Journal of Cotton Research

RESEARCH Open Access

# Genome-wide identification of *OSCA* gene family and their potential function in the regulation of dehydration and salt stress in *Gossypium hirsutum*

YANG Xiu[1†], XU Yanchao[1†], YANG Fangfang[1], MAGWANGA Richard Odongo[1,2†], CAI Xiaoyan[1], WANG Xingxing[1], WANG Yuhong[1], HOU Yuqing[1], WANG Kunbo[1], LIU Fang[1*] and ZHOU Zhongli[1*†]

## Abstract

**Background:** Cotton (*Gossypium hirsutum*) provides the largest natural fiber for the textile manufacturing industries, but its production is on the decline due to the effects of salinity. Soil salt-alkalization leads to damage in cotton growth and a decrease in yields. Hyperosmolality-gated calcium-permeable channels (*OSCA*) have been found to be involved in the detection of extracellular changes which trigger an increase in cytosolic free calcium concentration. Hyperosmolality-induced calcium ion increases have been widely speculated to be playing a role in osmosensing in plants. However, the molecular nature of the corresponding calcium ion channels remains unclearly. In this research work, we describe the *OSCA* genes and their putative function in osmosensing in plants by carrying out genome-wide identification, characterization and functional analysis of the significantly up-regulated OSCA gene, *GhOSCA1.1* through reverse genetics.

**Result:** A total of 35, 21 and 22 *OSCA* genes were identified in *G. hirsutum, G. arboreum*, and *G. raimondii* genomes, respectively, and were classified into four different clades according to their gene structure and phylogenetic relationship. Gene and protein structure analysis indicated that 35 *GhOSCA* genes contained a conserved RSN1_7TM (PF02714) domain. Moreover, the *cis*-regulatory element analysis indicated that the *OSCA* genes were involved in response to abiotic stress. Furthermore, the knockdown of one of the highly up-regulated genes, *Gh_OSCA1.1* showed that the virus-induced gene silenced (VIGS) plants were highly sensitive to dehydration and salinity stresses compared with the none VIGS plants as evident with higher concentration levels of oxidant enzymes compared with the antioxidant enzymes on the leaves of the stressed plants.

**Conclusion:** This study provides the first systematic analysis of the *OSCA* gene family and will be important for understanding the putative functions of the proteins encoded by the *OSCA* genes in cotton. These results provide a new insight of defense responses in general and lay the foundation for further investigation of the molecular role played by the OSCA genes, thereby providing suitable approaches to improve crop performance under salinity and drought stress conditions.

**Keywords:** *OSCA* gene family, *Gossypium hirsutum*, VIGS, Salt and dehydration stress

* Correspondence: liufcri@163.com; zhonglizhou@163.com
†Xiu YANG, Yanchao XU, Richard Odongo MAGWANGA and Zhongli ZHOU contributed equally to this work.
[1]Institute of Cotton Research, Chinese Academy of Agricultural Science (ICR, CAAS) / Research Base in Anyang Institute of Technology, State Key Laboratory of Cotton Biology, Anyang 455000, China
Full list of author information is available at the end of the article

## Background

Salt and dehydration stresses are the major forms of abiotic stress factors which limit the growth and development of the plant (Liu et al. 2010). A number of researchers have tried to explore the mechanism of salt and dehydration stress responses, although it is complicated (Nakashima and Yamaguchi-Shinozaki 2013; Qiu et al. 2011; Ullah and Sun 2018). Therefore, some potential signal pathways were proved in salt and dehydration stress response (Munns 2005; Zhu 2016). Moreover, a number of stress-responsive genes have been found to play a significant role in enhancing plants adaptation to various forms of abiotic stress factors such as drought and salinity stress (Magwanga et al. 2018). Furthermore, several investigations have been done in order to understand the plant's response or regulatory mechanism under salt and/or drought stress conditions (Deng et al. 2018; Sanchez-Barrena et al. 2004; Taji et al. 2004; Wu et al. 1996; Zhu et al. 2018; Zhu 2016). Salt-Overly-Sensitive (SOS) pathway was the first abiotic stress response signal pathway to be discovered in plants (Zhu 2000). Moreover, studies on the SOS pathways have shown that calcium ions are integral in the SOS salt-dehydrative responsive pathways in plants (Da and Ploy 2012; Siaud et al. 2010). In this pathway, the cytosolic calcium signal was sensed by the EF-hand calcium-binding protein (SOS3) under salt stress. Then, SOS3 interacts with and activates SOS2, a serine/threonine protein kinase (Ishitani et al. 2000). Previous studies showed that plants have development ABA-independent and ABA-dependent signal pathway to perceive and response to dehydration stress (Nakashima and Yamaguchi-Shinozaki 2013; Podia et al. 2018). Dehydration-responsive elements (DRE) play an important role in the ABA-independent pathway (Gupta et al. 2014; Pardo et al. 1998). The ABA-responsive element (ABRE) is involved in the ABA-dependent signal cascade pathway (Yoshida et al. 2014). However, the osmotic stress response is an important and common mechanism to regulated salt and dehydration stress, the mechanism underlying the early response to osmotic stress in plants remains undiscovered (Shavrukov 2012).

Hyperosmolality-induced change in $Ca^{2+}$ level was widely speculated to be involved in osmotic stress regulation in plants (Zhu 2002). The intracellular free calcium concentration is increased under dehydration and salt stress in plants (Knight et al. 1997; McAinsh and Pittman 2009). The hyperosmolality-induced free calcium concentration increase (OICI) is the very first process to mitigate the effects of osmotic stress (Knight et al. 1997). Furthermore, the osmotic stimuli-gated $Ca^{2+}$ permeable channels, osmosensors, and the regulated free calcium concentration have been observed in bacteria under osmotic stress (Árnadóttir and Chalfie 2010). Moreover, the *AtOSCA*, encoding a membrane protein, was involved in osmotic stress response as a hyperosmolality gated calcium-permeable channel in *Arabidopsis thaliana*. Fifteen and 11 *OSCA* family genes were identified in *Arabidopsis* and *Oryza sativa* (Kiyosue et al. 1994; Li et al. 2015), respectively. In *Arabidopsis*, early response to dehydration (ERD) genes were cloned and thought to be involved with dehydration-induced osmotic stress. *ERD*4 encodes a protein which contains a conserved DUF221 domain (Rai et al. 2012). The conserved DUF221 domain, including seven transmembrane regions, was renamed RSN1_7TM domain (PF02714) (Ganie et al. 2017). The previous study has shown that *OSCA* genes encode a protein, which contains a highly conserved RSN1_7TM domain (Camargo et al. 2007; Ganie et al. 2017; Rai et al. 2012; Shinozaki and Yamaguchi-Shinozaki 2000). Therefore, identifying the *OSCA* gene family will provide a potential resource to improve the deep understanding of regulation to dehydration and salt stress.

In this study, a total of 35, 21, 22 OSCA family members were identified in *Gossypium hirsutum*, *G. arboreum* and *G. raimondii*, respectively. Physical and chemical characteristics of the protein encoded by the *GhOSCA* genes were analysed. Phylogenetic relationships, chromosome location, gene and protein structure analysis were performed among these OSCAs. Furthermore, *OSCA* gene family member expansions were deeply analyzed for better understanding by performing the gene duplication events analysis. Expression levels in various organs/tissues and under dehydration and salt stress were analysis in our study. Gene silencing of *GhOSCA1.1* proved the potential function of the novel *OSCA* gene and its involvement in enhancing dehydration and salt-induced osmotic stress response in cotton. These results provide a new insight into defense responses in general and lay the foundation for future crop improvement.

## Materials and methods

### Plant material, dehydration and salt stress treatment

*G. hirsutum* var. *marie-galante 85* (MAR85) was selected for functional analysis of the GhOSCAs under dehydration and salt stress. The *G. hirsutum* accessions of MAR85 are known to be distributed in Guadeloupe and Guatemala, and were introduced from USDA-ARS Southern Agricultural Research Center in College Station, Texas, USA and perennially preserved in the National Wild Cotton Nursery (Sanya, Hainan), and managed by Institute of Cotton Research, Chinese Academy of Agricultural Sciences (ICR, CAAS). The seeds of MAR85 were first germinated at 28 °C in a 16 h light/8 h dark cycle and then transplanted in a normal hydroponic solution with a Hoagland solution for a

period of 3 weeks. After 3 weeks and with a fully expanded third leaf, the seedlings were exposed to salinity and drought stress, by adding 300 mmol·$L^{-1}$ of sodium chloride (NaCl) solution and 17% PEG6000, salinity and drought stress, respectively. The tissues examined were the roots and leaves, in which the samples were collected at 0 h, 3 h, 12 h, and 48 h after salt-alkali stress treatment. The samples were immediately frozen under −80 °C awaiting RNA extraction for RT-qPCR (quantitative real-time polymerase chain reaction) validation.

### Identification of OSCAs in *G. hirsutum*, *G. arboreum*, and *G. raimondii*

Genes and proteins annotated in *G. hirsutum, G. arboreum*, and *G. raimondii* were downloaded from the COTTONGEN database (https://www.cottongen.org/). For the two cotton genomes, *G. hirsutum* (AD) and *G. arboreum* (A), their annotations were obtained from the Cotton Research Institute, Nanjing Agricultural University website (http://mascotton.njau.edu.cn/) while the sequences for *G. raimondii* was obtained from phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). The *OSCA* genes family members of Arabidopsis and rice, which were used for identified candidate *OSCA* genes of *G. hirsutum*, *G. arboreum* and *G. raimondii*, were retrieved from UNIPROT (https://www.uniprot.org/). AtOSCAs and OsOSCAs were aligned with the protein sequences of the *G. hirsutum, G. arboreum* and *G. raimondii* with the default parameter by local BLASTP software. The conservative RSN1_7TM domain (PF02714) of the OSCA family was used to further reconfirm the candidate OSCAs of *G. hirsutum, G. arboreum* and *G. raimondii* by PFAM database (https://pfam.xfam.org/) and online CD-search tool of NCBI (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) (Marchler-Bauer et al. 2016). The biophysical characters of the encoded proteins were computed using the ExPASy ProtParam tool (http://us.expasy.org/tools/protparam.html). Prediction of the subcellular localization of the proteins encoded by the *OSCA* gene family using WoLFPSORT (https://wolfpsort.hgc.jp/).

### Mapping, phylogenetic tree construction, and gene structure analysis of the *OSCA* gene family

Mapping of *GhOSCA* genes was performed using Mapchart software (Voorrips 2002). The exon/intron structures of individual *OSCA* genes were determined by Gene Structure Display Server (GSDS 2.0) (Hu et al. 2014). Full-length sequences of GhOSCA proteins were first aligned with the ClustalX program (http://www.clustal.org/clustal2/) (Larkin et al. 2007), and the phylogenetic trees was constructed by using two methods, the neighbor-joining (NJ) method with 1 000 bootstrap replicas, and the Maximum likehood to validate the phylogentic tree (Fan et al. 2018; Kumar et al. 2016) and the Poisson model by using MEGA 7.0 software (http://www.megasoftware.net). Meanwhile, the orthologous gene pairs of GhOSCA in A, D genomes, At and Dt subgenomes were searched via InParanoid software (http://inparanoid.sbc.su.se/cgi-bin/index.cgi). Additionally, the $d_S$ and $d_N$ substitution rates were calculated with the PAL2NAL web server (http://www.bork.embl.de/pal2nal#RunP2N), which uses the CODEMAL program of PAML.

### RNA extraction and quantitative and real-time PCR

Results of RNA-seq were validated via quantitative real-time PCR (RT-qPCR) experiments and real-time PCR analyses were performed as the user manual of the TransScript II All-in-One First-Strand cDNA Synthesis SuperMix for PCR (TransGen Biotech) and the SYBR Premix Ex Taq II kit (Roche) described. The housekeeping gene was Ghactin7 (Forward sequence: 5'ATCCTCCGTCTTGACCTTG3'; Reverse sequence: 5'TGTC CGTCAGGCAACTCAT3'). The gene-specific primers designed using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) tool and primers are listed in Table 1. The experiments of quantitative real-time PCR were performed using three biological replicates for each tissue sample and at least three technical replicates of each biological replicate. The value of genes folds change was calculated using the $2^{-\Delta\Delta C}{}_T$ method.

### Vector construction and procedure for VIGS in cotton availability of supporting data

The TRV2 (Tobacco rattle virus) vectors construct TRV2: 00, TRV2:*CLA1* and TRV2:*GhOSCA1.1* which were prepared and introduced into *Agrobacterium tumefaciens* strain LBA4404. In order to monitor the silencing efficiency, the TRV2:*CLA1* vector was constructed as a visual marker. Primers were used to generate TRV2 vector forward sequence "GTGAGTAAGGTTACCGAATTCCAGCGTAATTGCAGGCAGTG" and reverse sequence "CGTGAGCTCGGTACCGGATCCGAACAGGTGTCACGGTA GCA". The Agrobacterium culture was Agroinfiltrated into two expanded cotyledons of 10-day-old soil-grown seedling of Marie-galante 85 (MAR85). The cotton seedlings were planted in a 26 °C and 16 h light/8 h dark cycle. At least 24 seedlings were inoculated for each construct. At 14 days after Agrobacterium inoculation when VIGS was established, the silenced seedlings were posted to salt and drought. At 20 days after salt-alkali stress treatment, the leaf samples were collected for expressed level, malondialdehyde (MDA), proline (PRO) and superoxide dismutase (SOD) assay.

**Table 1** Primers used in RT-qPCR analysis of salt and drought response genes

| Gene name | Forward primer(5′-3′) | Reverse primer(5′-3′) |
|---|---|---|
| *GhOSCA1.1* | CGAAACGCCAATCTCGAAGG | AGCCTCCGGTAAGGATTGTG |
| *GhOSCA3.1* | CTTTGCCTTGGGTTGGCTTG | ACAATATCGGGAGCGGAACC |
| *GhOSCA1.16* | CCCTGGCAGTGTGGACTATC | CAGCACCACTCTCGTACTGG |
| *GhOSCA3.3* | CACACGGCTCTTGAAGTTGC | TCTCAGGAATCAAGCTCGGC |
| *GhOSCA1.3* | ACACAACAATCCCGAGACCC | AAGGAATCGGTAGGTTCGGC |
| *GhOSCA1.2* | CGAAACGCCAATCTCGAAGG | AGCCTCCGGTAAGGATTGTG |
| *GhOSCA2.5* | GCGGAAAGTGTTGAGTGCAA | AAAGCAGCTGGACGTTCCTT |
| *GhOSCA2.12* | GCTCAAACTCAGCAGCACAC | TCCCTTGAACAGCAGTGACG |
| *GhOSCA3.2* | AAGAGTCCAAGGTTGTGGGC | CTGAAGTGGCATTCGGCAAG |
| *GhOSCA3.4* | CTCCGGAATCAGGCACTCAA | TGGTATAACAGCAGGGCACC |
| *GhOSCA1.4* | CTTGGCCTCGTTTATGCTGC | TGACGGAAGACGACGTAAGC |

### Determination of water loss rate, malondialdehyde, superoxide dismutase, and proline assays

After VIGS infusion at the three-leaf stage of the cotton seedlings growth stage, nine cotton leaves of similar size were taken from TRV2:00, TRV2:*CLA1* and TRV2:*GhOSCA1.1*, respectively. Leaves were cultured in an artificial climate incubator at 28 °C. Three repeats were set up. Each hour interval, the leaves were weighed, and the water loss rate of the isolated leaves was counted [Leaf sater loss rate (%) = (Leaf fresh weight−Leaf dry weight) *100%/Leaf fresh weight]. To detect the content of MDA and PRO and activity of SOD, leaves of MAR85 were collected after 48 h post to salt-alkali stress. The corresponding assay kits (Beijing Solarbio Science & Technology Co., Ltd.) were used for determining the content of MDA and PRO and the activity of SOD.

## Results

### Identification of *OSCA* genes family in the cotton genome

To explore members of the *OSCA* gene family in *G. hirsutum, G. arboreum* and *G. raimondii*, 16 AtOSCAs and 11 OsOSCAs protein sequences were used as a query to screen protein databases of *G. hirsutum, G. arboreum*, and *G. raimondii* genome. A total of 35, 21 and 22 candidate *OSCAs* of *G. hirsutum, G. arboreum* and *G. raimondii* were obtained, respectively. In previous studies, 15, 11, 10 and 21 *OSCA* genes were identified in *Arabidopsis*, rice, maize, and soybean, respectively (Gu et al. 2018). A large number of *OSCA* gene family members (Shan et al. 2005) in *G. hirsutum* may be related to the whole genome replication of cotton. But queerly, compared with the number of *OSCA* genes of diploid A and D genome donor species, *G. arboreum* (Magwanga et al. 2018) and *G. raimondii* (Magwanga et al. 2019b), the allotetraploid species *G. hirsutum* (Shan et al. 2005) showed fewer *OSCA* members. This result suggested that there was possible gene loss and/or as a result of chromosome rearrangement during the history of chromosome doubling and plant evolution. The results were in agreement with previous findings in other plant gene members such as the *LEA* genes, in which 157, 89 and 85 proteins encoded by the *LEA* genes were identified in *G. hirsutum*, *G. raimondii* and *G. arboreum*, respectively (Magwanga et al. 2018).

Furthermore, the *OSCA* genes of three different *Gossypium* species have various characteristics (Table 2). The length of the *OSCA* gene sequences ranged from 900 bp to 26 539 bp. The gene with the highest length of 26 539 had the highest level of intron interruption compared with all other members of the *OSCA* genes in *G. hirsutum*. The length of OSCA coding sequences ranged from 300 bp to 3 678 bp in three different cotton species. Interestingly, the length and number of OSCA introns are quite different in three *Gossypium* species. Above all, the various lengths of gene sequences among the *OSCA* gene family in cotton were the difference of intron structure. From Table 2, it can be found that the theoretical isoelectric point and molecular weight of OSCA protein have little difference, indicating that the physical and chemical properties of *OSCA* family genes have little difference. The isoelectric point (pI) of the majority of the GhOSCA proteins was alkaline except for *GhOSCA4.1*. The GRAVY values of the proteins were calculated as the sum of the hydropathy value of each residue, divided by the total number of the residues present in the sequences. Positive and negative GRAVY scores reflect hydrophobicity and hydrophilicity, respectively. Of all the three Gossypium species, the GRAVY scores of most GhOSCA proteins were positive, except *GhOSCA1.14* and *GhOSCA1.6* was negative, which indicated that most GhOSCA proteins were hydrophobic proteins. In addition, GhOSCAs contains multiple transmembrane

**Table 2** Physicochemical Properties of *OSCA* Gene

| Gene ID | Gene name | Chro. No | Gene length/bp | CDS length/bp | No. of amino acids/aa | MW (kDa) | pI | GRAVY | Subcellular localization | TM |
|---|---|---|---|---|---|---|---|---|---|---|
| Gh_A01G1022 | GhOSCA2.4 | $A_h$01 | 7 612 | 2 427 | 808 | 91 824.27 | 9.12 | 0.155 | Plasma membrane, | 9 |
| Gh_A05G1223 | GhOSCA3.1 | $A_h$05 | 3 281 | 2 178 | 725 | 82 096.29 | 9.22 | 0.292 | Plasma membrane | 10 |
| Gh_A05G1480 | GhOSCA1.1 | $A_h$05 | 3 655 | 2 313 | 770 | 87 946.96 | 9.2 | 0.162 | Plasma membrane | 10 |
| Gh_A05G1853 | GhOSCA2.6 | $A_h$05 | 5 707 | 2 235 | 744 | 85 367.08 | 9.36 | 0.213 | Nucleus | 11 |
| Gh_A05G3185 | GhOSCA2.8 | $A_h$05 | 4 572 | 2 148 | 715 | 81 461.11 | 8.14 | 0.29 | Plasma membrane | 11 |
| Gh_A06G1859 | GhOSCA2.1 | $A_h$06 | 1 977 | 1 197 | 398 | 45 096.6 | 8.6 | 0.218 | Plasma membrane | 4 |
| Gh_A06G1983 | GhOSCA3.3 | $A_h$06 | 2 640 | 2 178 | 725 | 82 653.21 | 9.24 | 0.249 | Plasma membrane | 9 |
| Gh_A07G1673 | GhOSCA1.5 | $A_h$07 | 6 194 | 2 307 | 768 | 88 108.01 | 8.96 | 0.189 | Plasma membrane | 8 |
| Gh_A08G0402 | GhOSCA1.10 | $A_h$08 | 10 521 | 2 271 | 756 | 86 578.39 | 8.97 | 0.201 | Plasma membrane | 11 |
| Gh_A08G1374 | GhOSCA1.14 | $A_h$08 | 3 247 | 1 836 | 611 | 70 529.57 | 9.2 | −0.116 | Plasma membrane | 5 |
| Gh_A10G0109 | GhOSCA1.3 | $A_h$10 | 4 272 | 2 319 | 772 | 87 912.93 | 9.01 | 0.162 | Plasma membrane | 9 |
| Gh_A11G1108 | GhOSCA1.12 | $A_h$11 | 6 928 | 2 223 | 740 | 85 096.93 | 9.42 | 0.235 | Plasma membrane | 10 |
| Gh_A11G1428 | GhOSCA1.8 | $A_h$11 | 3 463 | 2 370 | 789 | 89 986.06 | 9.17 | 0.193 | Plasma membrane | 9 |
| Gh_A11G1585 | GhOSCA4.1 | $A_h$11 | 2 409 | 2 409 | 802 | 90 638.37 | 6.27 | 0.204 | Plasma membrane | 7 |
| Gh_A12G1860 | GhOSCA1.16 | $A_h$12 | 5 926 | 2 427 | 808 | 93 172.78 | 9.27 | 0.088 | Plasma membrane | 8 |
| Gh_A13G0690 | GhOSCA2.10 | $A_h$13 | 5 580 | 2 148 | 715 | 81 740.83 | 8.9 | 0.34 | Plasma membrane | 9 |
| Gh_D01G1077 | GhOSCA2.5 | $D_h$01 | 7 701 | 2 448 | 815 | 92 658.29 | 9.14 | 0.142 | Plasma membrane | 9 |
| Gh_D03G1219 | GhOSCA2.3 | $D_h$03 | 900 | 462 | 153 | 17 508.91 | 9.75 | 0.572 | Plasma membrane | 4 |
| Gh_D04G0419 | GhOSCA2.9 | $D_h$04 | 4 534 | 2 034 | 677 | 76 933.92 | 7.5 | 0.349 | Plasma membrane | 10 |
| Gh_D05G1651 | GhOSCA1.2 | $D_h$05 | 3 620 | 2 313 | 770 | 87 927.8 | 9.07 | 0.161 | Plasma membrane | 10 |
| Gh_D05G2050 | GhOSCA2.7 | $D_h$05 | 5 563 | 2 202 | 733 | 84 190.8 | 9.17 | 0.241 | Nucleus | 11 |
| Gh_D05G2673 | GhOSCA2.12 | $D_h$05 | 3 818 | 2 133 | 710 | 80 902.88 | 8.62 | 0.225 | Plasma membrane | 10 |
| Gh_D05G3846 | GhOSCA3.2 | $D_h$05 | 3 276 | 2 178 | 725 | 82 107.23 | 9.22 | 0.287 | Plasma membrane | 10 |
| Gh_D06G0222 | GhOSCA2.2 | $D_h$06 | 4 705 | 2 238 | 745 | 85 103.71 | 9.34 | 0.225 | Plasma membrane | 10 |
| Gh_D06G1519 | GhOSCA3.4 | $D_h$06 | 2 650 | 2 178 | 725 | 82 484.13 | 9.27 | 0.279 | Plasma membrane | 9 |
| Gh_D07G1882 | GhOSCA1.6 | $D_h$07 | 26 539 | 3 678 | 1 225 | 140 180.38 | 5.85 | −0.089 | Plasma membrane, | 8 |
| Gh_D08G0493 | GhOSCA1.11 | $D_h$08 | 7 065 | 2 271 | 756 | 86 657.45 | 9.04 | 0.194 | Plasma membrane | 11 |
| Gh_D08G1669 | GhOSCA1.15 | $D_h$08 | 2 677 | 1 434 | 477 | 55 135.71 | 9.85 | 0.076 | Plasma membrane | 4 |
| Gh_D10G0112 | GhOSCA1.4 | $D_h$10 | 4 346 | 2 319 | 772 | 87 760.73 | 8.95 | 0.168 | Plasma membrane | 9 |
| Gh_D11G1258 | GhOSCA1.13 | $D_h$11 | 6 986 | 2 223 | 740 | 85 167.87 | 9.54 | 0.188 | Plasma membrane | 10 |
| Gh_D11G1580 | GhOSCA1.9 | $D_h$11 | 3 466 | 2 376 | 791 | 90 268.5 | 9.1 | 0.194 | Plasma membrane | 9 |
| Gh_D11G1742 | GhOSCA4.2 | $D_h$11 | 2 409 | 2 409 | 802 | 90 638.31 | 6.37 | 0.198 | Plasma membrane | 7 |
| Gh_D12G2031 | GhOSCA1.17 | $D_h$12 | 5 816 | 2 424 | 807 | 92 955.59 | 9.4 | 0.077 | Plasma membrane | 8 |
| Gh_D13G0811 | GhOSCA2.11 | $D_h$13 | 5 425 | 2 148 | 715 | 81 743.82 | 9.04 | 0.33 | Plasma membrane | 9 |
| Gh_Sca144594G01 | GhOSCA1.7 | scaffold | 300 | 300 | 99 | 11 379.51 | 7.79 | 0.706 | Chloroplast | 1 |
| Gorai.001G128900.1 | GrOSCA2.1 | $D_s$01 | 2 178 | 1 064 | 102 | 11.711 | 10.688 | 0.587 | Chloroplast | 2 |
| Gorai.001G215500.1 | GrOSCA1.1 | $D_s$01 | 7 158 | 3 080 | 768 | 87.993 | 8.988 | 0.18 | Plasma membrane | 8 |
| Gorai.002G137000.1 | GrOSCA2.2 | $D_s$02 | 9 231 | 2 862 | 808 | 91.773 | 9.198 | 0.156 | Plasma membrane | 9 |
| Gorai.004G055300.1 | GrOSCA1.2 | $D_s$04 | 5 021 | 2 362 | 586 | 67.115 | 9.246 | 0.244 | Plasma membrane | 9 |
| Gorai.004G180800.1 | GrOSCA1.3 | $D_s$04 | 3 437 | 1 880 | 511 | 58.929 | 10.403 | 0.12 | Plasma membrane | 5 |
| Gorai.007G134700.1 | GrOSCA1.4 | $D_s$07 | 7 654 | 2 816 | 740 | 85.254 | 9.888 | 0.183 | Plasma membrane | 10 |
| Gorai.007G171500.1 | GrOSCA1.5 | $D_s$07 | 3 609 | 2 515 | 788 | 89.983 | 9.369 | 0.173 | Plasma membrane | 7 |
| Gorai.007G191000.1 | GrOSCA4.1 | $D_s$07 | 2 968 | 2 968 | 802 | 90.652 | 6.976 | 0.196 | Plasma membrane | 7 |

**Table 2** Physicochemical Properties of *OSCA* Gene *(Continued)*

| Gene ID | Gene name | Chro. No | Gene length/bp | CDS length/bp | No. of amino acids/aa | MW (kDa) | pI | GRAVY | Subcellular localization | TM |
|---|---|---|---|---|---|---|---|---|---|---|
| Gorai.008G222200.1 | GrOSCA1.6 | $D_5$08 | 6 527 | 3 079 | 807 | 93.015 | 9.601 | 0.074 | Plasma membrane | 8 |
| Gorai.009G152500.1 | GrOSCA3.1 | $D_5$09 | 3 843 | 2 747 | 725 | 82.092 | 9.592 | 0.288 | Plasma membrane | 10 |
| Gorai.009G181100.1 | GrOSCA1.7 | $D_5$09 | 4 876 | 3 289 | 770 | 87.98 | 9.402 | 0.161 | Plasma membrane | 10 |
| Gorai.009G222800.1 | GrOSCA2.3 | $D_5$09 | 6 645 | 3 428 | 734 | 84.307 | 9.457 | 0.234 | Nucleus | 11 |
| Gorai.009G295700.1 | GrOSCA2.4 | $D_5$09 | 4 068 | 2 374 | 710 | 80.819 | 8.476 | 0.221 | Plasma membrane | 10 |
| Gorai.010G028700.1 | GrOSCA2.5 | $D_5$10 | 4 907 | 2 778 | 758 | 86.464 | 9.306 | 0.25 | Plasma membrane, | 9 |
| Gorai.010G051200.1 | GrOSCA2.6 | $D_5$10 | 1 914 | 893 | 105 | 11.8 | 9.562 | 0.35 | Plasma membrane | 1 |
| Gorai.010G139200.1 | GrOSCA2.7 | $D_5$10 | 1 543 | 675 | 105 | 11.854 | 9.579 | 0.389 | Plasma membrane, | 1 |
| Gorai.010G168400.1 | GrOSCA3.2 | $D_5$10 | 3 288 | 2 621 | 725 | 82.416 | 9.561 | 0.278 | Plasma membrane, | 9 |
| Gorai.010G206800.1 | GrOSCA2.8 | $D_5$10 | 1 966 | 858 | 105 | 11.828 | 9.579 | 0.345 | Plasma membrane | 1 |
| Gorai.011G012200.1 | GrOSCA1.8 | $D_5$11 | 5 358 | 3 017 | 772 | 87.761 | 9.275 | 0.167 | Plasma membrane | 9 |
| Gorai.011G134200.1 | GrOSCA2.9 | $D_5$11 | 976 | 1 998 | 105 | 11.828 | 9.579 | 0.345 | Plasma membrane | 1 |
| Gorai.012G051600.1 | GrOSCA2.10 | $D_5$12 | 5 526 | 2 745 | 715 | 81.339 | 8.538 | 0.297 | Plasma membrane | 11 |
| Gorai.013G089300.1 | GrOSCA2.11 | $D_5$13 | 6 843 | 2 821 | 715 | 81.828 | 9.017 | 0.333 | Plasma membrane | 9 |
| Cotton_A_27813 | GaOSCA2.4 | $D_5$10 | 5 108 | 2 247 | 748 | 85.668 | 9.42 | 0.168 | Nucleus | 10 |
| Cotton_A_26870 | GaOSCA3.1 | $D_5$08 | 2 643 | 2 106 | 701 | 79.736 | 9.552 | 0.282 | Plasma membrane | 9 |
| Cotton_A_26154 | GaOSCA3.2 | $D_5$10 | 3 290 | 2 178 | 725 | 82.116 | 9.547 | 0.291 | Plasma membrane | 10 |
| Cotton_A_25263 | GaOSCA2.2 | $D_5$08 | 3 961 | 2 259 | 752 | 86.418 | 9.29 | 0.267 | Plasma membrane | 11 |
| Cotton_A_22266 | GaOSCA2.6 | $D_5$10 | 2 580 | 357 | 118 | 13.129 | 4.672 | 0.07 | Chloroplast | 1 |
| Cotton_A_22163 | GaOSCA4.1 | $D_5$04 | 2 999 | 2 999 | 802 | 90.667 | 6.782 | 0.208 | Plasma membrane | 7 |
| Cotton_A_19862 | GaOSCA2.3 | $A_2$09 | 2 379 | 357 | 118 | 13.071 | 6.515 | 0.214 | Chloroplast | 1 |
| Cotton_A_18702 | GaOSCA1.1 | $A_2$01 | 6 179 | 2 349 | 782 | 89.858 | 8.768 | 0.188 | Plasma membrane | 8 |
| Cotton_A_18315 | GaOSCA2.8 | $A_2$12 | 5 664 | 3 237 | 715 | 81.423 | 8.041 | 0.302 | Plasma membrane | 11 |
| Cotton_A_12285 | GaOSCA1.4 | $A_2$04 | 3 474 | 2 367 | 788 | 89.803 | 9.306 | 0.182 | Plasma membrane | 9 |
| Cotton_A_08463 | GaOSCA1.9 | $A_2$12 | 11 662 | 4 659 | 1 552 | 177.449 | 7.745 | −0.245 | Plasma membrane | 9 |
| Cotton_A_08242 | GaOSCA1.5 | $A_2$04 | 6 890 | 2 223 | 740 | 85.169 | 9.769 | 0.223 | Plasma membrane | 10 |
| Cotton_A_07948 | GaOSCA1.6 | $A_2$06 | 9 663 | 3 000 | 999 | 113.318 | 10.059 | 0.051 | Plasma membrane | 8 |
| Cotton_A_04660 | GaOSCA1.7 | $A_2$10 | 3 655 | 2 313 | 770 | 87.917 | 9.57 | 0.163 | Plasma membrane | 10 |
| Cotton_A_04078 | GaOSCA2.7 | $A_2$10 | 3 822 | 2 145 | 714 | 81.288 | 8.562 | 0.216 | Plasma membrane | 10 |
| Cotton_A_03209 | GaOSCA1.2 | $A_2$03 | 7 085 | 2 271 | 756 | 86.574 | 9.222 | 0.203 | Plasma membrane | 11 |
| Cotton_A_29061 | GaOSCA2.1 | $A_2$07 | 6 765 | 2 199 | 732 | 83.592 | 9.864 | 0.205 | Plasma membrane | 9 |
| Cotton_A_29328 | GaOSCA1.3 | $A_2$03 | 4 128 | 2 418 | 805 | 92.665 | 9.643 | 0.066 | Plasma membrane | 7 |
| Cotton_A_29852 | GaOSCA1.8 | $A_2$11 | 2 752 | 1 344 | 447 | 51.723 | 10.307 | 0.084 | Plasma membrane | 4 |
| Cotton_A_40076 | GaOSCA2.5 | $A_2$10 | 6 579 | 3 296 | 715 | 81.868 | 9.017 | 0.33 | Plasma membrane | 9 |
| Cotton_A_41040 | GaOSCA2.9 | scaffold | 276 | 276 | 91 | 10.34 | 9.026 | 0.492 | Chloroplast | 1 |

Note: CDS length: Coding sequence length, MW: Relative molecular mass, pI: Isoelectric point, GRAVY: Grand average of hydropathy, TM: Transmembrane domain

domains. WoLF PSORT analysis found that most of OSCA family proteins were located in the plasma membrane, among which *GhOSCA2.4*, *GhOSCA3.3*, *GhOSCA1.14*, *GhOSCA1.8*, *GhOSCA2.5*, *GhOSCA2.12*, *GhOSCA1.6*, *GhOSCA1.15*, *GhOSCA1.13*, *GhOSCA1.9*, and *GhOSCA1.7* may be located in chloroplasts and mitochondria.

### Phylogenetic tree relationship and gene structure analysis of the *OSCA* gene family in cotton

To explore the phylogenetic relationship of the cotton *OSCA* genes family, a phylogenetic tree was constructed using sequence protein of the *OSCA* gene in three different cotton species and Arabidopsis and rice. Totally, 62 *OSCA* genes were divided into two

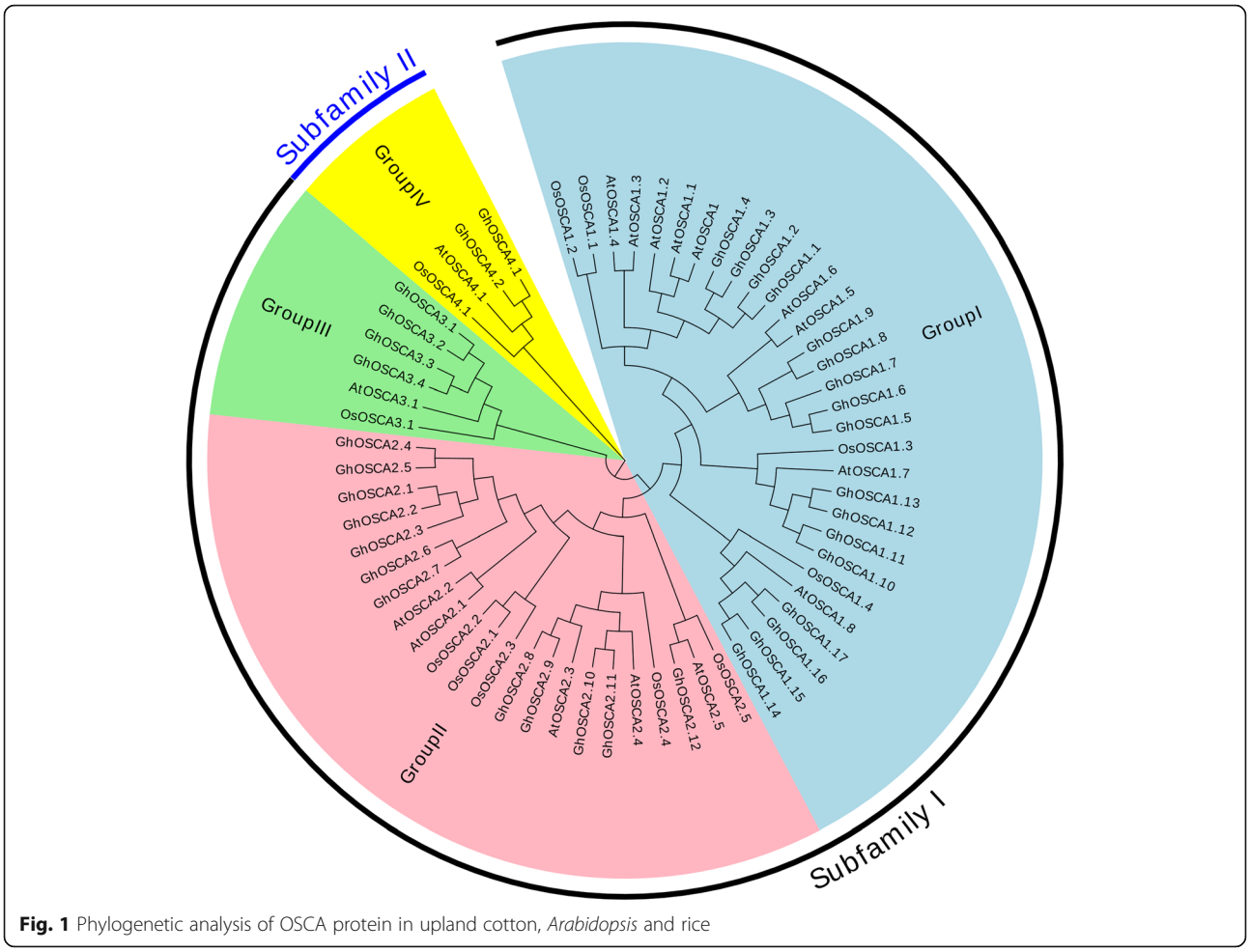


**Fig. 1** Phylogenetic analysis of OSCA protein in upland cotton, *Arabidopsis* and rice

subfamilies (Subfamily I and Subfamily II). Subfamily I contained three groups, and Subfamily II contained one group. Each group consists of at least one of cotyledonous plants Arabidopsis and monocotyledonous plant rice, which indicates that the differentiation time of the *OSCA* gene family is earlier than that of mono-and cotyledons (Fig. 1). The third and fourth groups of OSCA members were small, but they were retained throughout the evolution of species, suggesting a significant role in a biological process. From Fig. 2, it can be seen that the numbers of *G. arboreum* and *G. raimondii* of the *OSCA* family genes were similar, and the corresponding relationship is almost one-to-one, whereas in the *G. hirsutum* the *OSCA* family gene has a high number of amplification, which is in accordance with the species evolution relationship.

Through the genetic structure analysis some gene family evolution informations were obtained, and the difference between exon and intron distribution among the members of the OSCA family is compared (Fig. 3). The results showed that *G. hirsutum*, *G. arboreum*, and *G. raimondii OSCA* genes were divided into four groups according to the genetic structure, which was highly correlated with the classification based on the evolutionary tree. In the exon-intron composition mode, the same group is relatively similar and the difference is greater. This conserved genetic structure between genes in the same group is consistent with their close evolutionary relationship.

### Protein conserved domain and motility analysis of the *OSCA* gene family in *G. hirsutum*

Members of *GhOSCA* family highly conservative three-function domain structure, namely the late exocytosis and Cytosolic domain of 10 TM putative phosphate and Calcium-dependent channel. All members of the *GhOSCA* contained three conserved motifs except *GhOSCA1.7*, *GhOSCA2.3*, *GhOSCA2.8*, *GhOSCA2.9*, *GhOSCA2.12*, *GhOSCA3.2*, *GhOSCA3.3* and *GhOSCA3.4*, which had one conserved domain. We used the MEME software to analyzed conserved motifs in the *OSCA* gene family (Fig. 4). Through the analysis of the conservative motif of the

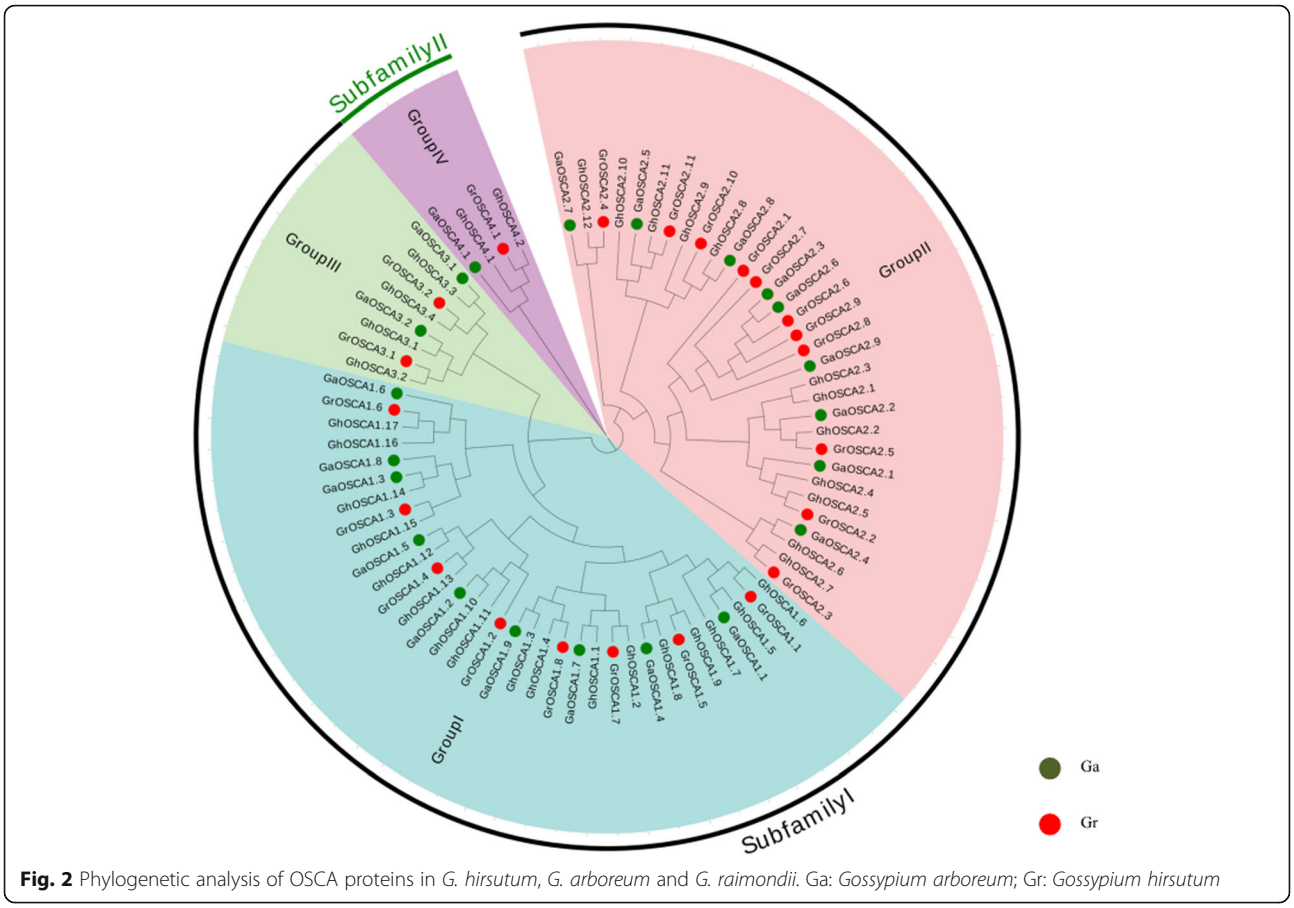


**Fig. 2** Phylogenetic analysis of OSCA proteins in *G. hirsutum*, *G. arboreum* and *G. raimondii*. Ga: *Gossypium arboreum*; Gr: *Gossypium hirsutum*

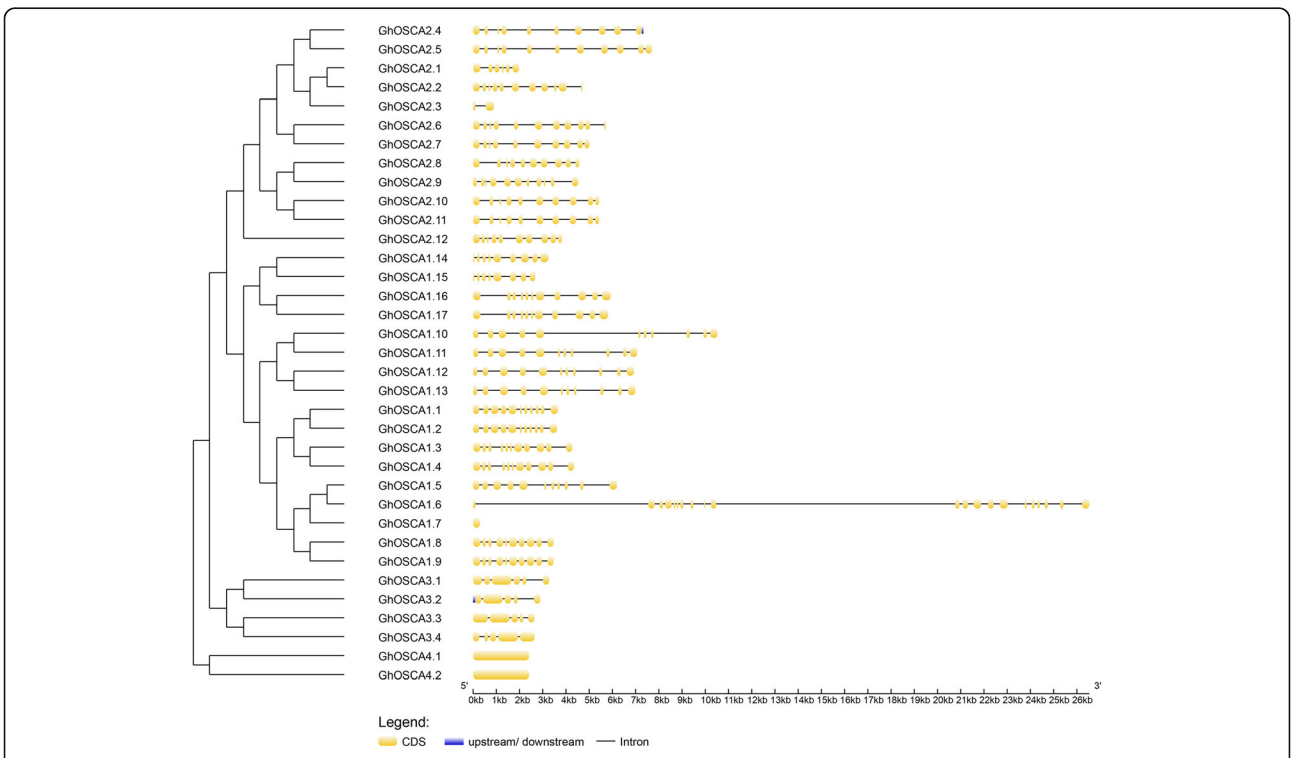


**Fig. 3** The gene structure of all members of *OSCA* gene family in *Gossypium hirsutum*

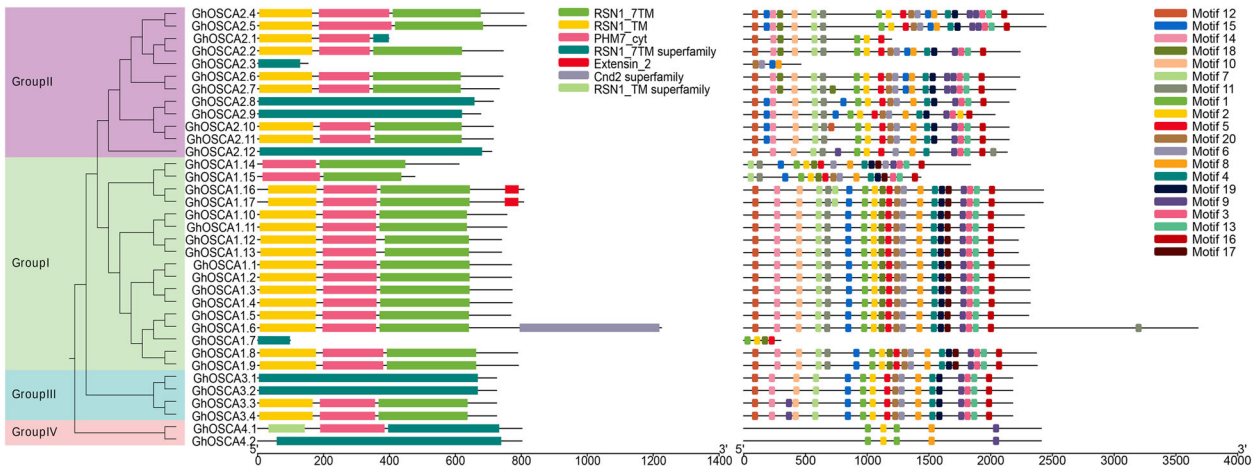


**Fig. 4** The protein structure analysis of all members of *OSCA* gene family in *G. hirsutum*

*OSCA* gene family, most members of the same group have a similar motif, suggesting that there are functional similarities in the same group. By multiple sequence alignment of amino acids, it was found that GhOSCA family protein had a high degree of sequence conservatism, especially calcium-dependent domain channel structure (Fig. 6). The protein sequences in the same group were highly conserved, but there were significant differences between groups, especially the Group IV of subfamily II and the three group sequences of the subfamily.

## Chromosome location and duplication analysis of the *GhOSCA* genes

To examine the genomic distribution of *OSCA* genes in *G. hirsutum* chromosomes, we investigated the chromosomal location of *GhOSCA* (Fig. 5). The result indicated that 31 *GhOSCA* genes were mapped onto 19 chromosomes, while four genes which could not obviously map to any chromosome were named *GhOSCA1.7*, *GhOSCA2.1*, *GhOSCA3.2*, *GhOSCA3.3*, respectively. We found the chromosomal location relatively uneven. Some

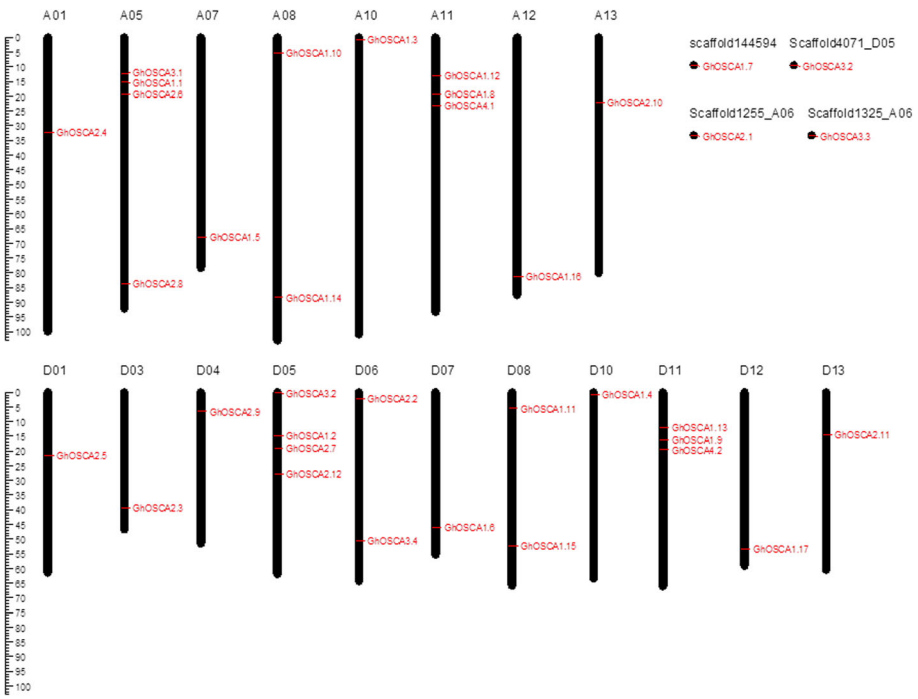


**Fig. 5** The distribution of *OSCA* genes on chromosomes

chromosomes and chromosome regions have a higher density of *GhOSCA* genes while others do not. Fourteen *GhOSCA* genes were located on At-subgenome chromosomes, respectively, on $A_h01$, $A_h05$, $A_h07$, $chrA_h08$, $A_h10$, $A_h11$, $chrA_h12$, $chrA_h13$. *GhOSCA1.7*, *GhOSCA2.1*, *GhOSCA3.2* and *GhOSCA3.3* were mapped to the scaffold, $A_h06$, $D_h05$, $A_h06$, respectively. The remaining *GhOSCA* genes were located in the Dt-subgenome chromosomes. Interestingly, many genes were located in clusters, especially at the top of chromosomes $A_h05$, $A_h11$, $D_h11$. For example, Chromosomes $A_h05$ had the largest number of *GhOSCA* genes, with four members of *GhOSCAs*. This unbalanced distribution of *GhOSCA* genes on chromosomes suggested that genetic variation existed in the evolutionary process.

Tandem and segmental duplication events are the main causes of gene-family expansion in *G. hirsutum*. Two or more genes located on the same chromosome, one following the other, confirms a tandem duplication event, while gene duplication on different chromosomes or within the same chromosome but not one following the other is designated a segmental duplication event. In order to understand potential gene duplication within the *G. hirsutum* genome, we analyzed the occurrence of tandem duplication and segmental duplication during the evolution of this gene family. According to whole genome analysis of gene duplication, we observed that 16 pairs of *GhOSCA* genes originating from segmental duplication, which deeply contributed to the expansion of the *GhOSCA* genes (Table 3). To calculate the evolutionary time of the *GhOSCA* gene family, synonymous ($d_S$) and non-synonymous ($d_N$) values were calculated using PAL2NAL. A $d_S/d_N$ value of 1 suggested neutral selection; a $d_S/d_N$ value of > 1 suggested positive selection; a $d_S/d_N$ value of < 1 suggested purifying selection. We found that all *GhOSCA* genes had $d_S/d_N$ values of less than 1, indicated that *GhOSCA* genes have evolved under the effect of purifying selection (Table 3).

### *Cis*-regulatory element analysis in the promoter regions of *GhOSCA* genes

An extensive analysis of 1 500 bp upstream promoter region of *GhOSCA* genes, we found that *cis*-regulatory element included ABA-responsive elements (ABREs), low-temperature responsive elements (LTRs), defense and stress-responsive elements (TC-rich repeats), salicylic acid responsive elements (TCA-elements), heat stress-responsive elements (HSEs), MeJA-responsive elements (TGACG-motifs and CGTCA-motifs), MYB-binding sites (MBS) (Table 4). However, ABREs, TCA-elements, and TGACG-motifs belong to plant hormone-responsive elements. ABREs, TCA-elements, and TGAC G-motifs are involved in ABA, SA and MeJA responsiveness, respectively. TCA-elements are the most abundant *cis*-regulatory hormone responsive element in the promoters of *GhOSCA* genes, as 27 gene members contained TCA-elements. Both CGTCA-motifs and TGAC G-motifs were involved in the SA reaction. In total, 17 members contained ABRE-elements. The other important type of *cis*-regulatory elements in the upstream regions of *GhOSCA* genes are the environmental stress-related elements. In total, four types of elements were

**Table 3** The dN/ds values for duplicate *GhOSCA* genes

| Paralogous | Amino acid sequence Identities (%) | $d_S$ | $d_N$ | $d_N/d_S$ | Duplicate | Purifying selection |
|---|---|---|---|---|---|---|
| GhOSCA1.1/1.2 | 98.7 | 0.018 8 | 0.006 2 | 0.329 1 | Segmental | Yes |
| GHOSCA1.3/1.4 | 98.96 | 0.069 4 | 0.005 6 | 0.080 3 | Segmental | Yes |
| GhOSCA1.5/1.6 | 95.31 | 0.067 5 | 0.034 2 | 0.506 9 | Segmental | Yes |
| GhOSCA1.8/1.9 | 97.08 | 0.029 3 | 0.012 9 | 0.440 4 | Segmental | Yes |
| GhOSCA1.10/1.11 | 98.55 | 0.019 3 | 0.006 9 | 0.359 | Segmental | Yes |
| GhOSCA1.12/1.13 | 97.16 | 0.019 8 | 0.013 6 | 0.687 1 | Segmental | Yes |
| GhOSCA1.14/1.15 | 89.1 | 0.048 3 | 0.037 5 | 0.778 2 | Segmental | Yes |
| GhOSCA1.16/1.17 | 97.89 | 0.027 5 | 0.008 | 0.292 3 | Segmental | Yes |
| GhOSCA2.1/2.2 | 98.24 | 0.007 8 | 0.007 7 | 0.988 5 | Segmental | Yes |
| GhOSCA2.4/2.5 | 98.76 | 0.026 8 | 0.004 5 | 0.168 6 | Segmental | Yes |
| GhOSCA2.6/2.7 | 96.73 | 0.039 6 | 0.018 4 | 0.465 1 | Segmental | Yes |
| GhOSCA2.8/2.9 | 94.24 | 0.079 1 | 0.029 8 | 0.376 3 | Segmental | Yes |
| GhOSCA2.10/2.11 | 98.6 | 0.024 3 | 0.007 | 0.288 6 | Segmental | Yes |
| GhOSCA3.1/3.2 | 98.48 | 0.021 3 | 0.007 3 | 0.344 4 | Segmental | Yes |
| GhOSCA3.3/3.4 | 98.34 | 0.023 | 0.008 3 | 0.359 4 | Segmental | Yes |
| GhOSCA4.1/4.2 | 99.25 | 0.019 1 | 0.004 2 | 0.219 1 | Segmental | Yes |

Note: $d_S$: synonymous values, $d_N$: non-synonymous

**Table 4** The *cis*-regulatory element analysis of *GhOSCA* promoters

| Gene name | ABRE | LTR | CGTCA-motif | TCA-element | TGAGG--motif | TC-rich repeats | HSE | MBS |
|---|---|---|---|---|---|---|---|---|
| GhOSCA1.1 | 4 | 3 | 0 | 1 | 0 | 2 | 0 | 1 |
| GhOSCA1.2 | 4 | 4 | 1 | 2 | 1 | 0 | 0 | 1 |
| GhOSCA1.3 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 1 |
| GhOSCA1.4 | 0 | 0 | 0 | 1 | 0 | 3 | 2 | 2 |
| GhOSCA1.5 | 0 | 2 | 1 | 1 | 1 | 1 | 1 | 0 |
| GhOSCA1.6 | 0 | 2 | 1 | 2 | 1 | 1 | 1 | 0 |
| GhOSCA1.8 | 0 | 0 | 1 | 3 | 1 | 1 | 4 | 2 |
| GhOSCA1.9 | 1 | 0 | 0 | 2 | 0 | 1 | 1 | 2 |
| GhOSCA1.10 | 0 | 0 | 5 | 1 | 5 | 2 | 3 | 2 |
| GhOSCA1.11 | 0 | 0 | 3 | 2 | 3 | 1 | 5 | 1 |
| GhOSCA1.12 | 2 | 2 | 0 | 1 | 0 | 1 | 2 | 0 |
| GhOSCA1.13 | 0 | 2 | 0 | 3 | 0 | 0 | 4 | 1 |
| GhOSCA1.14 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 2 |
| GhOSCA1.15 | 0 | 0 | 1 | 2 | 1 | 1 | 0 | 1 |
| GhOSCA1.16 | 0 | 1 | 1 | 5 | 1 | 1 | 2 | 1 |
| GhOSCA1.17 | 0 | 0 | 1 | 2 | 1 | 1 | 1 | 1 |
| GhOSCA2.1 | 0 | 0 | 3 | 0 | 3 | 1 | 3 | 0 |
| GhOSCA2.2 | 3 | 1 | 2 | 1 | 2 | 3 | 7 | 2 |
| GhOSCA2.3 | 1 | 0 | 0 | 4 | 0 | 2 | 3 | 2 |
| GhOSCA2.4 | 8 | 1 | 0 | 2 | 0 | 4 | 4 | 0 |
| GhOSCA2.5 | 8 | 1 | 0 | 1 | 0 | 3 | 2 | 0 |
| GhOSCA2.6 | 0 | 0 | 0 | 3 | 0 | 1 | 1 | 4 |
| GhOSCA2.7 | 1 | 0 | 0 | 4 | 0 | 1 | 0 | 3 |
| GhOSCA2.8 | 0 | 0 | 0 | 3 | 0 | 2 | 3 | 2 |
| GhOSCA2.9 | 0 | 0 | 0 | 2 | 0 | 3 | 2 | 1 |
| GhOSCA2.10 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 |
| GhOSCA2.11 | 0 | 1 | 1 | 3 | 1 | 3 | 2 | 0 |
| GhOSCA2.12 | 2 | 0 | 1 | 1 | 1 | 0 | 3 | 0 |
| GhOSCA3.1 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 |
| GhOSCA3.2 | 6 | 0 | 1 | 2 | 1 | 4 | 3 | 1 |
| GhOSCA3.3 | 6 | 0 | 1 | 0 | 1 | 0 | 4 | 0 |
| GhOSCA3.4 | 5 | 0 | 0 | 0 | 0 | 0 | 4 | 0 |
| GhOSCA4.1 | 0 | 2 | 2 | 0 | 2 | 2 | 2 | 1 |
| GhOSCA4.2 | 0 | 2 | 2 | 0 | 2 | 2 | 1 | 2 |

Note: *ABRE*: ABA responsive elements, *LTR*: Low-temperature-responsive, *TC-rich repeats*: Defense and stress-responsiveness elements, *TCA-elements*: Salicylic acid responsive elements, *HSEs*: Heat stress-responsive elements, *TGACG-motifs and CGTCA-motifs*: MeJA-responsive elements, *MBS*: MYB-binding sites

found that respond to four respective kinds of external environmental stresses. These were low-temperature-responsive (LTR), stress-responsive TC-rich repeats, heat-stress-responsive (HSEs) and drought responsive (MBSs). In total, 30 members contained TC-rich; 32 members contained HSEs; 26 members contained MBSs; and 17 members contained LTR-element. Among them, HSEs are the most enriched *cis*-regulatory element in all promoter sequences. We surmised that external environmental stress could induce the expression of *GhOSCA* genes through its response *cis*–regulatory element and further improve the resistance of plants to environmental stress.

### Expression profiling of the *GhOSCA* genes under drought and salinity stress conditions

Gene expression pattern is usually related to the gene's function. Previous studies have indicated that the *OSCA* gene plays an essential role in plant growth

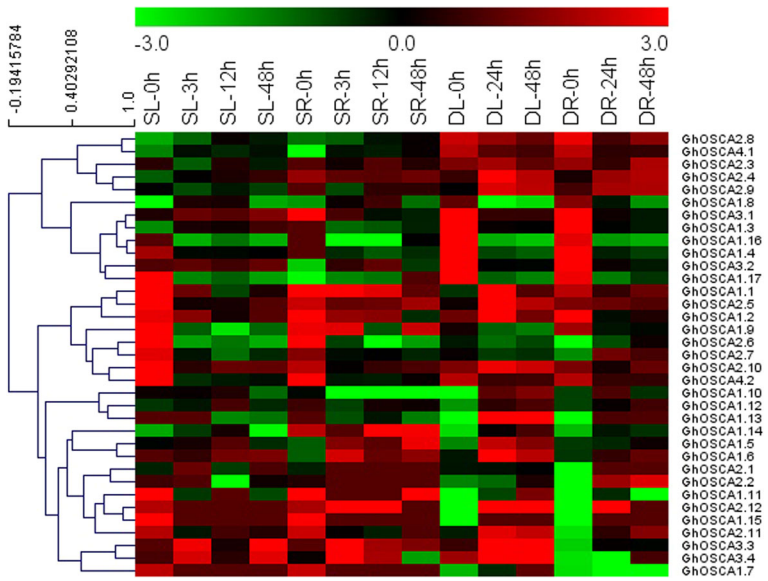


**Fig. 6** Expression analysis of *GhOSCA* genes in *G. hirsutum* under salt and drought stresses. The RNA-Seq expression profiles of *G. hirsutum* were used to identify the relative expression levels of *GhOSCA* genes. Levels of gene expression are depicted in different colors on the scale. Red color represents high expression and green color represents low expression

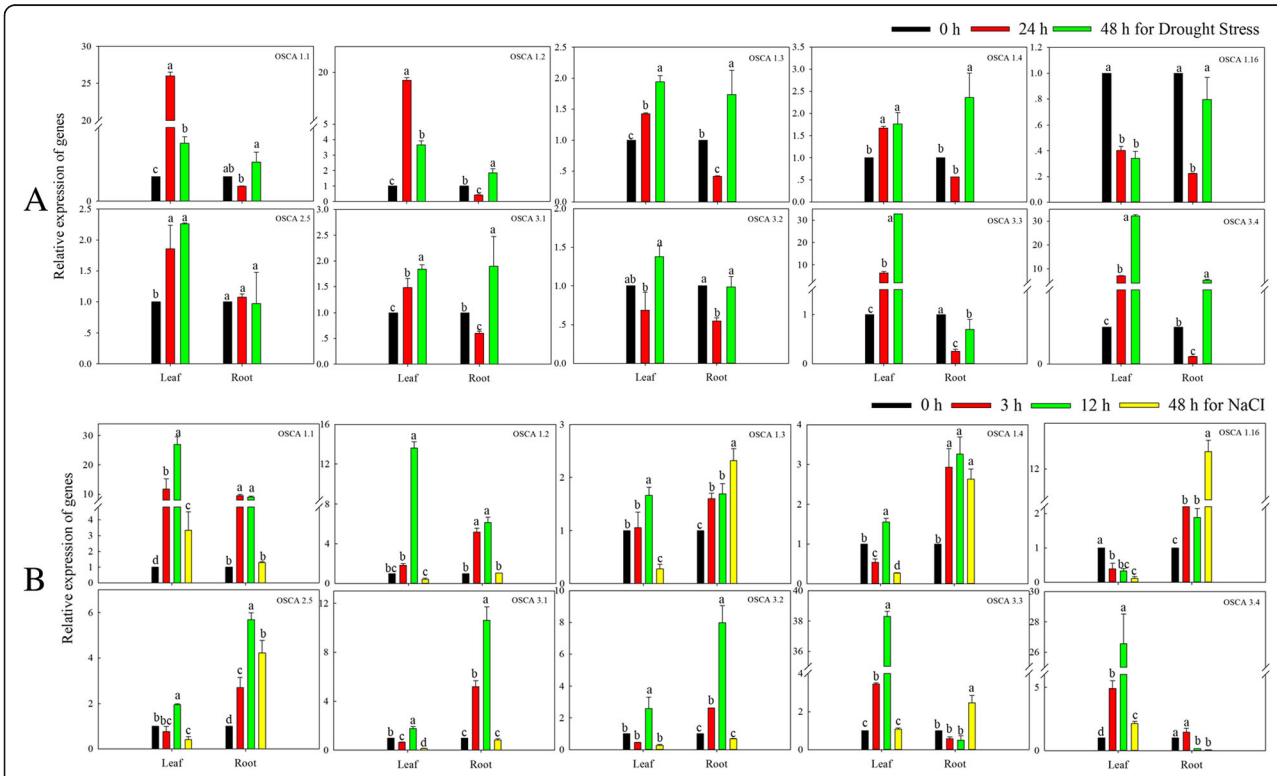


**Fig. 7** Expression analysis of 10 selected *GhOSCA* genes using quantitative real-time RT-PCR (RT-qPCR). (**a**) RT-qPCR analysis of the selected *GHOSCA* genes under drought stress conditions, imposed by adding 17% of PEG-6000. (**b**) RT-qPCR analysis of the selected *GHOSCA* genes under salt stress conditions, imposed by adding 300 mM of NaCl solution. The relative expression level of 10 selected *GhOSCA* genes was normalized to the reference gene histone 2 in different tissues. The transcripts in non-stressed was set as 1 for each gene in different tissues. The bars show the standard deviation of three technical repeats. Different letters indicate significant differences in the expression levels of the genes in tissues at different time, 0 h, 24h and 48h of drought stress exposure, while for salt stress conditions, samples were taken at 0h, 3h, 12h and 48h of post salt stress exposure (ANOVA; $P < 0.05$). 0 h: normal conditions

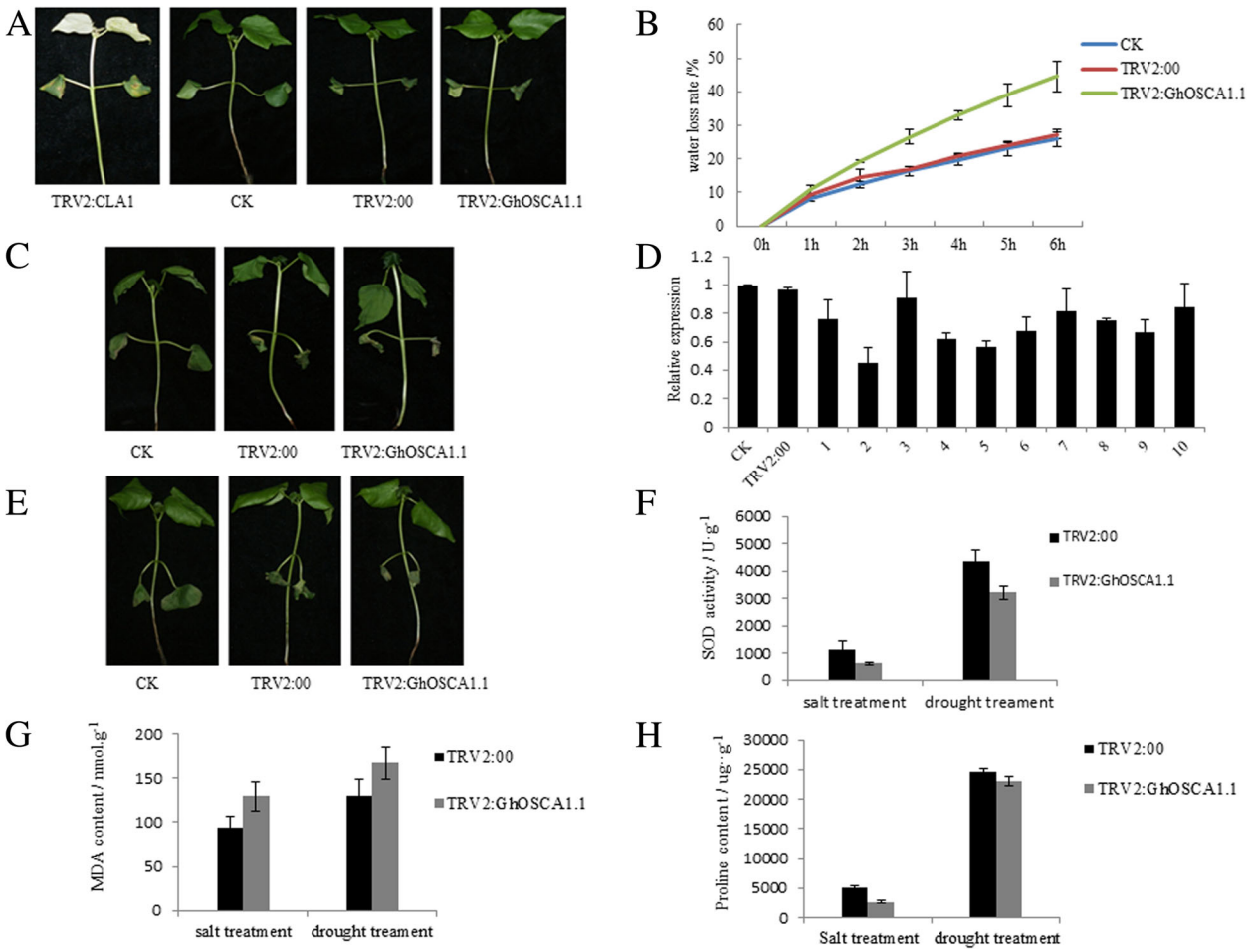


**Fig. 8** VIGS validates the function of *GhOSCA1.1* gene. **a:** The phenotypes of TRV2:CLA1, CK, TRV2:00 and TRV2:*GhOSCA1.1* seedlings, **b:** The water loss rate of CK, TRV2:00 and TRV2:*GhOSCA1.1* seedlings. **c:** the phenotypes of CK, TRV2:00 and TRV2:GhOSCA1.1 seedlings were observed 48 h after 17%PEG treatment. **d:** The silencing efficiency of GhOSCA1.1 gene in seedlings. **e:** The phenotypes of CK, TRV2:00 and TRV2:*GhOSCA1.1* seedlings were observed 48 h after salt stress treatment. **f:** The activity of SOD in TRV2:00 and TRV2:*GhOSCA1.1* seedlings after salt and drought stress treatment. **g:** The contents of MDA in TRV2:00 and TRV2:*GhOSCA1.1* seedlings after salt and drought stress treatment. **h:** The contents of PRO in TRV2:00 and TRV2:*GhOSCA1.1* seedlings after salt and drought stress treatment

and development. To understand the expression profiles of these 35 *GhOSCA* genes in *G. hirsutum*, we used transcriptome data to assess the expression pattern under salt and drought stress. In the environment of drought and salt stress, different genes showed different expression patterns in the roots and leaves (Fig. 6). The analysis revealed that 16 *GhOSCA* genes (*GhOSCA1.1/1.2/1.3/1.4/1.5/1.6/1.16/2.4/2.5/2.9/2.10/2.11/3.1/ 3.2/3.3/3.4*) responded to salt and drought stresses, whereas the expression of other genes was not significantly altered under different stresses. Of which 7 *GhOSCA* genes (*GhOSCA1.1/1.2/2.5/3.3/3.4/4.1/4.2*) were notably up-regulated under salt and drought treatment based on the transcriptome data, and were selected for further analysis by RT-qPCR (Fig. 7).

Under salt stress, some of the *GhOSCA* genes were found to exhibit a moderately high expression level in root and leaf tissues. In contrast, *GhOSCA1.1* and *GhOSCA1.2* transcript levels were higher in roots. Moreover, *GhOSCA2.2* and *GhOSCA2.1* exhibited significantly higher levels of expression in roots, whereas in leaves it showed very low expression. However, two genes, *GhOSCA3.1* and *GhOSCA3.2* displayed an up-regulation tissues of all plant materials analysed. Moreover, *GhOSCA1.3* and *GhOSCA1.4* were significantly up-regulated in roots, while *GhOSCA4.1* and *GhOSCA4.2* were not significantly expressed under salt stress.

The number of genes induced by drought treatment was higher than in salt treatment, and they showed different expression levels. Here, we found that most *GhOSCA* genes were up-regulated in all organs except *GhOSCA1.3, GhOSCA 1.4, GhOSCA 1.8, GhOSCA 1.9, GhOSCA 1.14, GhOSCA 1.16,* and *GhOSCA 1.17* which were down-regulated in most tissues. Moreover,

*GhOSCA3.3* and *GhOSCA3.4* were highly up-regulated in leaves, but exhibited differential expression pattern on root tissues. However, *GhOSCA1.16* and *GhOSCA1.8* were significantly up-regulated in leaves, but *GhOSCA3.1* and *GhOSCA3.2* showed insignificantly expression under drought stress.

### Increased salt and dehydration stress sensitivity in the *GhOSCA1.1* virus-induced gene silenced plants

To further investigate the functions of *GhOSCA1.1*, specific primers were designed for reverse genetics by adopting virus induced gene silencing (VIGS) method. Agrobacterium strain of LBA4404 was transformed with three vectors, TRV2:*CLA1*, TRV: 00 and TRV2:*GhOSCA1.1*, respectively. A relatively tolerant upland cotton, MAR85 was used, the vector containing the knocked gene, and the positively controlled vector (TRV:00) were infused to the seedlings cotyledons, and were allowed to grow under normal conditions till the emergence of the third true leaf under hydroponic condition. The plants infused with an albino mutant designated *CLA1–1* (for "cloroplastos alterados", or "altered chloroplasts") showed albino-like traits on their leaves. The CLA1–1 plants behave like wild-type in their capacity to etiolate and produce anthocyanins indicating that the light signal transduction pathway seems to be unaffected (Estévez et al. 2002). Albino leaves were observed in TRV2:*CLA1* inoculated seedlings after 7 days of inoculation (Fig. 8a). The appearance of the albino-like trait showed that the vector used was effective, and the results were in agreement to previous findings in which PDS has been used to monitor the effectiveness of the vector in the knockdown of cytochrome P450 genes in upland cotton (Magwanga et al. 2019b). The VIGS plants, the positively controlled and the wild types were exposed to drought and salt stress, and the VIGS plants ability to tolerate the effects of drought and salt stress were highly compromised. There was significantly higher rate of water loss on the leaves of *GhOSCA1.1* gene-silenced plants compared with the wild types and the positively controlled plants, the TRV2:00 infused plants (Fig. 8b). This result indicated that *GhOSCA1.1* gene might be related to drought resistance. The expression level of *GhOSCA1.1* was checked by RT-qPCR. Compared with TRV2:00 seedlings, the expression level of *GhOSCA1.1* was up-regulated in 10 (Ganie et al. 2017) gene-silencing seedlings after 20 days of inoculation (Fig. 8c). The difference was not observed between infected seedlings. This result suggested that lower expression levels of *GhOSCA1.1* could not alter the growth and development of cotton. Then, WT, TRV2:00 and TRV2:*GhOSCA1.1* seedlings were exposed to salt stress (300 $mmol{\cdot}L^{-1}$ NaCl) and dehydration stress. The leaves of TRV2:*GhOSCA1.1* seedlings were withered and wilting, compared with WT and TRV2:00 seedlings after 2 days of salt stress treatment (Fig. 8d). A similar morphological character was observed after dehydration stress (Fig. 8e). Additionally, compared with WT and TRV2:00 seedlings after 2 days of salt and drought stress treatment, the dehydration rate, proline, and the SOD content were significantly lower in the VIGS plants. On the contrary, the MDA was higher in TRV2:*GhOSCA1.1* seedlings (Fig. 8f). The higher concentration levels of the MDA in the leaf tissues of VIGS plants showed that the plants suffered more of oxidative stress compared with the wild types and the positively controlled plant under drought and salt stress conditions. The results obtained were in agreement with the previous findings in which the *Gh_A05G2067* (*GT-2*) knocked out plants registered higher concentration levels of MDA, hydrogen peroxide and significant reduction on the concentration level of catalase (CAT), peroxidase (POD) (Magwanga et al. 2019a). Therefore, these results suggested that *GhOSCA1.1* gene may improve salt and drought tolerance of cotton.

## Discussion

### Effects of abiotic stress on cotton growth and yield quality, and their response mechanism

Xinjiang has become the largest cotton planting area in China, but the soil salinity and water shortage are serious stresses, which greatly limit the production and improvement of cotton fiber quality and yield (Zhang et al. 2014). Therefore, surveying the endogenous salt-resistant genes in the whole genome of *Gossypium* is a practical and imperative way to provide a resource for further enhancing the salt and drought stress resistance. In the long evolutionary process, plants have evolved some shared biological processes in response to abiotic and biotic stress (Ahmed et al. 2013; Bihmidine et al. 2014; Podia et al. 2018; Qiu et al. 2011; Reguera et al. 2014; Shavrukov 2012). For instance, salt and drought stresses both induce osmotic stress in the plant (Shavrukov 2012). Similarly, homeostasis of cellular osmotic is responsible for ensuring that cotton grows and develops normally under salt and drought stress (Shi et al. 2014; Zhang et al. 2014). In previous studies, AtOSCA was found to be involved in osmotic stress response as a hyperosmolality gated calcium-permeable channel in *Arabidopsis thaliana* (Yuan et al. 2014). Moreover, the AtOSCA protein contains a conservative trans-membrane domain, which was also found among the *G. hirsutum* OSCA protein. These discoveries provide a new insight to investigate the *OSCA* gene family of *G. hirsutum* under salt and drought stress. Furthermore, carrying out the expression analysis of the *GhOSCAs* genes

under salt and dehydration stresses will facilitate selection of the potential target genes.

### Phylogenetic analysis of the proteins encoded by the *OSCA* genes in cotton and other plants

Upland cotton provides the largest natural fiber for the textile industry in the world. *G. hirsutum*, allotetraploid upland cotton, contains A-subgenome and D-subgenome. Gossypium, dicotyledonous plants, diverged from its relatives approximately 10–15 million years ago (MYA). Researchers thought that *G. arboreum* and *G. raimondii* are the donor species of A-subgenome and D-subgenome, respectively. The allopolyploid kinds of cotton emerged about 1–2 MYA due to an intergenomic hybridization event between A and D genomes (Flagel et al. 2012; Senchina et al. 2003; Shan et al. 2005). Therefore, studying the phylogenetic relationship of OSCAs in *G. arboreum, G. raimondii* and *G. hirsutum* will enhance the understanding of *OSCA* gene family diversification during the history of evolution and domestication. *OSCA* genes of dicotyledonous plant cotton, Arabidopsis, and monocotyledonous plant rice were divided into four clusters, which were named Group I-IV based on the phylogenetic tree (Fig. 1). This result is consistent with previous studies (Li et al. 2015; Yuan et al. 2014). Interestingly, every group included OSCAs of cotton, Arabidopsis, and rice, and OSCAs of dicotyledonous cotton and Arabidopsis were clustered closer than OSCAs of the monocotyledonous plant rice, which indicated that OSCA family Group I-IV split long before the separation of cotton, Arabidopsis and rice. What is more, *G. hirsutum* D-subgenome and *G. raimondii* have the closest relationship, and *G. hirsutum* A-subgenome and *G. arboreum* have the closest relationship, which further supported *G. arboreum* and *G. raimondii* is the donor species of A-subgenome and D-subgenome, respectively. The exception to this is that *GrOSCA2.1*, *GrOSCA2.6*, *GrOSCA2.7*, *GaOSCA2.3*, *GaOSCA2.6*, *GaOSCA2.9*, *GaOSCA2.8*, and *GaOSCA2.9* do not have a close relationship with any *OSCA* family gene of *G. hirsutum*. This result suggested gene-losing events haven occurred during the forming of allotetraploid upland cotton.

### Gene structure, *cis*-regulatory element and gene expression analysis

Protein structure and gene structure are closely related to gene function. Previous studies have shown that *OSCA* genes in most higher plants contain three conserved domains, namely late exocytosis (Pfam13967), cytosolic domain of 10 TM putative phosphate transporter (Pfam14703, DUF4463) and calcium-dependent channel (Pfam02714, DUF221)(Yuan et al. 2014). In this study, *GhOSCA1.7*, *GhOSCA2.1*, *GhOSCA2.3*, *GhOSCA2.12*, *GhOSCA2.8*, *GhOSCA2.9*, *GhOSCA3.1*, *GhOSCA3.2*, *GhOSCA4.1* and *GhOSCA4.2* which contain RSN1_7TM superfamily domain, without the RSN1_7TM domain. In addition, due to the long intron length of *GhOSCA1.6*, gene length (26.5 Kb) is much larger than other genes of the *OSCA* gene family in *G. hirsutum* and *GhOSCA1.6* contain a long Cnd2 super family domain. Those results suggested a more complex function of *GhOSCA1.6*. On the contrary, *OSCA1.1-OSCA1.5* protein structures were similar to that of AtOSCA, which suggested that these five *OSCA* genes were putatively involved in osmotic stress response as a hyperosmolality gated calcium-permeable channel. Furthermore, we found the same groups of the *GhOSCAs* to have similar gene structure, suggested the most conserved duplication events occurred during the *OSCA* gene family expansion in the same group.

Gene expression patterns can provide important clues to gene function, which is thought to be related to the differentiation of promoter regions (Xue et al. 2008). *Cis*-regulatory regulatory elements contained in gene' promoter regions play key role in conferring the developmental and environmental regulation of gene expression. In this research, members of the *OSCA* gene family contain a variety of environmental stress response elements, which can improve stress tolerance. There are more elements related to drought and ABA reaction, and fewer elements related to salt reaction. Based on the transcriptome results, we can find that *GhOSCA1.1*, *GhOSCA1.9*, *GhOSCA1.14*, *GhOSCA1.1*, *GhOSCA2.12* were up-regulated significantly, but analysis of *cis*-regulatory elements found that they did not contain saline-alkali stress response element. This result indicates that when plants are under saline-alkali stress, they induce the expression of other stress responsive elements, or hormone responsive elements, so to regulate the gene expression thereby improving their tolerance to saline-alkali stress.

### Knockdown of novel *OSCA* gene reveals their putative role in enhancing drought and salt stress in cotton

Dehydration and salt stress limited the cotton yield, although cotton is a typical plant with abiotic stress tolerance (Van Iersel and Oosterhuis 1996; Watanabe et al. 2000). Osmotic stress is an important phase to dehydration and salt stress response (Yuan et al. 2014). In the previous study, Osmoregulation occured during turgor-driven cell expansion of developing cotton fibers (Smart et al. 1998). Previously, $Ca^{2+}$ and calmodulin-dependent signal pathway regulate salt and dehydration tolerance response in the plant (Pardo et al. 1998; Saijo et al. 2000). Previous studies have shown that *AtOSCA* genes were expression in leaves, flowers, and roots in Arabidopsis (Yuan et al. 2014). In this study, expression levels of *GhOSCA* genes in three different accessions of *G. hirsutum* races were investigated under salt and dehydration stress by

RNA-seq. We found that *GhOSCA* genes expression pattern in the tissues analysis exhibited significant variation, and all the genes exhibited tissue specificity, which indicated that each member of the *GhOSCA* gene family played a specific role in different tissues/organs to regulate osmotic stress. Furthermore, we reconfirmed the transcriptional expression level by RT-qPCR. Interestingly, *GhOSCA1.1*, an orthologous gene pair to *AtOSCA*, was significantly up-regulated under salt and dehydration stress conditions, which demonstrated that *GhOSCA1.1* was a potential gene with significant role in enhancing salinity and dehydration tolerance in cotton.

TRV2 vector of *GhOSCA1.1* was constructed to investigate the salt and dehydration stress regulation by VIGS. The *GhOSCA1.1*-gene-silenced plant showed obvious wilting. Statistical analysis showed that the rate of water loss gradually increased VIGS-plants compared with their wild types. In particular, TRV2:*GhOSCA1.1* seedlings showed a significantly higher rate of water loss and MDA concentration after drought stress exposure, but lower SOD and POD activity than controlled and the TRV:00 infused seedlings, which indicated that the sensitivity of TRV2:*GhOSCA1.1* seedlings to drought and/or salt stresses was increased after post dehydration and salt stress treatment.

## Conclusions

A total of 78 *OSCA* genes were identified in the three cotton species, in which 35, 21 and 22 proteins encoded by the *OSCA* genes were obtained in *G. hirsutum, G. raimondii* and *G. arboreum,* respectively. The genes phylogenetically grouped into four groups, which were in agreement with the previous findings. The physiochemical properties of the proteins encoded by the *OSCA* genes showed that the majority of the protein encoded by the *OSCA* genes in cotton ranged from − 0.245 to 0.706, which implied their GRAVY values were less than 1, and thus were hydrophobic in nature. Moreover, segmental duplication was found to be the major evolutionary mechanism underlying the duplication of the various *OSCA* genes in cotton. RT-qPCR analysis of the *G. hirsutum OSCA* genes under drought and salinity stress conditions, showed that *Gh_A05G1480* (*GhOSCA1.1)* is evident by higher concentration levels of MDA and significant reduction in SOD and proline under drought and salt stress conditions, but when the gene was knocked down, the VIGS-plants showed increased sensitivity to drought and salt stress conditions. This study provides the first systematic analysis of *OSCAs* in cotton and provides a new insight of defense responses in general and lays the foundation for future crop improvement.

### Abbreviations

MDA: Malondialdehyde; OSCA: Hyper osmolality-gated calcium-permeable channels; PRO: Proline; SOD: Superoxide Dismutase; VIGS: Virus-induced gene silencing

### Acknowledgements

The authors are grateful for the help of the teachers and students of the Wild Cotton Project of the Institute of Cotton Research of Chinese Academy of Agricultural Sciences, particularly to Dr. Magwanga RO, for his help in proofreading and language editing of the manuscript.

### Authors' contributions

Zhou ZL, Liu F designed the study. Yang X and Xu YC wrote the manuscript and prepared Figs. 1, 2, 3, 4, 5, 6, 7 and 8 and Tables 1, 2, 3 and 4. Magwanga RO revised the manuscript; Yang X, Yang FF, Cai XY, Wang YH, Hou YQ carried out the experimental work and Yang X, Xu YC, Wang XX analyzed data. All authors reviewed the manuscript. All authors read and approved the final manuscript.

### Funding

This work was funded by the National Natural Science Foundation of China (31530053/31621005) and the National Key R&D Program (2016YFD0101401 /2017YFD0101601).

### Availability of data and materials

Not applicable.

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

[1]Institute of Cotton Research, Chinese Academy of Agricultural Science (ICR, CAAS) / Research Base in Anyang Institute of Technology, State Key Laboratory of Cotton Biology, Anyang 455000, China. [2]School Of Physical and Biological Sciences (SPBS), Main Campus, Jaramogi Oginga Odinga University of Science and Technology (JOOUST), 210-40601, Bondo, Kenya.

Received: 5 March 2019 Accepted: 10 July 2019
Published online: 31 July 2019

### References

Ahmed IM, Cao F, Zhang M, et al. Difference in yield and physiological features in response to drought and salinity combined stress during anthesis in Tibetan wild and cultivated barleys. PLoS One. 2013;8(10):e77869. https://doi.org/10.1371/journal.pone.0077869.

Árnadóttir J, Chalfie M. Eukaryotic mechanosensitive channels. Annu Rev Biophys. 2010;39:111–37. https://doi.org/10.1146/annurev.biophys.37.032807.125836.

Bihmidine S, Cao M, Kang M, et al. Expression of chlorovirus *MT325* aquaglyceroporin (aqpv1) in tobacco and its role in mitigating drought stress. Planta. 2014;240(1):209–21. https://doi.org/10.1007/s00425-014-2075-5.

Camargo SR, Cancado GMA, Ulian EC, et al. Identification of genes responsive to the application of ethanol on sugarcane leaves. Plant Cell Rep. 2007;26:2119. https://doi.org/10.1007/s00299-007-0430-8.

Da SR, Ploy M-C. Resistance acquisition via the bacterial SOS response: the inducive role of antibiotics. Medecine sciences: M/S. 2012;28(2):179–84. https://doi.org/10.1051/medsci/2012282016.

Deng F, Zhang X, Wang W, et al. Identification of *Gossypium hirsutum* long non-coding RNAs (lncRNAs) under salt stress. BMC Plant Biol. 2018;18(1):23. https://doi.org/10.1186/s12870-018-1238-0.

Estévez JM, Cantero A, Romero C, et al. Analysis of the expression of *CLA1*, a gene that encodes the 1-Deoxyxylulose 5-phosphate synthase of the 2- *C* -methyl-d-Erythritol-4-phosphate pathway in Arabidopsis. Plant Physiol. 2002; 124:95–104. https://doi.org/10.1104/pp.124.1.95.

Fan K, Li F, Chen J, et al. Asymmetric evolution and expansion of the NAC transcription factor in polyploidized cotton. Front Plant Sci. 2018;9:47. https://doi.org/10.3389/fpls.2018.00047.

Flagel LE, Wendel JF, Udall JA. Duplicate gene evolution, homoeologous recombination, and transcriptome characterization in allopolyploid cotton. BMC Genomics. 2012;13(1):302. https://doi.org/10.1186/1471-2164-13-302.

Ganie SA, Pani DR, Mondal TK. Genome-wide analysis of DUF221 domain-containing gene family in *Oryza* species and identification of its salinity stress-responsive members in rice. PLoS One. 2017;12(8):e0182469. https://doi.org/10.1371/journal.pone.0182469.

Gu XY, Wang P, Liu Z, et al. Genome-wide identification and expression analysis of the OSCA gene family in *Pyrus bretschneideri*. Canadian Journal of Plant Science. 2018;98(4):918-29. https://doi.org/10.1139/cjps-2017-0115.

Gupta K, Jha B, Agarwal PK. A dehydration-responsive element binding (DREB) transcription factor from the succulent halophyte *Salicornia brachiata* enhances abiotic stress tolerance in transgenic tobacco. Mar Biotechnol. 2014;16(6):657–73. https://doi.org/10.1007/s10126-014-9582-z.

Hu B, Jin J, Guo AY, et al. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2014;31(8):1296–7. https://doi.org/10.1093/bioinformatics/btu817.

Ishitani M, Liu J, Halfter U, et al. SOS3 function in plant salt tolerance requires N-myristoylation and calcium binding. Plant Cell. 2000;12(9):1667–78. https://doi.org/10.1105/tpc.12.9.1667.

Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K. 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 Mol Biol. 1994;25(5):791–8. https://doi.org/10.1007/BF00028874.

Knight H, Trewavas AJ, Knight MR. Calcium signalling in *Arabidopsis thaliana* responding to drought and salinity. Plant J. 1997;12(5):1067–78. https://doi.org/10.1046/j.1365-313X.1997.12051067.x.

Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–4. https://doi.org/10.1093/molbev/msw054.

Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23(21):2947–8. https://doi.org/10.1093/bioinformatics/btm404.

Li Y, Yuan F, Wen Z, et al. Genome-wide survey and expression analysis of the OSCA gene family in rice. BMC Plant Biol. 2015;15:261. https://doi.org/10.1186/s12870-015-0653-8.

Liu J, Guo WQ, Shi DC. Seed germination, seedling survival, and physiological response of sunflowers under saline and alkaline conditions. Photosynthetica. 2010;48(2):278–86. https://doi.org/10.1007/s11099-010-0034-3.

Magwanga RO, Kirungu JN, Lu P, et al. Genome wide identification of the trihelix transcription factors and overexpression of *Gh_A05G2067* (*GT-2* ), a novel gene contributing to increased drought and salt stresses tolerance in cotton. Physiol Plant. 2019a: 12920. https://doi.org/10.1111/ppl.12920.

Magwanga RO, Lu P, Kirungu JN, et al. Characterization of the late embryogenesis abundant (LEA) proteins family and their role in drought stress tolerance in upland cotton. BMC Genet. 2018;19:6. https://doi.org/10.1186/s12863-017-0596-1.

Magwanga RO, Lu P, Kirungu JN, et al. Knockdown of cytochrome P450 genes *Gh_D07G1197* and *Gh_A13G2057* on chromosomes D07 and A13 reveals their putative role in enhancing drought and salt stress tolerance in *Gossypium hirsutum*. Genes (Basel). 2019b;10:226. https://doi.org/10.3390/genes10030226.

Marchler-Bauer A, Bo Y, Han L, et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res. 2016;45(D1):D200–3. https://doi.org/10.1093/nar/gkw1129.

McAinsh MR, Pittman JK. Shaping the calcium signature. New Phytol. 2009;181(2):275–94. https://doi.org/10.1111/j.1469-8137.2008.02682.x.

Munns R. Genes and salt tolerance: bringing them together. The New Phytologist. 2005;167(3):645–63. https://doi.org/10.1111/j.1469-8137.2005.01487.x.

Nakashima K, Yamaguchi-Shinozaki K. ABA signaling in stress-response and seed development. Plant Cell Rep. 2013;32(7):959–70. https://doi.org/10.1007/s00299–013–1418-1.

Pardo JM, Reddy MP, Yang S, et al. Stress signaling through $Ca^{2+}$/calmodulin-dependent protein phosphatase calcineurin mediates salt adaptation in plants. Proc Natl Acad Sci. 1998;95(16):9681–6. https://doi.org/10.1073/pnas.95.16.9681.

Podia V, Milioni D, Martzikou M, et al. The role of *Arabidopsis thaliana RASD1* gene in ABA-dependent abiotic stress response. Plant Biol. 2018;20(2):307–17. https://doi.org/10.1111/plb.12662.

Qiu L, Wu D, Ali S, et al. Evaluation of salinity tolerance and analysis of allelic function of *HvHKT1* and *HvHKT2* in Tibetan wild barley. Theor Appl Genet. 2011;122(4):695–703. https://doi.org/10.1007/s00122–010–1479-2.

Rai A, Suprasanna P, D'Souza SF, et al. Membrane topology and predicted RNA-binding function of the 'early responsive to dehydration (ERD4)' plant protein. PLoS One. 2012;7(3):e32658. https://doi.org/10.1371/journal.pone0032658.

Reguera M, Bassil E, Blumwald E. Intracellular NHX-type cation/$H^+$ antiporters in plants. Mol Plant. 2014;7(2):261–3. https://doi.org/10.1093/mp/sst091.

Saijo Y, Hata S, Kyozuka J, et al. Over-expression of a single $Ca^{2+}$-dependent protein kinase confers both cold and salt/drought tolerance on rice plants. Plant J. 2000;23(3):319–27. https://doi.org/10.1046/j.1365-313x.2000.00787.x.

Sanchez-Barrena MJ, Martinez-Ripoll M, Zhu JK, et al. SOS3 (salt overly sensitive 3) from *Arabidopsis thaliana*: expression, purification, crystallization and preliminary X-ray analysis. Acta Crystallogr D Biol Crystallogr. 2004;60(7):1272–4. https://doi.org/10.1107/S0907444904008728.

Senchina DS, Alvarez I, Cronn RC, et al. Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. Mol Biol Evol. 2003;20(4):633–43. https://doi.org/10.1093/molbev/msg065.

Shan XH, Liu ZL, Dong ZY, et al. Mobilization of the active MITE transposons *mPing* and *pong* in rice by introgression from wild rice (*Zizania latifolia* Griseb.). Mol Biol Evol. 2005;22(4):976–90. https://doi.org/10.1093/molbev/msi082.

Shavrukov Y. Salt stress or salt shock: which genes are we studying? J Exp Bot. 2012;64(1):119–27. https://doi.org/10.1093/jxb/ers316.

Shi CC, Feng CC, Yang MM, et al. Overexpression of the receptor-like protein kinase genes *AtRPK1* and *OsRPK1* reduces the salt tolerance of *Arabidopsis thaliana*. Plant Sci. 2014;217:63–70. https://doi.org/10.1016/j.plantsci.2013.12.002.

Shinozaki K, Yamaguchi-Shinozaki K. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. Curr Opin Plant Biol. 2000;3(3):217–23. https://doi.org/10.1016/S1369-5266(00)80068-0.

Siaud N, Dubois E, Massot S, et al. The SOS screen in Arabidopsis: a search for functions involved in DNA metabolism. DNA Repair. 2010;9(5):567–78. https://doi.org/10.1016/j.dnarep.2010.02.009.

Smart LB, Vojdani F, Maeshima M, et al. Genes involved in osmoregulation during turgor-driven cell expansion of developing cotton fibers are differentially regulated. Plant Physiol. 1998;116(4):1539–49. https://doi.org/10.1104/pp.116.4.1539.

Taji T, Seki M, Satou M, et al. Comparative genomics in salt tolerance between Arabidopsis and Arabidopsis-related halophyte salt cress using Arabidopsis microarray. Plant Physiol. 2004;135(3):1697–709. https://doi.org/10.1104/pp.104.039909.

Ullah A, Sun H, Hakim, et al. A novel cotton WRKY gene, *GhWRKY6*-like, improves salt tolerance by activating the ABA signaling pathway and scavenging of reactive oxygen species. Physiol Plant. 2018;162(4):439–54. https://doi.org/10.1111/ppl.12651.

Van Iersel MW, Oosterhuis DM. Drought effects on the water relations of cotton fruits, bracts, and leaves during ontogeny. Environ Exp Bot. 1996;36(1):51–9. https://doi.org/10.1016/0098–8472(95)00041–0.

Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 2002;93(1):77–8. https://doi.org/10.1093/jhered/93.1.77.

Watanabe K, Tanaka T, Hotta Y, et al. Improving salt tolerance of cotton seedlings with 5-aminolevulinic acid. Plant Growth Regul. 2000;32(1):97–101. https://doi.org/10.1023/A:1006369404273.

Wu SJ, Ding L, Zhu JK. SOS1, a genetic locus essential for salt tolerance and potassium acquisition. Plant Cell. 1996;8(4):617–27. https://doi.org/10.1105/tpc.8.4.617.

Xue T, Wang D, Zhang S, et al. Genome-wide and expression analysis of protein phosphatase 2C in rice and Arabidopsis. BMC Genomics. 2008;9(1):550. https://doi.org/10.1186/1471-2164-9-550.

Yoshida T, Mogami J, Yamaguchi-Shinozaki K. ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. Curr Opin Plant Biol. 2014;21:133–9. https://doi.org/10.1016/j.pbi.2014.07.009.

Yuan F, Yang H, Xue Y, et al. OSCA1 mediates osmotic-stress-evoked $Ca^{2+}$ increases vital for osmosensing in Arabidopsis. Nature. 2014;514(7522):367. https://doi.org/10.1038/nature13593.

Zhang D, Jiang S, Pan J, et al. The overexpression of a maize mitogen- activated protein kinase gene (*ZmMPK5*) confers salt stress tolerance and induces defence responses in tobacco. Plant Biol. 2014;16(3):558–70. https://doi.org/10.1111/plb.12084.

Zhu G, Li W, Zhang F, et al. RNA-seq analysis reveals alternative splicing under salt stress in cotton, *Gossypium davidsonii*. BMC Genomics. 2018;19(1):73. https://doi.org/10.1186/s12864-018-4449-8.
Zhu JK. Genetic analysis of plant salt tolerance using Arabidopsis. Plant Physiol. 2000;124(3):941–8. https://doi.org/10.1104/pp.124.3.941.
Zhu JK. Salt and drought stress signal transduction in plants. Annu Rev Plant Biol. 2002;53(1):247–73. https://doi.org/10.1146/annurev.arplant.53.091401.143329.
Zhu JK. Abiotic stress signaling and responses in plants. Cell. 2016;167(2):313–24. https://doi.org/10.1016/j.cell.2016.08.029.