Chemical and Biological Technologies in Agriculture

RESEARCH

Open Access

Genome-wide analysis and characterization of the peptides containing tyrosine sulfation (*PSY*) gene family in *Triticum aestivum* L. unraveling their contributions to both plant development and diverse stress responses

Mahipal Singh Kesawat[1*], Bhagwat Singh Kherawat[2], Chet Ram[3], Swati Manohar[4], Santosh Kumar[5], Sang-Min Chung[6*], Sulaiman Ali Alharbi[7], Mohammad Javed Ansari[8] and Sangram K. Lenka[9*]

Abstract

Background Small-secreted peptides are increasingly recognized as a novel class of intracellular signal molecules, playing crucial roles in plant growth and development. However, the precise role and mechanism governed by peptides containing Tyrosine Sulfation (*PSY*) are still under investigation. Currently, there is a lack of accessible information concerning the *PSY* gene family in wheat.

Results Therefore, in this investigation, we identified 29 PSY genes in *Triticum aestivum*, with the aim of unraveling their significance in plant development processes and their response to a variety of stress conditions. Phylogenetic analysis showed that *TaPSY* genes clustered into five groups. Additionally, an analysis of the gene structure of *TaPSYs* displayed a conserved evolutionary path. The syntenic relationship demonstrated the 69 orthologous gene pairs in *T. dicoccoides, Ae. tauschii*, *T. turgidum,* and *H. vulgare,* respectively. Furthermore, the Ka/Ks analysis indicated that *TaPSY* genes have experienced purifying selection during their evolutionary processes. The promoters of *TaPSY* genes were found to contain numerous CAREs, and these elements are known to perform essential functions in various development processes, phytohormone responses, as well as defense and stress mechanisms. In addition, the identification of potential miRNAs targeting *TaPSY* genes was followed by an examination of their expression patterns across various tissues. Among the 29 *TaPSY* genes, twenty miRNAs were discovered to target eighteen of them. Moreover, *TaPSY* genes displayed a distinct expression across different tissues and stress conditions.

Conclusions Hence, these discoveries offer a significant reference point for forthcoming molecular investigations and hold promise for bolstering wheat yield and stress resilience through targeted genetic enhancements and strategic breeding approaches.

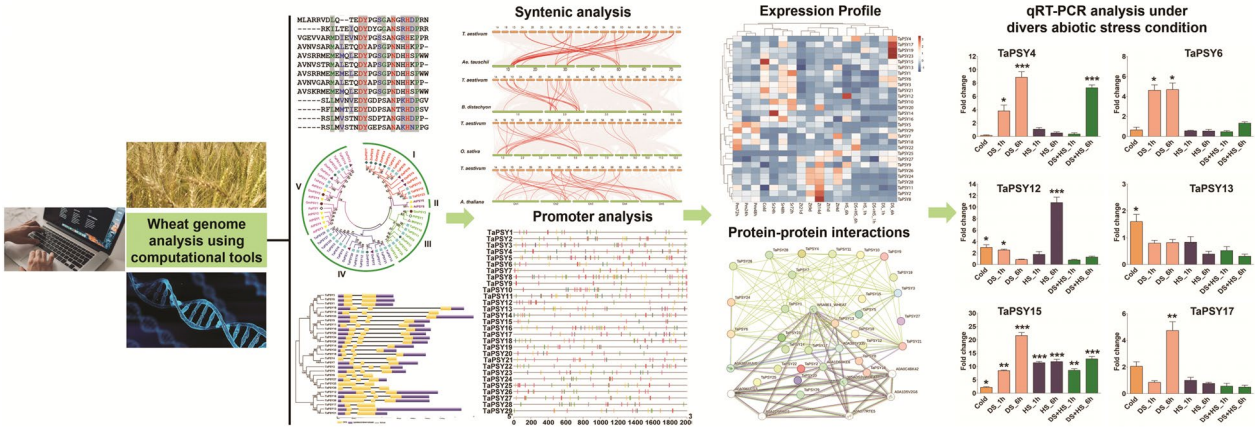
*Correspondence:
Mahipal Singh Kesawat
mahibiotech@snu.ac.kr
Sang-Min Chung
smchung@dongguk.edu
Sangram K. Lenka
sangram.lenka@gbu.edu.in
Full list of author information is available at the end of the article

Highlights

- ➢ Signaling peptides, particularly those with Tyrosine Sulfation (*PSY*) gene family members, have been extensively researched in Arabidopsis. However, there is scarce information available regarding *PSY* genes in other crops such as wheat.
- ➢ In this study, we have identified 29 *PSY* genes in wheat for the first time, shedding light on their significance in plant development and stress response.
- ➢ Phylogenetic analysis categorized *TaPSY* genes into five clusters, exhibiting conserved gene structures and notable purifying selection.
- ➢ Further, the presence of multiple cis-acting regulatory elements (CAREs) in *TaPSY* promoters and out of the 29 *TaPSY* genes, 18 *TaPSY* genes were targeting by miRNAs highlight their roles in various developmental and stress processes.
- ➢ These discoveries offer invaluable insights for future molecular investigations aimed at boosting wheat yield and enhancing stress tolerance through targeted genetic improvements.

Keywords *PSY*, Abiotic stress, Biotic stress, Cis-acting regulatory elements, Gene expression, Phylogenetic analysis, miRNAs and qRT-PCR

Graphical Abstract



Introduction

Small-secreted peptides are increasingly recognized as essential elements in cell to cell communication throughout various stages of plant development [1–4]. In the last decade, there has been a substantial increase in the total number of functionally characterized peptide signals, which now surpasses the total count of classical phytohormones [5–7]. Peptides are typically divided into two main classes based on the characteristics of N-terminal leader sequences: secreted peptides and non-secreted peptides. The cell to cell signaling is predominantly facilitated by secreted signals, however, evidence suggesting that non-secreted peptides also serve as cell to cell signals during plant defense responses [1, 6, 8]. Structurally, secreted peptide are classified into two primary categories: post-translationally modified small peptides (PTMP) and cysteine-rich peptides. These alterations are necessary to ensure the proper functioning of the mature peptides in their respective cellular roles [9]. Plant peptides containing sulfated tyrosine (*PSYs*) are part of the PTMP group. These peptides undergo at least one post-translational alteration, which may include hydroxyproline arabinosylation, tyrosine sulfation and proline hydroxylation [8, 9]. Numerous sulfated peptides have been identified in animals, whereas in plants, only one sulfated peptide has been discovered to date, which is phytosulfokine (*PSK*). *PSK*, consisting of five amino acids, is a secreted peptide containing two sulfated tyrosine residues [10, 11]. Apart from *PSK*, various other types of sulfated peptides have been identified in plants, such as root

growth meristem factors (RGFs), casparian strip integrity factors (CIF), and *PSY* [1, 12]. *PSY1*, a peptide isolated from Arabidopsis cell suspension culture, contains tyrosine sulfation. This is 18-amino acid glycopeptide, featuring an N-terminal signal peptide. At concentrations as low as nanomolar levels, it facilitates both cell differentiation and expansion within the root elongation zone [2]. These modifications are necessary for the maturation and functionality of the active peptide. In Arabidopsis, tyrosylprotein sulfotransferase initiates tyrosine sulfation [13, 14]. Tyrosine sulfation is crucial for maturation of *PSY*, as evidenced by severe developmental defects observed in tyrosylprotein sulfotransferase mutants. Moreover, these phenotype defects were restored upon treatment with *PSK*, *PSY1*, and *RGF1*[15]. These results indicate that sulfation plays a crucial role in ensuring the stability of the secreted peptide in the protoplast, while simultaneously increasing its binding affinity to its corresponding receptors [16, 17].

Signaling peptides have been shown to play roles in numerous biological processes, such as cell differentiation and expansion, the preservation of stem cell characteristics, abscission of floral organ, control of stomatal patterning, mediation of self-incompatibility, and initiation of defense responses and responses to various stress [5, 10, 18–25]. *PSY1* stands as the sole explored member of the PSY family thus far, having been expressed across various tissues throughout the plant's entire life cycle [1, 6]. In the *PSY1* signaling, there is a leucine-rich repeat receptor-like kinase referred to as the *PSY1* receptor, which serves as the primary ligand responsible for activating and phosphorylating the plasma membrane proton ATPases AHA2 [2, 26]. It has been enhanced the H^+ efflux, however, this response significantly reduced in the *psy1r* mutants but not totally lacking, which suggesting the existence of another *PSY* peptides. H + -ATPases generate proton transport, resulting in the creation of the proton drive force. This force, in turn, furnishes the necessary energy for powering other active transporters located in the plasma membrane. Plasma membrane H^+-ATPases have been implicated in stomata regulation, cell elongation and facilitate the plant acclimatize to diverse stress conditions [27, 28], hence, indicating a positive regulatory function of *PSY1* in cell proliferation, expansion and growth. It was reported that *PSY1* activation led to cell elongation in both the roots and hypocotyls [2, 26, 29]. In addition, *AtPSY1* plays a crucial role in response to plant defense [30]. The pathogen attacks activate *PSY1* signaling, which down-regulates genes implicated in salicylic acid signaling [30]. It has been demonstrated that *Xanthomonas oryzae* produces a sulfated peptide known as RaxX, which bears a significant similarity to the *PSY1* [29]. Recently, it has been found that PSK regulates nodulation in Lotus [16]. The expression level of *AtPSY1* was detected across all tissues in Arabidopsis, distinguishing it from other members of *AtPSY1* family. However, heightened expression level of *AtPSY1* was noted during late silique development, senescence and bolting stages. Additionally, a comparable expression pattern was observed for *AtPSY8*, with its expression level being notably elevated in root compared to other plant parts [4]. The expression level of *AtPSY2* was detected in the green parts of plant, with a notable increase observed during rosette development and bolting stages. *AtPSY5* showed lower level of expression in the root, while, *AtPSY2* might be participated in flower development [4].

Recent advance in DNA sequencing technology, there are increased in number of sequenced plant genome, which permit the genome wide analysis of *PSY* gene family in different crop species such as Arabidopsis [2, 4], Medicago [31] and soybean [32]. Small-secreted peptides are gaining recognition as growth regulators, with numerous examples playing pivotal roles in plant growth and development. Despite the annotation of over 1000 genes as putatively encoding secreted peptides in the Arabidopsis genome, only a small fraction have been confirmed to participate in specific cellular signaling pathways [6, 33]. In addition, Secreted peptides are now acknowledged as a novel class of intracellular signal molecules, orchestrating and specifying biological functions in plants. However, the precise role and mechanism governed by secreted peptides are still under investigation. Despite the crucial role played by *PSY* genes in a wide range of biological processes, they remain unexplored in recently sequenced crop genomes. Further, there is currently no available information on the *PSY* genes in wheat. The whole genome sequence of wheat (*Triticum aestivum*) was available, permitted us to perform a genome-wide survey of the *PSY* family members in wheat [34]. Wheat is one of the oldest and most important cereal crop [35–38]. Wheat supplies about 20% of the food calories for the world populations [39]. In addition, wheat serves as a as a crucial source of protein, carbohydrates, vitamins and minerals for both humans and animals [40–42]. Therefore, in this investigation, we conducted a genome-wide survey of the *PSY* gene family in wheat employing various computational tools. Furthermore, we examined the physicochemical characteristics, chromosomal mapping, gene structure, gene duplication events, motif composition, 3D structure, miRNA and expression patterns of *TaPSY* gene family members across various tissue and varied stress conditions. Thus, the discoveries presented in these findings offer a crucial point of reference for forthcoming molecular investigations. Through the application of genome editing

tools, the *TaPSY* genes and miRNAs highlighted can be manipulated, potentially leading to the creation of climate-resilient crops. This advancement holds promise for bolstering crop yield and fortifying resilience against environmental stress amidst the shifting global climate.

Materials and methods

Iidentification members of the *TaPSY* gene family in the wheat genome

Genomic data from both the Phytozome (https://phytozome-next.jgi.doe.gov/) and Ensembl Plants website (http://plants.ensembl.org/index.html) were collected to perform a genome-wide survey of the wheat genome. Two methodologies were utilized to identify putative *PSY* genes within the wheat genome. Initially, a local database of wheat protein sequences was established using BioEdit. Subsequently, BLASTp was employed, utilizing eight and seven PSY genes from Arabidopsis and rice, respectively, as queries against the local database. The identification of putative PSY genes in the wheat genome relied on a cutoff threshold of > 100-bit scores and an e-value of 10^ − 5. The BLASTp results were structured into a table format for further analysis. The protein sequences of PSYs from different plant species were acquired from Phytozome (https://phytozome-next.jgi.doe.gov/) and Ensembl (http://plants.ensembl.org/index.html) in the second method. A BLASTp search against the wheat proteome was performed, with a bit-score threshold set at > 100 and an e-value cutoff at 10^ − 5. Potential *PSY* candidates were identified based on both methods. Additionally, The Pfam database offered a Hidden Markov Model profile for the conserved domain of TaPSY, designated as IPR034430, through screening in the wheat genome database [43]. The identification of TaPSY family members was extended by utilizing the SMART databases [44] and NCBI-CDD [45]. Eventually, the protein sequences with PSY-related domains were extracted and named in sequence according to their positions on the chromosomes. The Isoelectric Point Calculator is employed for the analysis of both the pI and the MW of the TaPSY protein [46]. To predict the subcellular localization of proteins encoded by TaPSY, the PSORT and BUSCA tools were utilized [47, 48].

Phylogenetic tree, chromosomal distribution, gene duplication event and gene structure analysis of TaPSY genes

Protein sequences of *O. sativa, A. thaliana, G. max, Z. mays, P. trichocarpa, M. truncatula, P. patens,* and *T. aestivum PSY* were acquired from Ensembl Plants (https://plants.ensembl.org/index.html). Phylogenetic tree was created using MEGA 11 by aligning sequences with the ClustalW tool and employing the neighbor-joining method for tree construction. The reliability of the tree was evaluated using the bootstrap method with 1000 replicates [49]. To map them onto chromosomes, the chromosome localization of *TaPSY* genes was obtained from Ensembl Plants (http://plants.ensembl.org/biomart/martview). The mapping of *TaPSY* gene family members was conducted using PhenoGram [50]. To examine the gene duplication event and Ka/Ks value analyses were conducted using McScan tools [51] and TBtools [52]. Further, GSDS was utilized to visualize and analyze the gene organization of *TaPSY* genes [53].

Motif analysis, 3-D structure, *cis*-regulatory elements, GO enrichment analysis

The conserved motifs present in the TaPSY protein sequences were identified and visualized using the MEME webserver [54]. The 3-D structure of TaPSY was generated using the Phyre2 web server [55]. The analysis of promoter elements involved utilizing the 2000 bp sequence upstream of the *TaPSY* genes using the PlantCARE webserver [56]. The GO enrichment analysis of TaPSY proteins was carried out using Blast2GO [57] and agriGO program was utilized for the analysis [58].

Expression profile and identification of putative miRNA targets for *TaPSY* genes

Transcript per million values for different tissues and stress conditions were obtained from Wheat Expression Database (http://www.wheat-expression.com/). Further ClustVis [59] and TBtools [52] were employed to generate heatmaps and PCA plots. To identify feasible wheat miRNAs, 1063 mature miRNA sequences of wheat were downloaded from PmiREN (https://pmiren.com/download). The psRNATarget Server available was utilized to identify putative miRNA targets for the *TaPSY* gene family members (https://www.zhaolab.org/psRNATarget/analysis).

Plant growth conditions, stress treatments and qRT- PCR analysis

Wheat seeds were sown in plastic pots, and two-week-old plants were subjected to drought and heat stress (37 °C) for durations of 1 h and 6 h. The plants without stress were kept at 25 °C. The plant tissue were collected from both stressed plants and kept at − 80 °C. Further, RNA was extracted from both control and abiotic stressed plant tissues using the described method by [60, 61]. The iScript™ cDNA synthesis kit (Bio-Rad, Hercules, CA, USA) was utilized for synthesizing cDNA. The internal control utilized was wheat actin, and the qRT-PCR was conducted using the Applied Biosystems 7500 Fast Real-Time PCR system (Applied Biosystems). Each

qRT-PCR reaction was conducted with three technical replicates and replicated three times. The calculated fold change values were determined as described by [62, 63]. Subsequently, these calculated values were employed to generate the graph. The table presenting all primers utilized in this study can be found in Table S9.

Results

Genome-wide identification and evolutionary analysis of *TaPSY*

In this study, we employed several bioinformatics tools to identify a total of 29 *PSY* genes in the *T. aestivum* genome (Table 1; Table S1).

When compared to other crops, the genome of *T. aestivum* contains a relatively higher abundance of the *PSY* gene. For instance, *A. thaliana* (8), *Z. mays* (7), *O. sativa* (7), and *G. max* (3) (Table 2).

This phenomenon may be attributed to the hexaploid nature of wheat, which contributes to its larger genome size compared to other plant species. The 29 TaPSY genes exhibit distinct physicochemical features. The predictions suggest that the 29 TaPSY genes encode proteins ranging in length from 216 to 528 amino acids. The MW of PSY protein falls within the range of 7.87 (TaPSY18) to 19.36 (TaPSY20) kilo Dalton (Table 1). The pI values varied from 4.82 (TaPSY13) to 11.93 (TaPSY1). Notably, 17 out of the identified TaPSY proteins had a pI greater than 7, suggesting a prevalence of basic amino acids in the majority of TaPSY proteins. The GRAVY values ranged from -0.81 (TaPSY20) to 0.197 (TaPSY23) for the 29 TaPSY proteins that were assessed. Further, the GRAVY values for 26 TaPSY proteins (89.65%) were negative, with the exception of three TaPSY proteins: TaPSY15 (0.068), TaPSY19 (0.111), and TaPSY23 (0.206). These findings suggest that the majority of TaPSY proteins exhibit a highly hydrophilic nature. In addition, subcellular localization predictions indicate that most TaPSY proteins are found in both the extracellular space and organelle membranes. However, a smaller subset of TaPSY proteins are also localized in the chloroplast, nucleus, and plasma membrane (Table 1). Furthermore, we found a correlation between the MW of TaPSY proteins and their pI to examine the dispersal pattern of TaPSY proteins (Fig. S1). These findings unveiled a total of 29 TaPSY proteins are widely distributed, showing variations in their pI and molecular weights. Moreover, in order to gain insights into the evolutionary relationship between TaPSY and PSY genes in other crops, a phylogenetic tree was constructed using protein sequences from *T. aestivum*, *O. sativa*, *A. thaliana*, *G. max*, *Z. mays*, *M. truncatula, P. trichocarpa* and *P. patens* (Table S2). The PSY family was categorized into five distinct groups based on the generated phylogenetic tree (Fig. 1).

Group I comprises 6 members, whereas Group II, III, IV, and V contain 0, 6, 12, and 5 members, respectively (Fig. S2). In addition, the phylogenetic tree was constructed using the 29 TaPSY protein sequences and was further divided into three groups (Fig. S3).

The chromosomal distribution, gene duplication events and synteny analysis of *TaPSY* genes

To investigate gene duplication events in wheat, we performed a chromosomal mapping analysis of the identified *TaPSY* family genes. Utilizing the PhenGram webserver, we mapped the 29 *TaPSY* genes onto the 21 chromosomes of wheat. The results revealed that these *TaPSY* genes are distributed across 12 wheat chromosomes (Fig. 2A and Table 1).

Among the sub-genomes, the B and D sub-genomes possess the highest number of *TaPSY* genes (10 each), with the A sub-genomes following closely with 9 *TaPSY* genes (Fig. 2B). In addition, a single gene located on each of the chromosomes 2A, 2B, and 2D, whereas, two *TaPSY* genes were found on the chromosome 1A, 5A, 5B, and 5D (Fig. 2C). The three *TaPSY* genes were situated on the chromosome 1B and ID, respectively, while the four *TaPSY* genes were located on the chromosome 3A, 3B, and 3D (Fig. 2C). In contrast, none of the *TaPSY* genes were detected on the following chromosomes: 4A, 4B, 4D, 6A, 6B, 6D, 7A, 7B, and 7D. Therefore, these findings suggest that members of the *TaPSY* gene family are unevenly distributed across the chromosomes of wheat. In our study, we investigated gene duplication events within the *TaPSY* gene family members. The analysis aimed to identify any duplicated genes or gene families that might have arisen through duplication events in the wheat. The findings derived from this analysis can offer valuable insights into the evolutionary history and functional diversification of *TaPSY* genes. In this investigation, we identified nine duplicated gene pairs within the *PSY* gene family in wheat (Fig. 3; Fig. S3 and Table S3).

The phylogenetic tree of the *TaPSY* genes unveiled multiple instances of gene duplication events (Fig. S3). *TaPSY* gene family has undergone duplications during its evolutionary history (Fig. 3 and Table S3), leading to the emergence of multiple gene copies with potentially diverse functions. The gene duplication events are likely to have implicated to the expansion and functional diversification of the *TaPSY* gene family in wheat. These results indicate that the expansion of the *PSY* gene family in wheat predominantly arose from segmental and whole-genome duplications. In order to investigate the selection force affecting the duplicated *TaPSY* genes, we conducted Ka/Ks ratio calculations for the nine pairs of *TaPSY*

Table 1 General information and different biophysical characteristics of the peptides containing tyrosine sulfation (*PSY*) genes were predicted using various bioinformatics tools in wheat

Proposed gene name	Gene ID	Genomic location	Orientation	CDS length (bp)	Protein length (aa)	Molecular weight (KDa)	Isoelectric point (pI)	GRAVY	Predicted subcellular localization
TaPSY1	TraesCS1A02G305300	1A:497,902,609–497902674	Forward	930	309	10.85	11.93	−0.428	chloroplast
TaPSY2	TraesCS1A02G394200	1A:560,314,739–560314795	Forward	957	318	11.5	5.96	−0.308	extracellular space
TaPSY3	TraesCS1B02G316000	1B:540,366,273–540366338	Forward	930	309	10.97	11.04	−0.34	extracellular space
TaPSY4	TraesCS1B02G381100	1B:613,726,938–613,726,994	Reverse	813	270	9.83	7.54	−0.444	extracellular space
TaPSY5	TraesCS1B02G422500	1B:645,519,294–645,519,344	Reverse	948	315	11.62	10.15	−0.44	plasma membrane
TaPSY6	TraesCS1D02G305000	1D:402,438,480–402438545	Forward	921	306	10.59	11.06	−0.342	extracellular space
TaPSY7	TraesCS1D02G369000	1D:448,239,196–448,239,252	Reverse	741	246	8.96	8.64	−0.551	extracellular space
TaPSY8	TraesCS1D02G402300	1D:467,625,750–467625806	Reverse	948	315	11.49	7.08	−0.394	chloroplast
TaPSY9	TraesCS2A02G156700	2A:103,240,369–103240458	Forward	840	279	9.97	10.38	−0.402	plasma membrane
TaPSY10	TraesCS2B02G182200	2B:157,266,342–157,266,431	Forward	831	276	9.83	10.87	−0.3	extracellular space
TaPSY11	TraesCS2D02G162500	2D:106,961,619–106961708	Forward	840	279	10	10.38	−0.353	extracellular space
TaPSY12	TraesCS3A02G180700	3A:208,455,725–208455778	Forward	966	321	11.31	7.00	−0.431	organelle membrane
TaPSY13	TraesCS3A02G189500	3A:233,022,750–233022806	Reverse	966	321	11.09	4.82	−0.082	chloroplast
TaPSY14	TraesCS3A02G253100	3A:474,435,430–474436112	Reverse	651	216	7.91	10.28	−0.278	extracellular space
TaPSY15	TraesCS3A02G338900	3A:585,566,808–585566879	Reverse	903	300	10.38	6.33	0.068	extracellular space
TaPSY16	TraesCS3B02G210400	3B:246,504,378–246504431	Forward	1560	519	18.45	10.68	−0.675	extracellular space
TaPSY17	TraesCS3B02G218700	3B:262,078,578–262078634	Reverse	975	324	11.21	5.37	−0.042	organelle membrane
TaPSY18	TraesCS3B02G285000	3B:456,553,928–456,554,735	Reverse	651	216	7.87	9.97	−0.278	organelle membrane
TaPSY19	TraesCS3B02G370600	3B:582,752,177–582,752,248	Reverse	921	306	10.58	6.33	0.111	organelle membrane
TaPSY20	TraesCS3D02G185600	3D:170,931,290–170931343	Forward	1587	528	19.36	11.37	−0.81	extracellular space
TaPSY21	TraesCS3D02G193000	3D:183,135,367–183,135,423	Reverse	975	324	11.18	5.07	−0.138	organelle membrane
TaPSY22	TraesCS3D02G254000	3D:355,630,550–355631107	Reverse	651	216	7.91	10.28	−0.212	nucleus
TaPSY23	TraesCS3D02G332500	3D:444,713,299–444,713,370	Reverse	939	312	10.78	8.10	0.206	organelle membrane
TaPSY24	TraesCS5A02G228600	5A:444,595,611–444,595,691	Reverse	867	288	10.44	6.33	−0.142	extracellular space
TaPSY25	TraesCS5A02G298600	5A:505,624,665–505624721	Forward	1137	378	13.86	6.22	−0.142	organelle membrane
TaPSY26	TraesCS5B02G227400	5B:403,855,657–403856280	Reverse	867	288	10.59	8.03	−0.285	plasma membrane
TaPSY27	TraesCS5B02G297900	5B:480,299,291–480299347	Forward	894	297	10.73	5.39	−0.282	extracellular space
TaPSY28	TraesCS5D02G239000	5D:347,277,204–347277284	Forward	867	288	10.45	5.86	−0.236	organelle membrane
TaPSY29	TraesCS5D02G305400	5D:400,294,090–400294146	Forward	1380	459	17.03	7.88	−0.484	extracellular space

ID identity, *bp* base pair, *aa* amino acids, *pI* isoelectric point, *MW* molecular weight, *KDa* Kilo dalton

Table 2 The *PSY* genes encoded by various crop genomes

Crop species	Genome size	Coding genes	*PSY* genes
Triticum aestivum	17 Gb	107,891	29
Arabidopsis thaliana	135 Mb	27,655	8
Zea mays	2.4 Gb	39,591	7
Oryza sativa	500 Mb	37,960	7
Glycine max	1.15 Gb	55,897	3

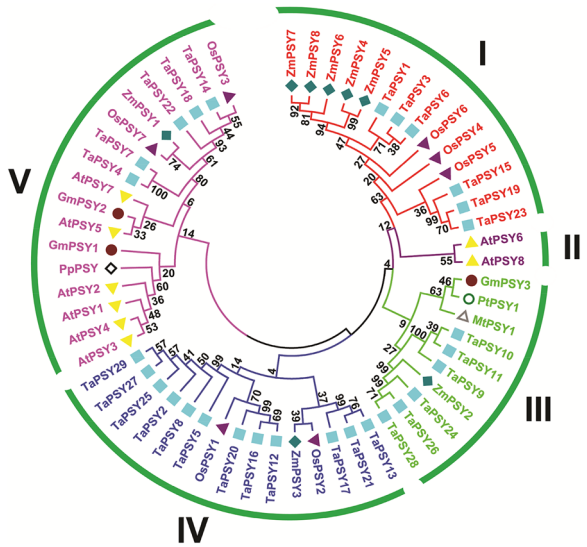


Fig. 1 The phylogenetic relationship of PSY proteins among various crop species including *T. aestivum* (29), *A. thaliana* (8), *O. sativa* (7), *Z. mays* (8), *G. max* (3), *P. trichocarpa* (1), *M. truncatula* (1) and *P. patens* (1). The phylogenetic tree was produced using MEGA11 through the neighbor-joining method, and bootstrap values were employed with 1000 replicates

genes (Table S3). The Ka/Ks value was found to be less than one for the eight *TaPSY* genes indicating that duplicated *TaPSY* genes have undergone purifying or negative selection. However, one gene pair (*TaPSY3*/*TaPSY6*) has shown Ka/Ks value more than one suggesting that this gene pair had gone through a positive selection. Overall, this finding indicates that the *TaPSY* gene family has evolved under purifying selection, ensuring the preservation of crucial traits in wheat. Further, we examined the syntenic relationships of *TaPSY* genes with those of other crop species, such as *B. distachyon, A. tauschii, A. thaliana*, and *O. sativa*. To identify orthologous gene pairs among genomes of different crop species, we employed MCScanX (Fig. 4 and Table S4).

Among *TaPSYs* and other *PSYs* in *Ae. tauschii, T. dicoccoides, T. turgidum*, and *H. vulgare*, we identified 19, 13, 13, and 24 orthologous genes, respectively. Furthermore, the comparison with *PSY* genes from *B. distachyon*, *Ae. tauschii,* and *O. sativa*, we found that 24, 25, and 22 *TaPSY* genes, respectively, exhibited collinearity. In addition, fewer *TaPSY* genes consist minimum two pairs of orthologous genes, for instance, *TaPSY1, TaPSY3, TaPSY6, TaPSY9, TaPSY10, TaPSY11, TaPSY13, TaPSY15, TaPSY17, TaPSY19, TaPSY21, TaPSY23, TaPSY24, TaPSY26* and *TaPSY28*, and these identified orthologous gene pairs may play a significant role in the evolution of the *PSY* gene family. In summary, these results collectively indicate that the *TaPSY* gene family may have originated from ancestral orthologous genes found in other crops.

Gene structure, conserved domain and 3-D structure analysis of *TaPSY* genes

Gene structure and motif analysis provides valuable insights into the conserved and evolutionary variances of *PSY* genes in wheat. Through this analysis, it was noted that the number of exons and introns varied in various subfamilies (Fig. 5).

This analysis also revealed that *TaPSY* gene family members exhibit slight variations in their gene structure (Fig. S4). *TaPSY* genes comprise 1–2 introns, for example, *TaPSY1, TaPSY3, TaPSY6, TaPSY13, TaPSY17, TaPSY21* and *TaPSY25* contain at least one intron, whereas majority of them contain a maximum of two introns such as *TaPSY2, TaPSY4, TaPSY5, TaPSY7, TaPSY8, TaPSY9, TaPSY10, TaPSY11, TaPSY12, TaPSY14, TaPSY15, TaPSY16, TaPSY18, TaPSY19, TaPSY20, TaPSY22, TaPSY23, TaPSY24, TaPSY26, TaPSY27, TaPSY28* and *TaPSY29*. Further, to understand the biological functions of *TaPSY* gene family members, we explored conserved motif analysis of TaPSY proteins by the MEME webserver. Lastly, we detected the 10 motifs within the TaPSY proteins (Fig. 6A, B).

The members of the TaPSY gene family were identified by the presence of the conserved PSY domain (IPR034430), and it was observed that all twenty-nine TaPSY proteins contain the PSY motif, which includes DY, N, H, and P domain. Further, an amino acid sequence alignment of TaPSY was performed, it was observed that all 29 TaPSY proteins contain a conserved PSY domain (Fig. 7 and Fig. S5A).

Furthermore, The Phyre2 webserver was utilized to determine the 3D structure of TaPSY proteins, aiming to understand their specific function in *T. aestivum* (Fig. S5B). Therefore, these findings would contribute to the comprehension and clarification of the exact role of TaPSY protein in regulating various signaling pathways associated with plant development processes and diverse environmental stimuli in wheat.

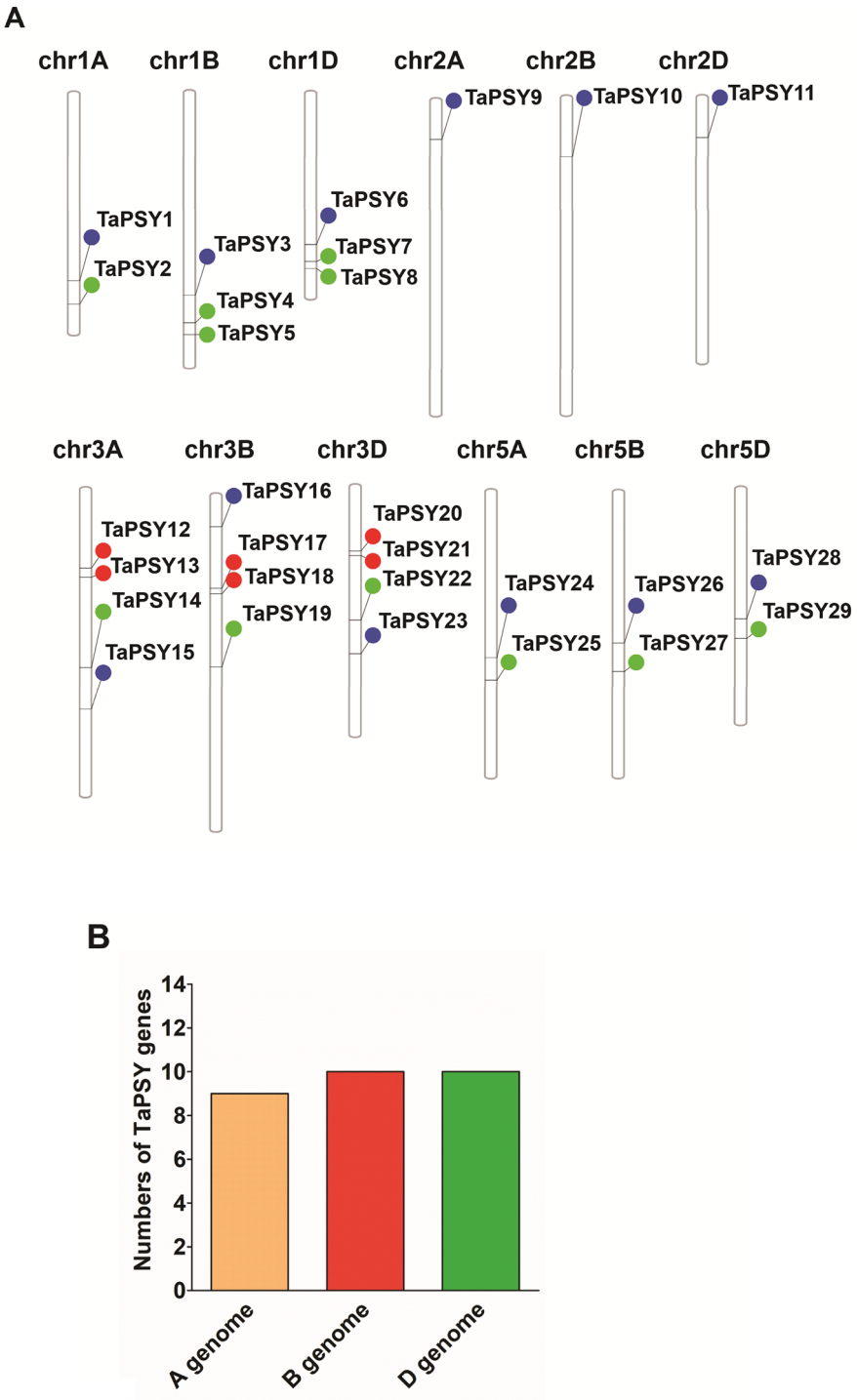


Fig. 2 Chromosomal locations of the *TaPSY* genes and their distribution across different wheat chromosomes and sub-genomes. **A** Schematic illustrations of the chromosomal allocation of *TaPSY* genes on wheat chromosomes, with the gene names indicated on the right side of each chromosome. The distinct colored circles on the wheat chromosomes specify the location of the *TaPSY* genes. The chromosome numbers are mentioned at the top of the chromosomes. **B** *TaPSY* genes are dispersed across the sub-genomes of wheat. **C** *TaPSY* genes are dispersed across the wheat chromosomes

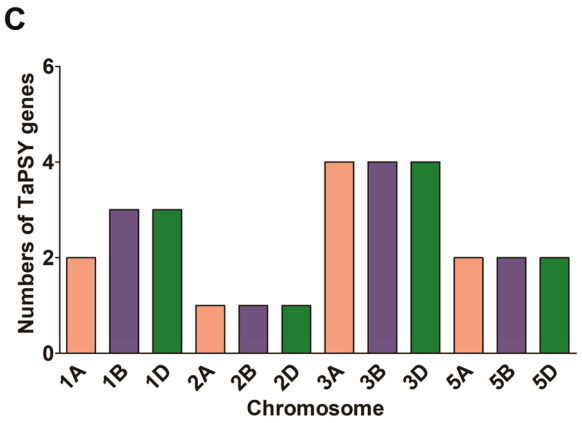


Fig. 2 continued

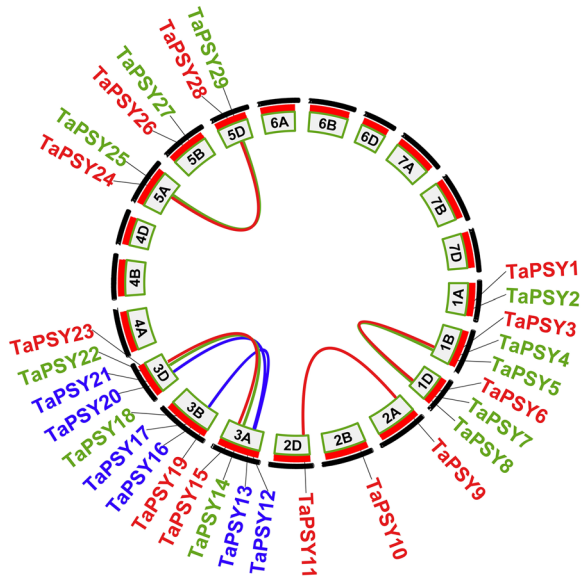


Fig. 3 Chromosomal allocation and duplicated *PSY* gene pairs were identified and analyzed in wheat. Duplicated *PSY* gene pairs and their relationships represented by distinct colors of lines. The Fig. was constructed by TB tools

Promoter element analysis of *TaPSY* genes

To understand the putative function of the *TaPSY* family genes, the examination of the 2000 bp upstream sequence of the *TaPSY* family genes was conducted using the PlantCARE online web server. During our analysis, we discovered numerous cis-regulatory elements within the 2000 bp upstream sequence of the *TaPSY* family genes. These elements encompassed various functional categories such as light response, phytohormones, circadian, cell cycle and seed-specific regulation, as well as stress response (Fig. 8A, B and Table S5).

The *TaPSY* genes were found to encompass five phytohormone responsive CAREs. These components comprise salicylic acid response element (SARE), MeJA response element (MeJARE), abscisic acid response element (ABRE), auxin response element (AuxRE), and gibberellin response element (GARE). The elements associated with light responses, MeJARE, ABRE, defense and stress responsiveness were predominantly found to in the *TaPSY* promoters (Fig. 8B). Hence, these results have shown that *TaPSY* genes might play a critical role in plant growth, development, and various stress conditions Furthermore, within the *TaPSY* genes, there are CARE elements associated with various functions, including endosperm expression, meristem expression, cell cycle regulation, circadian control, zein metabolism, and seed-specific regulation. The discovery of CAREs in *TaPSY* genes indicates that *TaPSY* genes might participate in diverse cellular processes. These results suggest that the *TaPSY* family genes may have a vital role in regulating plant development and stress responses by influencing multiple cis-regulatory elements in wheat. Collectively, these findings provide vital and valuable information for understanding the regulatory functions of these genes and this knowledge enhances our understanding of the complex regulatory mechanisms governing *TaPSY* gene expression and their significance in the overall physiology of wheat.

GO enrichment analysis of *TaPSY* genes

To gain a deeper understanding of the functions of TaPSY genes, we conducted GO enrichment analysis. The TaPSY genes were efficiently annotated and linked with GO terms through Blast2GO. Subsequently, these annotations were validated by eggNOG-Mapper and AgriGO (Fig. S6A-G, Fig. S7-S10, and Table S6-S8). This comprehensive annotation process helps to better

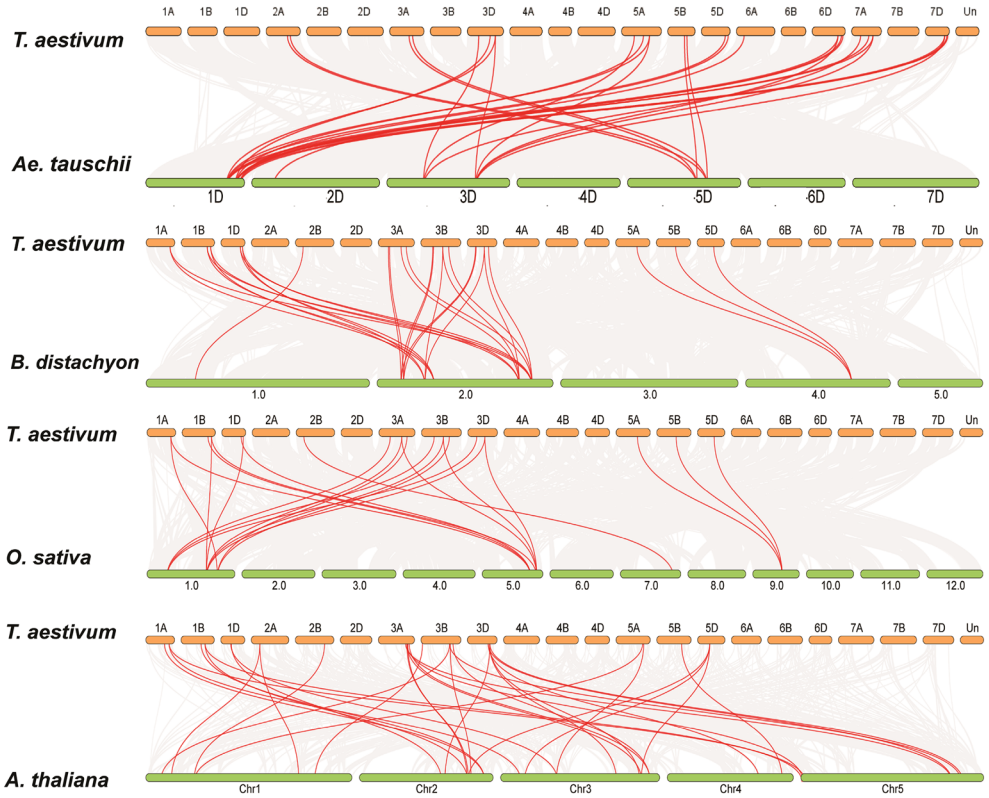


Fig. 4 Syntenic analysis of *TaPSY* genes among different crop species including *A. tauschii, B. distachyon, O. sativa* and *A. thaliana*. The gray outline in the background symbolizing the collinear blocks within *T. aestivum* and other crop species genomes, while red lines signify the orthologous gene pairs that have been identified between *T. aestivum* and other crop genomes

understand the functions and roles of *TaPSY* genes in various biological processes. The analysis of TaPSY genes revealed significant enrichment in several biological process categories, including response to salt (GO:1902074), stimulus (GO:0050896) and fluoride (GO:1902617) (Fig. S7). In addition, within the cellular component category, the TaPSY genes were found to be enriched in cellular anatomical organization (GO:0110165) and membrane (GO:0016020) (Fig. S8 and S9). In the molecular category, TaPSY genes demonstrated enrichment in several transporter activities, including aldonate transmembrane transporter activity (GO:0042879), carboxylic acid transmembrane transporter activity (GO:0046943) and gluconate transmembrane transporter activity (GO:0015128) (Fig. S10). Furthermore, subcellular localization also confirmed the same outcome (Table 1). Thus, these enrichments highlight the participation of *TaPSY* genes in diverse biological processes including stress response, signaling pathways, and membrane-associated functions, indicating their pivotal roles in plant development, and adaptation to environmental challenges.

Transcriptome profiling of *TaPSY* genes in various tissues and stress condtions

The expression profiles of *TaPSY* genes were extensively examined across various tissues and multifactorial stress conditions. This comprehensive analysis aimed to gain a better understanding of the functional roles of *TaPSY* genes. The expression values of the *TaPSY* gene family, TPM (transcripts per kilobase million), were obtained from the Wheat Expression Database (http://www.wheat-expression.com/). Subsequently, these expression values were utilized to construct heatmaps and principal component analysis (Fig. 9A, B, and Fig. 10).

We observed that 29 members of the *TaPSY* gene family displayed distinct expression patterns in different tissues and multifactorial stress conditions (Fig. 9A, B). In this study, we investigated the expression profiles of *TaPSY* genes in five different tissues and across three developmental stages. *TaPSY* genes exhibited a differential expression pattern in different tissues of wheat (Fig. 9A, B), for instance, the expression of *TaPSY4, TaPSY7, TaPSY13, TaPSY14, TaPSY15, TaPSY17, TaPSY18, TaPSY19, TaPSY21, TaPSY22, TaPSY23, TaPSY24, TaPSY26* and *TaPSY28* were highly elevated

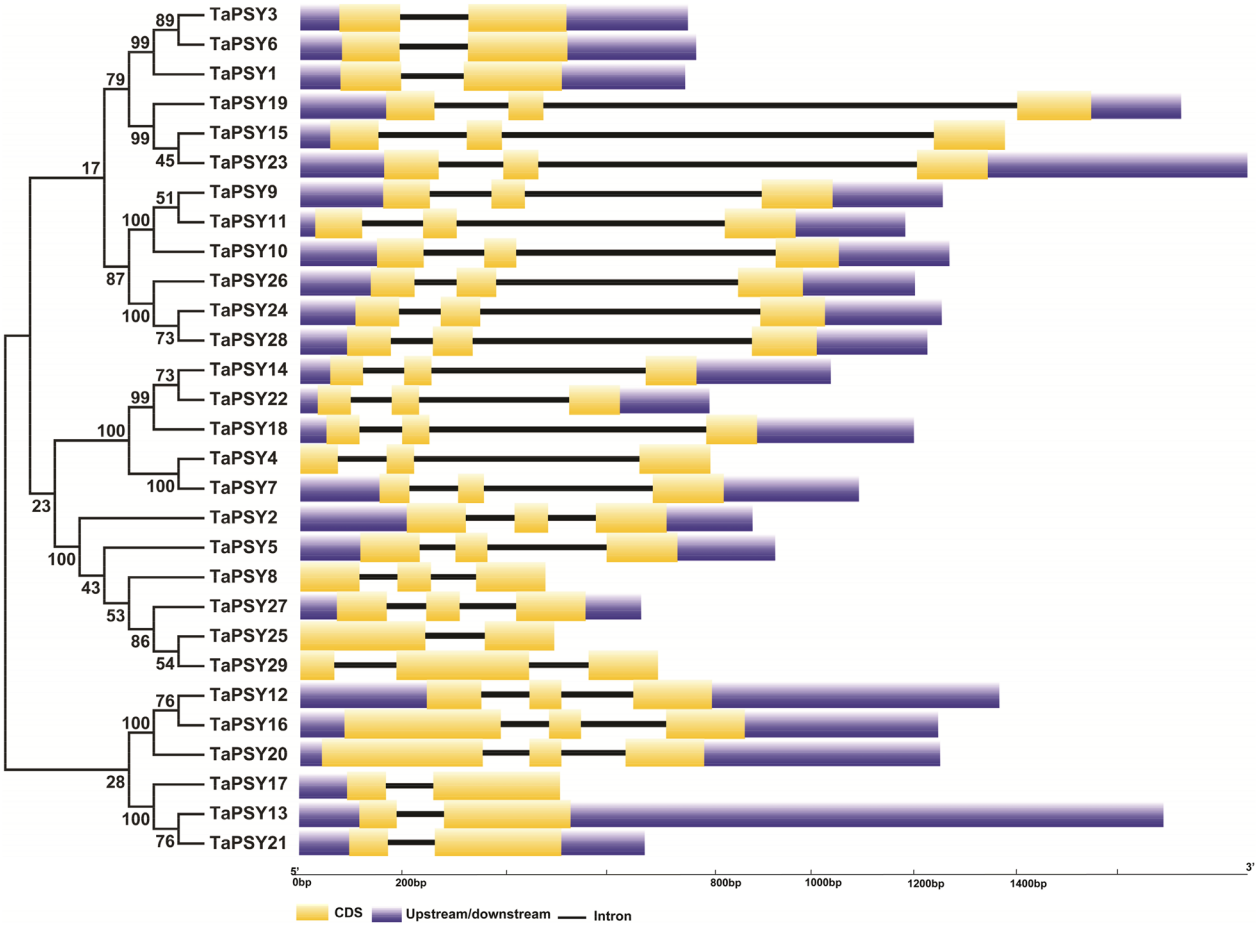


Fig. 5 The gene structure of the *TaPSY* genes. The yellow boxes indicate exons, while untranslated regions are point out by blue boxes and black lines denote introns. The length of the boxes and black lines corresponds to the actual length of the respective regions in the gene sequence

in stem_z32, while *TaPSY8, TaPSY9, TaPSY10, TaPSY11* and *TaPSY25* were up-regulated in stem_z65. Further, the expression levels of *TaPSY3* and *TaPSY6* were significantly raised in leaf_z71. The expression levels of *TaPSY5, TaPSY25, TaPSY27* and *TaPSY29* were induced root_z10, whereas *TaPSY29* in grain_z71. The results indicate that *TaPSY* genes may play a role in the development of various tissues in wheat.

Expression pattern of *TaPSY* were also studied in various stress conditions including powdery mildew, stripe rust, *septoria tritici* blotch, drought, cold and heat (Fig. 9B). The expression of *TaPSY5* in PM24h, *TaPSY5, TaPSY18* and *TaPSY29* in PM48h while, *TaPSY5, TaPSY18, TaPSY22* and *TaPSY29* were highly elevated in PM72h. The expression level of *TaPSY10* and *TaPSY14* in Sr24h, while, *TaPSY1, TaPSY3, TaPSY6, TaPSY10, TaPSY13, TaPSY16, TaPSY17* and *TaPSY21* in Sr48h and *TaPSY1, TaPSY10, TaPSY20,* and *TaPSY21* were highly induced in Sr72h. Further, the expression of *TaPSY7, TaPSY13, TaPSY24, TaPSY25, TaPSY26*and *TaPSY28* were significantly raised in Zt4d, whereas *TaPSY2, TaPSY24, TaPSY26* and *TaPSY28* were elevated in Zt9d. *TaPSY2, TaPSY8, TaPSY9, TaPSY11, TaPSY20* and *TaPSY28* expression level was increased in Zt14d. In addition, diverse environmental stress conditions also revealed differential transcript kinetics for *TaPSY* genes, for instance, the expression of *TaPSY1, TaPSY15* and *TaPSY21* were highly raised in cold. The expression of *TaPSY9, TaPSY12* and *TaPSY16* were up-regulated in HS_6h, while *TaPSY4, TaPSY6, TaPSY7, TaPSY17, TaPSY19* and *TaPSY23* shown increased expression level in DS_6h. Likewise, the expression of *TaPSY4, TaPSY18* and *TaPSY22* were induced in DS + HS_6h (Fig. 9B). Moreover, expression patterns of *TaPSY* genes were confirmed and validated using qRT-PCR. The qRT-PCR analysis revealed consistent and nearly identical results (Figs. 10, 11).

Taken together, these findings provide evidence that *TaPSY* genes are likely implicated in diverse

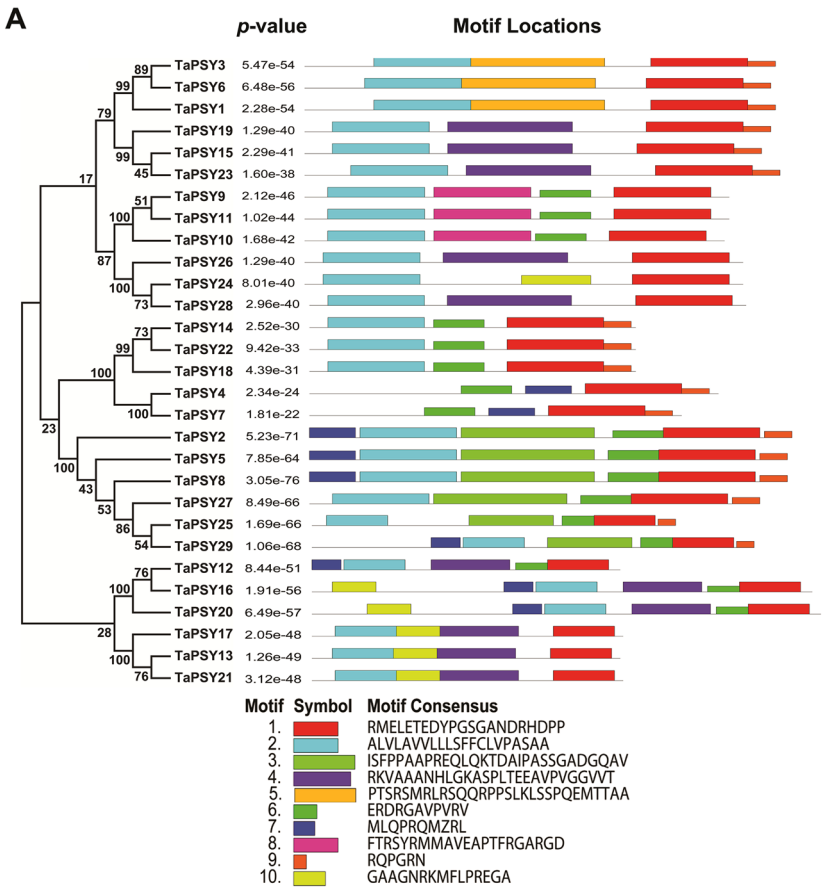


Fig. 6 The conserved motifs identified in TaPSY proteins. The conserved motifs were elucidating by MEME database. **A** The distinct colored boxes demonstrating diverse conserved motifs contain variable size and sequences. **B** Sequence logos represent the conservation and variability of amino acids at each position in the motif of the TaPSY proteins

Fig. 7 The TaPSY protein's amino acid sequence alignment displays the conserved PSY motif, which is indicated by a green box

developmental processes and various biotic and abiotic stress in wheat.

Identification of microRNA (miRNAs) and their *PSY* specific target genes in wheat

To identify potential wheat miRNAs targeting members of the *PSY* gene family, 1063 mature miRNA sequences from wheat were obtained from PmiREN. Subsequently, the Plant Small RNA Target Analysis Server was utilized to identify potential miRNAs targeting the *TaPSY* gene family members. Out of the 29 *TaPSY* genes, eighteen *TaPSY* genes including *TaPSY1, TaPSY2, TaPSY5, TaPSY9, TaPSY10, TaPSY11, TaPSY12, TaPSY16, TaPSY19, TaPSY20, TaPSY21, TaPSY23, TaPSY24, TaPSY25, TaPSY26, TaPSY27, TaPSY28* and *TaPSY29*, twenty miRNAs were found to target them, for example, *Tae-miR1120b, Tae-miR171n, Tae-miR2275p, Tae-miR390a, Tae-miR395a, Tae-miR395ai, Tae-MiR395bp, Tae-miR528a, Tae-miR530c, Tae-miR6196, Tae-miR9483, Tae-miR9657a, Tae-miR9661a, Tae-miR9670, Tae-miRN4309a, Tae-miRN4315, Tae-miRN4320a, Tae-miRN4375, Tae-miRN4402a* and *Tae-miRN45b* (Fig. 12; Table S10).

Further, we examined the expression pattern of miRNAs in various tissues, such as flower, grain, leaf, seed, seedling, spike, and whole plant (Table S11). The identified miRNAs displayed unique expression across different tissues in wheat. This suggests that these miRNAs may have significant roles in regulating the expression of *TaPSY* gene family members during various developmental processes in wheat (Fig. 13).

Thus, this outcome provide valuable insights into the specific functions of these miRNAs across various biological processes in wheat. Enhanced understanding of the regulatory roles of these miRNAs can deepen our comprehension of the molecular mechanisms governing wheat development, response to stress and other critical biological processes. This knowledge may have practical applications in improving wheat crop productivity and resilience to environmental challenges.

Analyzing the protein–protein interactions within the TaPSY family genes

Utilizing the STRING database, we established a protein–protein interaction network to investigate the interactions between TaPSYs and other proteins in wheat (Fig. 14 and Table S12).

According to the predictive results, we discerned thirteen TaPSYs that exhibit interactions with fifteen distinct wheat-specific proteins. Remarkably, TaPSY1 showed interactions with TaPSY4, TaPSY7, TaPSY9, TaPSY10, TaPSY11, TaPSY13, TaPSY17, TaPSY21, TaPSY24, TaPSY26, and TaPSY28. Similarly, TaPSY13 interacted with TaPSY3, TaPSY4, TaPSY6, TaPSY7, TaPSY9, TaPSY24, TaPSY26, TaPSY28, W5A9E1_WHEAT, and W5ADS2_WHEAT. Additionally, TaPSY17 had interactions with TaPSY3, TaPSY4, TaPSY6, TaPSY7, TaPSY9, TaPSY24, TaPSY26, TaPSY28, W5A9E1_WHEAT, and W5ADS2_WHEAT. These results offer valuable insights that can inspire further examinations aimed at elucidating the roles of *TaPSY* genes in diverse biological processes.

Discussion

Signaling peptides have been shown to play crucial roles in numerous cellular processes, such as cell differentiation and expansion, the preservation of stem cell characteristics, abscission of floral organ, control of stomatal patterning, mediation of self-incompatibility, and initiation of defense responses and responses to various stressors [5, 10, 18–25]. In this work, we identified 29 *PSY* genes in *Triticum aestivum* genome using the several bioinformatics tools (Table 1; Table S1). The 29 TaPSY genes exhibit distinct physicochemical features. The subcellular localization predictions indicate that most TaPSY proteins are found in both the extracellular space and organelle membranes. However, a smaller subset of TaPSY proteins are also localized in the chloroplast, nucleus, and plasma membrane. The 29 *TaPSY* proteins are widely

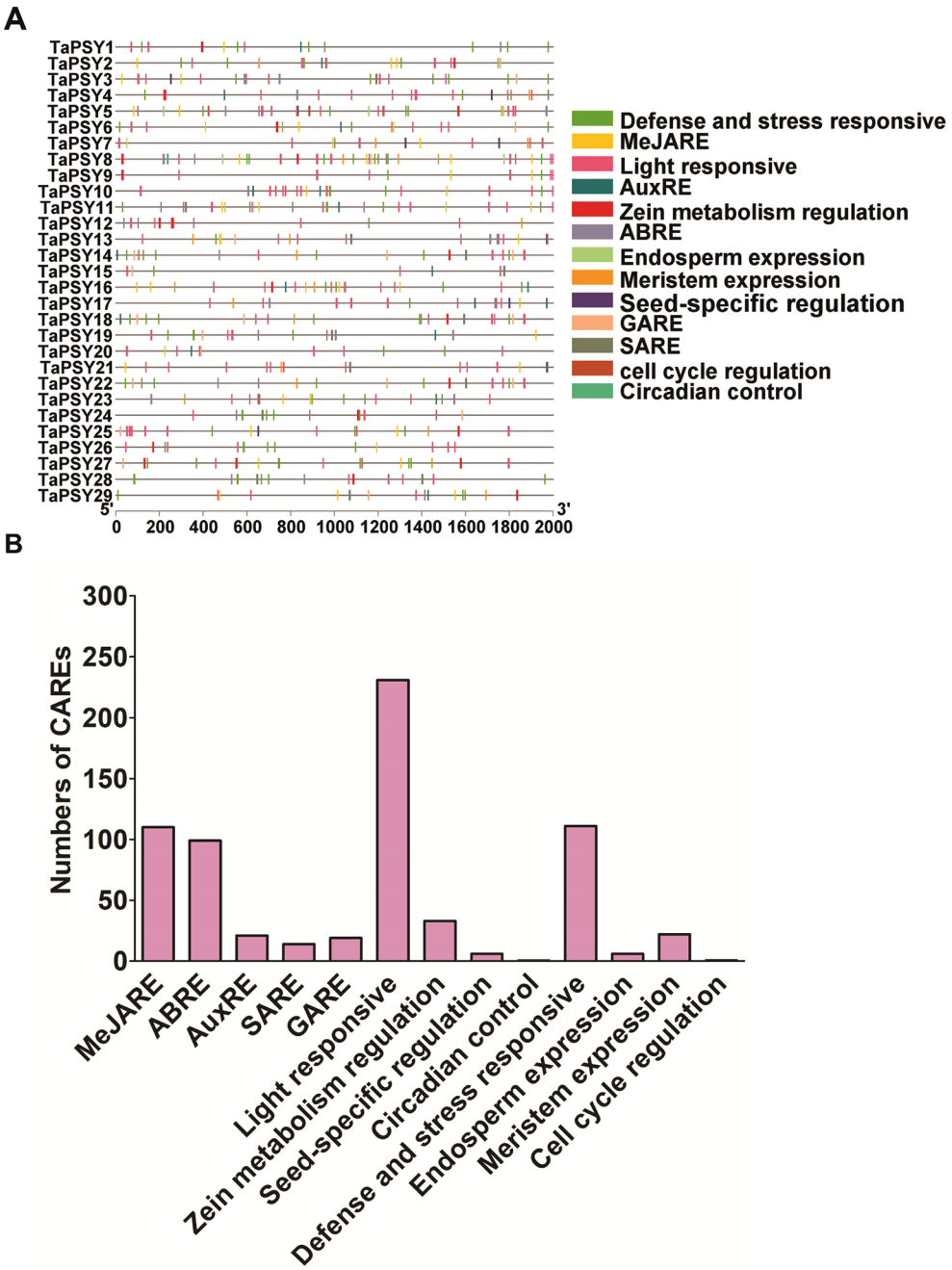


Fig. 8 Identified cis-regulatory element in the 2000 bp promoter region of the *TaPSY* genes. **A** Multiple CAREs were found in the *TaPSY* promoter represented by distinct colors. **B** The multiple CAREs found in *TaPSY* genes

distributed, showing variations in their pI and MW (Table 1). The broad spectrum of pI and MW observed in TaPSY proteins could potentially contribute to their diverse functional role in across various signalling pathways. The *PSY* family was categorized into five distinct groups based on the generated phylogenetic tree (Fig. 1). Group I comprises 6 members, whereas Group II, III, IV, and V contain 0, 6, 12, and 5 members, respectively (Fig. S2). The phylogenetic analysis has unveiled interesting patterns within the *TaPSY* gene family. Specifically, it revealed that cluster I and IV are specific to monocots, indicating their presence and evolution primarily in monocotyledonous plants, which include wheat. On the other hand, cluster II contains *PSY* genes specific to dicots,

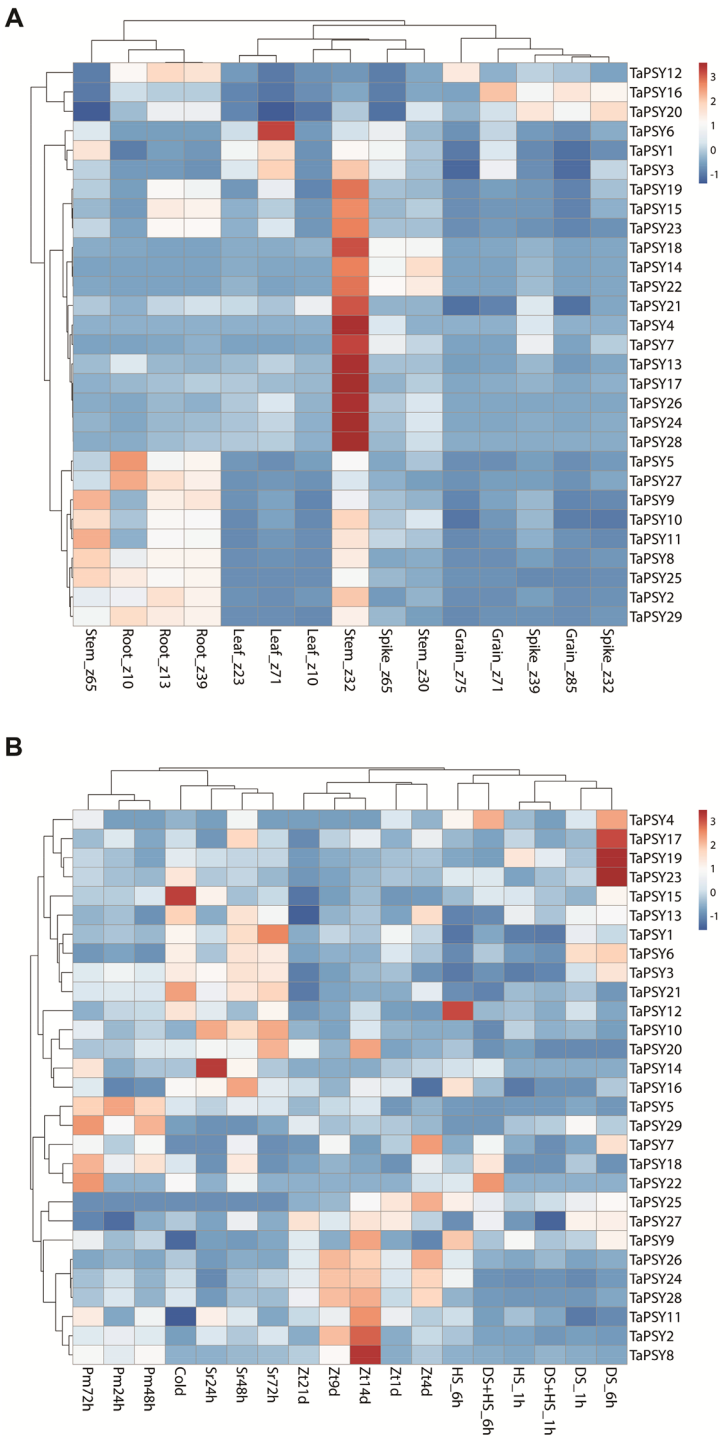


Fig. 9 Expression profiles of *TaPSY* genes. **A** The expression profiles of *TaPSY* in different tissues of wheat. **B** The expression profiles of *TaPSY* genes were investigated under various abiotic stress conditions in wheat. These conditions included drought stress (DS), heat stress (HS), and shared drought and heat stress (DS + HS). Additionally, the expression profiles were examined in response to *Zymoseptoria tritici* (Zt) infection, stripe rust (SR), and powdery mildew (PM). The expression levels were measured in terms of time, where "h" represents hours and "d" represents days

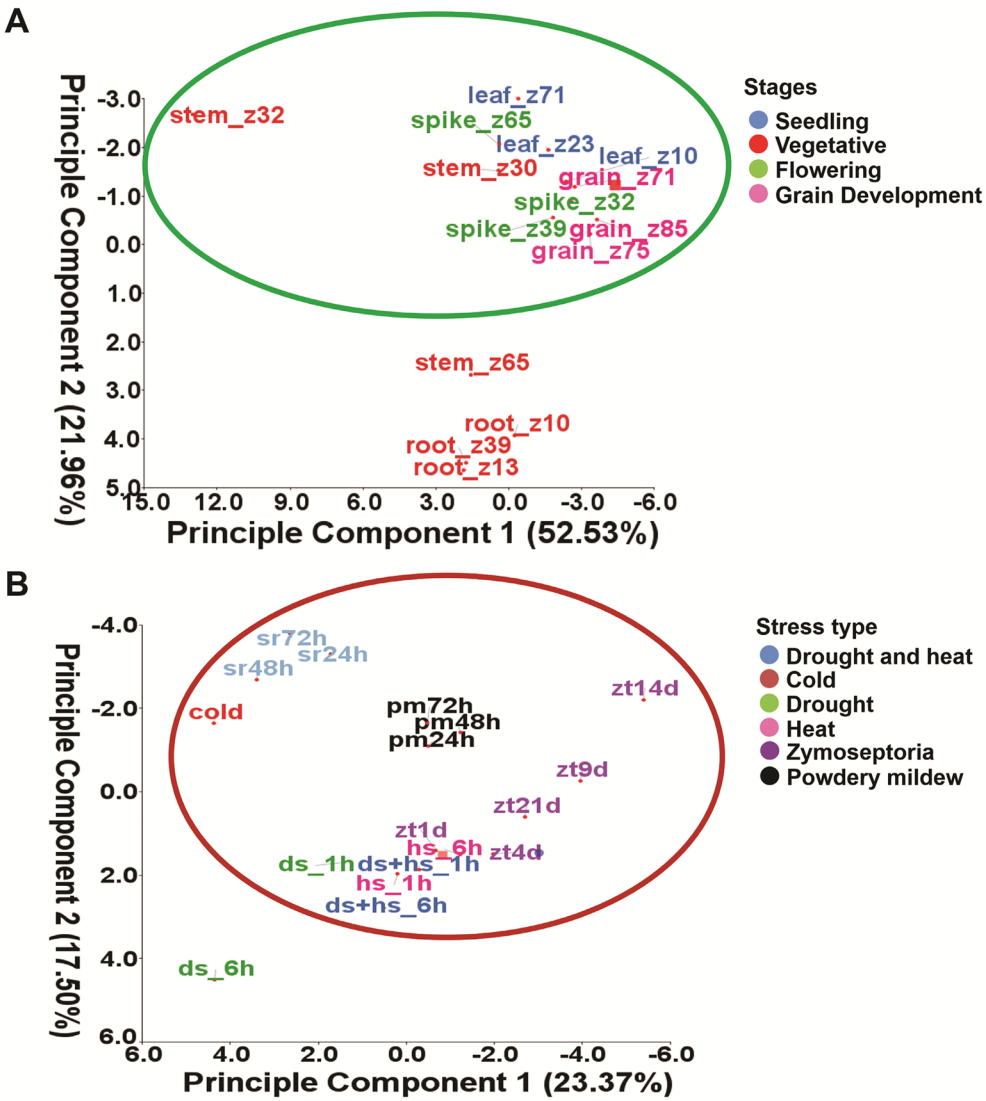


Fig. 10 PCA plots showing the different groups (**A**) Various tissues of wheat (**B**) *TaPSY* expression pattern in different stress conditions. Drought stress (DS), heat stress (HS), and shared drought and heat stress (DS + HS), *Zymoseptoria tritici* (Zt) infection, stripe rust (SR), and powdery mildew (PM). The expression levels were measured in terms of time, where "h" represents hours and "d" represents days

suggesting that these genes are prevalent and diversified in dicotyledonous plants (Fig. 1). The derivation of these type genes indicates that *PSY* genes might play an important role in morphogical development in the monocots and dicots [64–67]. Despite the clear distribution of *PSY* genes into well-known groups in *A. thaliana, O. sativa, Z. mays, G. max, P. trichocarpa M. truncatul* and *P. patens*. These results indicate that the expansion of the *PSY* gene family in wheat predominantly arose from segmental and whole-genome duplications. These types of duplications are known to play significant roles in the amplification and diversification of gene families, contributing to the adaptation and evolution of organisms over time [37, 65, 68]. The segmental and whole genome duplications likely provided ample opportunities for functional divergence and specialization within the *TaPSY* gene family, potentially leading to enhanced adaptability of wheat to different environmental conditions. These findings may provide a more comprehensive perspective on the evolution and regulation of *PSY* genes across different plant species and offer potential targets for further investigation of their roles in various biological processes and responses to environmental stimuli.

We mapped the 29 *TaPSY* genes onto the 21 chromosomes of wheat. The results revealed that these *TaPSY* genes are distributed across 12 wheat chromosomes

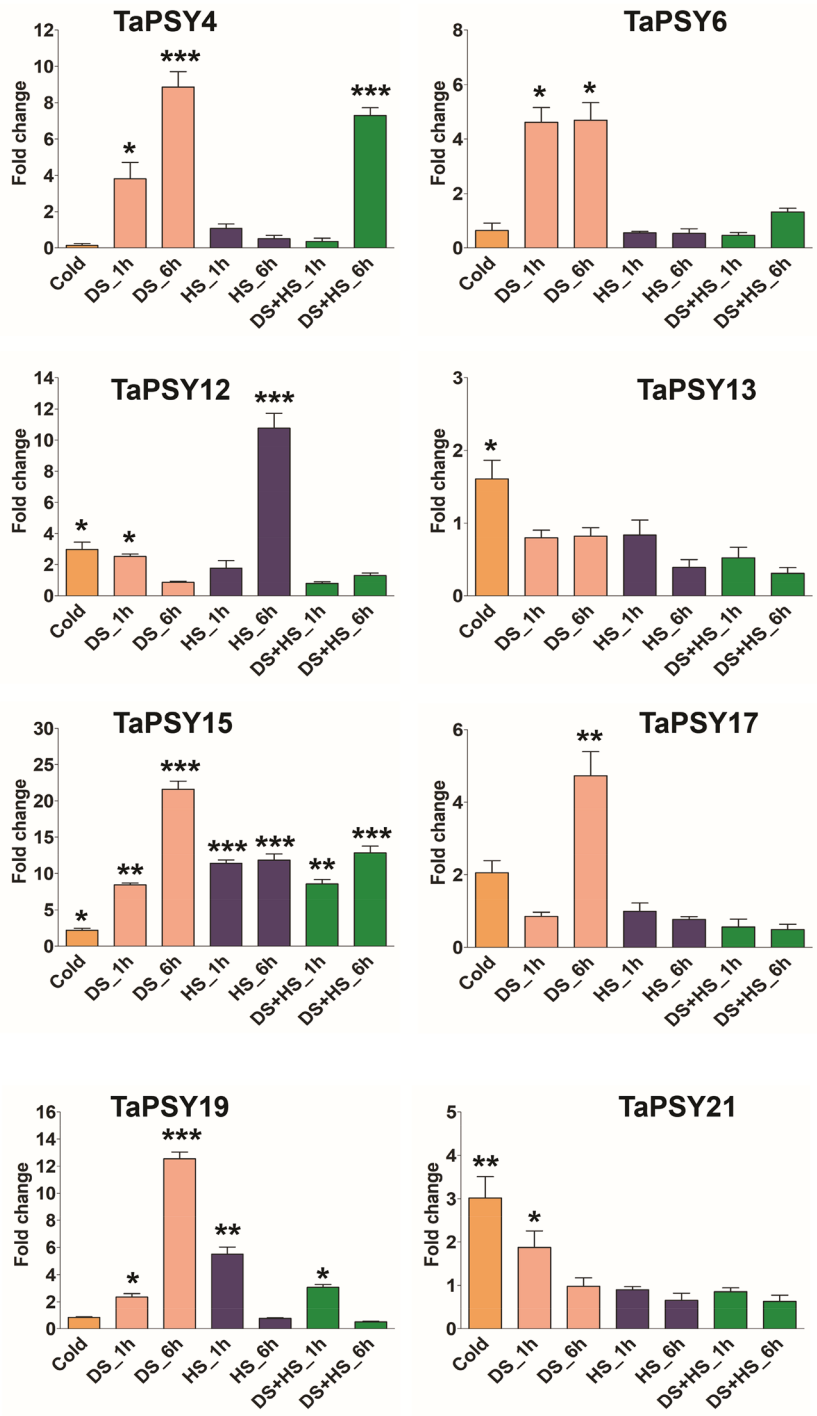


Fig. 11 qRT-PCR analysis was conducted to study the expression of *TaPSY* genes under various abiotic stress conditions in wheat. The asterisk symbol indicates significant differences in comparison to the control. The top bars represent the results of the Tukey HSD test at the < 0.05 and < 0.001 (***) levels, denoted as *$P < 0.05$ and ***$P < 0.001$, respectively. Drought stress (DS), heat stress (HS), and shared drought and heat stress (DS + HS). The expression levels were measured in terms of time, where "h" represents hours

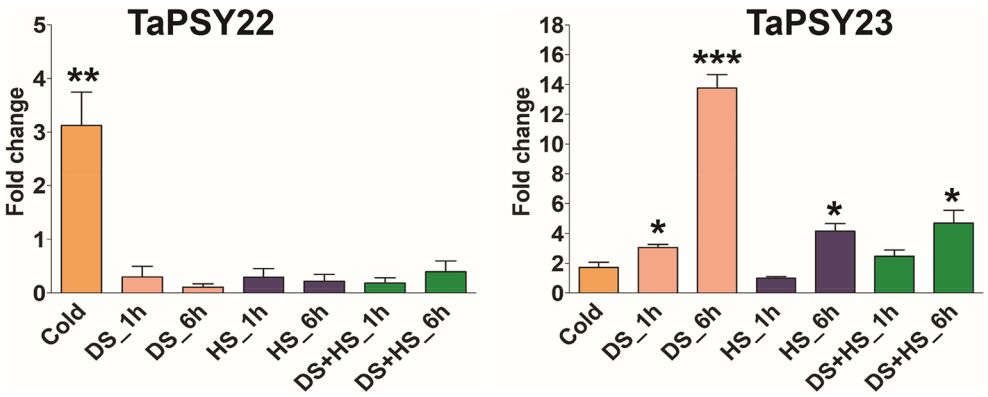


Fig. 11 continued

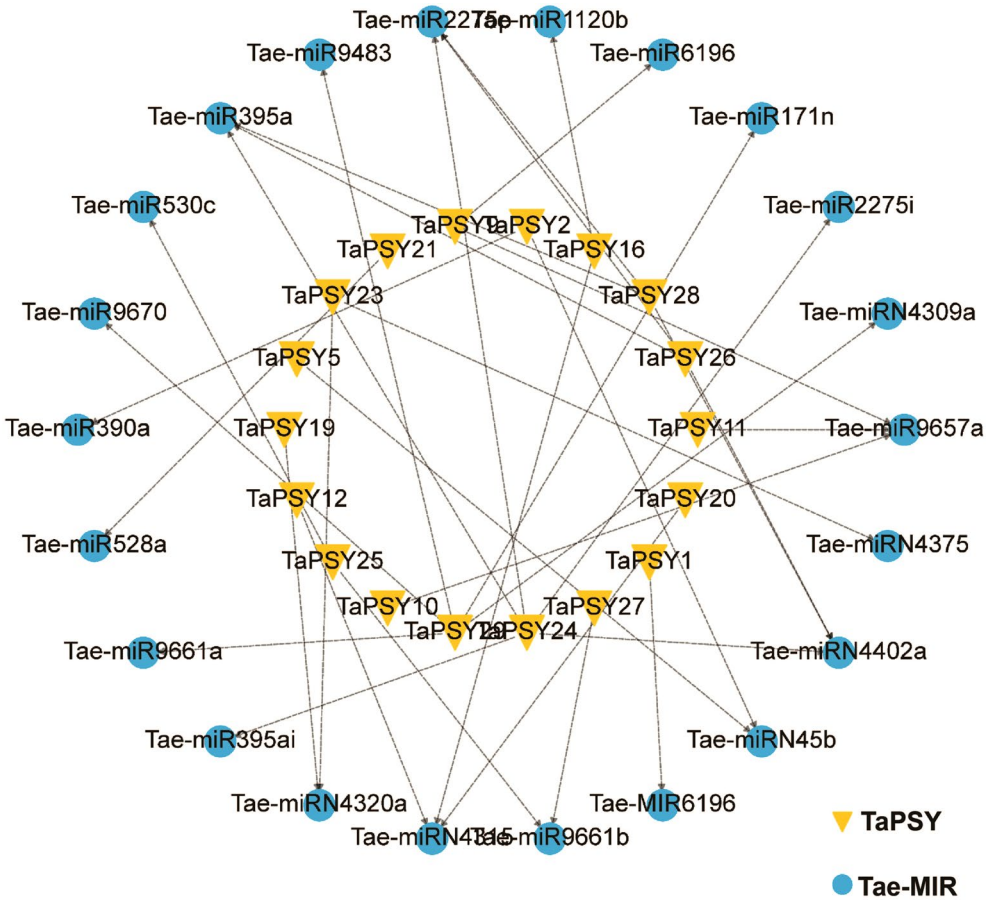


Fig. 12 The miRNA network includes specific targets within Wheat *PSY* gene family. The network features miRNAs and their corresponding target *TaPSY* genes, depicted as circles and triangles, respectively

(Fig. 2A and Table 1). We found gene clusters are present on Chr3A, Chr3B, and Chr3D (Fig. 2A, C). In wheat, analogous findings were noted within the PIN-FORMED (PIN), brassinazole-resistant (BZR), Proline-Rich Extensin-like Receptor Kinases (PERKs), and Aconitase (ACO) gene families [36, 37, 65, 66]. Therefore, the unequal distribution of *TaPSY* genes across the 12 chromosomes in wheat advocates the potential occurrence of gene addition and loss events through segmental and whole-genome duplication. Such events are

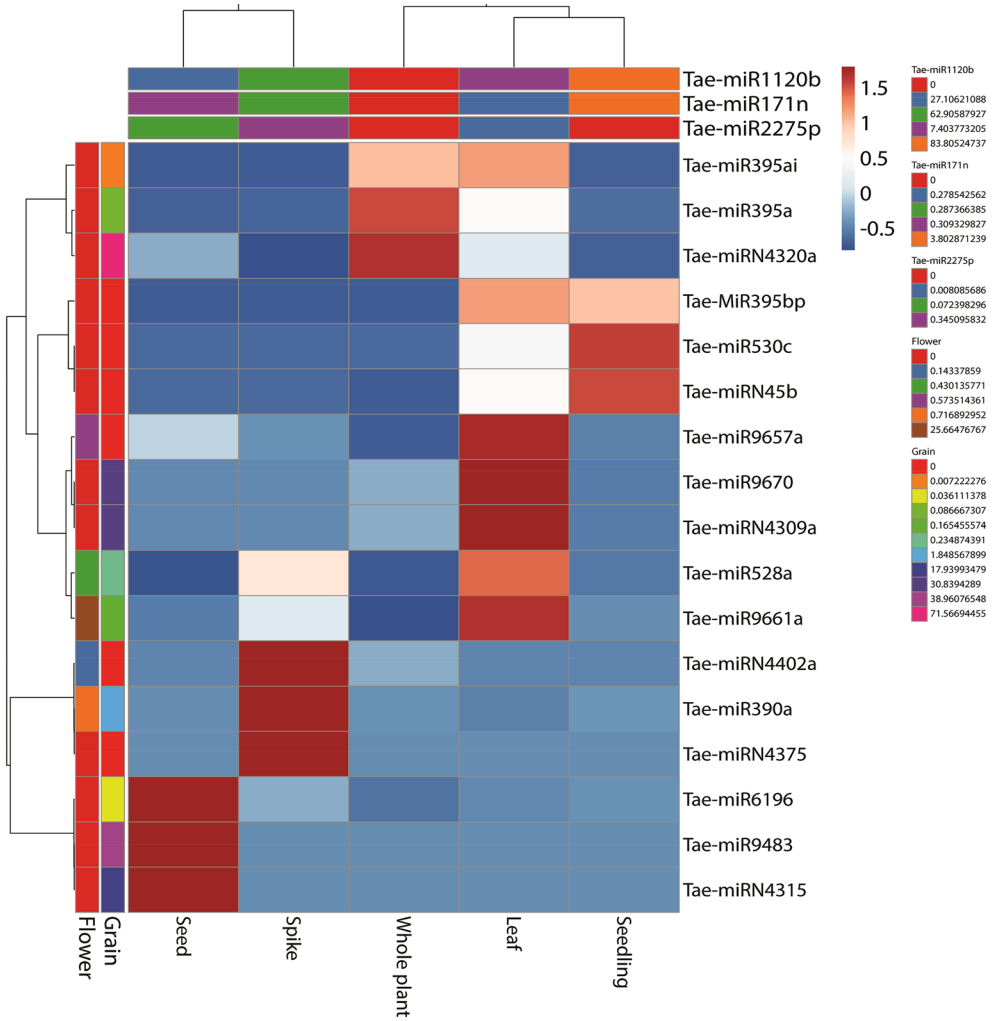


Fig. 13 Expression profile of potential miRNA and their *PSY* gene specific targets in Wheat. Heatmap displaying the expression pattern of miRNAs in different tissues and developmental stage in wheat

typical during the evolutionary history and can result in differences in gene numbers across different chromosomes. In this work, we identified nine duplicated gene pairs within the PSY gene family in wheat (Fig. 3; Fig. S3). The duplicated gene pairs are: *TaPSY3:TaPSY6, TaPSY15:TaPSY23, TaPSY9:TaPSY11, TaPSY24:TaPSY28, TaPSY14:TaPSY22, TaPSY4:TaPSY7, TaPSY25:TaPSY29, TaPSY12:TaPSY16* and *TaPSY13:TaPSY21*. These results indicate that the expansion of the PSY gene family in wheat predominantly arose from segmental and whole-genome duplications (Fig. 3 and Table S3). Further, The Ka/Ks value was less than one for the eight *TaPSY* genes indicating that duplicated *TaPSY* genes have undergone purifying or negative selection. However, one gene pair (*TaPSY3*/*TaPSY6*) has shown Ka/Ks value more than one suggesting that this gene pair had gone through a positive selection. Overall, this finding indicates that the *TaPSY* family genes has evolved under purifying selection, ensuring the preservation of crucial traits in wheat. Tandem repeats are responsible for generating gene clusters, whereas fragment repeats contribute to the emergence of homologous gene [69]. Further, we examined the syntenic relationships of *TaPSY* genes with those of other crop species, such as *B. distachyon, A. tauschii, A. thaliana*, and *O. sativa*. To identify orthologous gene pairs among genomes of different crop species, we employed MCScanX (Fig. 4 and Table S4). Among *TaPSYs* and other PSYs in *Ae. tauschii, T. dicoccoides, T. turgidum*, and *H. vulgare*, we identified 19, 13, 13, and 24 orthologous genes, respectively. Furthermore, the comparison with *PSY* genes from *B. distachyon, Ae. tauschii,* and *O. sativa*, we found that 24, 25, and 22 *TaPSY* genes, respectively, exhibited collinearity. In addition, fewer *TaPSY* genes consist minimum two pairs of orthologous genes, for

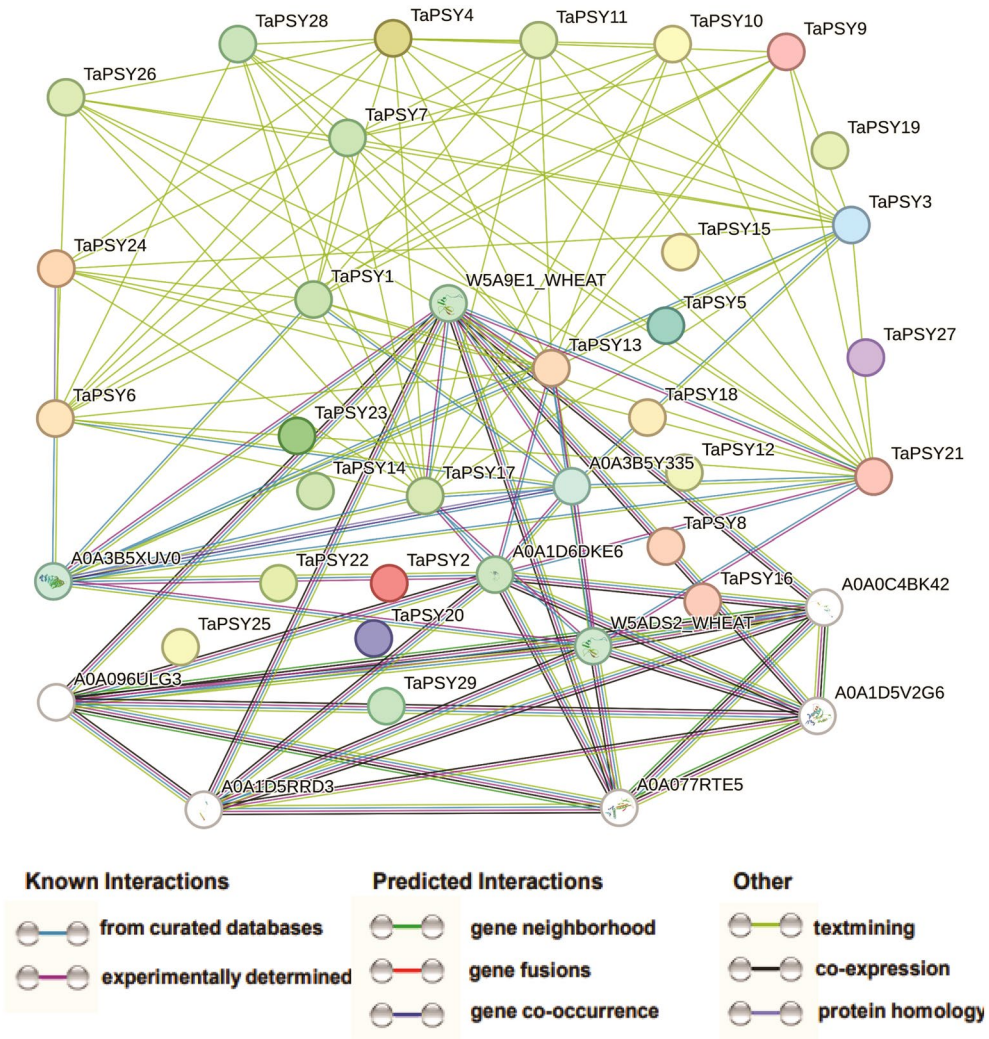


Fig. 14 Exploring protein–protein interactions among TaPSY proteins entails the examination of their connections. The protein–protein interaction network was constructed using STRING v9.1, where each protein is represented as a node, and their interactions are illustrated as edges. Additionally, the edges are color-coded to signify the type of evidence supporting each interaction

example, *TaPSY1, TaPSY3, TaPSY6, TaPSY9, TaPSY10, TaPSY11, TaPSY13, TaPSY15, TaPSY17, TaPSY19, TaPSY21, TaPSY23, TaPSY24, TaPSY26* and *TaPSY28* and these identified orthologous gene pairs may play a significant role in the evolution of the *PSY* gene family. In summary, these results collectively indicate that the *TaPSY* gene family may have originated from ancestral orthologous genes found in other crops.

Gene organization and motif analysis yield valuable comprehensions into the conserved and evolutionary variances of PSY genes in wheat. Through this analysis, it was noted that the number of exons and introns varied in various subfamilies (Fig. 5). This analysis also revealed that *TaPSY* gene family members exhibit slight variations in their gene structure (Fig. S4). *TaPSY* genes comprise 1–2 introns, for example, *TaPSY1*, *TaPSY3*, *TaPSY6*, *TaPSY13, TaPSY17, TaPSY21* and *TaPSY25* contain at least one intron, whereas majority of them contain a maximum of two introns such as *TaPSY2, TaPSY4, TaPSY5, TaPSY7, TaPSY8, TaPSY9, TaPSY10, TaPSY11, TaPSY12, TaPSY14, TaPSY15, TaPSY16, TaPSY18, TaPSY19, TaPSY20, TaPSY22, TaPSY23, TaPSY24, TaPSY26, TaPSY27, TaPSY28* and *TaPSY29*. Intron size plays a critical role in determining gene size. As an example, there exists a noteworthy contrast in gene size between the biggest gene, *TaPSY23* (1.8 kb), and the tiniest gene, *TaPSY21* (0.7 kb), primarily due to the disparity in their entire intron lengths (1.8 kb versus 0.7 kb). Many studies have emphasized the importance of introns in the evolutionary processes of different genes in crops

[70–72]. In plants, various gene families show diversity in the total number of introns, spanning from those with lesser, no introns or more introns [71, 72]. We speculate that the variation in the number of introns and exons could serve as a useful tool for documenting evolutionary history [73]. Further, to comprehend the precise functions of TaPSY gene family members, we explored conserved motif analysis of TaPSY proteins by the MEME webserver. Lastly, we detected the 10 motifs within the TaPSY proteins (Fig. 6A, B). Additionally, it was observed that Motif 1 and 2 exhibited high conservation across the majority of TaPSY proteins (Fig. 6A, B). An amino acid sequence alignment of TaPSY was performed, it was observed that all 29 TaPSY proteins contain a conserved contain PSY motif, which includes DY, N, H, and P domain (Fig. 7 and Fig. S5A). Furthermore, The Phyre2 webserver was utilized to determine the 3D structure of TaPSY proteins, aiming to understand their specific function in *T. aestivum* (Fig. S5B). Therefore, these findings would contribute to the comprehension and clarification of the exact role of TaPSY protein in regulating various signaling pathways associated with plant development processes and diverse environmental stimuli in wheat.

Cis-regulatory elements refer to noncoding DNA regions in the promoter that govern the transcription of adjacent genes [74–76]. In this investigation, we discovered numerous cis-regulatory elements within the 2000 bp upstream sequence of the *TaPSY* family genes. These elements encompassed various functional categories such as light response, phytohormones, circadian, cell cycle and seed-specific regulation, as well as stress response (Fig. 8A, B and Table S5). The *TaPSY* genes were found to encompass five phytohormone responsive CAREs. These components comprise SARE, MeJARE, ABRE, AuxRE, and GARE. The elements associated with light responses, MeJARE, ABRE, defense and stress responsiveness were predominantly found to in the *TaPSY* promoters (Fig. 8B). Hence, these results have shown that *TaPSY* genes might play a critical role in plant growth, development, and various stress conditions Furthermore, within the *TaPSY* genes, there are CARE elements associated with various functions, including endosperm expression, meristem expression, cell cycle regulation, circadian control, zein metabolism, and seed-specific regulation. The discovery of CAREs in *TaPSY* genes indicates that *TaPSY* genes might participate in diverse cellular processes. These results suggest that the *TaPSY* genes may have a vital role in regulating plant growth and stress responses by influencing multiple CAREs in wheat. Gene duplication has increased the count of gene family members under evolutionary force. Additionally, mutations within these genes have the potential to impact the expression patterns of gene family members [37, 77–79]. Signaling peptides have been demonstrated to be pivotal in various cellular processes, including cell differentiation and expansion, the preservation of stem cell characteristics, regulation of floral organ abscission, control of stomatal patterning, mediation of self-incompatibility, as well as initiation of defense responses and responses to diverse stressors [5, 10, 18–25]. The concept of gene expression blueprints provides an intriguing hypothesis that links the level of gene activity to their biological importance. In addition, the expression profile of *TaPSY* family genes were comprehensively studied in various tissues and multifactorial stress to gain a better understanding of the functional roles of *TaPSY* genes. In our investigation, we observed that 29 members of the *TaPSY* gene family displayed distinct expression patterns in different tissues, developmental stages, and under different biotic and abiotic stress conditions (Fig. 9A, B), for instance, the expression of *TaPSY4, TaPSY7, TaPSY13, TaPSY14, TaPSY15, TaPSY17, TaPSY18, TaPSY19, TaPSY21, TaPSY22, TaPSY23, TaPSY24, TaPSY26* and *TaPSY28* were highly elevated in stem_z32, while *TaPSY8, TaPSY9, TaPSY10, TaPSY11* and *TaPSY25* were up-regulated in stem_z65. Further, there was a notable increase in the expression levels of *TaPSY3* and *TaPSY6* in leaf_z71. The expression levels of *TaPSY5, TaPSY25, TaPSY27* and *TaPSY29* were induced root_z10, whereas *TaPSY29* in grain_z71. It was reported that *PSY1* activation led to cell elongation in both the roots and hypocotyls [2, 26, 29]. The overexpression of *AtPSKR1* in Arabidopsis alters growth patterns and cellular longevity [5]. It has been found that PSK regulates nodulation in Lotus [16]. The expression level of *AtPSY1* was detected across all tissues in Arabidopsis, distinguishing it from other *AtPSY1* members. However, heightened expression of *AtPSY1* was noted during late silique development, senescence and bolting stages. Additionally, a comparable expression pattern was observed for *AtPSY8*, with its expression level being notably elevated in root compared to other plant parts [4]. These results exhibited that *TaPSY* genes may participate in the development of different tissues in wheat. *TaPSY* family genes exhibited varying expression patterns in response to biotic stress conditions (Fig. 9B). The expression of *TaPSY5* in PM24h, *TaPSY5, TaPSY18* and *TaPSY29* in PM48h while, *TaPSY5, TaPSY18, TaPSY22* and *TaPSY29* were highly elevated in PM72h. In Sr72h, the expression of *TaPSY1, TaPSY10, TaPSY20*, and *TaPSY21* exhibited a significant increase. Similarly, in Zt4d, the expression levels of *TaPSY7, TaPSY13, TaPSY24, TaPSY25, TaPSY26*, and *TaPSY28* were markedly elevated. Notably, PSY1 is known to play a vital role in in response to plant defense [30]. The pathogen attacks activate PSY1 signaling, which down-regulates genes

implcated in salicylic acid signaling [30]. It has been demonstrated that *Xanthomonas oryzae* produces a sulfated peptide known as RaxX, which bears a significant similarity to the PSY1 [29]. Moreover, distinct transcript kinetics were observed for *TaPSY* genes in response to various environmental stresses, for example, the expression level of *TaPSY1, TaPSY15* and *TaPSY21* were highly raised in cold. The expression of *TaPSY9, TaPSY12* and *TaPSY16* were up-regulated in HS_6h, while *TaPSY4, TaPSY6, TaPSY7, TaPSY17, TaPSY19* and *TaPSY23* shown increased expression level in DS_6h (Fig. 9B). The GO analysis of *TaPSY* genes revealed significant enrichment in several biological process categories, including response to salt (GO:1902074), stimulus (GO:0050896) and fluoride (GO:1902617) (Fig. S7). In the cellular component category, the TaPSY genes were found to be enriched in cellular anatomical organization (GO:0110165) and membrane (GO:0016020) (Fig. S8 and S9). In the molecular category, TaPSY genes demonstrated enrichment in several transporter activities, including aldonate transmembrane transporter activity (GO:0042879), carboxylic acid transmembrane transporter activity (GO:0046943) and gluconate transmembrane transporter activity (GO:0015128) (Fig. S10). Thus, these enrichments highlight the participation of *TaPSY* genes in diverse biological processes including stress response, signaling pathways, and membrane-associated functions, indicating their pivotal roles in plant development, and adaptation to environmental challenges. Out of the 29 *TaPSY* genes, eighteen *TaPSY* genes including *TaPSY1, TaPSY2, TaPSY5, TaPSY9, TaPSY10, TaPSY11, TaPSY12, TaPSY16, TaPSY19, TaPSY20, TaPSY21, TaPSY23, TaPSY24, TaPSY25, TaPSY26, TaPSY27, TaPSY28* and *TaPSY29*, twenty miRNAs were found to target them, for example, *Tae-miR1120b, Tae-miR171n, Tae-miR2275p, Tae-miR390a, Tae-miR395a, Tae-miR395ai, Tae-MiR395bp, Tae-miR528a, Tae-miR530c, Tae-miR6196, Tae-miR9483, Tae-miR9657a, Tae-miR9661a, Tae-miR9670, Tae-miRN4309a, Tae-miRN4315, Tae-miRN4320a, Tae-miRN4375, Tae-miRN4402a* and *Tae-miRN45b* (Fig. 12; Table S10). Further, we examined the expression pattern of miRNAs in various tissues, such as flower, grain, leaf, seed, seedling, spike, and whole plant (Table S11). The identified miRNAs displayed unique expression across different tissues in wheat. This suggests that these miRNAs may have significant roles in regulating the expression of *TaPSY* gene family members during various developmental processes in wheat (Fig. 13). This knowledge may have practical applications in improving wheat crop productivity and resilience to environmental challenges. Moreover, thirteen TaPSY proteins exhibited interactions with fifteen distinct wheat-specific proteins. Remarkably, TaPSY1 exhibited interactions with TaPSY4, TaPSY7, TaPSY9, TaPSY10, TaPSY11, TaPSY13, TaPSY17, TaPSY21, TaPSY24, TaPSY26, and TaPSY28. Similarly, TaPSY13 was found to interact with TaPSY3, TaPSY4, TaPSY6, TaPSY7, TaPSY9, TaPSY24, TaPSY26, TaPSY28, W5A9E1_WHEAT, and W5ADS2_WHEAT. These results provide valuable insights that can inspire further investigations aimed at uncovering the roles of *TaPSY* genes in various biological processes. Therefore, the results have demonstrated that *TaPSY* gene family members potentially play an essential role in plant developmental processes and response to multifactorial stress in wheat. Consequently, these findings lay a robust foundation for further explorations aimed at unraveling the specific roles of *TaPSY* members in different tissues, responses to plant hormones, and diverse stress conditions in wheat.

Conclusions

In this study, we discovered 29 *TaPSY* genes within the wheat genome, and we further classified them into five subfamilies. Additionally, 29 *TaPSY* genes are distributed across 12 chromosomes of wheat, with 9 pairs of *TaPSY* involved in gene duplication events. Further, The Ka/Ks value was found to be less than one for the eight *TaPSY* genes indicating that duplicated *TaPSY* genes have undergone purifying or negative selection. The B and D sub-genomes comprise the highest *TaPSY* genes (10), followed by A sub-genomes (9). The *TaPSY* promoter region contains multiple CARE related to response to light, hormones and stress. The *TaPSY* family members showed a distinct expression pattern with variations in different tissues and under various stress conditions. Furthermore, we have identified putative candidate miRNAs targeting *TaPSY* genes and subsequently analyzed their expression profiles. Among the 29 *TaPSY* genes, eighteen were discovered to be targeted by twenty miRNAs. These discoveries provide a sturdy groundwork for investigations aimed at uncovering the specific roles of *TaPSY* family genes across various developmental stages, responses to plant hormones, and diverse stress conditions in wheat. Thus, the *TaPSY* genes and miRNAs we have identified hold potential for manipulation using genome editing tools. This could lead to the development of climate-smart crops with heightened resilience against environmental stress, particularly in the face of evolving global climate scenarios.

Abbreviations

PSY	Peptides containing Tyrosine Sulfation
CARE	Cis-acting regulatory elements
PTMS	Post-translationally modified small peptides
PSK	Phytosulfokine
RGFs	Root growth76 meristem factors
CIFs	Casparian strip integrity factors

miRNAs	Micro RNA
qRT-PCR	Quantitative real-time polymerase chain reaction
SARE	Salicylic acid response element
MeJARE	MeJA response element
ABRE	Abscisic acid responsive element
AuxRE	Auxin response element and
GARE	Gibberellin response element

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40538-024-00599-5.

Supplementary materials 1: **Fig. S1.** The plots of TaPSY genes display their kDa and pI. **Fig. S2.** *TaPSYs* are dispersed in a distinct group within the phylogenetic tree. **Fig. S3.** An evolutionary analysis of *TaPSY* genes was conducted. A phylogenetic tree was generated using MEGA11 with the NJ method and 1000 bootstrap replications. Duplicated gene pairs are indicated with a black asterisk. **Fig. S4.** The *TaPSY* gene family exhibits variations in the number of exons and introns. **Fig. S5.** The alignment and three-dimensional structure of TaPSY protein sequences were analyzed. **A** The *PSY* domain is highlighted with a green box. **B.** Predicted 3D structures of TaPSY proteins were generated. **Fig. S6**: Significant Blast2GO statistics based on BLAST search against the non-redundant proteins sequence database. **A**: Number of sequence with length, **B**: Hit coverage distribution, **C**: Sequence coverage distribution, **D**: Sequence similarity distribution, **E**: Data distribution, **F**: BLASTp hit species distribution, **G**: BLASTp top-hit species distribution. **Fig. S7**: The distribution of Gene Ontology terms in the TaPSY gene family was predicted using Blast2GO in Biological Process category. **Fig. S8**: The distribution of Gene Ontology terms in the TaPSY gene family was predicted using Blast2GO in Cellular Component category. **Fig. S9.** The distribution of Gene Ontology terms in the TaPSY gene family was predicted using AgriGO in Cellular Component category. **Fig. S10**: The distribution of Gene Ontology terms in the TaPSY gene family was predicted using Blast2GO in Molecular Function category.

Supplementary materials 2: **Table S1**: The genomic, CDS (coding sequence), protein, and promoter sequences of *TaPSY*. **Table S2**: A phylogenetic tree was generated using *PSY* proteins from *T. aestivum, A. thaliana, O. sativa, Z. mays, G. max, P. trichocarpa, M. truncatula* and *P. Patens*. **Table S3**: The Ka/Ks score and allocation of duplicated *PSY* genes in wheat. **Table S4**: Orthologous relationships of *TaPSY* genes with other *PSY* genes in *B. distachyon, Ae. tauschii, T. dicoccoides, T. turgidum, H. vulgare* and *O. sativa*. **Table S5**: The multiple cis-regulatory elements found in the promoter region of *TaPSY* genes, **Table S6**: The annotation of TaPSY genes using Blast2GO. **Table S7**: The annotation of TaPSY genes using AgriGo analysis. **Table S8**: The annotation of TaPSY genes using eggNOGmapper. **Table S9**: qRT-PCR primer sequences for *TaPSY* genes. **Table S10**: Identified potential miRNAs and their PSY specific target genes in wheat. **Table S11**: Expression profile of miRNAs in different tissues and their PSY specific target genes in wheat. **Table S12**: The network illustrating protein-protein interactions between TaPSY and other proteins in wheat.

Acknowledgements
MSK wishes to extend their acknowledgment to the Department of Genetics and Plant Breeding within the Faculty of Agriculture at Sri Sri University for facilitating the necessary infrastructure for conducting *insilico* analysis. This project was supported by Researchers Supporting Project number (RSP2025R5), King Saud University, Riyadh, Saudi Arabia.

Author contributions
M.S.K. and S.K.L. were responsible for the design and composition of the manuscript, and they also supervised the study. Valuable input to this study was provided by B.S.K., C.R., S.M., S.K., S.-M.C., S.A.A. and M.J.A. All the authors have reviewed and consented to the final published version of the manuscript.

Funding
Not applicable.

Availability of data and materials
Data is available in the manuscript and in the Supplementary Materials.

Declarations

Ethics approval and content to participate
Not applicable.

Consent for publication
Not applicable.

Institutional review board statement
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Genetics and Plant Breeding, Faculty of Agriculture, Sri Sri University, Cuttack 754006, Odisha, India. [2]Krishi Vigyan Kendra, Bikaner II, Swami Keshwanand Rajasthan Agricultural University, Bikaner 334603, Rajasthan, India. [3]ICAR-Central Institute for Arid Horticulture, Bikaner 334006, Rajasthan, India. [4]Department of Biotechnology, School of Life Sciences, Mahatma Gandhi Central University, East Champaran- 845404, Motihari, Bihar, India. [5]Department of Bioscience and Bioengineering, Indian Institute of Technology, 208016, Kanpur, Uttar Pradesh, India. [6]Department of Life Science, Dongguk University, Dong-gu-10326, Ilsan, Republic of South Korea. [7]Department of Botany and Microbiology, College of Science, King Saud University, 11451 Riyadh, Saudi Arabia. [8]Department of Botany, Hindu College Moradabad (Mahatma Jyotiba Phule Rohilkhand University Bareilly), Moradabad 244001, India. [9]Gujarat Biotechnology University, Gandhinagar 382355, Gujarat, India.

Received: 22 March 2024 Accepted: 19 May 2024
Published online: 18 June 2024

References

1. Matsubayashi Y. Posttranslationally modified small-peptide signals in plants. Annu Rev Plant Biol. 2014;65:385–413.
2. Amano Y, Tsubouchi H, Shinohara H, Ogawa M, Matsubayashi Y. Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in Arabidopsis. Proc Natl Acad Sci USA. 2007;104(46):18333–8.
3. Marshall E, Costa LM, Gutierrez-Marcos J. Cysteine-rich peptides (CRPs) mediate diverse aspects of cell–cell communication in plant reproduction and development. J Exp Bot. 2011;62(5):1677–86.
4. Tost AS, Kristensen A, Olsen LI, Axelsen KB, Fuglsang AT. The PSY peptide family—expression, modification and physiological implications. Genes. 2021;12(2):218.
5. Matsubayashi Y, Ogawa M, Kihara H, Niwa M, Sakagami Y. Disruption and overexpression of Arabidopsis phytosulfokine receptor gene affects cellular longevity and potential for growth. Plant Physiol. 2006;142(1):45–53.
6. Murphy E, Smith S, De Smet I. Small signaling peptides in Arabidopsis development: how cells communicate over a short distance. Plant Cell. 2012;24(8):3198–217.
7. Ryan CA, Pearce G, Scheer J, Moura DS. Polypeptide hormones. Plant Cell. 2002. https://doi.org/10.1105/tpc.010484.
8. Matsubayashi Y. Exploring peptide hormones in plants: identification of four peptide hormone-receptor pairs and two post-translational modification enzymes. Proc Jpn Acad B. 2018;94(2):59–74.
9. Tavormina P, De Coninck B, Nikonorova N, De Smet I, Cammue BP. The plant peptidome: an expanding repertoire of structural features and biological functions. Plant Cell. 2015;27(8):2095–118.
10. Matsubayashi Y, Hanai H, Hara O, Sakagami Y. Active fragments and analogs of the plant growth factor, phytosulfokine: structure–activity relationships. Biochem Biophys Res Commun. 1996;225(1):209–14.
11. Matsubayashi Y, Sakagami Y. Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of Asparagus officinalis L. Proc Natl Acad Sci USA. 1996;93(15):7623–7.

12. Doblas VG, Smakowska-Luzan E, Fujita S, Alassimone J, Barberon M, Madalinski M, et al. Root diffusion barrier control by a vasculature-derived peptide binding to the SGN3 receptor. Science. 2017;355(6322):280–4.
13. Komori R, Amano Y, Ogawa-Ohnishi M, Matsubayashi Y. Identification of tyrosylprotein sulfotransferase in Arabidopsis. Proc Natl Acad Sci USA. 2009;106(35):15067–72.
14. Moore KL. The biology and enzymology of protein tyrosine O-sulfation. J Biol Chem. 2003;278(27):24243–6.
15. Matsubayashi Y. Post-translational modifications in secreted peptide hormones in plants. Plant Cell Physiol. 2011;52(1):5–13.
16. Wang J, Li H, Han Z, Zhang H, Wang T, Lin G, et al. Allosteric receptor activation by the plant peptide hormone phytosulfokine. Nature. 2015;525(7568):265–8.
17. Kaufmann C, Sauter M. Sulfated plant peptide hormones. J Exp Bot. 2019;70(16):4267–77.
18. Ito Y, Nakanomyo I, Motose H, Iwamoto K, Sawa S, Dohmae N, et al. Dodeca-CLE peptides as suppressors of plant stem cell differentiation. Science. 2006;313(5788):842–5.
19. Fletcher JC, Brand U, Running MP, Simon R, Meyerowitz EM. Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. Science. 1999;283(5409):1911–4.
20. Schopfer CR, Nasrallah ME, Nasrallah JB. The male determinant of self-incompatibility in Brassica. Science. 1999;286(5445):1697–700.
21. Butenko MA, Patterson SE, Grini PE, Stenvik G-E, Amundsen SS, Mandal A, et al. Inflorescence deficient in abscission controls floral organ abscission in Arabidopsis and identifies a novel family of putative ligands in plants. Plant Cell. 2003;15(10):2296–307.
22. Hara K, Kajita R, Torii KU, Bergmann DC, Kakimoto T. The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. Genes Dev. 2007;21(14):1720–5.
23. Kesawat MS, Manohar S, Kherawat BS, Kumar S, Lenka SK, Parameswaran C, et al. Genome-wide survey of peptides containing tyrosine sulfation (PSY) gene family and potential PSY specific miRNA revealed their role in plant development and diverse stress conditions in rice (*Oryza sativa* L.). Plant Stress. 2024. https://doi.org/10.1016/j.stress.2024.100412.
24. Djordjevic MA, Mohd-Radzman NA, Imin N. Small-peptide signals that control root nodule number, development, and symbiosis. J Exp Bot. 2015;66(17):5171–81.
25. Marmiroli N, Maestri E. Plant peptides in defense and signaling. Peptides. 2014;56:30–44.
26. Fuglsang AT, Kristensen A, Cuin TA, Schulze WX, Persson J, Thuesen KH, et al. Receptor kinase-mediated control of primary active proton pumping at the plasma membrane. Plant J. 2014;80(6):951–64.
27. Falhof J, Pedersen JT, Fuglsang AT, Palmgren M. Plasma membrane H+-ATPase regulation in the center of plant physiol. Mol Plant. 2016;9(3):323–37.
28. Palmgren MG. Plant plasma membrane H+-ATPases: powerhouses for nutrient uptake. Annu Rev Plant Biol. 2001;52(1):817–45.
29. Pruitt RN, Joe A, Zhang W, Feng W, Stewart V, Schwessinger B, et al. A microbially derived tyrosine-sulfated peptide mimics a plant peptide hormone. New Phytol. 2017;215(2):725–36.
30. Mosher S, Kemmerling B. PSKR1 and PSY1R-mediated regulation of plant defense responses. Plant Signal Behav. 2013;8(5): e24119.
31. de Bang TC, Lundquist PK, Dai X, Boschiero C, Zhuang Z, Pant P, et al. Genome-wide identification of Medicago peptides involved in macronutrient responses and nodulation. Plant Physiol. 2017;175(4):1669–89.
32. Sin W-C, Lam H-M, Ngai S-M. Identification of diverse stress-responsive xylem sap peptides in soybean. Int J Mol Sci. 2022;23(15):8641.
33. Wang YH, Irving HR. Developing a model of plant hormone interactions. Plant Signal Behav. 2011;6(4):494–500.
34. Consortium IWGS, Appels R, Eversole K, Stein N, Feuillet C, Keller B, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science. 2018. https://doi.org/10.1126/science.aar7191.
35. Shiferaw B, Prasanna BM, Hellin J, Bänziger M. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. Food Secur. 2011;3:307–27.
36. Kesawat MS, Kherawat BS, Ram C, Singh A, Dey P, Gora JS, et al. Genome-wide identification and expression profiling of aconitase gene family members reveals their roles in plant development and adaptation to diverse stress in *Triticum aestivum* L. Plants. 2022;11(24):3475.
37. Kesawat MS, Kherawat BS, Singh A, Dey P, Routray S, Mohapatra C, et al. Genome-wide analysis and characterization of the proline-rich extensin-like receptor kinases (PERKs) gene family reveals their role in different developmental stages and stress conditions in wheat (*Triticum aestivum* L.). Plants. 2022;11(4):496.
38. Joshi A, Mishra B, Chatrath R, Ortiz Ferrara G, Singh RP. Wheat improvement in India: present status, emerging challenges and future prospects. Euphytica. 2007;157:431–46.
39. Gilland B. World population and food supply: can food production keep pace with population growth in the next half-century? Food Pol. 2002;27(1):47–63.
40. Šramková Z, Gregová E, Šturdík E. Chemical composition and nutritional quality of wheat grain. Acta Chim Slov. 2009;2(1):115–38.
41. Govindan V, Michaux KD, Pfeiffer WH. Nutritionally enhanced wheat for food and nutrition security. In: Reynolds MP, Braun H-J, editors. Wheat Improvement: Food Security in a Changing Climate. Cham: Springer International Publishing; 2022.
42. Kumar P, Yadava R, Gollen B, Kumar S, Verma RK, Yadav S. Nutritional contents and medicinal properties of wheat: a review. Life Sci Med Res. 2011;22(1):1–10.
43. Finn R, Griffiths-Jones S, Bateman A. Identifying protein domains with the Pfam database. Curr Protoc Bioinform. 2003. https://doi.org/10.1002/0471250953.bi0205s01.
44. Letunic I, Doerks T, Bork P. SMART 6: recent updates and new developments. Nucleic Acids Res. 2009. https://doi.org/10.1093/nar/gkn808.
45. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. Nucleic Acids Res. 2015;43(D1):D222–6.
46. Kozlowski LP. IPC–isoelectric point calculator. Biol Direct. 2016;11(1):1–16.
47. Nakai K, Horton P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci. 1999;24(1):34–5.
48. Savojardo C, Martelli PL, Fariselli P, Profiti G, Casadio R. BUSCA: an integrative web server to predict subcellular localization of proteins. Nucleic Acids Res. 2018;46(W1):W459–66.
49. Kumar S, Tamura K, Nei M. MEGA: molecular evolutionary genetics analysis software for microcomputers. Bioinformatics. 1994;10(2):189–91.
50. Wolfe D, Dudek S, Ritchie MD, Pendergrass SA. Visualizing genomic information across chromosomes with PhenoGram. BioData Min. 2013;6:1–12.
51. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012. https://doi.org/10.1093/nar/gkr1293.
52. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202.
53. Guo A-Y, Zhu Q-H, Chen X, Luo J-C. GSDS: a gene structure display server. Yi Chuan= Hereditas. 2007;29(8):1023–6.
54. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. Nucleic Acids Res. 2015;43(W1):W39–49.
55. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10(6):845–58.
56. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325–7.
57. Conesa A, Götz S. Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int J Plant Biol. 2008. https://doi.org/10.1155/2008/619832.
58. Du Z, Zhou X, Ling Y, Zhang Z, Su Z. agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res. 2010;38(suppl_2):W64–70.
59. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. Nucleic Acids Res. 2015;43(W1):W566–70.
60. Kesawat MS, Kim DK, Zeba N, Suh MC, Xia X, Hong CB. Ectopic RING zinc finger gene from hot pepper induces totally different genes in lettuce and tobacco. Mol Breed. 2018;38:1–24.
61. Kim DK, Kesawat MS, Hong CB. One gene member of the ADP-ribosylation factor family is heat-inducible and enhances seed germination in *Nicotiana tabacum*. Genes & Genom. 2017;39:1353–65.

62. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. Methods. 2001;25(4):402–8.
63. Narancio R, John U, Mason J, Spangenberg G. Selection of optimal reference genes for quantitative RT-PCR transcript abundance analysis in white clover (*Trifolium repens* L.). Funct Plant Biol. 2018;45(7):737–44.
64. Kesawat MS, Das BK, Bhaganagare GR, Manorama. Genome-wide identification, evolutionary and expression analyses of putative Fe–S biogenesis genes in rice (*Oryza sativa*). Genome. 2012;55(8):571–83.
65. Kesawat MS, Kherawat BS, Singh A, Dey P, Kabi M, Debnath D, et al. Genome-wide identification and characterization of the brassinazole-resistant (BZR) gene family and its expression in the various developmental stage and stress conditions in wheat (Triticum aestivum L.). Int J Mol Sci. 2021;22(16):8743.
66. Kumar M, Kherawat BS, Dey P, Saha D, Singh A, Bhatia SK, et al. Genome-wide identification and characterization of PIN-FORMED (PIN) gene family reveals role in developmental and various stress conditions in *Triticum aestivum* L. Int J Mol Sci. 2021;22(14):7396.
67. Lespinet O, Wolf YI, Koonin EV, Aravind L. The role of lineage-specific gene family expansion in the evolution of eukaryotes. Genome Res. 2002;12(7):1048–59.
68. Lawton-Rauh A. Evolutionary dynamics of duplicated genes in plants. Mol Phylogenetics Evol. 2003;29(3):396–409.
69. Hughes AL. The evolution of functionally novel proteins after gene duplication. Proc R Soc B. 1994;256(1346):119–24.
70. William Roy S, Gilbert W. The evolution of spliceosomal introns: patterns, puzzles and progress. Nat Rev Genet. 2006;7(3):211–21.
71. Roy SW, Penny D. Patterns of intron loss and gain in plants: intron loss–dominated evolution and genome-wide comparison of *O. sativa and A. thaliana*. Mol Biol Evol. 2007;24(1):171–81.
72. Liu H, Lyu HM, Zhu K, Van de Peer Y, Cheng ZM. The emergence and evolution of intron-poor and intronless genes in intron-rich plant gene families. Plant J. 2021;105(4):1072–82.
73. Lecharny A, Boudet N, Gy I, Aubourg S, Kreis M. Introns in, introns out in plant gene families: a genomic approach of the dynamics of gene structure. J Struct Funct Genomics. 2003;3:111–6.
74. Hernandez-Garcia CM, Finer JJ. Identification and validation of promoters and cis-acting regulatory elements. Plant Sci. 2014;217:109–19.
75. Roy AL, Sen R, Roeder RG. Enhancer–promoter communication and transcriptional regulation of Igh. Trends Immunol. 2011;32(11):532–9.
76. Kesawat MS, Kherawat BS, Katara JL, Parameswaran C, Misra N, Kumar M, et al. Genome-wide analysis of proline-rich extensin-like receptor kinases (PERKs) gene family reveals their roles in plant development and stress conditions in *Oryza sativa* L. Plant Sci. 2023;334: 111749.
77. Kesawat MS, Satheesh N, Kherawat BS, Kumar A, Kim H-U, Chung S-M, et al. Regulation of reactive oxygen species during salt stress in plants and their crosstalk with other signaling molecules—Current perspectives and future directions. Plants. 2023;12(4):864.
78. Heidari P, Abdullah FS, Poczai P. Magnesium transporter gene family: genome-wide identification and characterization in *Theobroma cacao*, *Corchorus capsularis*, and *Gossypium hirsutum* of family Malvaceae. Agronomy. 2021;11(8):1651.
79. Heidari P, Puresmaeli F, Mora-Poblete F. Genome-wide identification and molecular evolution of the magnesium transporter (MGT) gene family in *Citrullus lanatus* and *Cucumis sativus*. Agronomy. 2022;12(10):2253.

Publisher's Note