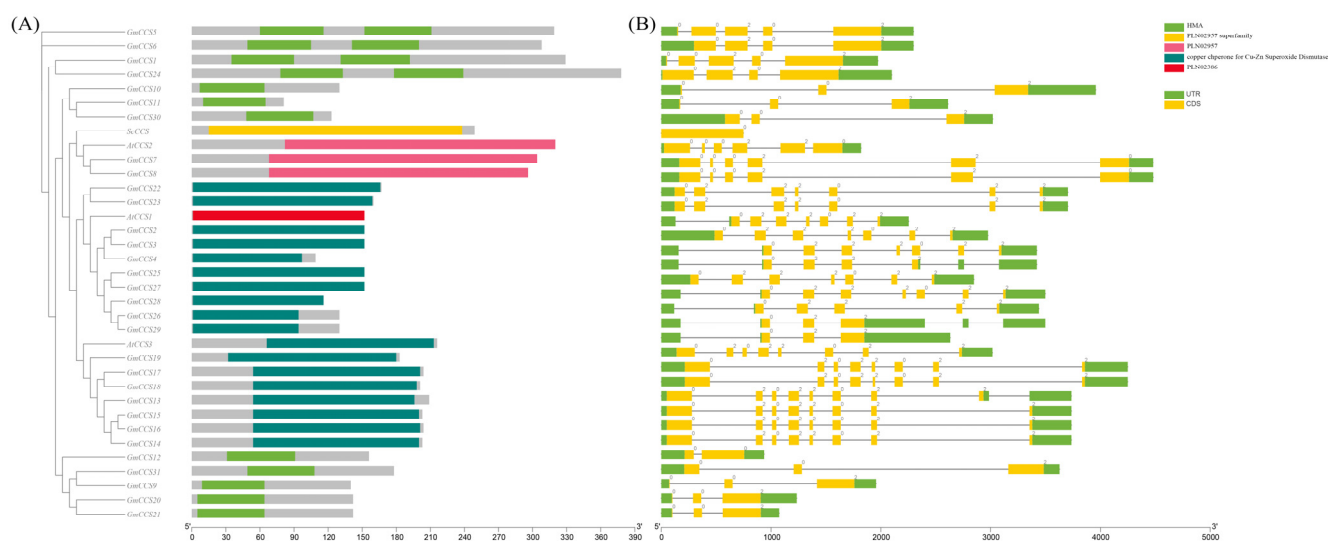


Figure 3. Gene structure, conserved domain analysis of GmCCS proteins from *Glycine (G.) max.* (A) The conserved domain in the GmCCS proteins. The green boxes indicate HMA domain (heavy-metal-associated), blue boxes respectively indicate Cu-zinc (Zn) superoxide dismutase and the grey box indicate non-specific functional area. (B) The UTR, CDS and intron organization of GmCCSs. The green boxes represent UTRs, the yellow boxes represent CDS and thin black lines represent introns.

In addition to the exon-intron pattern, the conserved domain could also be major for the various functions of *GmCCS* genes. Sequence analysis showed that the GmCCS proteins in soybean could be grouped as distinct clusters, more than half of all GmCCS proteins have the Cu-zinc (Zn) superoxide dismutase conserved domain (Figure 3A). Previous study showed that the CCS would recruit the copper ions to the copper-requiring enzyme Cu-zinc (Zn) superoxide dismutase (SOD1) whereas increasing the portion of CCS remarkably extends the survival of the transgenic mice administered with Cu^{II} (astm) [5]. It is worth noting that the twelve members of GmCCS proteins have heavy-metal-associated (HMA) domain, which directly binds AVR-Pik to activate plant defense during the stress of adversity [12].

2.3. The Motif, Cis-Regulatory Elements and Transcription Factor Analysis of GmCCSs

The MEME software was used to analyze the motifs of GmCCSs in *Glycine (G.) max.* Five different colored squares on the line represent different motifs in *GmCCS* members. The position on line indicates the composition of each CCS family member. Although various subgroups have different motifs, the *GmCCS*s members in the same subgroup have similar motif structures. For example, motif 4 is founded in subgroup I and subgroup IV (Figure 4A).

The analysis of promoter cis-elements can provide ideas for the tissue-specific expression and stress response modes of genes. The cis-regulatory elements of the 2000 bp promoter sequence of *GmCCS* genes were analyzed using PlantCARE. Twenty-two cis-regulatory elements were identified in the *GmCCS*s promoter sequence, including hormone-, stress- and photoresponsive element (Figure 4B). Light-responsive related elements were found in subgroup I and subgroup III, including TCT-motif, G-Box, GA-motif and GAGA-motif, TGA-box related to auxin-responsive [13]. Almost all members of GmCCS include stress-responsive elements, for instance, ABRE which response to ABA [14,15].

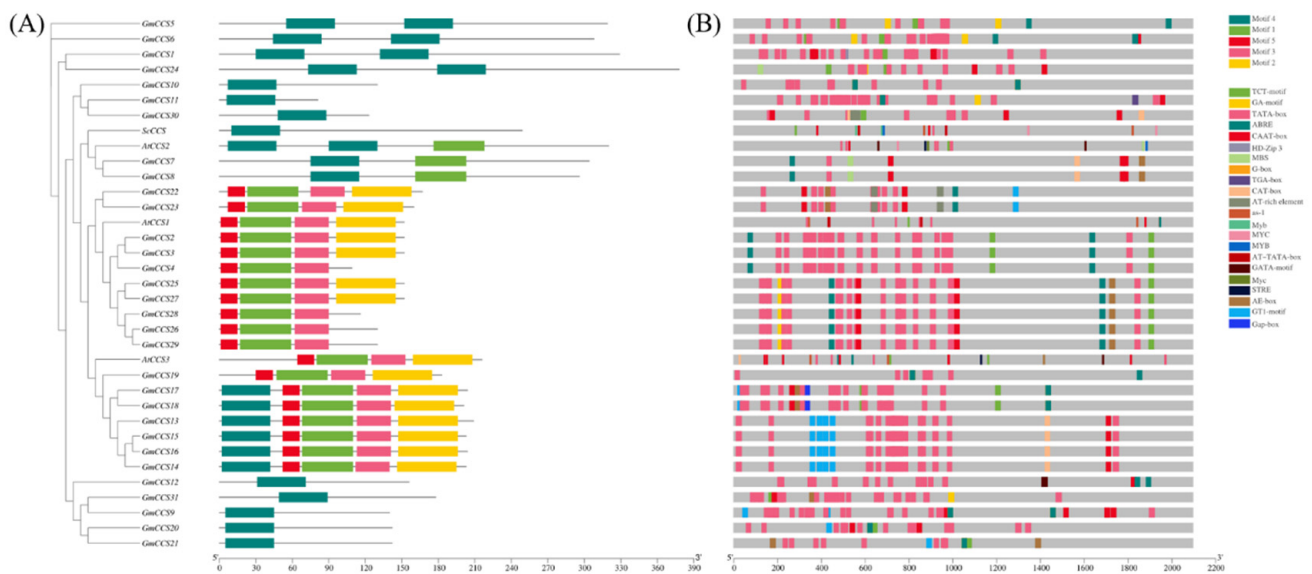


Figure 4. (A) The motifs in GmCCSs' protein. Boxes 1–5 represent motifs 1–5, the grey line indicate the non-specific functional area. (B) Visualization of hormone-responsive elements in GmCCSs promoters by TBtools, including position, kind and quantity of elements.

Promoter cis-elements combine with TFs to regulate the precise initiation and efficiency of transcription. Therefore, potential TFs that bind to the *GmCCSs* promoter were predicted by PlantTFDB (Figure 5). Almost all members were involved in the regulation of MADS, which responds to control plant flower development process [16–18]. DOF is involved to specific gene expression of stomatal guard cells and plant stress defense [19,20]. For instance, ERF, MYB, NAC, bZIP and WRKY transcription factors which correspond to biological stress and abiotic stress were found in 24 GmCCS members. ERF proteins were confirmed to control cell cycle and DNA repair by regulating genes. The PheE2F-PheDP complex exhibited diverse expression patterns in response to drought and salt treatment and diurnal cycles [21]. bZIP transcription factor regulates their target genes on their promoters and thereby direct adjusts plant developmental and abiotic stresses, such as drought, salinity, and cold. S1-bZIP is a unique one, which conserved upstream open reading frames mediate Sucrose-Induced Repression of Translation(SIRT), maintain sucrose homeostasis and withstand abiotic stresses in plants [22]. LcMYB2 promoted seed germination and root growth under drought and ABA treatments. LcMYB2 was confirmed in sheepgrass can regulate LcDREB2 expression by binding the promoter of LcDREB2, thereby, activating the expression of the osmotic stress marker genes in transgenic *Arabidopsis* [23].

2.4. Expression Analysis of GmCCS Genes in Different Tissues

To better understand the function of *GmCCS* genes in soybean (*Glycine (G.) max*), the expression profiles of *GmCCS* genes in different tissue, root, stem, leaf, flower and pod e.g., were analyzed using transcriptome data from the Phytozome (Figure 6A). The hierarchical clustering of the *GmCCS* genes was generated using a heatmap with the average fragments per kilobase of transcript per million mapped reads (FPKM) values in triplicate (Figure 6). 14 *GmCCS* genes exhibited high transcript abundance(FPKM > 5) in the root, followed by 11 *GmCCS* genes in the leaves, and 12 *GmCCS* genes in the hypocotyl, the expression of *GmCCS9* was not detected in any of the analyzed tissues. Overall, the expression levels of 8 *GmCCS* genes from I and II subfamilies accumulated (FPKM > 5) in all of the tested tissues, which suggests that these genes play a crucial role in tissue development.

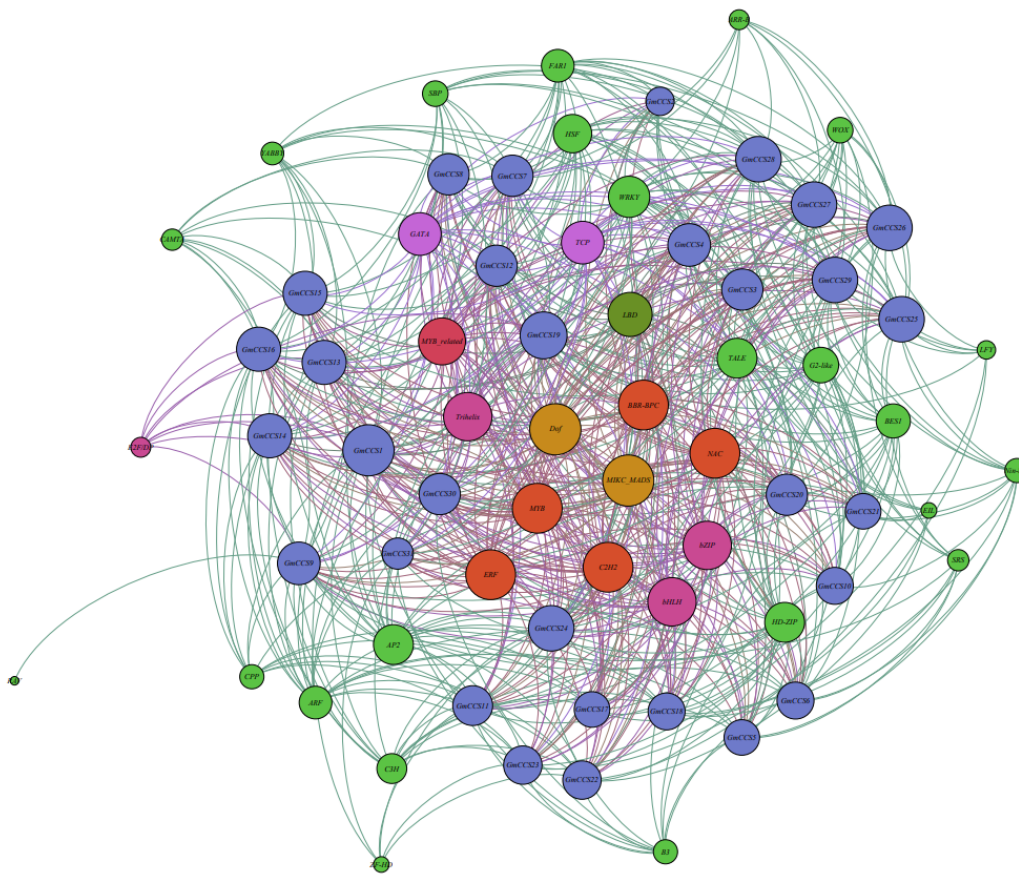


Figure 5. Interacting network among *GmCCS* family and their potential predicted TFs/protein. The blue circle represents the *GmCCS*s gene, the circles with different colors represent different TFs, the lines indicate the relationship between different TFs.

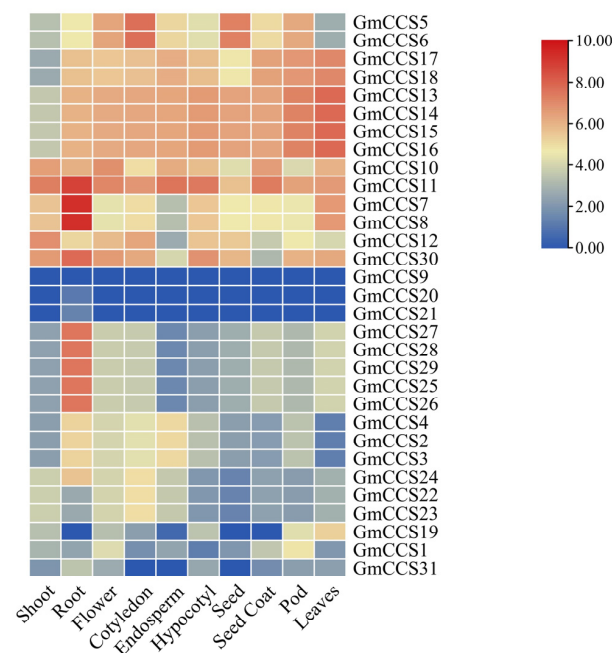


Figure 6. Heat map of expression profiles (in log₂-based FPKM) of soybean *GmCCS* subgroup genes in different tissues (root, root hairs, stem, leaf, nodules, flower and seed). The gene names are on the left and the tissue names are on the top of the figure, and the expression abundance of each transcript is represented by the color bar: Red, higher expression; blue, lower expression.

2.5. Expression Profiles of the Candidate *GmCCS* Genes under Different Abiotic Stresses

RT-qPCR was performed to analyze profiles of *GmCCS* genes under different abiotic stresses. The heatmap was created based on the RT-qPCR results. As shown in Figure 7A, transcript levels of *GmCCS7*, *GmCCS8* and *GmCCS24* were greatly induced by drought stress at some time points in both leaf and root. *GmCCS11* was greatly induced by drought stress at some time points only in leaf and *GmCCS5* was induced by drought stress at some time points only in root. These five *GmCCS* genes all respond to drought stress at transcript levels. The transcript level of the other *GmCCS* genes did not show significantly difference in drought stress during a series of treatment times. The transcript level of *GmCCS5* was greatly induced by NaCl stress at 2 h in leaf (Supplementary Figure S1). The transcript level of the other *GmCCS* genes did not show significantly difference upon NaCl stress during a series of treatment times (The fold change of gene expression was less than 1 fold.). Under NaHCO_3 stress treatment, transcript levels of *GmCCS14*, *GmCCS10*, *GmCCS22*, *GmCCS5* and *GmCCS6* were induced at 1 h in leaf (Supplementary Figure S2). However the fold changes of these five genes' expression were less than 1 fold. Above results indicated that *GmCCS5*, *GmCCS7*, *GmCCS8*, *GmCCS11* and *GmCCS24* might be involved in drought stress regulation, and *GmCCS5* might be involved in drought and salt stress regulation.

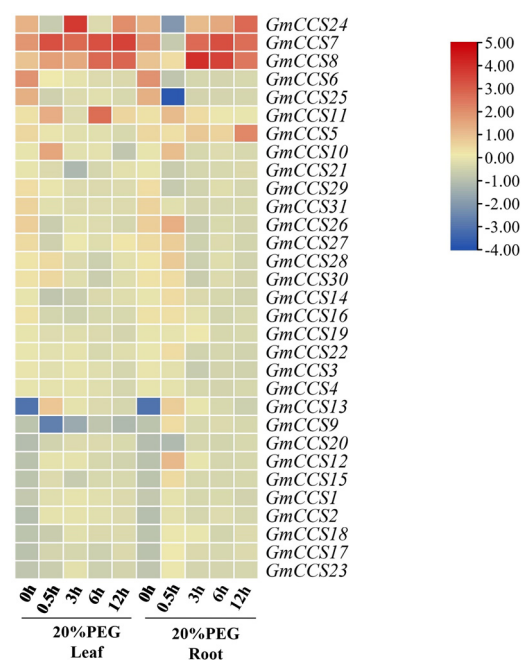


Figure 7. RT-qPCR analysis of *GmCCS* genes under drought stress. The heatmap was created based on the RT-qPCR results. The transcript abundance is represented by the color bar.

2.6. Overexpression of Abiotic Candidate *GmCCS* Genes Improved Drought Tolerance in Yeast Cells

To analyze whether these five abiotic responses *GmCCS* genes are involved in abiotic stress tolerance, *GmCCS5*, *GmCCS7*, *GmCCS8*, *GmCCS11* and *GmCCS24* were subcloned into pYES2 vector, and overexpressed in yeast cells respectively. Yeast cells transformed with an empty pYES2 vector was as the control. Spot assay grown on medium supplemented with sorbitol, NaCl, and NaHCO_3 was performed. Using sorbitol simulating drought stress, when the concentration of supplemented sorbitol was 0.4 M, expression of these five *GmCCS* genes dramatically enhanced sorbitol tolerance of yeast cells compared with control yeast cells (Figure 8). When sorbitol was supplemented with 0.8 and 1.6 M, expression of *GmCCS5*, *GmCCS7*, *GmCCS8* and *GmCCS11* dramatically enhanced sorbitol tolerance of yeast cells, however, transformants expressing *GmCCS5* did not (Figure 8). When sorbitol was supplemented with 2.0 M, transformants expressing these five *GmCCS* genes or the empty vector did not grow on the medium, indicating the concentration of sorbitol was too

high. Under NaCl and NaHCO₃ stress, expression of these five genes did not enhance NaCl or NaHCO₃ tolerance of yeast cells (Supplementary Figures S3 and S4). These five *GmCCS* genes all enhanced the drought tolerance of yeast, and *GmCCS7*, *GmCCS8*, *GmCCS11* and *GmCCS24* showed more stronger drought resistance.

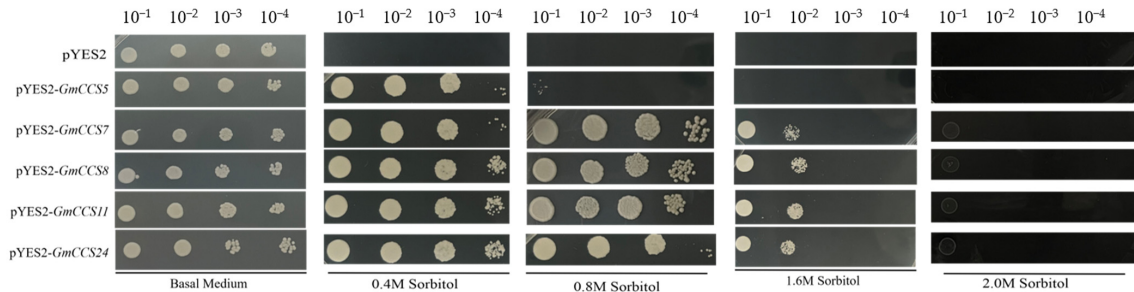


Figure 8. Drought tolerance test of candidate *GmCCS* genes in transformed yeast cells.

2.7. *GmCCS7* and *GmCCS24* Improve Drought Stress Tolerance in Transgenic Soybean Hairy Roots

To further confirm the drought stress tolerance of these *GmCCS* genes, a soybean transgenic hairy root system was used. Since *GmCCS7*, *GmCCS8*, *GmCCS11* and *GmCCS24* showed stronger drought resistance in yeast cells, we intended to further test the drought stress tolerance of these four genes. The amino acid identities of *GmCCS7* and *GmCCS8* were 97.37% and *GmCCS7* encoding protein had 8 more amino acids than *GmCCS8*, raising the possibility that *GmCCS7* represents the function of *GmCCS8*. Thus, *GmCCS7*, *GmCCS11* and *GmCCS24* were overexpressed in soybean hairy roots for drought stress tolerance analysis.

Accordingly, we overexpressed three *GmCCS*s under the control of the CaMV 35S promoter in the soybean hair roots. Compared to control hairy roots, transgenic roots had a significant increase of *GmCCS7*, *GmCCS11* and *GmCCS24* expression, respectively (Figure 9A). Soybean seeding plants with *GmCCS7*-OE and *GmCCS24*-OE hair roots performed better than the control, however *GmCCS11*-OE did not (Figure 9B). Survival rates (%) were calculated after drought stress treatment. As shown in Figure 9C, survival rates were 67.3% and 73.2% for soybean seeding plants with *GmCCS7*-OE and *GmCCS24*-OE root, 13.9% for seeding plants with *GmCCS11*-OE, and 27.3% for control plants. The survival rates of soybean seeding plants with *GmCCS7*-OE and *GmCCS24*-OE root were significantly higher than the control plants, however, survival rates of soybean seeding plants with *GmCCS11*-OE root were not. These results indicated that *GmCCS7* and *GmCCS24* improve drought stress tolerance in soybean, but *GmCCS11* does not.

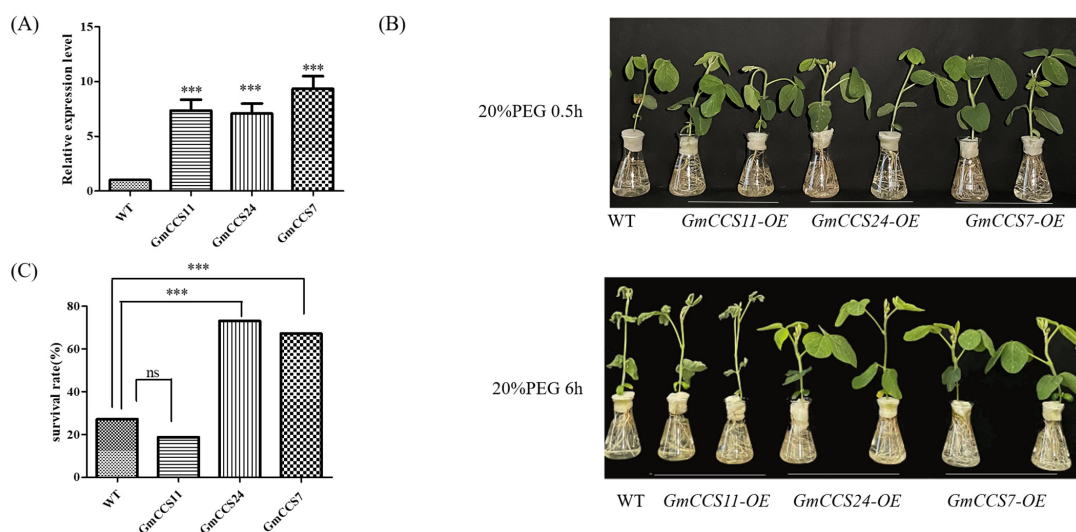


Figure 9. Cont.

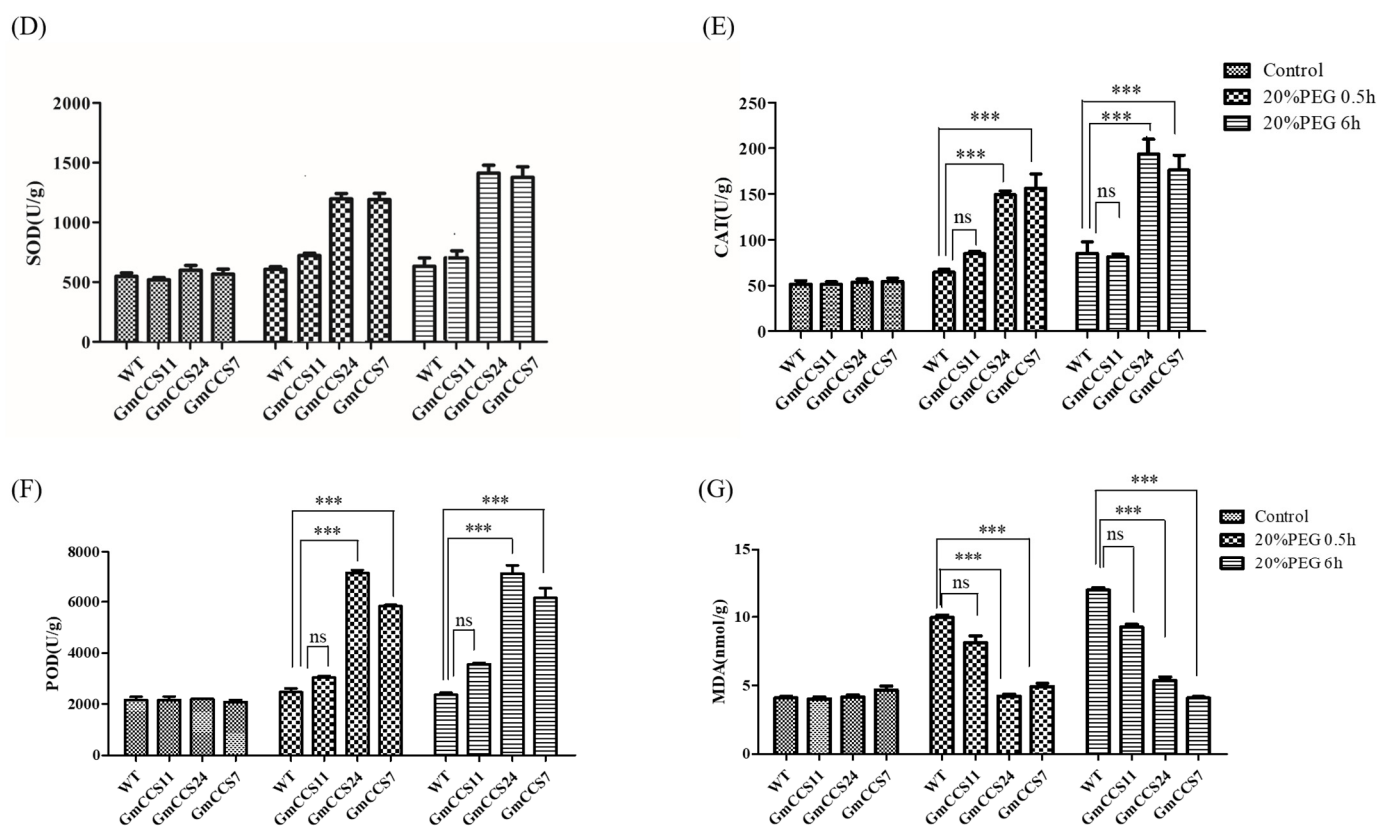


Figure 9. Soybean hairy roots expressing *GmCCS7/GmCCS24* showed improved drought stress tolerance. (A) RT-qPCR analysis of *GmCCS7*, *GmCCS11* and *GmCCS24* transcript levels in transgenic hairy roots. Transcript levels relative to *TUA5* were represented in each treatment; error bars indicate standard deviation (n = 3 plants). (B,C) Performance and survival rates of soybean seeding plants with transgenic hairy roots overexpressing *GmCCS7/GmCCS11/GmCCS24* under drought stress; error bars indicate standard deviation (n = 15 plants). (D–G) The antioxidant enzymes (POD, SOD and CAT) and the MDA content analysis under drought stress; error bars indicate standard deviation (n = 10 plants). The ns indicate none significant. One-tailed student's *t*-test was used to generate the *p* values. Stars indicate significant difference (***) compared to control plants (expressing empty vector).

To explore the mechanism of *GmCCS7*, *GmCCS11* and *GmCCS24* in drought stress tolerance, the activities of antioxidant enzymes (SOD, CAT, POD) were examined. As shown in Figure 9D–F, the SOD, CAT, and POD activities of three *GmCCS*-*OE* transgenic soybean roots showed no differences compared with control plants without drought treatment. Under drought stress conditions, the activities of these three enzymes in *GmCCS7*-*OE* and *GmCCS24*-*OE* transgenic soybean roots were significantly higher than those in control plants. The activities of CAT in *GmCCS11*-*OE* transgenic soybean roots were higher than those in control soybean roots, however, the activities of SOD and POD in *GmCCS11*-*OE* transgenic soybean roots were comparable to the control plants. *GmCCS11* could enhance the activities of antioxidant enzymes CAT, but not SOD and POD, raising the possibility of why *GmCCS11* did not improve drought stress tolerance in soybean. MDA content which represents the degree of cell membrane damage was assessed, as shown in Figure 9G, MDA contents of *GmCCS7*-*OE* and *GmCCS24*-*OE* transgenic soybean roots were lower than the control plants. Thus, the degree of damage to the cell membrane damage of *GmCCS7*-*OE* and *GmCCS24*-*OE* transgenic soybean roots was lower than the control. These results indicated that *GmCCS7* and *GmCCS24* enhanced drought stress tolerance by elevating the activities of antioxidant enzymes and decreasing cell membrane damage.

3. Discussion

Drought is an important factor that seriously affects the production and quality of soybean, especially for China which possesses a population base. The copper chaperone for superoxide dismutase (CCS) has been identified as a key factor in the integration of copper/zinc superoxide dismutase (Cu/Zn SOD) in yeast and mammals [24,25]. In *Arabidopsis* (*Arabidopsis thaliana*), *AtCCS* has been proven as the activation of Cu/Zn SOD activity [26]. However, studies to identify the biological function of *GmCCS* genes family under abiotic stress were little. In our study, a total of 31 *GmCCS* family genes were identified in this study in the soybean by comparative analysis with the *Arabidopsis* CCS gene family. The *GmCCS*s protein containing heavy-metal-associated (HMA) domain and copper chaperone for copper-zinc superoxide dismutase conserved domain, these two conserved domains were necessary for *GmCCS* genes in different plants to perform biological function. Rice HMA plant protein 19, which directly binds AVR-Pik to activate plant defense during the stress of adversity [12]. Cu chaperone CCS achieves appropriate installation of Cu within the SOD and activation of SOD, which in turn mediates degeneration in the fetal motor neuron disorder amyotrophic lateral sclerosis. *AtCCS* gene was responsive to the activation of all three types of CuZnSOD activity in *Arabidopsis*. To date, very little is known about the functional development of copper chaperones in soybean, in our study, the essential information about *GmCCS* family members were systematically and specifically elaborated.

We identified 31 *GmCCS* family members, were through two rounds of BLASTP based on 3 *Arabidopsis* CCS. Gene duplication is the most common evaluation mechanism for gene family expansion, which plays an important role in evolution by facilitating the generation of new genes and gene functions. There are three main evolutionary modes of gene duplication: segmental duplication (fragment duplications), tandem duplication, and translocation events. Tandem and fragment duplications are the major gene replication events in plant gene family expansion. Herein, several gene clusters were found on all 17 chromosomes of the common bean, which is probably due to several gene duplication events (Figure 1). CCS protein plays a key role in various cellular processes, participates in plant growth, development and stress responses [9]. The study on the function of CCS protein will help to elucidate the mechanism by that plants respond to environmental stresses and provide the theoretical basis for environmental adaptation in plants. Recently, *MaCCS* proteins have been found in *Musa acuminata* cv. *Tianbaojiao*, is involved in Abiotic and Hormonal Stress Responses [27]. 31 *GmCCS* genes in soybean and classified into four subfamilies, *AtCCS2* and *ScCCS* were distributed in Groups I, which contains 9 *GmCCS* family members. It implies that there are similar functions and more closely related in the evolution of CCS gene (Figure 2). More than half of all *GmCCS* proteins have the Cu-zinc (Zn) superoxide dismutase conserved domain. And twelve members of *GmCCS* proteins have heavy-metal-associated (HMA) domain, which directly binds AVR-Pik to activate plant defense (Figure 3).

The genes in each subgroup had similar motifs, gene structures, and cis-regulatory elements, supporting the reliability of the subfamily classification. The arrangement and number of exons and introns can provide important information about evolutionary relationships. Exon number plays a vital role in gene family evolution, while introns regulate gene expression. Therefore, gene function can be elucidated by analyzing gene structure. Generally, closely related genes have similar exon-intron arrangements with similar protein structures. In this study, the CCS genes in the same subgroup have the same number of exons and introns, indicating that they may have similar functions (Figure 3). This difference in the number of exons and introns of different genes in the same subfamily may be due to the functional diversity of genes during evolution.

Analysis of the arrangement and number of protein motifs in each CCS further demonstrated the categorization of the CCS family. The results showed that the structures were highly similar within each subfamily but different in each conserved region of the 5 motifs. Promoter analysis of CCS genes showed that CCS associated with stress (MBS, ARE, MRE and TC-rich), the light response (TCT-motif, G-Box, GA-motif and GAGA-motif, TGA-box) and hormone-related (ABR) genes broadly existed (Figure 4B). Promoter cis-elements combine with TFs to regulate the precise initiation and efficiency of transcription. ERF, MYB, NAC, bZIP and WRKY transcription factors which correspond to biological stress and abiotic stress were found in 24 *GmCCS* members (Figure 5). These results indicated that *GmCCS* genes may also be involved in various stresses and the regulation of different biological processes.

GmCCS5, *GmCCS7*, *GmCCS8*, *GmCCS11* and *GmCCS24* were significantly expressed in leaf and root in the present study and these genes are important regulators in the response to drought stress (Figures 6 and 7). These *GmCCS* genes were distributed in the Group I through phylogenetic analysis (Figure 2), and these genes contain heavy-metal-associated (HMA) domain and copper chaperone for copper-zinc superoxide dismutase conserved domain (Figure 3). Moreover, the plants which transformed with *GmCCS7-OE* and *GmCCS24-OE* showed a lesser wilting phenotype than the negative control plants under drought stress, *GmCCS11-OE* can not withstand drought stress (Figure 9A). On the basis of these collective reasons, transgenic soybean overexpressing *GmCCS7* and *GmCCS24* exhibited an increase in antioxidant enzymes under drought tolerance.

4. Materials and Methods

4.1. Identification of the *GmCCS* Gene Family

The whole genome sequence and annotation files of *Glycine (G.) max* were downloaded from the laboratory Wm82.a2.v1 (Phytozome; <https://phytozome-next.jgi.doe.gov/>) (accessed on 24 March 2022). The details of genome annotations were uploaded as supplemental files, including gene sequences, CDS, protein sequences and gff3 files. AtCCSs protein sequences were downloaded from TAIR (TAIR; <http://www.arabidopsis.org>). BLAST searches using TBtools (Auto Blast Two Sequences Set E-value = 1×10^{-5}) were performed using AtCCSs protein sequences as query sequences (Chen et al., 2020). Putative *GmCCS*s obtained by BLAST searches were then confirmed by NCBI-CDD and Pfam searches for the presence of the heavy-metal-associated domain (IPR006122) or copper chaperone for Superoxide dismutase domain (IPR024134). ExpASy-proparam was used to predict the coding sequence (CDS) length, predicted isoelectric (pI) and molecular weights (MWs) of all *GmCCS*s, Plant-mPLo 2.0 was used to predict their subcellular localization.

4.2. Gene Structure, Motif and Conserved Domain Analysis

The conserved motifs of the genes were analyzed by the MEME program with the following parameters: optimal motif width was set to 30–70, the number of repetitions was set to zero or one, the maximum number of motifs was set to identify 10 motifs, and visualized using TBtools. Gene structure and functional domains were analyzed and visualized using NCBI Batch CD-Search and TBtools.

4.3. Phylogenetic, Cis-Acting Element, Chromosomal Location Analysis, and Interacting Protein Prediction

CCS family protein sequences (Supplement Table S2) of *Arabidopsis*, *Glycine (G.) max*, *Saccharomyces cerevisiae* were downloaded from Phytozome database. A maximum likelihood (ML) phylogenetic tree was constructed using MEGA 7.0, with the following parameters: Poisson correction, pairwise deletion, and bootstrap (1000 replicates). The 2000 bp promoter fragments upstream of start codon were obtained using TBtools [28]. The *cis*-acting element searches were performed using the putative *cis*-element databases PlantCare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) and visualized using TBtools after the statistical screening. The chromosomal locations of *GmCCS*s were obtained from

the NCBI database and visualized using TBtools. The interacting protein with GmCCSs was performed using PlantTFDB (<http://planttfdb.cbi.pku.edu.cn/>).

4.4. Tissue Expression Profiling Analysis of GmCCS Gene Family in Soybean

Transcription data were obtained from NCBI to analyze the tissue expression patterns of GmCCS gene family. Gene expression level was calculated by FPKM values and illustrated using TBtools.

4.5. Plant Materials and Treatments

Soybean cultivar Williams 82 was used to analyze the expression patterns of GmCCS genes under abiotic stresses. Seedlings were grown in pots in the greenhouse with 16 h-light/8 h-dark photoperiod, 28/20 °C day/night temperatures, and 60% relative humidity. Salt/NaHCO₃ stress treatment: soybean seedling was irrigated with 300 mM NaCl and 75 mM NaHCO₃. The leaves and roots of seedlings were sampled at 0 h, 0.5 h, 1 h, 2 h, 4 h, 6 h, 9 h and 12 h time points during treatments. Drought stress treatment: Watering was halted. The moisture of soil was monitored. When the moisture of soil decreased to 30%, the leaves and roots of seedlings were sampled at 0 h, 0.5 h, 1 h, 2 h, 4 h, 6 h, 9 h and 12 h time points during drought treatments.

4.6. RNA Extraction and RNA Extraction and Reverse Transcription-Quantitative PCR (RT-qPCR) Analysis

Total RNA was extracted from soybean leaves and roots of seedlings using the manufacturer's protocol (Qiagen, Hilden, Germany). The isolated RNA was then subjected to reverse transcription by using the SuperScriptTM III Reverse Transcriptase kit. RT-qPCR was performed on each cDNA sample with the SYBR Green Master Mix on an ABI 7500 sequence detection system (Applied Biosystems, Waltham, Massachusetts, MA, USA). The measured Ct values were converted to relative copy-numbers using the $\Delta\Delta C_t$ method. Relative fold enrichment was calculated by normalizing the amount of a target DNA fragment against a reference gene *TUA5*. Three fully independent biological replicates ($n = 3$ plants) were obtained and subjected to RT-qPCR in technical triplicates. Raw data were standardized as described previously [29].

4.7. Stress Tolerance Analysis in Yeast

The coding sequence of GmCCSs was PCR amplified using cDNA from Williams 82 leaves with the primer pair listed in Supplemental Table S3. The PCR products were then inserted in the pYES2 vector. The resulting construct was then introduced into yeast strain INVSc1. Stress tolerance test in yeast was performed according to Gautam [30]. Yeast cells containing the vectors pYES2 (control) or pYES2-GmCCSs were grown in SC-URA (without uracil) medium with shaking (200 rpm) at 28 °C for 24 h, and their OD₆₀₀ was adjusted with SC-URA to 1, 0.1, 0.01, 0.001, respectively. For the salt stress test, one microliter of each diluted culture was spotted onto SC-URA medium supplemented with 0.8, 1.0, 1.1, 1.3 M NaCl; For the osmotic stress test, the yeast cells were cultured in 0.4, 0.8, 1.6, 2.0 M sorbitol. For the alkali stress test, the yeast cells were cultured on SC-URA medium supplemented with 0.2, 0.4, 0.6, 0.8 M NaHCO₃ at 28 °C. Transgenic yeast harboring empty vector pYES2 was used as the control. The experiments were repeated three times.

4.8. Stress Tolerance Analysis in Transgenic Soybean Hairy Roots

The full-length GmCCS7/GmCCS11/GmCCS24 coding sequence was inserted downstream of the CaMV 35S promoter and fused to the GFP, in the pCAMBIA1300 vector respectively. The resulting construct was then introduced into *A. rhizogenes* strain K599. To generate transformed soybean hairy roots, the soybean cultivar Dongnong50 (DN50) was used for *A. rhizogenes*-mediated transformation [31]. Seeds were germinated under a 16 h light/8 h dark photoperiod at 25 °C in a humidity chamber. After a week, healthy

plants were injected with *A. rhizogenes* strain K599 harboring pCAMBIA1300 (CK) or pCAMBIA1300-GmCCSs construct vectors. The infected plants were then transferred to the chamber and kept under high humidity until hairy roots grew out of the infection sites to 2–5 cm in length [32]. The original roots were removed before subjecting to stress treatment, then the seedlings with 2–5 cm hairy roots were transferred into Hoagland solution with 20% (*w/v*) PEG 6000 for 6 h drought stress treatment. Then the plants were transferred to fresh medium without PEG 6000. Survival rates (%) were calculated from the number of surviving plants per total plants tested. Three independent experiments were conducted with individual samples containing 15 plants. Antioxidant enzyme (SOD, CAT, POD) activity measurement was performed as described previously [33].

5. Conclusions

In this study, 31 *GmCCS* genes were identified from the soybean. The *GmCCS* members were divided into four subgroups based on the analysis of the composition, phylogenetic tree, gene structure, motifs, cis-regulatory elements, and collinearity. Members in each subgroup had similar analysis results. Transcriptome data analysis showed that *GmCCS* genes were specifically expressed in different tissues. RT-qPCR showed that five *GmCCS* genes responded to drought stresses at transcript levels. Overexpression of *GmCCS7* and *GmCCS24* increased drought tolerance in soybean hair root and yeast.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms24065154/s1>.

Author Contributions: Conceptualization, S.J., X.B. and H.Z.; methodology, S.J. and J.Z.; software, Y.Z. and F.C.; validation, S.J., R.F. and Y.H.; formal analysis, S.J.; data curation, S.J.; writing—original draft preparation, S.J.; writing—review and editing, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was mainly funded by National Natural Science Foundation of China, grant number 31971830, Strategic Priority Research Program of the Chinese Academy of Sciences, grant number XDA28070302.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the Supplementary Materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Minussi Winck, J.E.; Sarmiento, L.F.V.; Zanon, A.J.; Librelon, S.S.; Garcia, A.; Streck, N.A. Growth and transpiration of soybean genotypes with HaHB4(R) transcription factor for drought tolerance. *Physiol. Plant* **2022**, *174*, e13557. [[CrossRef](#)] [[PubMed](#)]
2. Imran, M.; Latif Khan, A.; Shahzad, R.; Aaqil Khan, M.; Bilal, S.; Khan, A.; Kang, S.M.; Lee, I.J. Exogenous melatonin induces drought stress tolerance by promoting plant growth and antioxidant defence system of soybean plants. *AoB Plants* **2021**, *13*, plab026. [[CrossRef](#)] [[PubMed](#)]
3. Casareno, R.L.; Waggoner, D.; Gitlin, J.D. The copper chaperone CCS directly interacts with copper/zinc superoxide dismutase. *J. Biol. Chem.* **1998**, *273*, 23625–23628. [[CrossRef](#)] [[PubMed](#)]
4. Rosen, D.R.; Siddique, T.; Patterson, D.; Figlewicz, D.A.; Sapp, P.; Hentati, A.; Donaldson, D.; Goto, J.; O'Regan, J.P.; Deng, H.X.; et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **1993**, *362*, 59–62. [[CrossRef](#)]
5. Furukawa, Y.; Tokuda, E. Does wild-type Cu/Zn-superoxide dismutase have pathogenic roles in amyotrophic lateral sclerosis? *Transl. Neurodegener.* **2020**, *9*, 33. [[CrossRef](#)]
6. Kirby, K.; Jensen, L.T.; Binnington, J.; Hilliker, A.J.; Ulloa, J.; Culotta, V.C.; Phillips, J.P. Instability of Superoxide Dismutase 1 of *Drosophila* in Mutants Deficient for Its Cognate Copper Chaperone. *J. Biol. Chem.* **2008**, *283*, 35393–35401. [[CrossRef](#)]
7. Sagasti, S.; Bernal, M.A.; Sancho, D.; Del Castillo, M.B.; Picorel, R. Regulation of the chloroplastic copper chaperone (CCS) and cuprozinc superoxide dismutase (CSD2) by alternative splicing and copper excess in *Glycine max*. *Funct. Plant Biol.* **2014**, *41*, 144–155. [[CrossRef](#)]

8. Fukuoka, M.; Tokuda, E.; Nakagome, K.; Wu, Z.; Nagano, I.; Furukawa, Y. An essential role of N-terminal domain of copper chaperone in the enzymatic activation of Cu/Zn-superoxide dismutase. *J. Inorg. Biochem.* **2017**, *175*, 208–216. [[CrossRef](#)]
9. Grasso, M.; Bond, G.J.; Kim, Y.J.; Boyd, S.; Matson Dzebo, M.; Valenzuela, S.; Tsang, T.; Schibrowsky, N.A.; Alwan, K.B.; Blackburn, N.J.; et al. The copper chaperone CCS facilitates copper binding to MEK1/2 to promote kinase activation. *J. Biol. Chem.* **2021**, *297*, 101314. [[CrossRef](#)]
10. Benson, D.M., Jr.; Yu, J.; Becknell, B.; Wei, M.; Freud, A.G.; Ferketich, A.K.; Trotta, R.; Perrotti, D.; Briesewitz, R.; Caligiuri, M.A. Stem cell factor and interleukin-2/15 combine to enhance MAPK-mediated proliferation of human natural killer cells. *Blood* **2009**, *113*, 2706–2714. [[CrossRef](#)]
11. Boyd, S.D.; Calvo, J.S.; Liu, L.; Ullrich, M.S.; Skopp, A.; Meloni, G.; Winkler, D.D. The yeast copper chaperone for copper-zinc superoxide dismutase (CCS1) is a multifunctional chaperone promoting all levels of SOD1 maturation. *J. Biol. Chem.* **2019**, *294*, 1956–1966. [[CrossRef](#)] [[PubMed](#)]
12. Maidment, J.H.R.; Franceschetti, M.; Maqbool, A.; Saitoh, H.; Jantasuriyarat, C.; Kamoun, S.; Terauchi, R.; Banfield, M.J. Multiple variants of the fungal effector AVR-Pik bind the HMA domain of the rice protein OsHIPP19, providing a foundation to engineer plant defense. *J. Biol. Chem.* **2021**, *296*, 100371. [[CrossRef](#)]
13. Cheng, Y.; Cheng, L.; Hu, G.; Guo, X.; Lan, Y. Auxin and CmAP1 regulate the reproductive development of axillary buds in Chinese chestnut (*Castanea mollissima*). *Plant Cell Rep.* **2022**, 1–10. [[CrossRef](#)] [[PubMed](#)]
14. Yamaguchi-Shinozaki, K.; Shinozaki, K. Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends Plant Sci.* **2005**, *10*, 88–94. [[CrossRef](#)] [[PubMed](#)]
15. Soma, F.; Takahashi, F.; Yamaguchi-Shinozaki, K.; Shinozaki, K. Cellular Phosphorylation Signaling and Gene Expression in Drought Stress Responses: ABA-Dependent and ABA-Independent Regulatory Systems. *Plants* **2021**, *10*, 756. [[CrossRef](#)]
16. Wang, Y.; Zhang, J.; Hu, Z.; Guo, X.; Tian, S.; Chen, G. Genome-Wide Analysis of the MADS-Box Transcription Factor Family in *Solanum lycopersicum*. *Int. J. Mol. Sci.* **2019**, *20*, 2961. [[CrossRef](#)]
17. Li, H.; Li, Y.; Zhang, X.; Cai, K.; Li, Y.; Wang, Q.; Qu, G.; Han, R.; Zhao, X. Genome-wide identification and expression analysis of the MADS-box gene family during female and male flower development in *Juglans mandshurica*. *Front. Plant Sci.* **2022**, *13*, 1020706. [[CrossRef](#)]
18. Zhao, J.; Gong, P.; Liu, H.; Zhang, M.; He, C. Multiple and integrated functions of floral C-class MADS-box genes in flower and fruit development of *Physalis floridana*. *Plant Mol. Biol.* **2021**, *107*, 101–116. [[CrossRef](#)]
19. Ruta, V.; Longo, C.; Lepri, A.; De Angelis, V.; Occhigrossi, S.; Costantino, P.; Vittorioso, P. The DOF Transcription Factors in Seed and Seedling Development. *Plants* **2020**, *9*, 218. [[CrossRef](#)]
20. Cheng, Z.; Hou, D.; Liu, J.; Li, X.; Xie, L.; Ma, Y.; Gao, J. Characterization of moso bamboo (*Phyllostachys edulis*) Dof transcription factors in floral development and abiotic stress responses. *Genome* **2018**, *61*, 151–156. [[CrossRef](#)]
21. Li, L.; Shi, Q.; Li, Z.; Gao, J. Genome-wide identification and functional characterization of the PheE2F/DP gene family in Moso bamboo. *BMC Plant Biol.* **2021**, *21*, 158. [[CrossRef](#)] [[PubMed](#)]
22. Wang, H.; Zhang, Y.; Norris, A.; Jiang, C.Z. S1-bZIP Transcription Factors Play Important Roles in the Regulation of Fruit Quality and Stress Response. *Front. Plant Sci.* **2021**, *12*, 802802. [[CrossRef](#)] [[PubMed](#)]
23. Zhao, P.; Hou, S.; Guo, X.; Jia, J.; Yang, W.; Liu, Z.; Chen, S.; Li, X.; Qi, D.; Liu, G.; et al. A MYB-related transcription factor from sheepgrass, LcMYB2, promotes seed germination and root growth under drought stress. *BMC Plant Biol.* **2019**, *19*, 564. [[CrossRef](#)] [[PubMed](#)]
24. Himelblau, E.; Mira, H.; Lin, S.J.; Culotta, V.C.; Penarrubia, L.; Amasino, R.M. Identification of a functional homolog of the yeast copper homeostasis gene ATX1 from Arabidopsis. *Plant Physiol.* **1998**, *117*, 1227–1234. [[CrossRef](#)]
25. Wong, P.C.; Waggoner, D.; Subramaniam, J.R. Copper chaperone for superoxide dismutase is essential to activate mammalian Cu/Zn superoxide dismutase. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 2886–2891. [[CrossRef](#)]
26. Chu, C.C.; Lee, W.C.; Guo, W.Y.; Pan, S.M.; Chen, L.J.; Li, H.M.; Jinn, T.L. A copper chaperone for superoxide dismutase that confers three types of copper/zinc superoxide dismutase activity in Arabidopsis. *Plant Physiol.* **2005**, *139*, 425–436. [[CrossRef](#)]
27. Feng, X.; Chen, F.; Liu, W.; Thu, M.K.; Zhang, Z.; Chen, Y.; Cheng, C.; Lin, Y.; Wang, T.; Lai, Z. Molecular Characterization of MaCCS, a Novel Copper Chaperone Gene Involved in Abiotic and Hormonal Stress Responses in *Musa acuminata* cv. Tianbaojiao. *Int. J. Mol. Sci.* **2016**, *17*, 441. [[CrossRef](#)]
28. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)]
29. Willems, E.; Leyns, L.; Vandesompele, J. Standardization of real-time PCR gene expression data from independent biological replicates. *Anal. Biochem.* **2008**, *379*, 127–129. [[CrossRef](#)]
30. Gautam, R.; Ahmed, I.; Shukla, P.; Meena, R.K.; Kirti, P.B. Genome-wide characterization of ALDH Superfamily in Brassica rapa and enhancement of stress tolerance in heterologous hosts by BrALDH7B2 expression. *Sci. Rep.* **2019**, *9*, 7012. [[CrossRef](#)]
31. Kereszt, A.; Li, D.; Indrasumunar, A.; Nguyen, C.D.; Nontachaiyapoom, S.; Kinkema, M.; Gresshoff, P.M. Agrobacterium rhizogenes-mediated transformation of soybean to study root biology. *Nat. Protoc.* **2007**, *2*, 948–952. [[CrossRef](#)] [[PubMed](#)]

32. Shi, W.Y.; Du, Y.T.; Ma, J.; Min, D.H.; Jin, L.G.; Chen, J.; Chen, M.; Zhou, Y.B.; Ma, Y.Z.; Xu, Z.S.; et al. The WRKY Transcription Factor GmWRKY12 Confers Drought and Salt Tolerance in Soybean. *Int. J. Mol. Sci.* **2018**, *19*, 4087. [[CrossRef](#)] [[PubMed](#)]
33. Ning, P.; Liu, C.; Kang, J.; Lv, J. Genome-wide analysis of WRKY transcription factors in wheat (*Triticum aestivum* L.) and differential expression under water deficit condition. *PeerJ* **2017**, *5*, e3232. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.