Genome-wide analysis of maize OSCA family members and their involvement in drought stress

Shuangcheng Ding, Equal first author, Xin Feng, Equal first author, Hewei Du, Hongwei Wang, Corresponding Author

1 Agricultural College, Yangtze University, Jingzhou, China
2 Hubei Collaborative Innovation Center for Grain Industry, Yangtze University, Jingzhou, China
3 College of Life Science, Yangtze University, Jingzhou, China

Background. Worldwide cultivation of maize is often impacted negatively by drought stress. Hyperosmolality-gated calcium-permeable channels (OSCA) have been characterized as osmosensors in Arabidopsis. However, the involvement of members of the maize OSCA (ZmOSCA) gene family in response to drought stress is unknown. It is furthermore unclear which ZmOSCA gene plays a major role in genetic improvement of drought tolerance in Maize.

Methods. We predicted the protein domain structure and transmembrane regions by using the NCBI CDD database and TMHMM server separately. The phylogeny tree was built by Mega7. We used the mixed linear model in TASSEL to perform the family-based association analysis.

Results. In this report, 12 ZmOSCA genes were uncovered in the maize genome by a genome-wide survey and analyzed systematically to reveal their synteny and phylogenetic relationship with the genomes of rice, maize and sorghum. These analyses indicated a relatively conserved evolutionary history of the ZmOSCA gene family. Protein domain and transmembrane analysis indicated that most of the 12 ZmOSCAs shared similar structures with their homologs. The result of differential expression analysis under drought at various stages, as well as the expression profiles in 15 tissues, revealed a functional divergence of ZmOSCA genes. Notably, the expression level of ZmOSCA4.1 being up-regulated in both seedlings and adult leaves. Notably, the association analysis between genetic variations in these genes and drought tolerance was detected. Significant associations between genetic variation in ZmOSCA4.1 and drought tolerance were found at the seedling stage. Our report provides a detailed analysis of the ZmOSCAs in the maize genome. These findings will contribute to future studies on the functional characterization of ZmOSCA proteins in response to water deficit stress, as well as understanding the mechanism of genetic variation in drought tolerance in maize.
Genome-wide analysis of maize OSCA family members and their involvement in drought stress

Shuangcheng Ding\textsuperscript{1,2,4}, Xin Feng\textsuperscript{1,4}, Hewei Du\textsuperscript{3}, Hongwei Wang\textsuperscript{1,2}

1 Agricultural College, Yangtze University, Jingzhou, China
2 Hubei Collaborative Innovation Center for Grain Industry, Yangtze University, Jingzhou, China
3 College of Life Science, Yangtze University, Jingzhou, China
4 These authors contributed equally to this work

Corresponding Author:
Hongwei Wang
Nanhuan road 1, Jingzhou district, Jingzhou, Hubei, China
Email address: wanghw@yangtzeu.edu.cn
Abstract

Background. Worldwide cultivation of maize is often impacted negatively by drought stress. Hyperosmolality-gated calcium-permeable channels (OSCA) have been characterized as osmosensors in Arabidopsis. However, the involvement of members of the maize OSCA (ZmOSCA) gene family in response to drought stress is unknown. It is furthermore unclear which ZmOSCA gene plays a major role in genetic improvement of drought tolerance in Maize.

Methods. We predicted the protein domain structure and transmembrane regions by using the NCBI CDD database and TMHMM server separately. The phylogeny tree was built by Mega7. We used the mixed linear model in TASSEL to perform the family-based association analysis.

Results. In this report, 12 ZmOSCA genes were uncovered in the maize genome by a genome-wide survey and analyzed systematically to reveal their synteny and phylogenetic relationship with the genomes of rice, maize and sorghum. These analyses indicated a relatively conserved evolutionary history of the ZmOSCA gene family. Protein domain and transmembrane analysis indicated that most of the 12 ZmOSCAs shared similar structures with their homologs. The result of differential expression analysis under drought at various stages, as well as the expression profiles in 15 tissues, revealed a functional divergence of ZmOSCA genes. Notably, the expression level of ZmOSCA4.1 being up-regulated in both seedlings and adult leaves. Notably, the association analysis between genetic variations in these genes and drought tolerance was detected. Significant associations between genetic variation in ZmOSCA4.1 and drought tolerance were found at the seedling stage. Our report provides a detailed analysis of the ZmOSCAs in the maize genome. These findings will contribute to future studies on the functional characterization of ZmOSCA proteins in response to water deficit stress, as well as understanding the mechanism of genetic variation in drought tolerance in maize.
INTRODUCTION

Drought stress is a key limiting environmental factor which can eventually cause food and societal problems, and therefore numerous studies have been devoted to unraveling the mechanisms of drought resistance in plants. Studies over the past decades have found that plant responses to drought stress mainly include the perception, regulation, and transmission of signals through various pathways, as well as the regulation of stress-responsive gene expression (Ingram and Bartels 1996; Zhu 2002; Bartels and Sunkar 2005). As a consequence, physiological and morphological modifications result to address the stress. During the signal perception, especially during osmotic changes, calcium concentrations have been found to be elevated (Knight et al. 1997; McAinsh and Pittman 2009). Importantly, the increase in the concentration of Ca\(^{2+}\) occurs within five seconds, which might be the earliest detectable event in plants under stress (Knight et al. 1997). During the last three decades, studies have shown that Ca\(^{2+}\) is an important second messenger in the signal transduction process when plants respond to biotic and abiotic stress (Reddy 2001; Hepler 2005; Defalco et al. 2009). Ca\(^{2+}\)-permeable channels have been proposed to function as osmosensors in bacteria and animals stimulated by osmotic/mechanical stress (Booth et al. 2007; Arnadottir and Chalfie 2010). Therefore, it can be inferred that Ca\(^{2+}\) is crucial for both sensing external osmotic stress and activating many signal transduction pathways.

Studies were carried out to find genes regulating early response to stress, especially the concentrations of calcium. When investigating the gene expression profile in Arabidopsis subjected to drought conditions, 16 early responses to dehydration (ERD) genes were reported to be identified after one hour of drought stress (Kiyosue et al. 1994). Amongst the 16 ERDs, ERD4, harboring a highly conserved DUF221 domain (Pfam accession: 02714), was conserved between species (Camargo et al. 2007; Liu et al. 2009; Rai et al. 2012). Recently, using a gene screening strategy based on calcium imaging, an Arabidopsis hyperosmolality-gated calcium-permeable channel (OSCA) mutant was identified; osca1 showed a low hyperosmolality-induced Ca\(^{2+}\) increase (OICI), and OSCA1 was further characterized to be an osmosensor in Arabidopsis.
Interestingly, *OSCA1* also contains the DUF221 domain. The results disclose that DUF221 may participate in osmotic adjustment.

Association analysis, based on linkage disequilibrium (LD) and excavation of genetic variation, have been utilized as a novel strategy for dissecting the genetic basis of complex traits in crops (Yu and Buckler 2006). Drought tolerance is a complex and intrinsic trait, and the identification of favorable alleles for drought tolerance in maize through association studies, including candidate gene association analysis, is an ever-growing research area. Due to the rapid decrease in LD, association analysis in the maize genome could screen genetic variations at a single gene level. Although several genetic variations were found to enhance drought resistance (Liu et al. 2013; Mao et al. 2015; Wang et al. 2016; Xiang et al. 2017), allelic variations underlying drought tolerance still need to be identified to facilitate molecular breeding.

It has been shown that *OSCA1* acts as an osmosensor in *Arabidopsis* (Yuan et al. 2014), and the role of the *OSCA* gene family is essential for plants to respond to stress. Recently, genome-wide analysis of genes containing the DUF221 domain were performed in rice to understand its possible roles (Li et al. 2015; Ganie et al. 2017). Therefore, in this study, we characterized the *OSCA* family members in the maize genome and analyzed the phylogenetic and syntenic relationship amongst these *OSCA*s. In addition, we studied the expression profiles of *OSCA*s under drought stress and performed a family-based genome-wide association study. These results could be applied for further functional research of *ZmOSCA*s and raise our understanding of the roles of plant *OSCA*s in drought stress.

**MATERIALS & METHODS**

**Identification of OSCA protein-coding genes in the maize genome**

Conserved OSCA domain DUF221 (Pfam accession: 02714) from the Pfam database (Finn et al. 2014) was used to build the Hidden Markov Model-based searches (HMMER:...
http://hmmer.janelia.org/) and scanned against the maize genome (genome assembly: AGPv3) and the sorghum genome (genome assembly: V3). All the retrieved sequences were curated using the NCBI Conserved Domain Database (CDD) (www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchlerbauer et al. 2011) to determine whether the proteins harbor the DUF221 domain or not.

Genic structure and phylogenetic relationships analysis

The protein sequences of the identified maize and sorghum OSCAs were downloaded from Phytozome v10.0. The protein sequence of previously identified OSCA protein-coding genes in the Arabidopsis and rice genome (Yuan et al. 2014; Li et al. 2015) were also downloaded from Phytozome v10.0. To show the exon/intron structure, the coding sequence of each OSCA gene were aligned to its corresponding genomic sequence and then a schematic representation was generated using GSDS 2.0 (http://gsds.cbi.pku.edu.cn) (Hu et al. 2015). To construct a phylogenetic tree of the identified OSCA proteins in rice, Arabidopsis, and sorghum, and maize genome, ClustalW (Thompson et al. 1994) was used to align multiple protein sequences of OSCA. The phylogenetic tree was constructed using this alignment output based on a neighbor-joining method in MEGA7 (Kumar et al. 2016) using the following parameters: poisson correction, pairwise deletion, uniform rates and bootstrap (1000 replicates).

Prediction of transmembrane region

TMHMM Server V2 was used to predict the transmembrane region (TMs) of ZmOSCA s, and the prediction of TMs was manually edited according to the TMs of OSCA1.1, OSCA1.2 and OsOSCA1.2 (Jojoa-Cruz et al. 2018; Liu et al. 2018; Maity et al. 2018; Murthy et al. 2018; Zhang et al. 2018).

Gene duplication analysis

Chromosomal locations of ZmOSCA s were collected from the maize genome (assemble version:
AGPv3) in Phytozome v10.0. The duplicated segmental blocks and gene pairs were analyzed on the Plant Genome Duplication Database (available online: http://chibba.agtec.uga.edu) with a display distance of 100 kilobases.

Expression profile analysis

Seedlings growth conditions and drought treatments of B73 (maize inbred line) were conducted according to Wang (Wang et al. 2016). Hydraulically cultured 3-leaf stage seedlings were put on a plate and subjected to dehydration (40-60% relative humidity and 28°C). As for the effect of drought implied on adult leaf, drought stress was applied by withholding water after the eight-leaf stage (V8), with well-watered plants (soil water content 40%) as control. The middle section of flag leaf which came from three replicates, were collected at the twelve-leaf stage (V12), the fourteen-leaf stage (V14), the sixteen-leaf stage (V16) and the silking stage (R1) for both drought stressed and well-watered plants as control. Leaf samples from at least 3 replicates were frozen by liquid nitrogen and then stored at -80°C before RNA isolation.

Raw RNA was extracted from leaf samples using TRI Reagent (Invitrogen, USA) according to the product manuals. The relative expression of ZmOSCAs was quantified using quantitative real time-PCR (qRT-PCR). qRT-PCR was tested in 96-well plates using an ABI7500 Real-Time PCR Systems (Applied Biosystems, USA). The PCR reaction system consists of 1 μl cDNA, 200 nM primers, and 5 μl SYBR Premix Ex Taq II (Takara, China), and the reaction volume was 10μl. The PCR reaction was performed with the following conditions: 10 min at 94°C, 40 cycles of 15 s at 94°C, and 30 s at 60°C. The internal control was the expression of ZmUbi-2 (UniProtKB/TrEMBL; ACC: Q42415). The quantification method used was 2^-ΔCT, and the variation in the expression was derived from three biological replicates.

Association Analysis

Association analysis for ZmOSCAs was conducted by using a mapping population containing
367 maize inbred lines and corresponding drought resistance phenotypes reported previously (Wang et al. 2016). The mapping population contains 556k single nucleotide polymorphism (SNP) markers, and the minor allele frequency (MAF) of each marker was greater than or equal to 0.05. All identified ZmOSCAs harbored 168 SNPs in the coding region and both the 5'- and 3’-untranslated region (UTR). Three statistical models including the general linear model (GLM) model, adjusting the first two principal components (PC<sub>2</sub>), and the mixed linear model (MLM) model (incorporating PC<sub>2</sub> and a Kinship matrix) were selected to identify the SNPs significantly associated with drought resistance by using the TASSEL4.0 program (Yu et al. 2006; Bradbury et al. 2007).

RESULTS

Identification of the OSCA family members in maize

We carried out a systematic genome-wide screen of putative OSCA genes in maize. Initially, a Hidden Markov Model (HMM) search was performed against the maize genome (genome version: AGPv3.0) utilizing the DUF221 domain (Pfam accession: 02714). Ultimately, 12 genes were found to be ZmOSCAs in the maize genome and named according to the Arabidopsis orthologs (Table 1). The physical location of each ZmOSCA in the genome was identified according to physical coordination provided by MaizeGDB (https://www.maizegdb.org). Twelve ZmOSCA genes were distributed unevenly on all the ten chromosomes except chromosomes 2, 4, 7 and 10, without any clustering. Chromosomes 1 and 3 possessed as many as three ZmOSCA genes (the largest number of OSCA genes on a chromosome), and chromosomes 5 and 8 equally contained two genes. In contrast, chromosomes 6 and 9 each only harbored one ZmOSCA. The exon number of ZmOSCAs varied from 1 to 11. Approximately 75% (8/12) of the ZmOSCA genes contained more than eight exons, and only 17% (2/12) of the genes had less than two exons.

Phylogenetic relationship analysis of OSCA genes
In order to elucidate the phylogenetic relationships among OSCAs, a neighbor-joining tree of ZmOSCA proteins and the corresponding orthologs from rice, sorghum, and *Arabidopsis* was built, and the tree was based on the alignment of full-length OSCA proteins (Fig. 1; Table S1). As shown on the phylogenetic tree, 49 OSCA genes can be classified into four main classes, clades 1, 2, 3 and 4 (Fig. 1). All *Arabidopsis* and rice OSCAs fell in the same class or clade as previously reported, which is in agreement with previous work (Yuan et al. 2014). Interestingly, proteins derived from dicotyledonous *Arabidopsis* clustered separately from those of three monocot plants. Also, we found that some proteins from rice, sorghum and maize displayed pairwise correspondence (blue box in Fig 1), not only indicating that these genes are phylogenetically conserved among these species, but demonstrating that maize and sorghum possessed a closer phylogenetic relationship compared to maize and rice, conforming to the perspective that sorghum is a closer relative to maize than rice.

Throughout the phylogenetic tree, clades 1, 2, 3 and 4 contain 21, 20, four and four OSCA proteins, respectively (Fig. 1). Amongst the 49 OSCAs, there were 5, 5, 1, and 1 ZmOSCAs in clades 1, 2, 3 and 4 respectively. However, for the rice and sorghum genomes, there were 4, 5, 1 and 1 OSCAs in clades 1, 2, 3 and 4 respectively (Table 2). Obviously, compared with rice and sorghum, maize possessed one more OSCA in clade 1 (Fig 1). Meanwhile, we detected the nodes that lead to maize-, sorghum- and rice-specific clades (red circles in Fig. 1). These nodes denote the divergence point between maize, sorghum, and rice, and therefore reveal the most recent common ancestral (MRCA) genes before the split. After MRCA analysis, there were 11 clades (indicated by red circles in Fig. 1), and nearly all the clades were constituted by one gene from the rice, maize, and sorghum genomes. However, there was just one clade that contained two genes from the maize genome (indicated by the red arrow in Fig. 1). The results showed that between the maize, sorghum, and rice genomes, the OSCA gene family shared a similar evolutionary history, whilst the maize genome possessed a duplication event leading to a gain in one gene (indicated by red arrow in Fig. 1) after the maize, sorghum and rice genomes split.
Synteny of OSCA4s in maize, sorghum and rice genomes

Gene collinearity comparative genomics indicates homologous gene function and phylogenetic relationships amongst several species and even the genome organization of extinct ancestral species. Thus, we investigated the colinearity of OSCA genes in rice, sorghum, and maize genomes. Initially, gene colinearity data were retrieved from the Plant Genome Duplication Database using ZmOSCA genes as anchors, then the chromosomal segment containing multiple homologous genes across species was defined as a genomic syntenic block. After this analysis, 10 ZmOSCAs were found to have collinearity, and the syntenic members or collinear genes in rice and sorghum were shown in Fig. 2. As a result, eight chromosomal segments containing OSCA were identified as being evolutionally conserved between rice, maize, and sorghum; these include ZmOSCA1.2, ZmOSCA1.3, ZmOSCA1.4, ZmOSCA2.2, ZmOSCA2.3, ZmOSCA2.4, ZmOSCA2.5, and ZmOSCA4.1 (Fig. 2, Table S2). In addition, genes within the eight syntenic block amongst the maize, sorghum, and rice genomes were identical to the phylogenetic analysis (Fig. 1), confirming the accuracy of our analysis. Syntenic blocks of ZmOSCA1.1b and ZmOSCA2.1b were observed between maize and sorghum genomes respectively, but no syntenic blocks of these genes were found between maize and rice genomes. Meanwhile, no collinear segments of ZmOSCA1.1a and ZmOSCA3.1 were found in the genome of maize, rice, and sorghum. The results suggest that most of the ZmOSCAs existed before the divergence of species, but some ZmOSCAs may have originated from duplication of the maize genome.

The conserved domain of ZmOSCAs.

CDD was used to identify protein domains contained in ZmOSCAs. It was found that most ZmOSCAs harbored three domains except that ZmOSCA1.1a, ZmOSCA3.1, and ZmOSCA4.1 contained two domains (Fig. 3). It is noteworthy that DUF221 is located at the C-terminal of all ZmOSCAs, indicating that this domain is not only necessary but also relatively conservative compared with other domains in the gene family. It was reported that there were 11 TMs in the OSCA protein sequences (Jojoa-Cruz et al. 2018; Liu et al. 2018; Maity et al. 2018; Murthy et al.
We found that half of the 12 ZmOSCAs had exactly 11 transmembrane regions (Fig. 3). A multiple sequence alignment was used to show the presence of TMs in DUF221 of ZmOSCAs (Fig. 4). It was found that protein domain DUF221 contained a different number of TMs, while protein domain pfam13967 contained a fixed number of TMs. However, no TMs were detected in the protein domain pfam14703 in ZmOSCAs. Different ZmOSCA members contained 2-7 TMs in the DUF221 region. Most of the ZmOSCAs contained at least eight TMs with two exceptions. ZmOSCA1.1a and ZmOSCA3.1 had fewer TMs than others, suggesting that a deletion event occurred for ZmOSCA1.1a and ZmOSCA3.1.

Expression profiles of ZmOSCA genes under drought stress

We studied the expression profiles of the 12 ZmOSCA genes in 15 different tissues using the reported transcriptomic data (Sekhon et al. 2011). As shown in Fig. 5A, expression patterns of ZmOSCA genes varied greatly. Expression values of ZmOSCA1.1a, ZmOSCA1.1b, ZmOSCA1.2, and ZmOSCA4.1 were relatively higher than the rest among the tissues tested. Previous studies have shown that OSCAs had key roles in different aspects of plant development, especially in stress responses (Li et al. 2015). In order to obtain further insight into the roles of ZmOSCAs in drought tolerance, their expression profiles were monitored by quantitative real-time PCR analyses using three-week-old leaves of maize seedlings treated with drought stress for five hours or 24 hours. As shown in Fig. 5B, in the genotype of B73, more than half of expression level of the ZmOSCAs were regulated by drought stress, including six ZmOSCAs that were up-regulated significantly and one down-regulated ZmOSCA; the others were hardly responsive to drought stress. In addition, genes showing up- or down-regulation were found unfavorably in each of the four major clades. The results showed that proteins from different clades might have overlapping and/or antagonistic functions in the regulation of drought responses in plants. Amongst the six up-regulated ZmOSCA genes, in comparison to normal growth conditions, the expression of ZmOSCA2.4 was induced more than six-fold, that being the largest folds change in relative expression levels (Fig. 5B). Five ZmOSCAs (ZmOSCA1.1b, ZmOSCA1.2, ZmOSCA1.4,
ZmOSCA1.1, and ZMOSCA4.1) were up-regulated over two-fold in response to drought (Fig. 5B).

Additionally, a slightly upregulated expression was observed in some ZmOSCA s following drought treatment for five hours or 24 hours; these include ZmOSCA1.1a, ZmOSCA2.2, ZmOSCA2.4, ZmOSCA2.5, and ZmOSCA3.1. In contrast to the other ZmOSCA s, the expression level of ZmOSCA2.3 decreased about three-fold after drought treatment (Fig. 5B). Relative expression of ZmOSCA s were also detected in adult leaves at four growth stages, V12, V14, V16, and R1. Drought stress for the adult leaves was applied by withholding water at the V8 stage, while the corresponding control was well-watered (soil water content 40%) plants. As shown in Fig. 6, five ZmOSCA s were up-regulated at least one stage, and one ZmOSCA was down-regulated at two stages under drought. Across the four stages, the relative expression levels of ZmOSCA2.1, ZmOSCA2.2, and ZmOSCA4.1 were upregulated consistently. Taken together, the results clearly show the functional divergence of ZmOSCA s responding to drought stress in maize seedlings and adult leaves. Collectively, the data demonstrated that different ZmOSCA s showed variable expression patterns under drought stress, while the expression value of ZmOSCA4.1 was highly expressed in 15 tissues and up-regulated in both seedling and adult leaf by drought stress.

Association analysis of genetic variations in ZmOSCA s with drought tolerance

In order to further inquire whether the genetic variation in ZmOSCA s was associated with drought tolerance, a family-based association analysis was performed for these genes. At the seedling stage, the drought tolerance of the mapping population was assessed by evaluating seedling survival rate under severe drought stress. Based on previously reported methods and data (Liu et al. 2013; Wang et al. 2016), genetic polymorphism was characterized as the presence and number of SNP markers in each of the 12 ZmOSCA s. All the 12 identified ZmOSCA s were identified as polymorphic, with 14 SNPs on average for each identified gene (Table 3). As a consequence, ZmOSCA4.1 was identified as the second most polymorphic gene, with 34 SNPs in this mapping population. Subsequently, three statistical models were used to find out significant
associations between genotype and phenotype. In brief, a general linear model (GLM), with the first two principal components (PC2), and mixed linear model (MLM) incorporating both PC2 and a kinship matrix correcting the effect of cryptic relatedness were used in the associations. GLM method was used to conduct the single-marker analysis. PC2, via the first two principal components on SNP data, was utilized to adjust spurious associations resulted from population structure. MLM was regarded as an effective method for controlling false positives in association analysis (Yu and Buckler 2006; Yu et al. 2006). Subsequently, candidate gene association analysis identified significant associations between genetic variation of ZmOSCA4.1 with drought tolerance under different models with a P-value ≤ 0.01 (Table 3). The results appeared to show that ZmOSCA4.1 might be an important candidate gene for maize drought tolerance.

DISCUSSION

Previous study has indicated that OSCA1, containing the domain DUF221, is an osmosensor in Arabidopsis (Yuan et al. 2014). Subsequently, studies have been done to identify OSCA family members in Arabidopsis and rice, which lead to the identification of 15 AtOSCA s and 11 OsOSCA s respectively (Yuan et al. 2014; Li et al. 2015). However, no detailed analysis has been performed for the OSCA gene family in maize, especially its expression profile under drought stress. Moreover, it is not yet known which ZmOSCA is directly associated with the diversity of drought tolerance in maize. Therefore, we carried out this study to address these questions. Our findings will provide fundamental factors that may be used to facilitate the genetic enhancement of drought tolerance in maize, as well as increase our understanding of the role of this gene family under drought.

In our study, we identified 12 ZmOSCA s in maize, and then we systematically analyzed their phylogenetic and syntenic relationship with rice, maize and sorghum genomes (Fig. 1; Fig. 2). After MRCA analysis, we found that the Arabidopsis genome possessed the largest number of OSCAs, whilst the maize genome possessed three gene losses compared to Arabidopsis and one
gene gain compared to the rice genome. The number of OSCAs within the maize, sorghum, and rice genomes did not vary much, indicating that the majority of OSCAs in maize, sorghum, and rice genomes undergo relatively conserved evolutionary history after their divergence. As shown in Fig. 1, 12 maize genes were clustered into four clades as reported previously (Yuan et al. 2014). Gene numbers in clades 2, 3 and 4 were identical between maize, sorghum and rice genomes. As shown in Fig. 1, the maize genome possessed one more OSCA, namely ZmOSCA1.1a, having the fewest number of exons in clade 1. But, in the synten analysis, we did not find the synten blocks related to ZmOSCA1.1a, suggesting that ZmOSCA1.1a may derive from duplication after the divergence of the maize, sorghum and rice genomes and possess functional differentiation from other ZmOSCAs. Meanwhile, we found that, despite ZmOSCA3.1, OsOSCA3.1 and Sobic.002G034500 were clustered together (Fig. 1), ZmOSCA3.1 was absent from the synten analysis (Fig. 2). Interestingly, ZmOSCA3.1 was constituted by two exons. The results suggest that ZmOSCA3.1 may exit before the maize, sorghum and rice genomes split, but evolved in a different direction from OsOSCA3.1 and Sobic.002G034500. Previous studies showed that there were 11 TMs in the OSCA protein sequences (Jojoa-Cruz et al. 2018; Liu et al. 2018; Maity et al. 2018; Murthy et al. 2018; Zhang et al. 2018), and each OSCA protein in the Arabidopsis genome contained 11 TMs (Jojoa-Cruz et al. 2018). However, only half of the 12 ZmOSCAs had exactly 11 transmembrane regions, demonstrating that ZmOSCAs might possess a greater genetic variation during the evolution. Protein domain and transmembrane regions analysis also revealed that ZmOSCA1.1a and ZmOSCA3.1 contained fewer domains and TMs compared with others (Fig. 3; Fig. 4). Obviously, deletions in ZmOSCA1.1a and ZmOSCA3.1 occurred independently during maize evolution. However, both ZmOSCA1.1a and ZmOSCA3.1 did not respond to drought treatment, suggesting that they may not have a function in drought resistance.

To our knowledge, although the relationship between OSCA proteins and stresses has been reported (Zhao et al. 2015), the dynamic drought-responsive expression patterns of ZmOSCAs were still obscure. Expression pattern analysis of ZmOSCAs helped us to understand their
possible functions and offer a thorough foundation for future functional studies. Generally, $ZmOSCA$s exhibited differential expression under drought stress in maize seedlings and adult leaves, not only amongst sub-groups but also amongst members within the same sub-groups, suggesting that these $ZmOSCA$s may have diverse functions. Our results showed that the relative expression levels of six $ZmOSCA$s were significantly up-regulated and expression of one $ZmOSCA$ was clearly down-regulated (Fig 5), indicating that these genes might serve as key mediators of drought stress responses. In other research, the relative expression of $OsOSCA1.1$, $OsOSCA1.2$, $OsOSCA2.1$, $OsOSCA2.4$, $OsOSCA2.5$, and $OsOSCA4.1$ were up-regulated by PEG treatment (Li et al. 2015). Interestingly, in this study, we found that six $ZmOSCA$s, including $ZmOSCA1.1b$, $ZmOSCA1.2$, $ZmOSCA1.4$, $ZmOSCA2.1$, $ZmOSCA2.4$, and $ZmOSCA4.1$, could be up-regulated by drought stress (Fig 5). When we analyzed the expression pattern of $ZmOSCA$s in adult leaves at four stages, $ZmOSCA2.1$, $ZmOSCA2.2$ and $ZmOSCA4.1$ were found to be up-regulated consistently (Fig. 6) at both seedlings and adult leaf. Notably, the relative expression level of the rice orthologs of $ZmOSCA2.1$, $ZmOSCA2.2$ and $ZmOSCA4.1$ were also up-regulated by drought stress (Li et al. 2015), demonstrating that these genes may share a conserved function.

To date, no study is known of which detects the association between genetic variations in $ZmOSCA$s and drought tolerance. To answer this question, the genetic polymorphism amongst all of the 12 $ZmOSCA$s were analyzed. Under the three models, $ZmOSCA4.1$ was consistently the most significantly associated with drought tolerance ($P$-value ≤ 0.01, Table 3). Furthermore, in response to drought stress, the expression of $ZmOSCA4.1$ was induced, indicating that this gene might be engaged in resistance to drought stress. Further detailed analysis of biochemical and molecular functions of $ZmOSCA4.1$ will contribute to our understanding of $OSCA$ function.

**CONCLUSIONS**

In this report, 12 $ZmOSCA$ genes were uncovered in the maize. Synteny and phylogenetic
relationship analyses were performed, and it was found that the ZmOSCA gene family shared a conserved evolutionary history. Protein domain and transmembrane analysis indicated that most of the 12 ZmOSCAs shared similar structures with their homologs. Differential expression analysis under drought at various stages, as well as the expression profiles in 15 tissues, revealed a functional divergence of ZmOSCA genes. Notably, the expression level of ZmOSCA4.1 being up-regulated in both seedlings and adult leaves. Importantly, significant associations between genetic variation in ZmOSCA4.1 and drought tolerance were found at the seedling stage. Our research will enhance our understanding of the role of ZmOSCAs under drought.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Du for his critical reading the manuscript.

REFERENCES

Arnadottir J, Chalfie M. 2010. Eukaryotic mechanosensitive channels. Annual review of biophysics 39:111-137 DOI 10.1146/annurev.biophys.37.032807.125836

Bartels D, Sunkar R. 2005. Drought and Salt Tolerance in Plants. Critical reviews in plant sciences 24:23-58

Booth IR, Edwards MD, Black S, Schumann U, Miller S. 2007. Mechanosensitive channels in bacteria: signs of closure? Nature reviews. Microbiology 5:431-440 DOI 10.1038/nrmicro1659

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633-2635 DOI 10.1093/bioinformatics/btm308

Camargo SR, Cancado GM, Ulian EC, Menossi M. 2007. Identification of genes responsive to the application of ethanol on sugarcane leaves. Plant cell reports 26:2119-2128 DOI 10.1007/s00299-007-0430-8
DeFalco TA, Bender KW, Snedden WA. 2009. Breaking the code: Ca^{2+} sensors in plant signalling. *Biochemical Journal* 425:27-40 DOI 10.1042/bj20091147

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M. 2014. Pfam: the protein families database. *Nucleic Acids Research* 42:D222-230 DOI 10.1093/nar/gkt1223

Ganie SA, Pani DR, Mondal TK. 2017. Genome-wide analysis of DUF221 domain-containing gene family in *Oryza* species and identification of its salinity stress-responsive members in rice. *PLoS One* 12:e0182469 DOI 10.1371/journal.pone.0182469

Hepler PK. 2005. Calcium: a central regulator of plant growth and development. *Plant Cell* 17:2142-2155 DOI 10.1105/tpc.105.032508

Hu B, Jin JP, Guo AY, Zhang H, Luo JC, Gao G. 2015. GSRS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31:1296-1297 DOI 10.1093/bioinformatics/btu817

Ingram J, Bartels D. 1996. The Molecular Basis of Dehydration Tolerance in Plants. *Annual review of plant physiology and plant molecular biology* 47:377-403 DOI 10.1146/annurev.arplant.47.1.377

Jojoa-Cruz S, Saotome K, Murthy SE, Tsui CCA, Sansom MSP, Patapoutian A, Ward AB. 2018. Cryo-EM structure of the mechanically activated ion channel OSCA1.2. *eLife* 7:e41845 DOI 10.7554/eLife.41845

Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K. 1994. Cloning of cDNAs for genes that are early-responsive to dehydration stress (ERDs) in *Arabidopsis thaliana* L.: identification of three ERDs as HSP cognate genes. *Plant Molecular Biology* 25:791-798 DOI 10.1007/bf00028874

Knight H, Trewavas AJ, Knight MR. 1997. Calcium signalling in *Arabidopsis thaliana* responding to drought and salinity. *Plant Journal* 12:1067-1078 DOI 10.1046/j.1365-313x.1997.12051067.x

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* 33:1870-1874 DOI
Li Y, Yuan F, Wen Z, Li Y, Wang F, Zhu T, Zhuo W, Jin X, Wang Y, Zhao H, Pei ZM, Han S. 2015. Genome-wide survey and expression analysis of the OSCA gene family in rice. BMC Plant Biology 15:261 DOI 10.1186/s12870-015-0653-8

Liu S, Wang X, Wang H, Xin H, Yang X, Yan J, Li J, Tran LS, Shinozaki K, Yamaguchi-Shinozaki K, Qin F. 2013. Genome-wide analysis of ZmDREB genes and their association with natural variation in drought tolerance at seedling stage of Zea mays L. PLoS Genetics 9:e1003790 DOI 10.1371/journal.pgen.1003790

Liu X, Wang J, Sun L. 2018. Structure of the hyperosmolality-gated calcium-permeable channel OSCA1.2. Nat Commun 9:5060 DOI 10.1038/s41467-018-07564-5

Liu YH, Li HY, Shi YS, Song YC, Wang TY, Li Y. 2009. A Maize Early Responsive to Dehydration Gene, ZmERD4, Provides Enhanced Drought and Salt Tolerance in Arabidopsis. Plant Molecular Biology Report 27:542-548 DOI 10.1007/s11105-009-0119-y

Maity K, Heumann J, McGrath AP, Kopcho NJ, Hsu P-K, Lee C-W, Mapes JH, Garza D, Krishnan S, Morgan GP, Medrano-Soto A, Hendargo KJ, Klose T, Rees SD, Saier MH, Piñeros M, Komives EA, Schroeder JI, Chang G, Stowell MHB. 2018. Cryo-EM Structure of OSCA1.2 from Oryza sativa: Mechanical basis of hyperosmolality-gating in Plants. bioRxiv505453 DOI 10.1101/505453

Mao H, Wang H, Liu S, Li Z, Yang X, Yan J, Li J, Tran LS, Qin F. 2015. A transposable element in a NAC gene is associated with drought tolerance in maize seedlings. Nature Communications 6:8326 DOI 10.1038/ncomms9326

Marchlerbauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, Deweescott C, Fong JH, Geer LY, Geer RC, Gonzales NR. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Research 39:D225-D229

McAinsh MR, Pittman JK. 2009. Shaping the calcium signature. New Phytologist 181:275-294 DOI 10.1111/j.1469-8137.2008.02682.x

Murthy SE, Dubin AE, Whitwam T, Jojoo-Cruz S, Cahalan SM, Mousavi SAR, Ward AB,
Patapoutian A. 2018. OSCA/TMEM63 are an evolutionarily conserved family of mechanically activated ion channels. *eLife* 7:e41844 DOI 10.7554/eLife.41844

Rai A, Suprasanna P, D'Souza SF, Kumar V. 2012. Membrane topology and predicted RNA-binding function of the 'early responsive to dehydration (ERD4)' plant protein. *PLoS One* 7:e32658 DOI 10.1371/journal.pone.0032658

Reddy AS. 2001. Calcium: silver bullet in signaling. *Plant Science* 160:381-404

Sekhon RS, Lin H, Childs KL, Hansey CN, Buell CR, de Leon N, Kaeppler SM. 2011. Genome-wide atlas of transcription during maize development. *Plant Journal* 66:553-563 DOI 10.1111/j.1365-313X.2011.04527.x

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22:4673-4680

Wang X, Wang H, Liu S, Ferjani A, Li J, Yan J, Yang X, Qin F. 2016. Genetic variation in *ZmVPP1* contributes to drought tolerance in maize seedlings. *Nature Genetics* 48:1233-1241 DOI 10.1038/ng.3636

Xiang Y, Sun X, Gao S, Qin F, Dai M. 2017. Deletion of an Endoplasmic Reticulum Stress Response Element in a *ZmPP2C-A* Gene Facilitates Drought Tolerance of Maize Seedlings. *Molecular Plant* 10:456-469 DOI 10.1016/j.molp.2016.10.003

Yu J, Buckler ES. 2006. Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology* 17:155-160 DOI 10.1016/j.copbio.2006.02.003

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38:203-208 DOI 10.1038/ng1702

Yuan F, Yang H, Xue Y, Kong D, Ye R, Li C, Zhang J, Theprungsirikul L, Shrift T, Krichilsky B, Johnson DM, Swift GB, He Y, Siedow JN, Pei ZM. 2014. *OSCA1* mediates osmotic-stress-evoked Ca$^{2+}$ increases vital for osmosensing in Arabidopsis. *Nature* 514:367-
Zhang M, Wang D, Kang Y, Wu JX, Yao F, Pan C, Yan Z, Song C, Chen L. 2018. Structure of the mechanosensitive OSCA channels. *Nat Struct Mol Biol* **25**:850-858 DOI 10.1038/s41594-018-0117-6

Zhao X, Xu M, Wei R, Liu Y. 2015. Expression of OsCAS (Calcium-Sensing Receptor) in an Arabidopsis Mutant Increases Drought Tolerance. *PLoS One* **10**:e0131272 DOI 10.1371/journal.pone.0131272

Zhu J. 2002. Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology* **53**:247-273

**FIGURE LEGENDS**

**Figure 1.** Phylogenetic tree and gene structure of predicted OSCA genes from maize, rice, sorghum, and Arabidopsis. (A) The phylogenetic tree of 49 full-length OSCA protein sequences from four species. The names used for OSCA gene in rice and Arabidopsis are according to Li et al., 2010. Each node was labeled with bootstrap values from 1,000 replicates. Genes contained in a blue-box were regarded as direct orthologous genes among species. (B) Position of exons, introns, and (untranslated region) UTR in the OSCA genes. Introns are indicated by lines, exons by black boxes, UTR by gray box.

**Figure 2.** Schematic diagram of syntenic chromosomal segments containing ZmOSCA genes between the sorghum, rice, and maize genomes. *Zm, Os,* and *Sb* stand for chromosomes in maize, rice, and sorghum respectively. The homologous chromosomal regions of different genomes are connected by black dotted lines Each OSCA homologous gene pair is linked by a red line. Yellow, blue and brown boxes represent the homologous regions within the maize, rice and sorghum genome.
Figure 3. Protein domains and transmembrane regions in ZmOSCAs. Protein domains and transmembrane regions were predicted by CDD search and TMHMM. Black lines indicate the transmembrane regions. The protein domains were colored according to the color legend.

Figure 4. Transmembrane region of the DUF221 in ZmOSCAs. DNAm and TMHMM were used to perform multiple sequence alignments and predict transmembrane region. TM1-TM7 are the transmembrane regions that are marked by black lines. Regions between red lines

Figure 5. Expression profiles of ZmOSCAs. (A) Expression profiles of the 12 ZmOSCA genes in fifteen tissues illustrated by a heat map. Different colors in the heat map represent the normalized gene expression values indicated by the scale bar. The color scale ranged from green, representing low expression, and passes through yellow and finally to red, representing high expression. The gray color shows missing data. (B) Expression profile of ZmOSCA genes in maize B73 seedlings under drought. Total RNA was extracted from the third leaves before and after the drought treatments. 5 and 24 hours, represent the collection time points, responding to relative leaf water content (RLWC) of 70% and 58%, respectively. Transcript levels of ZmUbi-2 were used as an internal control for data normalization. The represented mean and SD derived from three biological replicates. (t-test, *p ≤ 0.05, **p ≤ 0.01).

Figure 6. Expression patterns of ZmOSCA genes in maize B73 adult leaves under drought. The drought treatment started from the (eight leaves) V8 stage, while the control was well-watered. Total RNA was collected at V12, V14, V16 and R1 stage for both drought stressed and well-watered plants. Transcript levels of ZmUbi-2 transcript levels were used served as an internal control for data normalization. Data represent the represented mean and SD derived from of
three biological replicates. (*-test, *$p \leq 0.05$, **$p \leq 0.01$).

Table 1. Detailed information for twelve ZmOSCA genes in the Zea mays L. genome.

Table 2. Numbers and classification of OSCA genes in Arabidopsis, rice, maize, and sorghum.

Table 3. Number of significantly associated genetic variation in ZmOSCA genes with drought tolerance at the seedling stage.

Supporting Information

Table S1. Gene and protein sequences used to construct phylogenetic tree.

Table S2. Collinear segments between maize, rice and sorghum genomes.
Figure 1

Phylogenetic tree and gene structure of predicted OSCA genes from maize, rice, sorghum, and Arabidopsis

(A) The phylogenetic tree of 49 full-length OSCA protein sequences from four species. The names used for OSCA gene in rice and Arabidopsis are according to Li et al., 2010. Each node was labeled with bootstrap values from 1,000 replicates. Genes contained in a blue-box were regarded as direct orthologous genes among species. (B) Position of exons, introns, and (untranslated region) UTR in the OSCA genes. Introns are indicated by lines, exons by black boxes, UTR by gray box.
Manuscript to be reviewed
Figure 2

Schematic diagram of syntenic chromosomal segments containing ZmOSCA genes between the sorghum, rice, and maize genomes.

Zm, Os, and Sb stand for chromosomes in maize, rice, and sorghum respectively. The homologous chromosomal regions of different genomes are connected by black dotted lines. Each OSCA homologous gene pair is linked by a red line. Yellow, blue and brown boxes represent the homologous regions within the maize, rice and sorghum genome.
Protein domains and transmembrane regions in ZmOSCAs.

Protein domains and transmembrane regions were predicted by CDD search and TMHMM. Black lines indicate the transmembrane regions. The protein domains were colored according to the color legend.
Figure 4

Transmembrane region of the DUF221 in ZmOSCAs.

DNAman and TMHMM were used to perform multiple sequence alignments and predict transmembrane region. TM1-TM7 are the transmembrane regions that are marked by black lines. Regions between red lines
Figure 5

Expression profiles of *ZmOSCAs*.

**(A)** Expression profiles of the 12 *ZmOSCA* genes in fifteen tissues illustrated by a heat map. Different colors in the heat map represent the normalized gene expression values indicated by the scale bar. The color scale ranged from green, representing low expression, and passes through yellow and finally to red, representing high expression. The gray color shows missing data. **(B)** Expression profile of *ZmOSCA* genes in maize B73 seedlings under drought. Total RNA was extracted from the third leaves before and after the drought treatments. 5 and 24 hours, represent the collection time points, responding to relative leaf water content (RLWC) of 70% and 58%, respectively. Transcript levels of *ZmUbi-2* were used as an internal control for data normalization. The represented mean and SD derived from three biological replicates. (*t*-test, *p* ≤ 0.05, *p* ≤ 0.01).
Figure 6

Expression patterns of ZmOSCA genes in maize B73 adult leaves under drought.

The drought treatment started from the (eight leaves) V8 stage, while the control was well-watered. Total RNA was collected at V12, V14, V16 and R1 stage for both drought stressed and well-watered plants. Transcript levels of ZmUbi-2 transcript levels were used served as an internal control for data normalization. Data represent the represented mean and SD derived from of three biological replicates. (t-test, $^\dagger p \leq 0.05$, $^\ddagger p \leq 0.01$).
Table 1 (on next page)

Detailed information for twelve ZmOSCA genes in the Zea mays L. genome.
Table 1. Detailed information for twelve *ZmOSCA* genes in the *Zea mays* L. genome.

| Gene name     | Gene identifier | Chromosome | Protein length (aa) | ORF (bp) | Number of exons | class |
|---------------|-----------------|------------|---------------------|----------|-----------------|-------|
| ZmOSCA1.1a    | GRMZM2G064189   | 3          | 327                 | 984      | 4               | 1     |
| ZmOSCA1.1b    | GRMZM2G021194   | 3          | 768                 | 2307     | 11              | 1     |
| ZmOSCA1.2     | GRMZM2G456000   | 8          | 768                 | 2307     | 11              | 1     |
| ZmOSCA1.3     | GRMZM2G181206   | 6          | 748                 | 2247     | 11              | 1     |
| ZmOSCA1.4     | GRMZM2G128641   | 1          | 810                 | 2433     | 11              | 1     |
| ZmOSCA2.1     | GRMZM2G163059   | 3          | 586                 | 1761     | 8               | 2     |
| ZmOSCA2.2     | GRMZM2G409093   | 1          | 765                 | 2298     | 10              | 2     |
| ZmOSCA2.3     | GRMZM2G164470   | 5          | 749                 | 2250     | 10              | 2     |
| ZmOSCA2.4     | GRMZM2G039186   | 1          | 699                 | 2100     | 10              | 2     |
| ZmOSCA2.5     | GRMZM2G039186   | 5          | 249                 | 750      | 2               | 3     |
| ZmOSCA3.1     | GRMZM2G162253   | 5          | 796                 | 2391     | 1               | 4     |
Table 2 (on next page)

Numbers and classification of OSCA genes in *Arabidopsis*, rice, maize, and sorghum.
Table 2. Numbers and classification of OSCA genes in *Arabidopsis*, rice, maize, and sorghum.

| Species             | clade1 | clade2 | clade3 | clade4 | In Total |
|---------------------|--------|--------|--------|--------|----------|
| *Arabidopsis thaliana* | 8      | 5      | 1      | 1      | 15       |
| *Oryza sativa*      | 4      | 5      | 1      | 1      | 11       |
| *Zea mays*          | 5      | 5      | 1      | 1      | 12       |
| *Sorghum bicolor*   | 4      | 5      | 1      | 1      | 11       |
Table 3 (on next page)

Number of significantly associated genetic variation in ZmOSCA genes with drought tolerance at the seedling stage.
Table 3. Number of significantly associated genetic variation in ZmOSCA genes with drought tolerance at the seedling stage.

| Gene ID         | Gene name    | Polymorphic number* | GLM $P \leq 0.01$ | PC$_2^*$ | PC$_2+K^*$ |
|-----------------|--------------|---------------------|------------------|---------|-----------|
| GRMZM2G064189   | ZmOSCA1.1a   | 4                   | 0                | 0       | 0         |
| GRMZM2G021194   | ZmOSCA1.1b   | 22                  | 0                | 0       | 0         |
| GRMZM2G456000   | ZmOSCA1.2    | 11                  | 4                | 0       | 0         |
| GRMZM2G181206   | ZmOSCA1.3    | 14                  | 0                | 0       | 0         |
| GRMZM2G128641   | ZmOSCA1.4    | 42                  | 7                | 0       | 0         |
| GRMZM2G163059   | ZmOSCA2.1    | 12                  | 2                | 0       | 0         |
| GRMZM2G409093   | ZmOSCA2.2    | 3                   | 0                | 0       | 0         |
| GRMZM2G164470   | ZmOSCA2.3    | 2                   | 0                | 0       | 0         |
| GRMZM2G039186   | ZmOSCA2.4    | 13                  | 0                | 0       | 0         |
| GRMZM2G402708   | ZmOSCA2.5    | 4                   | 0                | 0       | 0         |
| GRMZM2G162253   | ZmOSCA3.1    | 7                   | 0                | 0       | 0         |
| GRMZM2G059891   | ZmOSCA4.1    | 34                  | 10               | 2       | 3         |

*MAF (Minor Allele Frequency) $\geq$ 0.05.