BIOINFORMATIC ANALYSES OF OSCA GENE FAMILY IN OPIUM POPPY

ABSTRACT

In environmental conditions such as drought and salinity that cause hyperosmolality in cells, extracellular Ca2+ are transferred into the cell via a group of channel proteins in the membrane, and as a result, the intracellular Ca2+ concentration increases. Calcium ions within the cell are seconder messengers and initiate a series of signal pathways that will form the cell's defense against hyperosmolality. One of the protein families that transport the extracellular Ca2+ to the intracellular region is hyperosmolality-gated calcium-permeable channels (OSCA). Papaver somniferum L., from which opium is obtained from its capsule and oil from its seeds, is an industrial plant. In this study, the OSCA gene family were genome wide analyzed using bioinformatic tools. According to results, 16 PsOSCA proteins were identified and distributed unevenly on ten chromosomes, and located all in the plasma membrane, most in the endoplasmic reticulum and vacuole, while a few were in the extra cellular membrane (PsOSCA3 and -7), Golgi body (PsOSCA2 and -5), and chloroplast (PsOSCA4). All PsOSCA proteins were thermally stable and the hydrophobic. But 9 of them were unstable in a test tube. The promotors of 16 PsOSCA genes had different number and type the light-responsive, environmental stress-responsive, hormone-responsiveness, and development-related cis-regulatory elements, suggesting that they have specialized functions. In additions, most of PsOSCA had unique motif distributions and gene structures, suggesting their specialized diverse functions. PsOSCA9 were the most interacted protein. It interacted with nine different proteins which are responsible for mainly ribonucleotide binding, signaling receptor activity, ion binding, cellular protein modification process, protein phosphorylation, phosphate-containing compound metabolic process, signal transduction, cell surface receptor signaling pathway, response to stimulus, protein serine/threonine kinase activity, cell communication, and cellular anatomical entity. Also, PsOSCA 2 and PsOSCA11 were interacted with A0A4Y7K9T3 which is responsible for immune response-regulating signaling pathway, negative and positive regulation of response to biotic stimulus, regulation of cell death. Consequently, this study can be useful for understanding of stress regulatory mechanisms of the OSCA gene family members in opium poppy.

Keywords: Papaver somniferum, OSCA gene family, bioinformatics

INTRODUCTION

Papaver somniferum (opium poppy) is an industrial plant containing alkaloids such as morphine and codeine and has effects of pain relief, cough suppressor euphoria, sleepiness and addiction. In addition, poppy seeds yield 45-50% oil for human consumption, and this oil is high in vitamin E and other tocopherols (Lewkowitsch and Warburton, 1914). Poppy seeds are produced mostly in five different countries of the world, including India, North Macedonia, Palestine, Serbia and Turkey (Faostat, 2021). Unfortunately, it is inevitable that poppy production and yield will be adversely affected by unfavorable conditions such as drought and salinity as a result of climate change. Therefore, it is very important to understand and develop a plant defense system against adverse environmental conditions.

Hyperosmolality is one of the stress factors that negatively affect plant growth and development. Calcium-permeable channels potentially act as osmosensors under osmotic stress. These channels are opened to allow calcium transportation from extracellular to intracellular (Murthy et. al. 2018). The elevation of the concentration of intracellular calcium triggers serial signal transduction, thus Ca2+ acts as a second messenger system.

The hyperosmolality-gated calcium-permeable channels (OSCA) is one of the ion channels and act as hyperosmotic stress sensor (Liu et. al. 2018). Since the OSCA gene family has potential roles in stress resistance against hyperosmolality in plants, it has been investigated in detail in some plant species (Li et. al. 2015; Yin et. al. 2021; Miao et. al. 2022; Shan et. al. 2023). OSCA proteins contain three conserved domains, namely late exocytosis (pfam13967), cytosolic domain of 10TM putative phosphate transporter (pfam14703), and calcium-dependent channel (pfam02714) (Swarbreck et. al. 2013).

In this study, it was aimed to carry out a comprehensive genome- wide analysis of OSCA gene family in P. somniferum, including identification of family members, phylogenetic relationship, subcellular localization, chromosome localization, gene structural features, cis-elements and protein-protein interactions.

MATERIAL and METHODS

Identification and Chromosomal Localization of OSCA Genes

To identify the OSCA genes, Arabidopsis thaliana OSCA protein sequences thrived from TAIR (https://www.arabidopsis.org/) were utilized as query sequences and blasted (BlastP) against the sorghum genome with an E-value of 1e-5 using NCBI (https://www.ncbi.nlm.nih.gov/). For identifying chromosomal positions, CDS sequences, and gene sequences of OSCA genes, the NCBI database (https://www.ncbi.nlm.nih.gov/) was utilized. The chromosomal distribution of

OSCA genes was mapped using the TBtools program (Chen et al., 2020). The physical and chemical parameters of OSCA proteins were predicted with Protparam (https://web.expasy.org/protparam/).

Analysis of the Phylogeny, Gene Structure, Conserved Protein Motif, and Synteny

Using default settings, CLUSTALW (https://www.genome.jp/toolsbin/clustalw) was used to align the OSCA protein sequences of *A. thaliana* and *S. bicolor* (Larkin et al. 2007). The phylogenetic trees were created using the MEGA-11 software's neighbour-joining approach (Saitou and Nei 1987; Tamura et al. 2021), and a bootstrap consensus tree was derived from 1000 replications (Felsenstein 1985). The MEME programme (http://memesuite.org/tools/meme) was used to obtain the conserved motifs of OSCA proteins. The Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/) was applied to show the introns–exons (Hu et al. 2015). The cis-elements on the promoters were detected using plantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

Identification of Protein-Protein Interactions

The protein-protein interactions were investigated (Szklarczyk et al., 2019) using STRING platform V11.5. Disconnected nodes in the network were concealed using a medium confidence value of 0.400 and no more than 10 interactions to display.

RESULTS and DISCUSSION

The first genomic sequence in *P. somniferum* was published in 2018 (Hu et. al. 2018). With the publication of genome sequencing, genome-wide identification of important stress-related gene families in this species will be possible. A gene family encode a group of genes which have similar amino acids sequence and have related but specialized tasks. In this study, hyperosmolality related gene family, OSCA, were detailed bioinformatic analyzed. Several studies have proved the special tasks of OSCA proteins in plant growth, development and response to abiotic stress (Ding et. al. 2019; Cai et. al. 2022; Chakraborty et. al. 2023; Kaur et. al. 2022). A total of 16 OSCA proteins in opium poppy were detected by searching three conserved Pfam domains (pfam13967, pfam14703, pfam02714), and redundant sequences were eliminated. The OSCA proteins were then numbered from PsOSCA1 to PsOSCA16 according to their chromosomal locations. Previous studies showed that the number of *PsOSCA* genes differed in several species, for example, 15 in Arabidopsis (Liu et. al. 2018), 11 in rice (Li et. al. 2015), 12 in maize (Ding et. al. 2019), 13 in four leguminous species (Chakraborty et. al. 2023), 14 in barley (Cai et. al. 2022), 21 in soybean (Liu, et. al. 2022), 35 in cotton (Yang et. al. 2019), and 42 in bread wheat (Tong et. al. 2021).

The physiochemical features of the PSOSCA proteins were given in (Table 1). The amino acid length ranged from 645 amino acids (PsOSCA8) to 775 amino acids (PsOSCA1, -10). The protein with the smallest and largest molecular weight (kDa) was PsOSCA6 (73.12 kDa) and PsOSCA10 (7-88.78 kDa) respectively. The gravy index scores ranged from 0.048 (PsOSCA1) to 0.457 (0PsOSCA8), indicating the hydrophobic nature of all these proteins. The aliphatic index lay from 99.43 (PsOSCA1) to 110.94 (PsOSCA9), indicating they are considered thermally stable. Also, the isoelectric point (pI) also varied from 8.65 (PsOSCA7) to 9.55 (PsOSCA12). Out of the 16 SbWAKs/SbWAKLs, 9 were unstable, with an instability index greater than 40.

Table 1. Physiochemical properties of PsOSCA proteins

Proteins	Protein accession number (NCBI)	Length of amino acids	Molecular weight kDa	pI	Grand average of hydropathicity (GRAVY)	Aliphatic index	Instability index
PsOSCA1	XP_026408256.1	775	88.77	8.92	0.048	99.43	39.72
PsOSCA2	XP_026404934.1	719	81.56	9.01	0.248	104.67	37.16
PsOSCA3	XP_026455562.1	782	88.70	8.97	0.248	101.34	43.53
PsOSCA4	XP_026453924.1	760	86.12	8.92	0.176	102.11	45.53
PsOSCA5	XP_026381252.1	649	73.42	9.08	0.426	115.15	37.00
PsOSCA6	XP_026381253.1	648	73.12	9.13	0.412	110.20	39.70
PsOSCA7	XP_026388439.1	691	78.64	8.65	0.312	104.93	46.91
PsOSCA8	XP_026386289.1	645	72.94	9.08	0.457	114.19	38.27
PsOSCA9	XP_026386291.1	648	73.20	8.83	0.424	110.94	39.13
PsOSCA10	XP_026392738.1	775	88.78	8.99	0.066	100.31	41.31
PsOSCA11	XP_026394975.1	719	81.35	9.16	0.268	104.55	36.27
PsOSCA12	XP_026401920.1	776	84.49	9.55	0.177	104.25	49.23
PsOSCA13	XP_026410511.1	766	88.67	9.53	0.123	101.41	48.39
PsOSCA14	XP_026418766.1	717	81.84	9.10	0.143	99.92	50.93
PsOSCA15	XP_026422721.1	765	87.73	8.73	0.135	101.74	46.00
PsOSCA16	XP_026425391.1	717	82.03	8.97	0.146	99.92	51.30

To investigate the evolutionary relationship between OSCAs of *A. thaliana* and *P. somniferum*, a phylogenetic tree was drawn using the Neighbor-joining method with 1000 replicates (Figure 1). The phylogenetic tree was grouped into four main groups. Group I was further three sub-grouped, consisting of 8 PsOSCA proteins (PsOSCA3, -5, -6, -7, -8, -9, -14, -16) and AtOSCA2.1, -2.2, -2.3, -2-.4, -2.5 proteins. The other group included AtOSCA1.1, -1.2, -1.3, -1,4, -1.5, 1.6) and six PsOSCA (-1, -4, -10, 12, -13, -15). AtOSCA4.1 and AtOSCA 3.1 are

separated the other proteins and PsOSCA2 and PsOSCA11 are closer them from the other PsOSCAs.

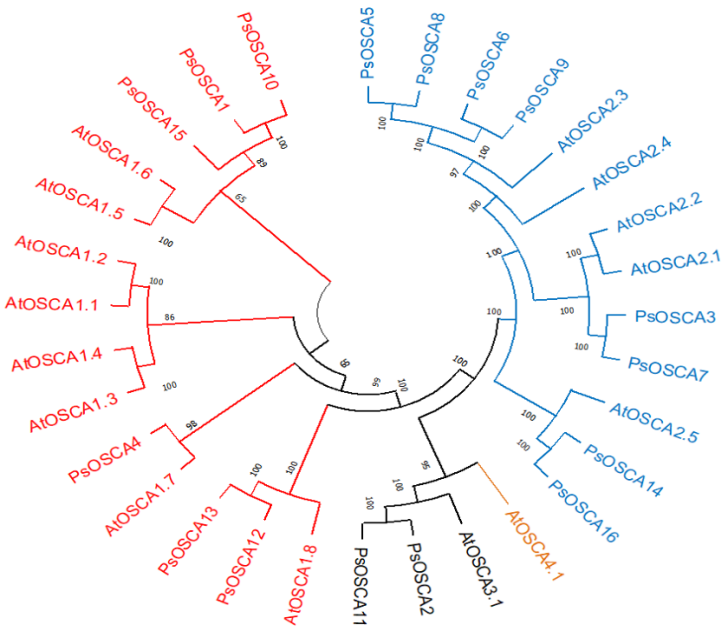


Figure 1. Phylogenetic relationship of OSCA gene family from *Arabidopsis thaliana and P. somniferum.*

Sixteen Ps*OSCA* genes was distributed on all chromosomes except chromosome 2 (Figure 2). The highest number of genes were found on chromosomes 5, while chromosome 7, 8, 9, 10 and 11 contained 2 genes, representing the least number of genes on a chromosome.

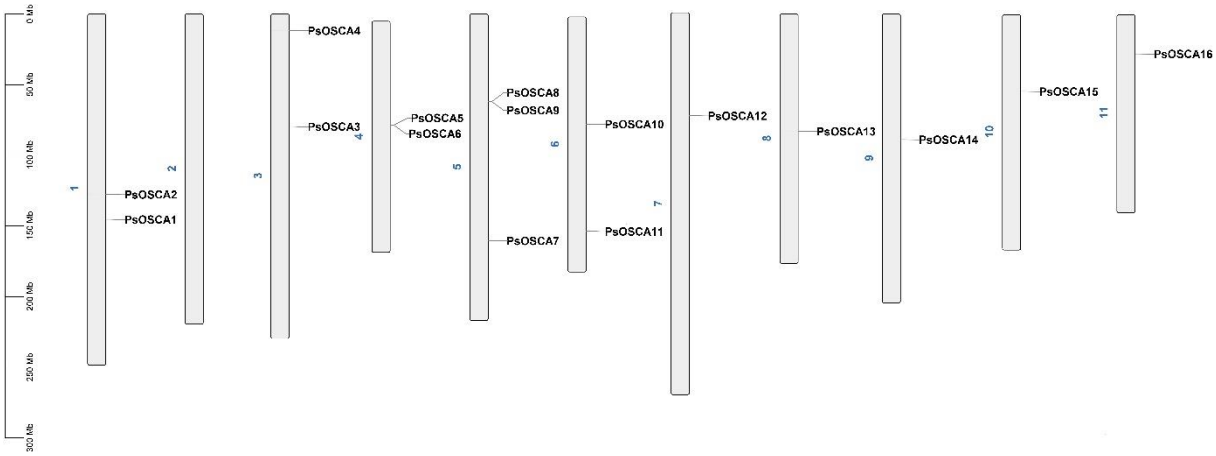
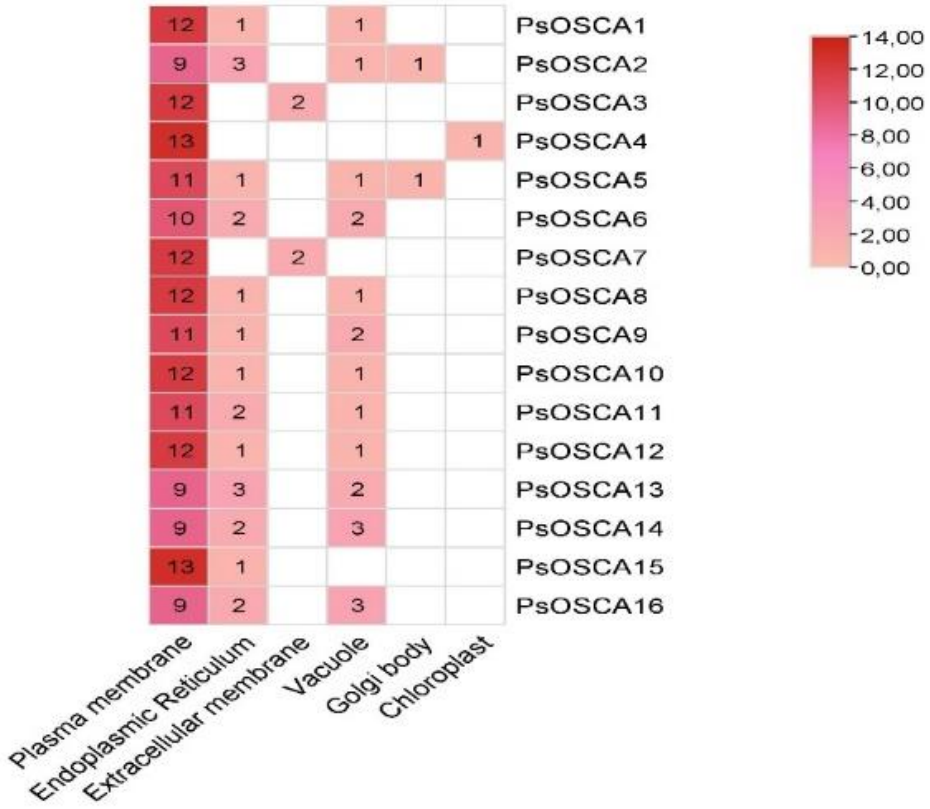


Figure 2. Chromosomal localization of the Ps*OSCA* genes in *P. somniferum* chromosomes. Chromosome numbers are shown beside each chromosome

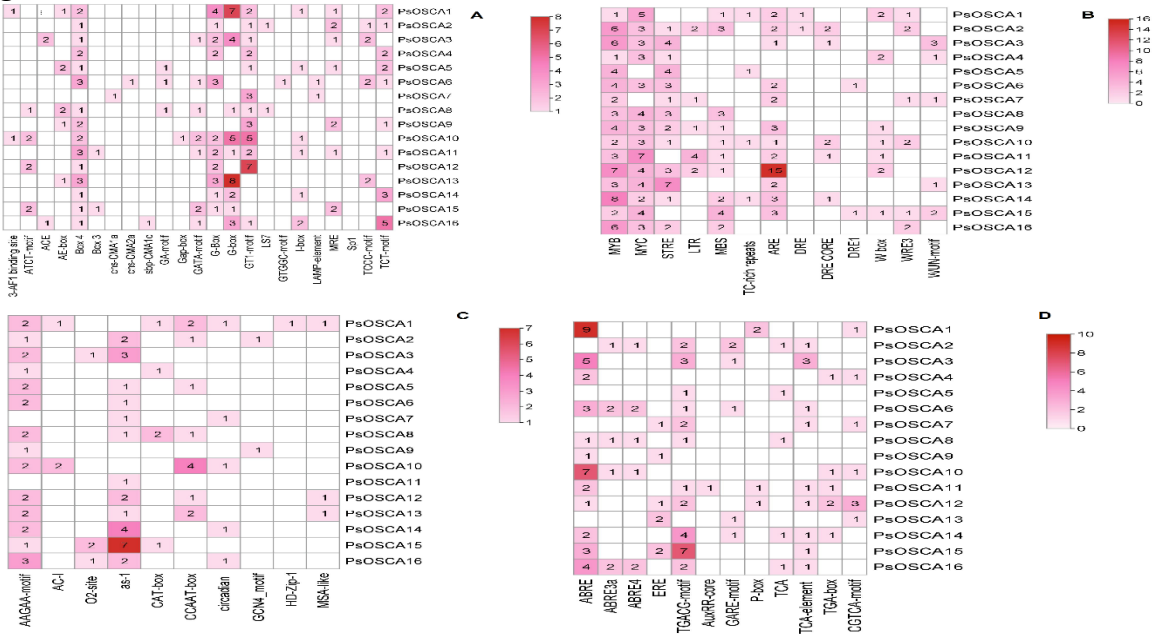
While all PsOSCA proteins were found in the plasma membrane (Table 2). Then, PsOSCAs were mostly found in the endoplasmic reticulum and vacuole. While there are only two PsOSCA proteins each in the extracellular and Golgi body, chloroplast contained only PsOSCA4 protein.

Table 2. Subcellular localization of PsOSCA proteins



Cis-regulatory elements regulate the transcription of genes in vicinity and have important roles in development (Wittkopp et. al. 2012). In this study, the light-responsive elements, environmental stress-responsive elements, hormone-responsive elements, and development-related elements were predicted in the promotor of *PsOSCA* genes. 23 light-responsive elements were detected (Table 3a). G-box, box 4, g-box and GT1 motif were most abundant. Considering environmental stress response element, MYB binding site (drought inducibility), MYC (drought, salt, and stress response), STRE (stress response), MBS (drought inducible) and ARE (anaerobic induction) were predominant. Of the 10 development-related cis-elements detected, two (AAGAA-motif; the endosperm-specific negative expression and as-1; root-specific expression) were the most abundant. Finally, 12 cis-regulotary elements were also identified as hormone-responsive elements. These findings showed that PsOSCAs are important for growth and development.

Table 3. A. Light responsive elements, B. Environmental stress response elements, C. Development response elements, D. Hormone response elements on the promotors of *PsOSCA* genes.



A total of 10 conserved motifs were identified in PsOSCA-1, -2, -4, -10, -11, -12, -14, -15 and -16 (Figure 3b). The least number of motifs were found in PsOSCA3. Only PsOSCA13 lacked one motif (motif 9) and PsOSCA3, 5, -6, -7, -8, -9 did not contain motif number 10. Phylogenetic group distribution (Figure 3a) of PsOSCA proteins did not showed concordance with motif patterns.

The structure analysis of *PsOSCA* genes was performed to reveal exon/intron distribution (Figure 3c). Interestingly, the *PsOSCA* genes classified into the same subgroups have different numbers and lengths of intron/exon, indicating their specialized diverse functions. *PsOSCA1, -4, -12* contained the highest number of exons (11) while *PsOSCA11,* -13 had lowest number of exon (2). The remaining number of exons ranged from 4 to 10 amongst *PsOSCA* genes. Additionally, all *PsOSCA* genes had a varied length of UTR region. The longest and shortest gene length was observed in *PsOSCA7* and *PsOSCA8* respectively.

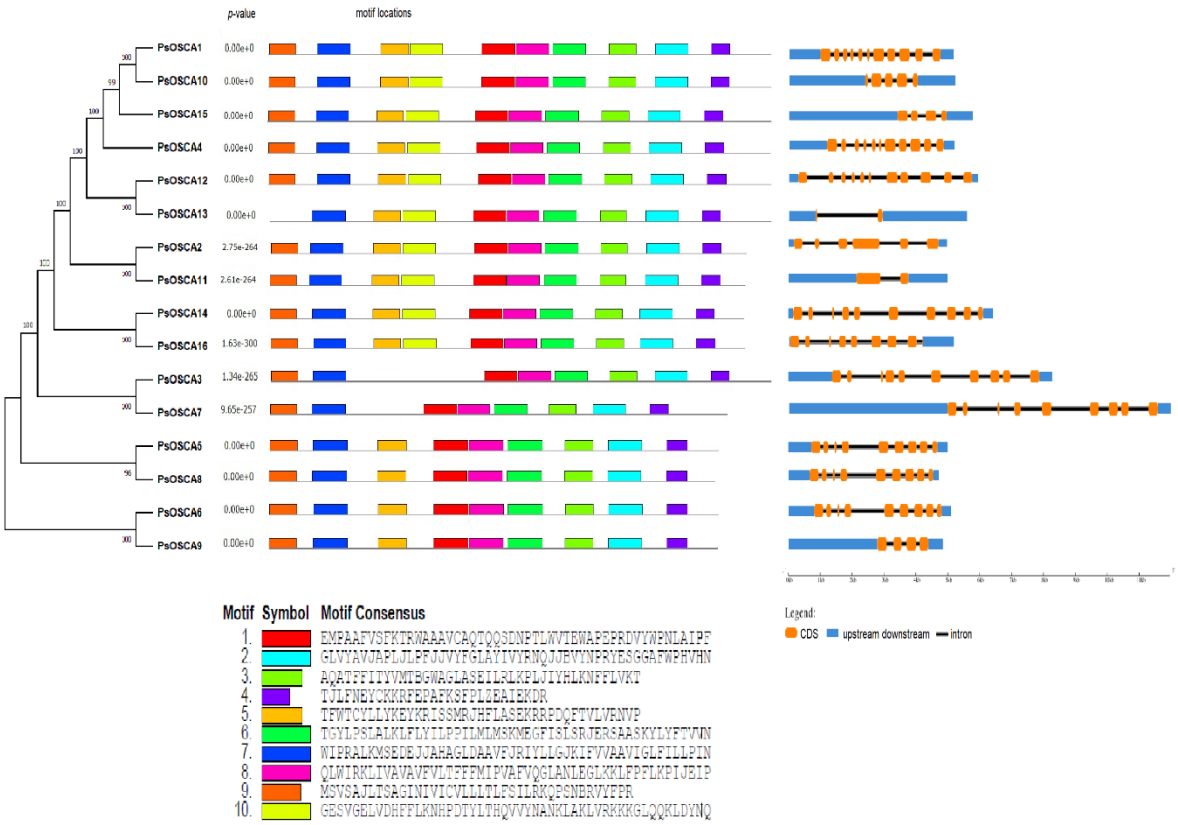


Figure 3. (a) The phylogenetic tree of PsOSCA gene family was constructed with 1000 bootstrap using the neighbor-joining algorithm, (b) the motifs of the SbOSCA protein sequences (c) The gene structure of PsOSCAs. The blue boxes, orange boxes and black lines represent upstream/downstream regions, exons, and introns, respectively.

To estimate the potential role of the PsOSCAs, a protein-protein interaction was assed. Only three PsOSCA (PsOSCA9, -2, -11) proteins were predicted to have interacted with various proteins (Figure 4). PsOSCA9 were the most interacted protein and interacted with nine different, many of which contain protein kinase domains. Seven proteins (A0A4Y7IVB7, A0A4Y7L054, A0A4Y7I5Q0, A0A4Y7IZI5, A0A4Y7KA48, A0A4Y7KRX7, A0A4Y7L172) that interact with PsOSCA9 have been found to have roles related to adenyl nucleotide binding, ribonucleotide binding, small molecule binding, signaling receptor activity, ion binding, anion binding, organic cyclic compound binding, carbohydrate derivative binding, heterocyclic compound binding, cellular protein modification process, protein phosphorylation, phosphate-containing compound metabolic process, signal transduction, cell surface receptor signaling pathway, response to stimulus, protein serine/threonine kinase activity, cell communication, cellular anatomical entity, ATP binding. Interestingly, PsOSCA 2 and PsOSCA11 were interacted with A0A4Y7K9T3 which is responsible for activation of innate immune response, immune response-regulating signaling pathway, response to molecule of bacterial origin,

positive and negative regulation of immune system process, negative and positive regulation of response to biotic stimulus, response to bacterium, regulation of cell death.

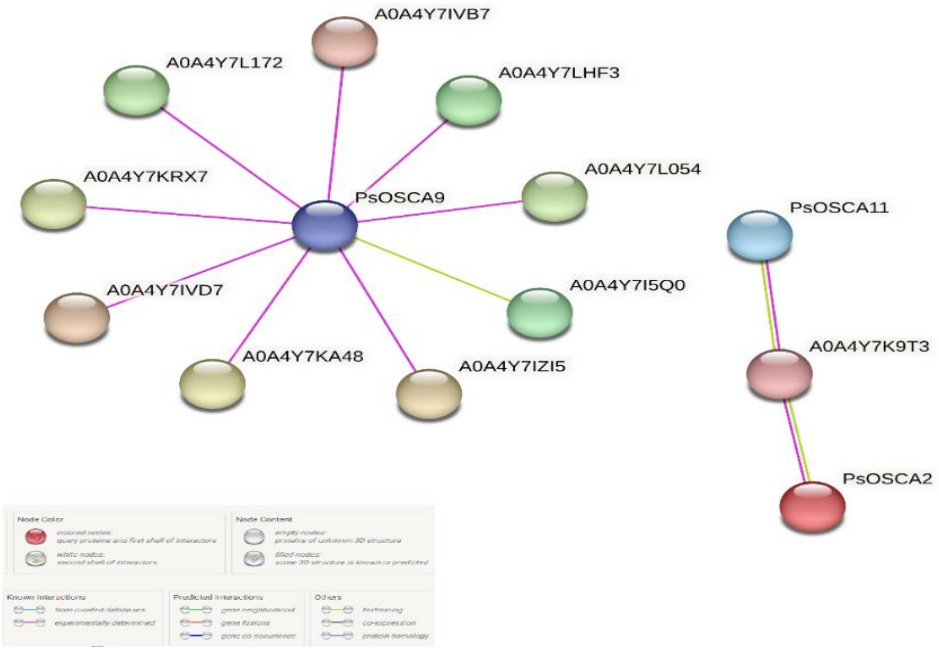


Figure 4. Protein-protein interactions of PsOSCA proteins

For the first time, a bioinformatic examination of the opium poppy OSCA gene family's structure and evolution was performed in this work. However, functional investigation of OSCA genes has to be performed in plant growth, development, and stress response, especially hyperosmolality stress response.

REFERENCES

Cai, Q., Wang, Y., Ni, S., Mu, J., Liu, M., Wang, Y., & Zhao, Y. (2022). Genomewide identification and analysis of the OSCA gene family in barley (Hordeum vulgare L.). *Journal of Genetics*, *101*(2), 34.

Chakraborty, S., Gangwar, R., Zahra, S., Poddar, N., Singh, A., & Kumar, S. (2023). Genome-wide characterization and comparative analysis of the OSCA gene family and identification of its potential stress-responsive members in legumes. *Scientific Reports*, *13*(1), 5914.

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., & Xia, R. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular plant*, *13*(8), 1194-1202.

Ding, S., Feng, X., Du, H., & Wang, H. (2019). Genome-wide analysis of maize OSCA family members and their involvement in drought stress. *PeerJ*, *7*, e6765.

Faostat, https://www.fao.org/faostat/en/

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *evolution*, *39*(4), 783-791.

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., & Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, *31*(8), 1296-1297.

Hu, Y., Zhao, R., Xu, P., Jiao, Y. 2018. The genome of opium poppy reveals evolutionary history of morphian pathway. Genomics Proteomics Bioinformatics 16(6):460-462.

Julius Lewkowitsch, George H. Warburton (1914). George H. Warburton (ed.). Chemical technology and analysis of oils, fats and waxes. Vol. 2 (5 ed.). Macmillan.

Kaur, A., Sharma, A., Madhu, Dixit, S., Singh, K., & Upadhyay, S. K. (2022). OSCA genes in bread wheat: Molecular characterization, expression profiling, and interaction analyses indicated their diverse roles during development and stress response. *International Journal of Molecular Sciences*, *23*(23), 14867.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., ... & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *bioinformatics*, *23*(21), 2947-2948.

Li, Y., Yuan, F., Wen, Z., Li, Y., Wang, F., Zhu, T., ... & Han, S. (2015). Genome-wide survey and expression analysis of the OSCA gene family in rice. *BMC plant biology*, *15*(1), 1-13.

Liu, C., Wang, H., Zhang, Y., Cheng, H., Hu, Z., Pei, Z. M., & Li, Q. (2022). Systematic characterization of the OSCA family members in soybean and validation of their functions in osmotic stress. *International Journal of Molecular Sciences*, *23*(18), 10570.

Liu, X., Wang, J., & Sun, L. (2018). Structure of the hyperosmolality-gated calcium-permeable channel OSCA1. 2. *Nature Communications*, *9*(1), 5060.

Miao, S., Li, F., Han, Y., Yao, Z., Xu, Z., Chen, X., ... & Wang, A. (2022). Identification of OSCA gene family in Solanum habrochaites and its function analysis under stress. *BMC genomics*, *23*(1), 1-14.

Murthy, S. E., Dubin, A. E., Whitwam, T., Jojoa-Cruz, S., Cahalan, S. M., Mousavi, S. A. R., & Patapoutian, A. (2018). OSCA/TMEM63 are an evolutionarily conserved family of mechanically activated ion channels. *Elife*, *7*, e41844.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, *4*(4), 406-425.

Shan, F., Wu, Y., Du, R., Yang, Q., Liu, C., Wang, Y., ... & Chen, Y. (2023). Evolutionary analysis of the OSCA gene family in sunflower (Helianthus annuus L) and expression analysis under NaCl stress. *PeerJ*, *11*, e15089.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., ... & Mering, C. V. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research*, *47*(D1), D607-D613.

Swarbreck, S. M., Colaço, R., and Davies, J. M. (2013). Update on plant calcium-permeable channels. *Plant Physiol.* 163, 1514–1522. doi: 10.1104/pp.113.220855

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, *38*(7), 3022-3027.

Tong, K., Wu, X., He, L., Qiu, S., Liu, S., Cai, L., ... & Chen, J. (2021). Genome-wide identification and expression profile of OSCA gene family members in Triticum aestivum L. *International Journal of Molecular Sciences*, *23*(1), 469.

Wittkopp, P., Kalay, G. *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet* **13**, 59–69 (2012). https://doi.org/10.1038/nrg3095

Yang, X., Xu, Y., Yang, F., Magwanga, R. O., Cai, X., Wang, X., ... & Zhou, Z. (2019). Genome-wide identification of OSCA gene family and their potential function in the regulation of dehydration and salt stress in Gossypium hirsutum. *Journal of Cotton Research*, *2*, 1-18.

Yin, L., Zhang, M., Wu, R., Chen, X., Liu, F., & Xing, B. (2021). Genome-wide analysis of OSCA gene family members in Vigna radiata and their involvement in the osmotic response. *BMC Plant Biology*, *21*, 1-12.