BMC Genomics

COMMENT

Open Access

Functional characterization of 2,3-oxidosqualene cyclase8 in *Taraxacum mongolicum*: overexpression enhances taraxasterol biosynthesis and antioxidant capacity

Lu Yang[1†], Huan He[1†], Guangzhi Jiang[1†], Jiayin Wang[1], Pingping Du[1], Lele Liu[1], Hongbin Li[1], Zhuang Meng[1], Haitao Shen[1], Shuangquan Xie[1], Fei Wang[1*] and Quanliang Xie[1*]

Abstract

Background Taraxasterol, a pentacyclic triterpenoid compound, has been widely used in traditional and modern medicine because of its pharmacological properties such as anti-inflammatory, antioxidative and antitumor effects. 2,3-oxidosqualene cyclase (OSC) is highly important for the generation of phytosterols and triterpenoid compounds and the structural diversity of natural products. However, the specific role of *TmOSCs* in the taraxasterol biosynthesis pathway has not yet been precisely resolved.

Results In this study, 10 *TmOSC* gene family members were identified via the *Taraxacum mongolicum* (dandelion) genome and classified into three subgroups. Phylogenetic and collinearity analyses revealed evolutionary conservation among OSC proteins from *Asteraceae* species. RNA-seq data analysis revealed that *TmOSC8* and *TmOSC10* were highly expressed in nutrient-containing tissues. In addition, methyl jasmonate (MeJA) and abscisic acid (ABA) significantly induced the expression of *TmOSC3*, and the change in its relative expression was consistent with the taraxasterol content. The relative expression level of the *TmOSC8*-overexpressioning line was significantly increased by approximately 20 fold compared with that of the wild type, and the content of taraxasterol was significantly increased to 3 fold greater than that of the wild type.

Conclusion This study systematically analyzed the evolutionary characteristics and expression patterns of the *TmOSC* gene family, suggested potential key roles of *TmOSC3* and *TmOSC8* in sterol synthesis and stress response, and provided a preliminary theoretical basis for the metabolic engineering of medicinal components in dandelion.

†Lu Yang, Huan He and Guangzhi Jiang contributed equally to this work.

*Correspondence:
Fei Wang
feiw@shzu.edu.cn
Quanliang Xie
xiequanliang001@shzu.edu.cn

Full list of author information is available at the end of the article

Keywords Dandelion, OSC, Taraxasterol, MeJA, ABA

Introduction

Dandelion is a perennial herb in the *Asteraceae* family that is widely distributed in China. It is valued in traditional medicine for its diverse secondary metabolites, which exhibit significant biological activities [1]. These metabolites underlie not only have medicinal properties but also play crucial roles in environmental adaptation and defense [2, 3]. Advances in biotechnology have led to increased focus on the functions of secondary dandelion metabolites and their regulatory mechanisms [4]. Taraxasterol, a pentacyclic triterpenoid compound, is predominantly found in the roots of dandelion. Its molecular formula is $C_{30}H_{50}O$, and its molecular structure is highly similar to that of steroid hormones containing cyclopentane and polyhydrophenanthrene [5]. Triterpenoid compounds are synthesized through two pathways: the mevalonate (MVA) pathway in the cytoplasm and the methylerythritol phosphate (MEP) pathway in the plastid [6]. Taraxasterol enhances cellular antioxidant capacity and regulates the expression of apoptosis- and autophagy- related genes by activating the nuclear factor erythroid-derived 2-like 2 (Nrf2) signaling pathway, thereby alleviating aflatoxin B1-induced liver injury [7]. Bao demonstrated that taraxasterol inhibits the growth of hepatocellular carcinoma by increasing the expression of histidine triad nucleotide binding protein (HINT), which regulates gene and cell cycle protein expression. They also suggested that taraxasterol might be a candidate for the treatment of human hepatocellular carcinoma [8]. Many natural active ingredients in plants, such as taraxasterol, have anti-inflammatory, anti-oxidative [9] and antitumor properties [10] and are less likely to induce drug resistance [11]. Taraxasterol can be used alongside antibiotics to increase their efficacy and reduce antibiotic use, indicating broad application prospects. Given the important medicinal value of taraxasterol and its low abundance, research on the biosynthesis and metabolic engineering of plant-specific triterpenoids is becoming increasingly important.

Oxidosqualene cyclase (2,3-oxidosqualene cyclase, OSC), as a key enzyme in the triterpene metabolic pathway, catalyzes the formation of a variety of triterpene skeletons, which can be further modified to form a wide range of biologically active triterpenes [12]. *OSCs* play an important role in the sterol synthesis pathway, therefore, by regulating the activity and expression level of oxidized squalene cyclase, the synthesis yield and type of sterol can be effectively regulated [13]. In several studies, the *OSC* gene family was initially investigated in numerous plants. Thirty-nine *OSC* genes which are highly conserved, have been identified in the heterozygous hexaploid plant. *Triticum aestivum* (*T. aestivum*), and some members of this gene family are associated with powdery mildew resistance [14]. In the medicinal plant *Melia toosendan* (*M. toosendan*), six *OSC* genes were identified by analyzing the expression profile of *OSC* genes in relation to metabolites, of which *MtOSC1* plays a crucial role in catalyzing the generation of cynarin biosynthesis [15]. It has also been reported in *Phellodendron amurense* [16], *Momordica charantia* (*M. charantia*) [17], *Rosa rugosa* (*R. rugosa*) [18] and *Oryza sativa* (*O. sativa*) [9]. Although these studies provide important clues to the mechanism of triterpene synthesis, the specific synthesis mechanism of taraxasterol remains unclear, and studies on the specific biosynthesis pathways of *OSCs* for active ingredients in plants, such as dandelion, are extremely limited. Secondary metabolites, which are ubiquitously distributed in medicinal plants, are primarily categorized into three classes: phenolic acids, terpenoids, and nitrogenous compounds. Environmental stresses, including salt stress, drought stress, mechanical wounding, and pathogen infection, have been shown to trigger the accumulation of these metabolites [19, 20]. Among these, MeJA has been shown to induce the accumulation of secondary metabolites in several plant species [21, 22]. For example, MeJA has been shown to increase taraxasterol content in dandelion root suspension cells [23], induce more effectively in *Astragalus membranaceus* hairy roots [24], and promote triterpene ganoderic acid (GAs) biosynthesis in *Ganoderma lucidum* [25]. However, the link between *TmOSC* genes and MeJA-induced taraxasterol synthesis in dandelion remains unclear.

Oxidative stress accompanies all stages of biosynthesis, and when an organism is subjected to oxidative stress, overexpressed reactive oxygen species (ROS) and reactive nitrogen radicals cannot be scavenged in a timely manner, thus disrupting overall biosynthetic homeostasis [9]. Taraxasterol has been reported to have significant therapeutic effects in the treatment of acute pneumonia by lowering the lung index, ameliorating lung pathological damage, and decreasing lipid peroxidation and malondialdehyde (MDA) production [10]. Taraxasterol can also enhance immune system function and antioxidant capacity by increasing the activity of glutathione reductase (GSH), thus preventing cigarette smoke (CS)-induced pneumonia in mice [26]. In addition, taraxasterol can attenuate lipid peroxidation, reduce the serum MDA content, and increase GSH and superoxide dismutase (SOD) activities in immunologically hepatic-injured mice, thereby alleviating immune hepatic injury by reducing oxidative stress [9]. Taraxasterol also inhibited the production of ROS in a mouse model of colitis, and

the strength of its inhibitory effect was positively correlated with its concentration [27]. This study investigated the origin and gene function of *TmOSC* genes in depth, revealing their specific role in taraxasterol biosynthesis by combining dandelion secondary metabolite function with abiotic stress adaptation mechanisms.

In summary, this study utilized a range of techniques, including bioinformatics analysis, gene expression profiling, molecular biology techniques, and physiological and biochemical assays. By mining the genomic data of dandelion, we aimed to identify *TmOSC* gene family members, analyze their properties and conserved motifs, and decode their expression patterns across tissues and under abiotic stress via transcriptome sequencing and quantitative real-time PCR (qRT−PCR). We will also validate the effects of *TmOSCs* on taraxasterol synthesis and stress resistance through gene overexpression. This study provides a theoretical basis for the molecular breeding of crops with high taraxasterol contents, enhances their antioxidant capacity and yield, and provides insights into stress-resistant breeding in other crops.

Methods

Plant material and hormone treatment

Plants and seeds of dandelion were provided by the College of Life Sciences, Shihezi University. In the laboratory, the plants were cultured in a greenhouse (24 ± 2 °C; photoperiod: 16 h light/8 h dark) for three months. Uniformly grown dandelion plants were precultured in Hoagland's medium for 2 days. After 2 days, the plants were treated with 500 μM MeJA or 100 μM ABA, and five time periods of 0, 3, 6, 12, and 24 h were selected, with three biological replicates performed for each group. Using 3-month-old dandelion plants as experimental materials, their root and leaf tissue samples were collected after treatments at the respective time points. The samples were immediately frozen in liquid nitrogen and stored at -80℃ for subsequent analysis of relative gene expression and compound detection. *Nicotiana benthamiana* (*N. benthamiana*) were grown in a plant culture room at 24 ℃ under a 16/8 h (light/dark) photoperiod for one month and then used for subcellular location analysis.

TmOSC gene family identification

To identify the *TmOSC* gene, we first obtained the latest reference genome information [28]. Using the AtOSC protein [29] as a query sequence, we performed a BLASTP search against the *dandelion* genome (Genome Warehouse database, accession number: GWHBCHG00000000) to retrieve OSC protein sequences, with an e-value threshold set to 1×10^{-5}, generating a candidate dataset of TmOSC proteins [30]. The hidden Markov models (HMMs) PF13243 and PF13249 for OSC proteins were downloaded from Pfam (http://pfam.xfam.org/, accessed on 30 June 2025), and the Hmm-search tool [31] with default parameters was used to screen the entire dandelion proteome. The filtered sequences were subsequently subjected to BLAST analysis, and the *OSC* gene sequences with an e-value < 10^{-5} and the longest transcripts were selected for subsequent analysis. Using the HMM-search template in TBtools software, all OSC members with conserved superfamily domains were identified from the protein data, with an e-value threshold of 10^{-5}. To validate domain integrity, redundant sequences lacking complete OSC superfamily domains (SQHop_cyclase_C and SQHop_cyclase_N) were filtered out via NCBI-CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/, accessed on 30 June 2025) [32], InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/, accessed on 30 June 2025), and SMART (https://smart.embl.de/smart/batch.pl, accessed on 30 June 2025) [33], finally yielding the *TmOSC* gene family members.

The molecular weight (MW) and isoelectric point (pI) of OSC proteins were predicted via ExPaSy (https://www.expasy.org/, accessed on 30 June 2025) [34] and WoLF PSORT (https://www.genscript.com/wolf-psort.html, accessed on 30 June 2025) [35], whereas while the subcellular localization of TmOSC proteins was analyzed via the Cell-PLoc 2.0 tool (http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/, accessed on 30 June 2025) [36]. The secondary structure of the TmOSC proteins (Figure S1) was predicted and analyzed via the SOPMA server (https://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html, accessed on 30 June 2025). Finally, the three-dimensional structure of the protein was predicted via homology modeling, modeled via SWISS-MODEL (https://swissmodel.expasy.org/, accessed on 30 June 2025) [37], and visualized via PyMOL software (Figure S2).

TmOSC protein and its gene structure, conserved structural domains and phylogenetic analysis

The motif composition and distribution of the TmOSC proteins were characterized via MEME (https://meme-suite.org/meme/tools/meme, accessed on 30 June 2025) [38] to resolve their conserved motifs. The parameters were set as follows: the optimal motif width was set to ≥ 6 and ≤ 50; and the number of motifs was set to 10. The signal peptides and transmembrane structural domains of the TmOSC proteins were predicted via Signal 6.0 (https://services.healthtech.dtu.dk/services/SignalP-6.0/, accessed on 30 June 2025) [39]. The sequences were obtained from TAIR (https://www.arabidopsis.org/, accessed on 30 June 2025), the Ensembl database (http://plants.ensembl.org/index.html, accessed on 30 June 2025) and the Genome Archive database (https://ngdc.cncb.ac.cn/gwh/, accessed on 30 June 2025) and the National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/, accessed on 30 June 2025). The

same methodology used for the identification of *TmOSC* gene family members was used to screen for *OSC* members. To explore evolutionary relationships, multiple sequence comparisons of OSC proteins from different species obtained from the screening were performed via Clustal W software [40, 41], and then a phylogenetic tree was subsequently constructed via the maximum likelihood estimation (ML) method via MEGA version 11 software [42]. To assess the reliability of the phylogenetic tree, the bootstrap value was set to 1000 repetitions. Evolutionary trees were visualized and annotated via the online software iTOL (https://itol.embl.de/, accessed on 30 June 2025) [43].

Collinearity analysis of *OSC* genes and promoter analysis of *TmOSC* genes

The TBtools software was used to retrieve the *OSC* gene ID files of dandelion, *Taraxacum koksaghyz (T. koksaghyz)*, *Helianthus annuus* (*H. annuus*), *Lactuca sativa* (*L. sativa*), and *Artemisia argyi* (*A. argyi*). The Gene Location Visualization plugin was subsequently employed to perform visual chromosomal mapping of *OSC* genes in each species. Gene duplication events were analyzed via MCScanX with default parameters to investigate the collinearity of *OSC* genes among these species [44]. The 2000 bp upstream of the start codon (ATG) was intercepted from the dandelion reference genome, and the *cis*-acting elements in the promoter region were predicted via the online tool PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). The HeatMap plugin of TBtools was used for visual analysis of the *cis*-acting elements in the *TmOSC* genes.

Overexpression and subcellular localization of *TmOSCs*

The full-length coding sequence (CDS) of the target genes (*TmOSC3*,*TmOSC8* and *TmOSC10*), isolated from *dandelion*, were amplified by PCR using gene-specific primers containing Sma Ⅰ and Sal Ⅰ enzyme cleavage sites. The amplified fragments were inserted into the pCAMBIA1300-GFP vector to construct the *TmOSC*-eGFP fusion expression vector, which was verified by colony PCR (primer-targeted vector insertion ligation) and Sanger sequencing, The verified recombinant vector was transformed into *Agrobacterium tumefaciens* GV3101 by freeze-thaw method for subsequent infection experiments on dandelions. After agroinfection, the infected dandelion explants were transferred to selective regeneration medium supplemented with 25 mg/L hygromycin B and subcultured every 14 days. Following 8–10 weeks of continuous selection and cultivation, resistant shoots were induced and rooted to obtain putative transgenic plantlets.

To confirm successful transgene integration and expression, molecular identification was conducted on resistant plants. For genomic PCR verfication, total genomic DNA from young leaves of putative transformants was amplified with specific primers flanking the *TmOSC*-eGFP fusion fragment; a target band confirmed stable exogenous gene integration into the dandelion genome. For qRT-PCR verification, total RNAfrom the same leaves was reverse-transcrbed into cDNA, and gene-specific primers targeting the *TmOSC* gene, eGFP tag and endogenous Actin reference gene were used to quantify exogenous *TmOSC* expression (Table S6). Only plants positive for both assays were identified as transgenic lines.

In order to conduct the transient overexpression test, the CDS sequence with the stop codon removed was reorganized and transformed according to the above method. The positive clones were resuspend in the infiltration buffer (10 mM MES, 10mM $MgCl_2$, 200 μM acetylsyringone, pH 5.6) until the OD was 0.6. Cultivate in the dark at room temperature for 2 h, and then permeate the bacterial suspension on the abaxial surface of 4-week-old *N. benthamiana* leaves. After incubation in the dark for 48 h, fluorescent signals in the epidermal cells of *N. benthamiana* leaves were observed via a laser confocal microscope (Nikon, Tokyo, Japan). The pCAMBIA1300-35 S-NES-mCherry plasmid (source: Pytbio, China) was used as a cytoplasmic marker in the experiment.

RNA isolation and related gene expression analysis

Total RNA from the samples was extracted for qRT−PCR via TRIzol reagent (Invitrogen, Carlsbad, CA, USA). cDNA synthesis was performed via the FastQuant First Strand cDNA Synthesis Kit (Tiangen, Beijing, China). A LightCycler 480 Real-Time PCR system (Roche, Basel, Switzerland), a qPCR kit (Vazyme) and a Roche LightCycler instrument were used for qRT−PCR. Three biological replicates were used for each treatment. qRT−PCR primer design was performed via Prime5 software, version v5.00 (PREMIER Biosoft, San Francisco, CA, USA) (Table S6). *TmACTIN* was used as an internal reference gene and the data were analyzed via the $2^{-\Delta\Delta ct}$ method [45]. Statistical analysis was performed via SPSS 22.0 software (SPSS Inc., Chicago, IL, USA). Statistical differences between measurements at different times and between treatments were analyzed via Duncan's multiple extreme difference test. Differences were considered significant at the $p < 0.05$ probability level.

RNA-Seq analysis and determination of plant physiological indicators

The raw RNA sequencing raw data for leaves, roots, and flowers in the dandelion (SRP512321, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1117572/) and the raw RNA sequencing data for RNA treated with ABA for

different time periods (SRP387218, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA860343/) were obtained from the NCBI's publicly available data. The raw data were converted to fastq format via the SRA Toolkit v2.9. Subsequently, Trimmomatic-0.39 was employed to perform quality trimming on the raw reads. After quality checking, the clean reads were aligned to the dandelion reference genome via StringTie v2.1.3 to calculate gene expression levels. Gene expression was quantified as log2 (transcripts per million (TPM) + 1)), where TPM represents transcripts per kilobase of exon model per million mapped reads. On the basis of the average TPM values from three biological replicates, a gene expression heatmap was generated via TBtools software.

Catalase (CAT),, peroxidase (POD), MDA and SOD were measured via the Physiological Indicator Assay Kit (Solarbio Beijing, China). The results are presented as the means ± standard deviations (SD) and were analyzed and plotted via Duncan's multiple range test ($^{*}p < 0.05$) via GraphPad Prism software (version 8.0.2). Histochemical staining for ROS in young leaf sections of dandelion was performed using a solution of diaminobenzidine (DAB) and nitro-tetrazolium blue chloride (NBT, p-Nitro-Blue tetrazolium chloride). The relative intensities of the DAB staining were calculated as described in a previous study [46] via the ImageJ software to measure the relative intensity of DAB and NBT staining.

Extraction and content analysis methods for total triterpenes and taraxasterol

The total triterpenes content was measured via a ferric thiophosphate color reagent prepared by dissolving 2.5 g of $FeCl_3$ in 87% concentrated phosphoric acid, diluting it to 100 mL, and then mixing 8.0 mL of this stock solution with concentrated sulfuric acid to 100 mL. The mixture was vigorously shaken with the extract, and allowed to stand for 15 min, after which the absorbance was measured at 538 nm. Analysis of taraxasterol was performed according to high-performance liquid chromatography (HPLC) standard protocols [47].

Results

Identification and classification of *TmOSC* gene family members in dandelion

Through AtOSC protein sequence alignment, motif-based screening, and transcript deduplication, we identified 10 *OSC* family members in the dandelion genome (Table 1). According to the order of *OSC* gene localization on chromosomes, named *TmOSC1-TmOSC10* (Table S1), the *TmOSCs* were distributed on three of the eight chromosomes (Chr3, Chr7, and Chr8), with the majority of the members located in the Chr7 gene cluster, *TmOSC1* was located in the Chr3 gene cluster, whereas *TmOSC10* was distributed in the Chr8 gene cluster. We predicted and analyzed the structure of the TmOSC proteins, and the secondary structures of the members were dominated by α-helical structures, which accounted for 42.66% to 48.11% of the total (Figure S1). The diversity of protein tertiary structures originates from the different combinations of secondary structures, and there was a high similarity between the TmOSC7 and TmOSC8 proteins in the conformation formed by spatial folding (Figure S2). The predicted results of the trans-membrane structural domain analysis revealed one trans-membrane structural domain in each of the TmOSC3 and TmOSC6 proteins, whereas the remaining proteins did not have trans-membrane structures (Figure S3). The physicochemical properties of each member of the *TmOSC* were subsequently analyzed, including the isoelectric point of the protein sequence, molecular weight, number of amino acids, open reading frame, and stability constants. The molecular weights of the *TmOSCs* ranged from 69.46 kDa to 92.80 kDa, and the number of amino acids ranged from 607 bp (*TmOSC2*) to 804 bp (*TmOSC5*). The full length of the open reading frame ranged to 1824 bp (*TmOSC2*) to 2415 bp (*TmOSC5*). In addition, the isoelectric points of the *TmOSC* members ranged from 5.54

Table 1 Overview and molecular characterization of the members of the *TmOSCs* in dandelion

Gene name	Gene ID	pI	MW (kDa)	aa	ORF/bp	Instability index	Aliphatic index	GRAVY	Predicted location
TmOSC1	*GWHPBCHG024441*	6.28	86.19	758	2277	37.52	83.40	-0.25	Cyto
TmOSC2	*GWHPBCHG054796*	6.19	69.46	607	1824	44.44	73.79	-0.33	Cyto
TmOSC3	*GWHPBCHG054802*	6.36	77.01	671	2016	41.38	77.76	-0.35	Cyto
TmOSC4	*GWHPBCHG054809*	6.00	86.91	758	2277	38.30	76.33	-0.38	Cyto
TmOSC5	*GWHPBCHG054822*	6.37	92.80	804	2415	37.42	73.91	-0.45	Cyto
TmOSC6	*GWHPBCHG055152*	5.55	73.33	638	1917	44.26	76.90	-0.37	Cyto
TmOSC7	*GWHPBCHG055195*	5.54	88.60	770	2313	40.45	79.16	-0.42	Cyto
TmOSC8	*GWHPBCHG055200*	5.73	88.80	769	2310	39.02	78.86	-0.43	Cyto
TmOSC9	*GWHPBCHG058221*	5.86	87.16	760	2283	47.54	78.57	-0.30	Cyto
TmOSC10	*GWHPBCHG067314*	5.74	87.98	766	2301	41.20	77.42	-0.37	Cyto

pI Isoelectric point, *MW* Molecular weight, *aa* Number of amino acids, *ORF* Open reading frame, *GRAVY* Grand average of hydropathicity, *Cyto* Cytoplasm

(*TmOSC7*)-6.37 (*TmOSC5*) and all the members were acidic proteins (pI < 7). Four proteins, namely, *TmOSC1*, *TmOSC4*, *TmOSC5*, and *TmOSC8*, were more stable, with an instability index of < 40.0. The predicted lipid indices of the family members revealed that the TmOSC proteins were highly stable, with the highest lipid index of 83.4 for *TmOSC1* and the lowest of 73.79 for *TmOSC2*. The total mean grand average of hydrophilicity (GRAVY) of the *TmOSC* family members was less than 0, indicating that the TmOSC proteins were hydrophilic. Subcellular localization predictions revealed that all proteins were localized in the cytoplasm (Table 1).

Analysis of conserved motifs, domains, and gene structures of *TmOSC* genes in the dandelion

Protein structure analysis revealed that TmOSCs consisted of a total of three subgroups (Fig. 1. A), TmOSC2, 3, 4, and 5 were from the same subgroup, but TmOSC2 lacked motif 10 and TmOSC3 lacked motif 5; TmOSC6, 10, 7, and 8 were from the same subgroup, and the distribution positions of motifs in the three genes were roughly similar, but TmOSC6 still lacked motif 10. TmOSC1 and TmOSC9 were from the same subgroup, and there were two motif 6 s in TmOSC1. The conserved motifs indicated that all the TmOSCs contained eight motifs except for motif 10 and motif 5 (Fig. 1. B). The exon number of the TmOSCs ranged from 13 to 20 exons, and the fixed exon number of the SQHop_cyclase_C and/or SQHop_cyclase_N structural domains indicated that the protein structure of the TmOSCs is conserved (Fig. 1. CD). The majority of TmOSCs structural domains comprising 3-8-6-5 motifs (corresponding to SQHop_cyclase_C) and 7-2-1-9-4 motifs (corresponding to SQHop_cyclase_N) were shown to be highly aligned in the motif analysis (Table S2).

To explore the evolutionary process of the TmOSC proteins, the amino acid sequences were compared, with a focus three highly conserved sequences: the QXXXXXW motif, DCT (A/G) E motif, and MWC (Y/H) CR (Figure S4). Alignment analysis revealed that the TmOSC members contained 3 to 5 highly conserved repetitive QXXXXXW motifs.

Phylogenetic analysis of TmOSC proteins

To clarify the evolutionary relationships of OSC proteins across diverse plant species, we constructed a phylogenetic tree incorporating dandelion and several reference organisms: model plants (O. sativa, N. benthamiana, Arabidopsis thaliana (A. thaliana)) and closely related Asteraceae species (L. sativa, H. annuus, T. koksaghyz). This

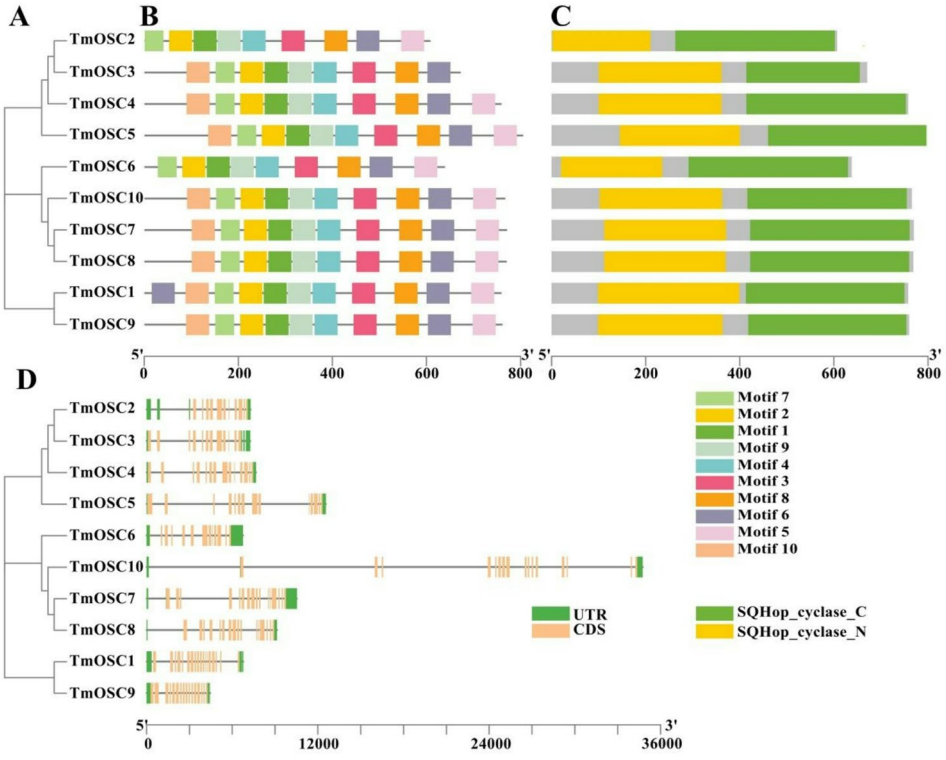

Fig. 1 Gene structure, conserved structural domains and protein motifs of the *TmOSC* gene family. **A** Phylogenetic tree constructed on the basis of the full-length sequences of the TmOSC proteins. **B** The conserved motifs of the TmOSC proteins include 10 motifs distinguished by specific colors, which are ordered to correspond to their positions in the individual protein sequences. **C** Conserved N-terminal and C-terminal structural domains of TmOSC. **D** Intron–exon organization of the *TmOSC* gene. Pink boxes represent CDSs; green boxes represent UTRs

comparative analysis enabled a comprehensive understanding of OSC protein divergence and conservation within the Asteraceae family and across broader plant lineages (Table S3). On the basis of sequence similarity and phylogenetic analysis, the OSC proteins were mainly divided into four main clades, with 26, 31, 12, and 20 proteins in evolutionary Clades 1–4 (Fig. 2. A). TmOSC1 was clustered in Clade 1, TmOSC9 was clustered in Clade 2, TmOSC2, TmOSC3, TmOSC4, and TmOSC5 were focused in Clade 3 and TmOSC6, TmOSC7, TmOSC8, and TmOSC10 were focused in Clade 4. OsOSCs were all clustered in Clade 1, whereas most AtOSC proteins were distributed in Clade 3, indicating a more conserved evolutionary relationship. In contrast, OSC proteins from L. sativa, H. annuus, and T. koksaghyz—species more closely related to dandelion—were more closely grouped with TmOSCs, indicating a closer interspecific evolutionary relationship among their OSC proteins. Phylogenetic analysis of OSC gene family members from dandelion and published non-Asteraceae species (R. rugosa, M. toosendan, T. aestivum, and M. charantia) revealed a five-clade phylogenetic tree (Fig. 2. B). Species-specific conservation was evident: most T. aestivum OSCs clustered in Clade 4, whereas TmOSCs predominantly occupied Clade 3. These results suggest that TmOSCs exhibit greater evolutionary distance from OSCs in non-Asteraceae families, highlighting lineage-specific diversification within the Asteraceae.

Collinearity analysis of the *TmOSC* genes

To assess the covariate distribution relationships in dandelion and the *Asteraceae* relatives *T. koksaghyz*, *H. annus*, and *L. sativa*, we analyzed the chromosomal localization distributions of the *OSC* gene family in different species (Figure S5) and their evolutionary relationships (Fig. 3). The results revealed that the distribution of *OSCs* on each chromosome was heterogeneous, and in diploid dandelion, the *TmOSCs* were distributed on three chromosomes, Chr3, Chr7, and Chr8. In addition, the *OSC* genes were distributed on three or more chromosomes in species such as *O. sativa*, and dandelion which presented three pairs of collinear *OSC* genes with *T. koksaghyz* and four, three, and five collinear genes with *L. sativa*, *H. annus*, and *A. argyi*, respectively, and all four species presented colinearities with *TmOSC1* and *TmOSC10* (Table S4).

Analysis of *cis*-acting elements of the *TmOSC* gene family

Promoter *cis*-acting elements are important transcription initiation binding regions that play significant roles in the regulation of gene expression. Multiple classes of *cis*-acting elements were identified in the 2000 bp region upstream of the promoters of the 10 *TmOSC* genes (Table S5). These elements were grouped into five broad categories: light-responsive, hormone-responsive, environment-responsive, developmentally relevant, and promoter-associated binding sites (Fig. 4. A). Among them, promoter-associated binding sites accounted for the largest proportion (47.68% of all elements), environmental factors accounted for 10.98%, hormone-responsive sites

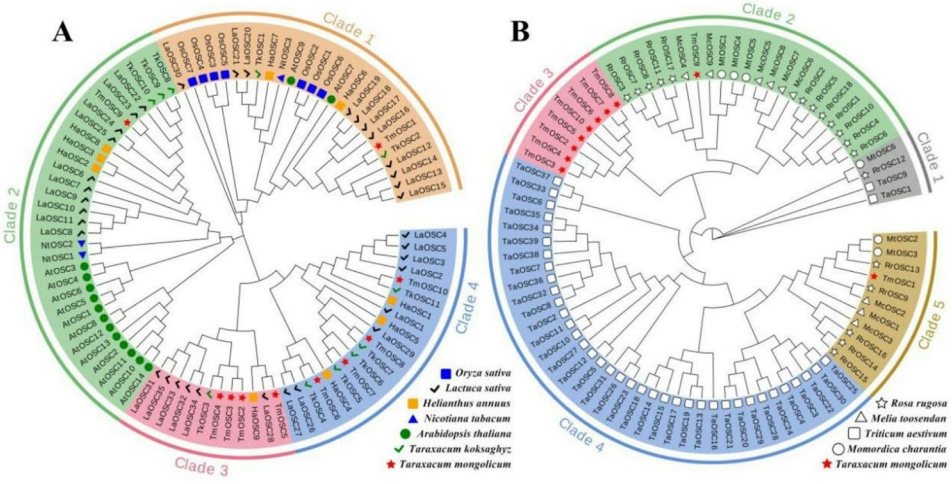

Fig. 2 Phylogenetic tree of OSCs among different species. **A** Phylogenetic analysis of the OSC proteins of *O. sativa*, *L. sativa*, *H. annus*, *N. benthamiana*, *A. thaliana*, *T. koksaghyz* and dandelion. Phylogenetic trees were drawn via the ML method with 1000 bootstraps. 89 genes were classified into four clades (1–4) and identified with different colors. The blue rectangles represent *O. sativa*, the black check marks represent *L. sativa*, the orange rectangles represent *H. annuus*, the blue triangles represent *N. benthamiana*, the green circles represent *A. thaliana*, the green check mark represent *T. koksaghyz* and the red stars represent dandelion. **B***R. rugosa*, *M. toosendan*, *T. aestivum*, and *M. charantia* OSC proteins were phylogenetically analyzed with TmOSC proteins. Phylogenetic trees were drawn via the maximum likelihood method with 1000 bootstraps. Where the black line star box represents R. *rugosa*, the black line triangle box represents *M. toosendan*, the black line rectangle box represents *T. aestivum* and the black line circle box represents *M. charantia*

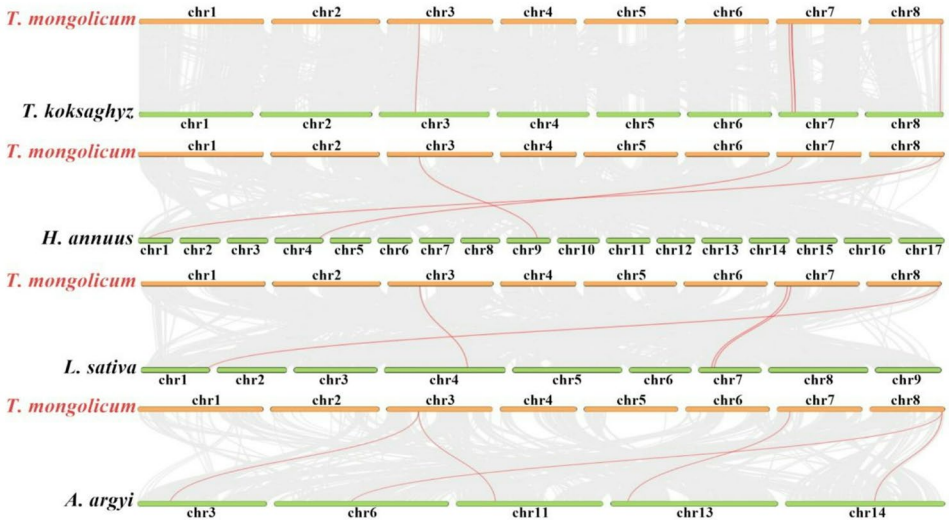

Fig. 3 Collinearity of *OSC* genes in dandelion and other species. The gray line in the background represents the collined portion of the dandelion and other plant genomes. The highlighted red lines represent collinear *OSC* gene pairs

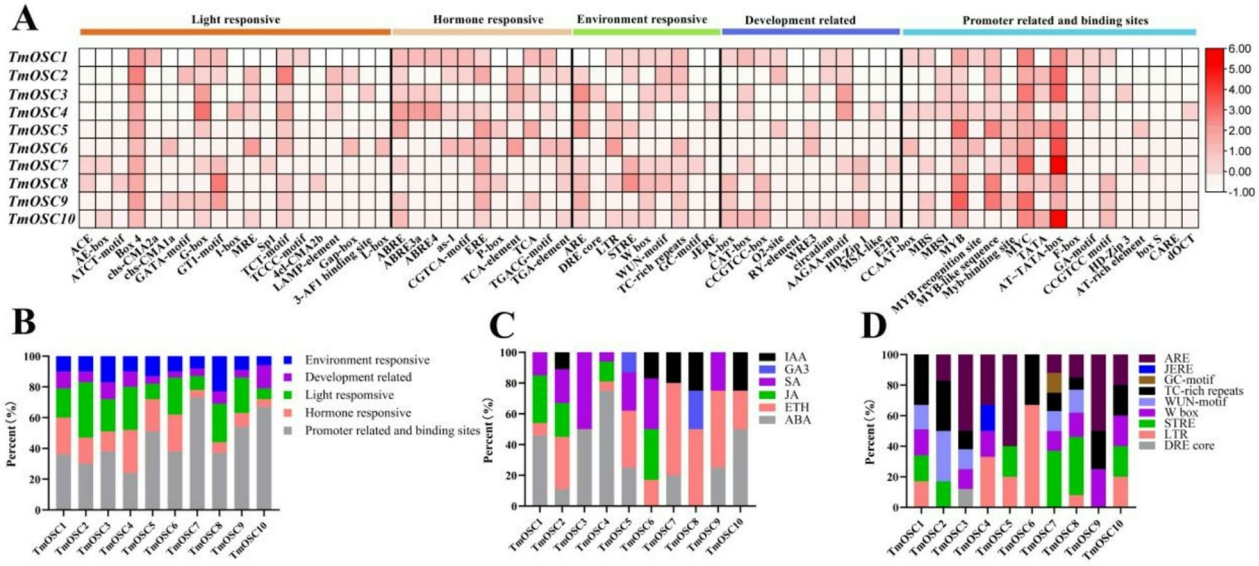

Fig. 4 Analysis of *cis*-acting elements in *TmOSC* genes. **A** Different colors on the heatmap indicate the number of different *cis*-acting elements in each *TmOSC* gene. **B** Different colors on the stacked bar graph indicate the percentage of each *cis*-acting element in different *TmOSC* genes. **C** Different colors on the stacked bar graph indicate the percentages of different hormone components in different *TmOSC* genes. **D** Different colors on the stacked bar graph indicate the percentage of environmentally responsive components in different *TmOSC* genes

accounted for 13.72%, and hormone-responsive elements accounted for a large proportion (Fig. 4. B); most of the genes in all the family members contained ABA-responsive elements, ABRE, ABRE3a, and ABRE4, as well as MeJA-responsive elements, TGACG-motif and CGTCA-motif, whereas *TmOSC6* and *TmOSC8* do not contain response elements for ABA, and only *TmOSC1*, *TmOSC2*, *TmOSC4* and *TmOSC6* contain response elements for MeJA (Fig. 4. C). The environmental response element ARE is present in several *OSC* gene family members and plays a key role in the antioxidant defense mechanism (Fig. 4. D). These results suggest that *TmOSC* family members may play a role in plant growth and metabolic regulation.

Subcellular localization of *TmOSC* expression

To further explore the expression of TmOSC proteins in cells, three genes tentatively identified as potential key candidates through transcriptome data analysis of dandelion for subsequent transient expression assays (Figure S6. B). We then performed subcellular localization to verify the specific locations where the proteins

exert their functions. The *TmOSC* gene was ligated to the pCAMBIA1300-eGFP vector via the primer OSCs-T (Table S6) with the terminator removed, after which it was coinfected with cytoplasmic markers and transiently expressed in *N. benthamiana*. The green fluorescence of the *TmOSC*-pCAMBIA1300-eGFP fusion protein and the red fluorescence of the cytoplasmic markers in the merged region indicated that *TmOSC3*, *TmOSC8*, and *TmOSC10* were located in the cytoplasm and distributed in the nucleus (Fig. 5), which is in agreement with the results of the previous predictions of subcellular localization.

Analysis of the gene expression pattern and physiological indicators of *TmOSC*

To further explore the function of the *TmOSC* genes, we treated dandelion with MeJA and designed primers (Table S6) to detect the expression of *OSC* genes in dandelion root tissues using qRT−PCR. The revealed that the expression of *TmOSC3*, *TmOSC4*, *TmOSC5*, *TmOSC6*, *TmOSC7*, *TmOSC8*, and *TmOSC9* was significantly upregulated in the root system after MeJA induction, with *TmOSC3* being the most upregulated gene, which is hypothesized to play a key role in the MeJA regulatory pathway. The addition of exogenous hormones activated the expression of these genes to different degrees, and *TmOSC1*, *TmOSC2* and *TmOSC10* were gradually downregulated, all of which reached their lowest values at 12 h, and then tended to be regulated. According to a related study, the identification and quantification of taraxasterol in dandelion after 500 μM MeJA treatment revealed that the concentration of taraxasterol reached higher values at 3 h and 6 h, followed by a gradual decrease in the content (Figure S6. A), which was consistent with the changes in total sterol content (Fig. 6. B). Interestingly, the *TmOSC3*, *TmOSC4*, *TmOSC6*, *TmOSC8* and *TmOSC10* genes presented the greatest expression at 3 h, and the *TmOSC5* and *TmOSC9* genes presented the greatest expression at 6 h (Fig. 6. A). A comparison of the transcriptome data revealed that the expression of *TmOSC8* and *TmOSC10*

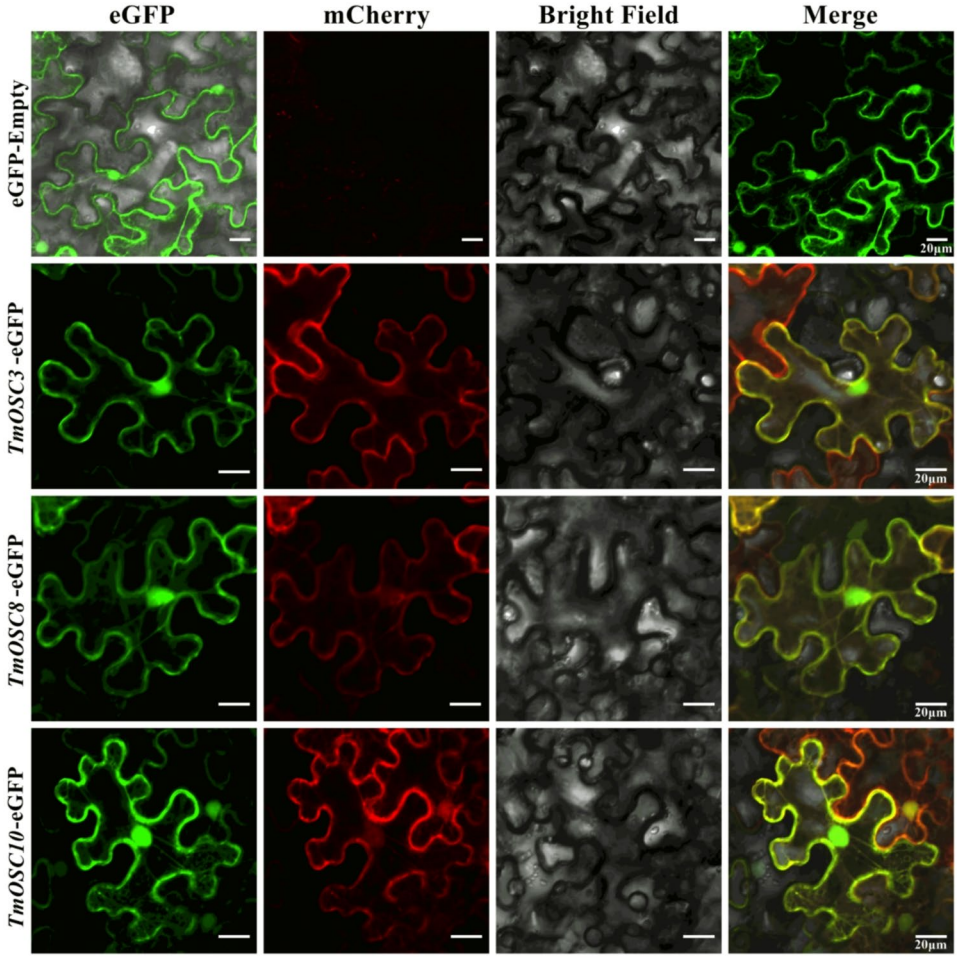

Fig. 5 Subcellular localization. eGFP was used as a control for comparison. All proteins were transiently expressed in *N. benthamiana* epidermal cells, and the GFP signal was observed after 48 h of incubation. The cytoplasm was observed with mCherry-labeled cytoplasmic markers. Scale bar, 20 μm

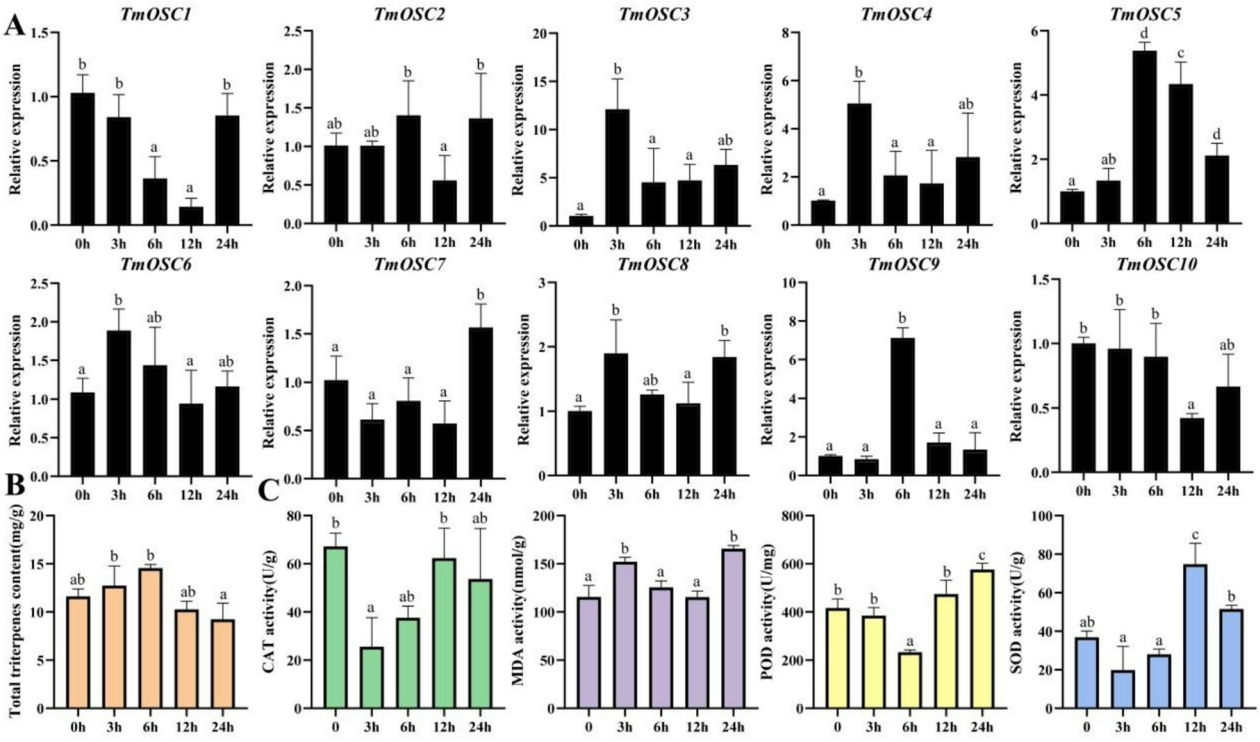

Fig. 6 Gene expression patterns and physiological indices of TmOSCs under MeJA treatment. **A** Relative gene expression levels of TmOSCs; (**B**) total sterol content; (**C**) CAT activity, MDA content, POD activity and SOD activity. The data are expressed as means ± SD of three biological replicates. The error bars represent the means ± SE of three independent replicates. Values with different letters within the same column are significantly different at $p < 0.05$, as determined via Duncan's multiple range test

was significantly greater than that of the other genes in the dandelion leaves, roots and flowers (Figure S6. B). These results suggest that these genes are involved in the regulatory pathway of MeJA metabolism and regulate the growth and development process in specific tissues.

To investigate the extent of cellular damage and the activation of the antioxidant defense system in dandelion under MeJA treatment, we determined the CAT, MDA, POD and SOD contents of dandelion roots (Fig. 6. C) and leaf (Figure S7. A). Tissues subjected to different treatments. The results revealed that under MeJA treatment, the CAT and SOD activities in dandelion roots peaked at 12 h, and the expression of MDA and POD peaked at 24 h, with all four physiological indices showing a decreasing and then increasing trend. Interestingly, the changes in the CAT, POD, and SOD contents in the leaves analyzed under MeJA treatment were consistent with those observed in the roots (Figure S7).

In addition, we treated dandelion root tissues with ABA and examined *TmOSC* gene expression (Fig. 7), which revealed that most family members were significantly upregulated by ABA induction. The upregulation of *TmOSC3* and *TmOSC4* in the root system was greater, which was consistent with the transcriptome data (Figure S6. B, Table S7). Interestingly, the expression levels of *TmOSC1*, *TmOSC2*, *TmOSC6*, *TmOSC7*, *TmOSC8*, and *TmOSC10* in roots all exhibited a pattern of first decreasing and then increasing, but then decreasing again at 24 h, with the lowest expression observed at 3 h. Additionally, the total sterol content in dandelion roots reached its lowest value at 3 h, at only 5 mg/g. To evaluate cellular damage in ABA-treated dandelion, antioxidant enzyme activities and lipid peroxidation were analyzed (Fig. 7. C). CAT activity gradually decreased with prolonged ABA treatment but recovered at 24 h. POD activity decreased initially decreased but then increased, mirroring the trend observed in *TmOSC* gene expression. SOD activity peaked at 6 h, with no significant changes compared with that at 0 h at the other time points. The MDA content increased sharply at 3 h before declining, indicating transient membrane damage followed by repair. The changes in the CAT, POD, and SOD contents in the leaves under ABA treatment were consistent with those expressed in the root, indicating that members of the *TmOSC* family are involved in the regulation of the ABA metabolic pathway to a certain extent. In conclusion, the three genes *TmOSC3*, *TmOSC8*, and *TmOSC10* may serve as key targets and have an impact on the synthesis of taraxasterol.

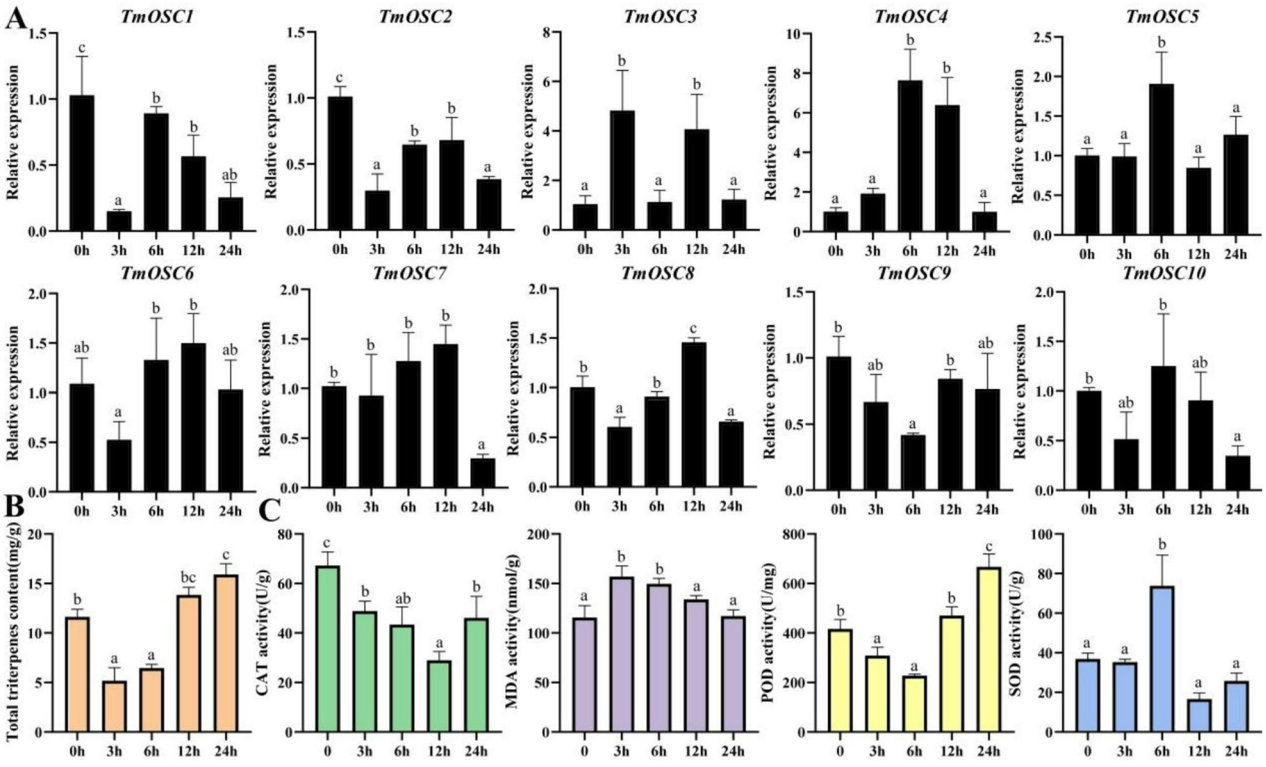

Fig. 7 Gene expression patterns and physiological indices of *TmOSCs* under ABA treatment. **A** Relative gene expression levels of *TmOSCs*; (**B**) total sterol content; (**C**) CAT activity, MDA content, POD activity and SOD activity. The data are expressed as means ± SD of three biological replicates. The error bars represent the means ± SE of three independent replicates. Values with different letters within the same column are significantly different at $p < 0.05$, as determined via Duncan's multiple range test

Over-expression of *TmOSC8* gene promotes taraxasterol accumulation in dandelion

Induction of MeJA induction promotes triterpenoid accumulation in dandelion, suggesting that MeJA may regulate triterpenoid biosynthesis by activating related signaling pathways. As the *OSC is* serves as a rate-limiting enzyme in sterol biosynthesis and, based on its high root expression (Figure S6. B), strong homology to cycloartenol synthases [13], and conserved DCTAE motif, *TmOSC8* may serve as a key regulator in the sterol biosynthesis pathway and its cyclization diversifies terpenes, we speculate that *TmOSC8* may govern the biosynthesis of taraxasterol in the dandelion. To elucidate whether *TmOSC8* affects the production of taraxasterol, three independent *TmOSC8*-overexpression lines (OE-2, OE-7, and OE-16) were selected for further study (Fig. 8. A, Figure S8). The results of the qRT−PCR assay revealed that the transcript level of the *TmOSC8* gene in the roots of the transgenic lines was significantly greater than that in the control. Compared with those of the WT strains, the *TmOSC8* gene transcript levels of the OE2 and OE16 strains increased approximately 20-fold (Fig. 8. B). Compared with the control, the OE7 strain presented as much as a 60-fold increase in the *TmOSC8* gene transcript level, and the taraxasterol content was increased from 0.032 to 0.118 mg/g, which was 3.68-fold greater than that of the control (Fig. 8. C, Table S8).

To investigate the extent of antioxidant damage to dandelion cause by overexpression of the *TmOSC8* gene, we determined the CAT, MDA, POD, and SOD contents in the roots of the OE2, OE7, and OE16 overexpression lines (Fig. 8. D), and the results revealed that the contents of three enzymes, CAT, POD, and SOD, were significantly greater than those in the control, while the MDA content was significantly lower than taht in the control. To investigate whether overexpression of *TmOSC8* could eliminate ROS to a certain extent, we used DAB and NBT staining and showed that the staining density in the young leaves of the overexpression lines was significantly lower than that in the wild type lines, indicating that overexpression of *TmOSC8* could eliminate ROS from the plant body (Fig. 8. E, F). These physiological indices, to some extent, indicate the increase in the cellular antioxidant capacity of dandelion after overexpression of *TmOSC8* gene.

Discussion

Dandelion has attracted much attention in the fields of botany and pharmacology because of its rich medicinal and edible value. Its biological properties and secondary

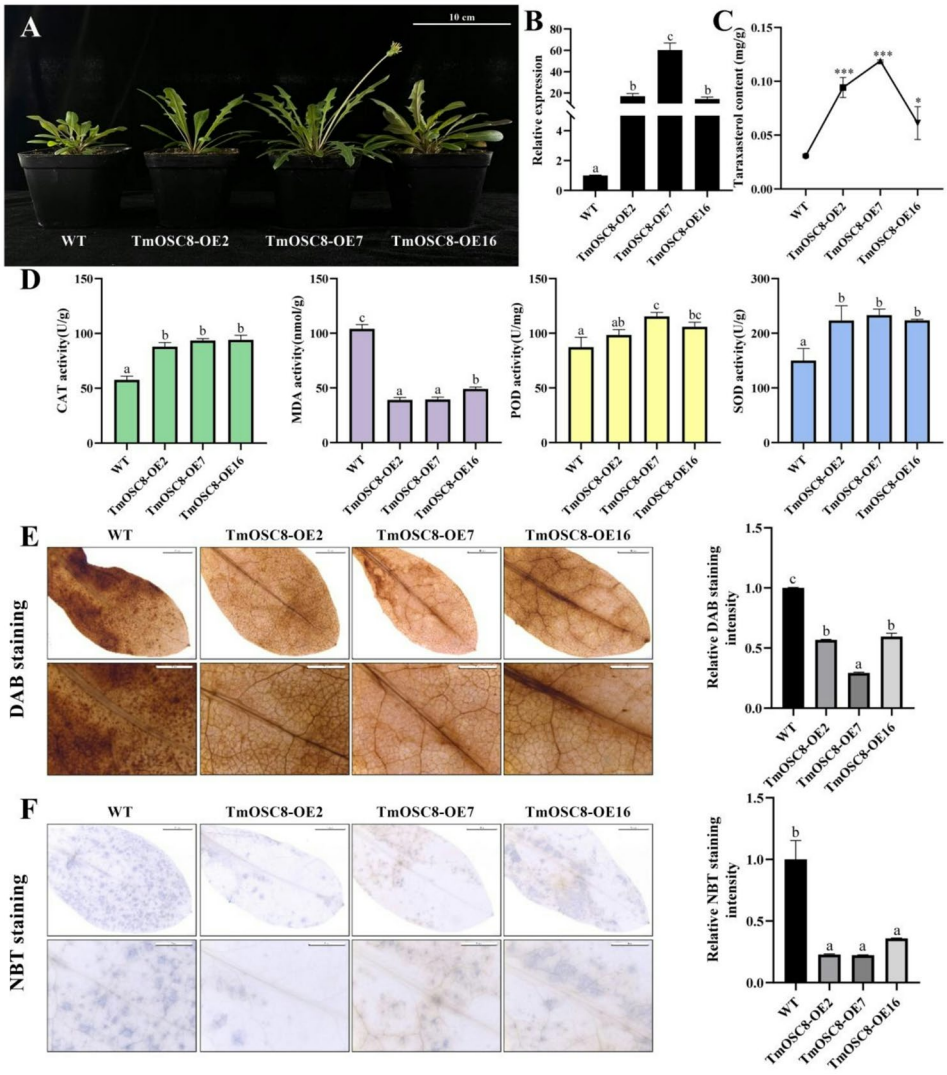

Fig. 8 The overexpression of *TmOSC8* in the dandelion transgenic lines resulted in an increase in taraxasterol. **A** Phenotype of *TmOSC8* in three-month-old dandelion of the over-expression lines. **B** Expression of the *TmOSC8* gene in the wild type and overexpression lines OE2, OE7, and OE16, with Actin used as an internal reference gene. **C** Determination of the content of taraxasterol in the wild type and OE2, OE7, and OE16 strains. **D** Determination of CAT activity, MDA content, POD activity, and SOD activity in the roots of the WT and overexpression strains OE2, OE7, and OE16. **E**, **F**: In situ assay of ROS in the young leaves of WT/*TmOSC8*-OE cattails. In suit ROS assay. DAB (**E**) and NBT (**F**) staining were used to analyze ROS production. The data are expressed as means ± SD of three biological replicates. The error bars represent the means ± SE of three independent replicates. Values with different letters within the same column are significantly different at $p < 0.05$, as determined via Duncan's multiple range test

metabolite biosynthesis pathways, especially the terpenoid synthesis network, have become the core direction of current research. In this study, we focused on the genome-wide identification and sequence characterization of members of the *TmOSC* gene family in dandelion, aiming to reveal the regulatory mechanisms of the family members in response to ABA and MeJA treatment, and to provide new perspectives for analyzing the sterol synthesis pathway. *OSCs* play key roles in the biosynthesis of plant terpenoids, catalyzing the oxidative conversion of squalene into a wide range of terpenoids including sterols, triterpenoids, and sesquiterpenoids. Members of the *OSC* gene family may be involved in the biosynthesis of different terpenoids through functional differentiation among members of dandelion. This study involved a genome-wide analysis that identified a total of 10 *OSC* gene family members (Table 1). Sequence analysis revealed that all the members harbored conserved motifs typical of the *OSC* family. These include the DCTAE motif, which is involved in substrate binding and protonation; the MWCYCR motif, which determines substrate folding modes and product types; and the QXXXXXW motif, which affects enzyme stability and product profiles [48]. *OSCs* belong to a superfamily of enzymes that are highly conserved in plants. Its conserved region is capable of catalyzing the protonation, cyclization,

rearrangement, and deprotonation of 2,3-oxidosqualene to form various triterpene skeletons, such as α - amyrin [49].

In the phylogenetic tree (Fig. 2), OSC proteins from dandelion clustered closely with *Asteraceae* relatives such as *L sativa* and *H annuus*, indicating a common ancestry and shared selective pressures in triterpenoid biosynthesis [50]. This conservation likely reflects critical functional domains essential for metabolic roles. Conversely, distant clustering with non-*Asteraceae* species (*A. thaliana*, *T. aestivum*) highlights family-specific divergence driven by ecological adaptations and metabolic requirements [51]. Notably, *TmOSCs* lack intragroup collinearity suggesting frequent chromosomal rearrangements [52]. However, the widespread cross-species covariance of OSC proteins between dandelion and other *Asteraceae* species implies that *Asteraceae* plants may have preserved conserved chromosomal segments or regulatory modules along their coevolutionary clade. These genetic elements likely act as key determinants for maintaining the copy number and functional stability of *OSC* genes [6] The above results provide important clues for resolving the evolutionary trajectory of the *TmOSC* gene family and its functional differentiation mechanism, especially the dual pattern of intragroup rearrangement and interfamily conservation, which may constitute a key evolutionary strategy for adapting to environmental stresses and maintaining terpene synthesis diversity in dandelion.

Cis-acting elements in the promoter region play a key role in regulating plant growth, development, and biotic and abiotic stress resistance. These include hormone response elements including ABA (ABRE, ABRE3a, ABRE4), MeJA (TGACG motif, CGTCA motif), and ethylene (ERE) response elements, which may participate in various hormone signaling pathways [53, 54]. Consistent data from the referenced transcriptome and expression profiles show that *TmOSC3* was the most significantly upregulated member under ABA stress, while *TmOSC8* and *TmOSC10* exhibited markedly higher expression levels in leaves, roots and flowers than other genes. Their distinct expression patterns suggest that these genes are potential key candidates involved in stress responses, prompting us to further investigate their subcellular function location. We treated dandelion with MeJA at different time points and found that the expression of *TmOSC* genes in roots changed dynamically, with a 10-fold increase in the expression of *TmOSC3* at 3 h, a 5-fold increase in *TmOSC4* during the same period, and a 7-fold increase in *TmOSC9* at 6 h (Fig. 6. A). Notably, the promoter regions of these genes (within 2,000 bp) respond directly to MeJA. The scarcity of elements suggests that MeJA may regulate *OSC* gene expression through an indirect pathway, which may involve the mediation of other signaling molecules, transcription factors, or signaling pathways [47] and likewise does not exclude posttranscriptional regulation [55]. These findings align with previous reports showing that transcription factors can modulate triterpenoid synthesis [47], suggesting the complexity of the regulatory network of plant secondary metabolism. As an important plant signaling molecule, MeJA can regulate *OSC* genes through indirect or direct pathways. MeJA alters secondary metabolism gene expression in *Melissa officinalis* [56], which is consistent with our findings in dandelion, where MeJA induced dynamic *TmOSC* expression (Fig. 6). MeJA treatment also triggered oxidative stress in dandelion root, evidenced by increased MDA levels (Fig. 6. C), which aligns with the ROS-mediated membrane damage reported in other species [57, 58]. Notably, the MDA content in dandelion roots exhibited a transient increase followed by a decrease, which was negatively correlated with the activities of antioxidant enzymes (CAT, POD, and SOD) (Fig. 6. C). This biphasic response occurs in *O. sativa*, where low MeJA levels activate defense pathways, whereas high MeJA levels exacerbate oxidative damage via antioxidant enzyme inhibition [59]. The initial increase in MDA likely reflects unmanaged ROS production, whereas subsequent enzyme activation, with SOD as the first-line defense, mitigates stress, demonstrating the adaptive antioxidant network of dandelion [60]. Mechanistically, MeJA may bind intracellular receptors to indirectly modulate *OSC* transcription indirectly, possibly via transcription factors or posttranscriptional regulation [50], given the scarcity of MeJA-responsive elements in *TmOSC* promoters. The temporal coordination of antioxidant enzymes, such as the priming of SOD followed by the synergism of CAT and POD, underscores plants' multilayered defense strategies against oxidative stress, balancing ROS scavenging and metabolic adaptation. These findings increase our understanding of the dual role of MeJA's in stress signaling and triterpenoid biosynthesis regulation.

The biosynthesis of triterpenoids begins with the generation of isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) via the mevalonate pathway (MVA). The construction of their core skeletons relies on cyclization reactions catalyzed by *OSCs*. To date, 152 *OSCs* from different plants have been reported to catalyze the formation of 31 different triterpene skeletons. Among them, β-coumarinol synthase and lupulinol synthase, among others, are the most widely functionally characterized isoforms, dominating the structural diversity of plant triterpenes [6]. The *TcOSC1* enzyme from dandelion catalyzes the generation of a variety of triterpenoids via heterologous expression in yeast [61]. The tissue-specific downregulation of latex duct-specific *OSCs* was achieved via RNA interference technology, and rubber weed plants with reduced triterpenoid contents

were successfully cultivated to reduce triterpenoids in natural rubber by approximately 67% [62]. As a natural product with strong physiological, pharmacological and biological activities extracted from the dandelion, its content is extremely low, making it important to increase its content. In transgenic lines overexpressing TaMYC2, the taraxasterol content in dandelion was significantly increased from 0.08 mg/g to 0.16 mg/g [7]. Under 500 μM MeJA stress in dandelion roots, it was found that the changes in taraxasterol content were consistent with the changes in total sterol content (Fig. 6, Figure S6). Previous studies have shown that *L. sativa* OSCs synthesize pentacyclic triterpenes (α/β-amyrin, lupeol) [63], and that these triterpenes possess defensive and anti-inflammatory activities [64]. Given the sequence and structural homology between *TmOSC* and these *OSCs*, we infer that *TmOSC* also triggers the cyclization of 2,3-oxidosqualene in dandelion to generate conserved triterpene backbones, which are further modified into species-specific defensive metabolites via specialized enzymes. Thus, TmOSC likely plays a core role in triterpene biosynthesis, providing critical metabolic defense for dandelion against biotic stresses. This study revealed that overexpression of the *TmOSC8* gene triggered the plant's antioxidant defense system. Antioxidant enzymes (POD, SOD, and CAT) significantly increased, whereas the ROS content and MDA (a lipid membrane peroxidation product) level were significantly lower than those in the WT (Fig. 8). Overexpression of the *TmOSC8* gene was associated with enhanced catalytic conversion of oxidized squalene, which contribute to the accumulation of taraxasterol. In this study, through the combined screening via transcriptome analysis and stress treatment, *TmOSC3*, *TmOSC8*, and *TmOSC10* were identified as the core members of taraxasterol, and the correlation between their expression patterns and hormone responses provided potential targets for sterol synthesis. In the future, we need to combine yeast two-hybridization and metabolic flow analysis to elucidate the signaling intermediates through which MeJA regulates *OSC* gene transcription and the real-time regulatory mechanism of posttranscriptional modification of enzyme activity.

On the basis of these findings, we have gained a more comprehensive understanding of the *TmOSC* gene family and determined that it may not only be involved in the synthesis of taraxasterol, but also respond to a variety of biotic and abiotic stresses. Notably, most analyses in this study rely on bioinformatics predictions; except for the preliminary association between *TmOSC8* overexpression and taraxasterol accumulation, the actual biological functions of other *TmOSC* family members have not been validated by in vitro (e.g., enzyme activity assay) or in vivo (e.g., gene knockout/knockdown) functional experiments, and their predicted functions await confirmation.

Conclusions

In this study, the *OSC* gene family of dandelion was comprehensively analyzed, and a total of 10 *TmOSC family* members were identified. The basic physicochemical properties, three-dimensional structural model, phylogeny, conserved structural domains, gene structure, chromosomal localization, and *cis*-acting elements of the *TmOSCs* were also analyzed via genomic and bioinformatics techniques. Phylogenetic and homology analyses of OSC proteins in different plants provide a reference for the evolution of the *OSC* gene family in dandelion. The tissue expression patterns of the *TmOSC* genes under MeJA and ABA treatments were also analyzed, and the increase in taraxasterol was associated with the overexpression of the *TmOSC8* gene suggesting a potential role in dandelion. This study provides useful information for the study of the *OSC* gene family in dandelion, and these findings will contribute to a better understanding of the biological functions and molecular mechanisms of *OSCs* in medicinal plants and lay a preliminary foundation for functional studies and molecular breeding for genetic improvement in dandelion.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-026-12593-2.

Supplementary Material 1. Table S1-S8

Supplementary Material 2. Figure S1-S8

Acknowledgements
Not applicable.

Authors' contributions
Conceptualization, L.Y., H.H., J.G., F.W. and Q.X.; writing—original draft preparation, L.Y., H.H. and J.G.; date curation, J.W., P.D. and, L.L.; formal analysis, H.L., Z.M. and H.S.; methodology, S.X.; writing—review and editing, Q.X. and F.W.; resources, Q.X.; supervision, Q.X. and F.W. All authors have read and agreed to the published version of the manuscript.

Funding
This research was funded by the National Natural Science Foundation of China, grant number 32560079 and the Key Laboratory Project of Xinjiang Phytomedicine Resources Utilization Ministry of Education, grant number XPRU202201 and Xinjiang Production and Construction Corps Guiding Technology Program Project, grant number 2023ZD050.

Data availability
All data on the results of this study can be found in the paper and its supplemental material published online.

Declarations

Ethics approval and consent to participate
This study on plants, including the collection of plant material, complied with relevant institutional, national, and international guidelines and legislation.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Xinjiang Production and Construction Corps Key Laboratory of Oasis Town and Mountain- basin System Ecology, Key Laboratory of Xinjiang Phytomedicine Resource Utilization, Ministry of Education, College of Life Sciences, Shihezi University, Shihezi 832003, China

Received: 12 October 2025 / Accepted: 23 January 2026
Published online: 19 February 2026

References

1. Grauso L, Emrick S, de Falco B, Lanzotti V, Bonanomi G. Common dandelion: a review of its botanical, phytochemical and Pharmacological profiles. Phytochem Rev. 2019;18(4):1115–32.
2. Liu Y, Shi Y, Zou J, Zhang X, Zhai B, Guo D, et al. Extraction, purification, structural features, biological activities, modifications, and applications from *Taraxacum mongolicum* polysaccharides: a review. Int J Biol Macromol. 2024;259(2):129193.
3. Liu J-B, Xie J, Wang H-Q, Lai S-T, Shao H-J, Chen R-Y, Kang J. Four sesquiterpenes isolated from *Taraxacum Mongolicum*. J Asian Nat Prod Res. 2023;25(11):1038–43.
4. Liu Y, Sun J, Dai X, Zhao Z, Wang X, Zhang G. Response of soil bacterial community structure and function under two salt-tolerant plants in a coastal saline soil area of Eastern Hebei Province of China. Int J Phytorem. 2021;24(8):842–54.
5. Thapliyal S, Ahmad MI. Standardization of roots of *Taraxacum officinale*. Environ Conserv J. 2023;24(2):327–33.
6. Xue Z, Duan L, Liu D, Guo J, Ge S, Dicks J, ÓMáille P, Osbourn A, Qi X. Divergent evolution of oxidosqualene cyclases in plants. New Phytol. 2011;193(4):1022–38.
7. Sang R, Ge B, Li H, Zhou H, Yan K, Wang W, et al. Taraxasterol alleviates aflatoxin B1-induced liver damage in broiler chickens via regulation of oxidative stress, apoptosis and autophagy. Ecotox Environ Safety. 2023;251:114546.
8. Bao T, Ke Y, Wang Y, Wang W, Li Y, Wang Y, Kui X, Zhou Q, Zhou H, Zhang C, et al. Taraxasterol suppresses the growth of human liver cancer by upregulating Hint1 expression. J Mol Med. 2018;96(7):661–72.
9. Xue Z, Tan Z, Huang A, Zhou Y, Sun J, Wang X, Thimmappa RB, Stephenson MJ, Osbourn A, Qi X. Identification of key amino acid residues determining product specificity of 2,3-oxidosqualene cyclase in *Oryza* species. New Phytol. 2018;218(3):1076–88.
10. Xueshibojie L, Duo Y, Tiejun W. Taraxasterol inhibits cigarette smoke-induced lung inflammation by inhibiting reactive oxygen species-induced TLR4 trafficking to lipid rafts. Eur J Pharmacol. 2016;789:301–7.
11. Ajmire PV, Chavhan SA, Thete PV, Bakal RL. Pharmacognosy, phytochemistry, Pharmacology and clinical applications of *Taraxacum officinale*. J Pharm Phytochem. 2021;10(3):165–71.
12. Thimmappa R, Geisler K, Louveau T, O'Maille P, Osbourn A. Triterpene biosynthesis in plants. Annu Rev Plant Biol. 2014;65(1):225–57.
13. Wang J, Guo Y, Yin X, Wang X, Qi X, Xue Z. Diverse triterpene skeletons are derived from the expansion and divergent evolution of 2,3-oxidosqualene cyclases in plants. Crit Rev Biochem Mol Biol. 2021;57(2):113–32.
14. Guo CF, Xiong XC, Dong H, Qi XQ. Genome-wide investigation and transcriptional profiling of the oxidosqualene cyclase (OSC) genes in wheat (*Triticum aestivum*). J Syst Evol. 2021;60(6):1378–92.
15. Lian X, Zhang X, Wang F, Wang X, Xue Z, Qi X. Characterization of a 2,3-oxidosqualene cyclase in the Toosendanin biosynthetic pathway of *Melia Toosendan*. Physiol Plant. 2020;170(4):528–36.
16. Zhang M, Wang Y, Zhang W, Li M, Song W, Zheng B, et al. Genome-wide identification and functional characterization of 2, 3-oxidosqualene cyclase genes in Phellodendron amurense. Ind Crop Prod. 2024;208:117721.
17. Han Y, Yang Y, Li Y, Yin X, Chen Z, Yang D, et al. Genome-wide identification of OSC gene family and potential function in the synthesis of Ursane- and oleanane-type triterpene in momordica charantia. Int J Mol Sci. 2021;23(1):196.
18. Wang J, Wang P, Xu M, Chen Y, Feng L. Systematic identification and analysis of OSC gene family of *Rosa rugosa* thunb. Int J Mol Sci. 2022;23(22):13884.
19. Huang X-Q, Dudareva N. Plant specialized metabolism. Curr Biol. 2023;33(11):R473–8.
20. Lanier ER, Andersen TB, Hamberger B. Plant terpene specialized metabolism: complex networks or simple linear pathways? Plant J. 2023;114(5):1178–201.
21. Zhai Q, Li C. The plant mediator complex and its role in jasmonate signaling. J Exp Bot. 2019;70(13):3415–24.
22. Yu X, Zhang W, Zhang Y, Zhang X, Lang D, Zhang X. The roles of methyl jasmonate to stress in plants. Funct Plant Biol. 2019;46(3):197–212.
23. Sharma K, Zafar R. Optimization of Methyl jasmonate and β-cyclodextrin for enhanced production of Taraxerol and taraxasterol in (*Taraxacum officinale* Weber) cultures. Plant Physiol Biochem. 2016;103:24–30.
24. Jiao J, Gai Q-Y, Wang W, Luo M, Zu Y-G, Fu Y-J, Ma W. Enhanced Astragaloside production and transcriptional responses of biosynthetic genes in *Astragalus Membranaceus* hairy root cultures by elicitation with Methyl jasmonate. Biochem Eng J. 2016;105:339–46.
25. Ren A, Qin L, Shi L, Dong X, Mu DS, Li YX, Zhao MW. Methyl jasmonate induces Ganoderic acid biosynthesis in the basidiomycetous fungus *Ganoderma lucidum*. Bioresour Technol. 2010;101(17):6785–90.
26. Dall'Acqua S, Catanzaro D, Cocetta V, Igl N, Ragazzi E, Giron MC, Cecconello L, Montopoli M. Protective effects of ψ taraxasterol 3-*O*-myristate and arnidiol 3-*O*-myristate isolated from *Calendula officinalis* on epithelial intestinal barrier. Fitoterapia. 2016;109:230–5.
27. Zhao X, Guo B, Sun W, Yu J, Cui L. Targeting squalene epoxidase confers metabolic vulnerability and overcomes chemoresistance in HNSCC. Adv Sci. 2023;10(27):e2206878.
28. Chen M, Ma Y, Wu S, Zheng X, Kang H, Sang J, Xu X, Hao L, Li Z, Gong Z, et al. Genome warehouse: A public repository housing Genome-Scale data. Genom Proteom Bioinf. 2021;19(4):584–9.
29. Shibuya M, Xiang T, Katsube Y, Otsuka M, Zhang H, Ebizuka Y. Origin of structural diversity in natural triterpenes: direct synthesis of *seco*-Triterpene skeletons by oxidosqualene cyclase. J Am Chem Soc. 2007;129(5):1450–5.
30. He H, Wang J, Meng Z, Dijkwel PP, Du P, Shi S, et al. Genome-wide analysis of the *SRPP/REF* gene family in *Taraxacum kok-saghyz* provides insights into its expression patterns in response to ethylene and methyl jasmonate treatments. Int J Mol Sci. 2024;25(13):6864.
31. Wang C, Wang X, Zhou P, Li C. Genome-wide identification and characterization of RdHSP genes related to high temperature in *Rhododendron Delavayi*. Plants. 2024;13(13):1878.
32. Yang M, Derbyshire MK, Yamashita RA, Marchler-Bauer A. NCBI's conserved domain database and tools for protein domain analysis. Curr Protoc Bioinf. 2019;69(1):e90.
33. Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2018;46(D1):D493–6.
34. Gasteiger E, Gattiker A, Hoogland C, Ivanyi A, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003;31(13):3784–8.
35. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 2007;35(Web Server):W585–7.
36. Chou K-C, Shen H-B. Cell-PLoc: a package of web servers for predicting subcellular localization of proteins in various organisms. Nat Protoc. 2008;3(2):153–62.
37. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296–303.
38. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202–8. (Web Server).
39. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G, Nielsen H. SignalP 6.0 predicts all five types of signal peptides using protein Language models. Nat Biotechnol. 2022;40(7):1023–5.
40. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genmome Biol. 2015;16(1):157.
41. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genmome Biol. 2019;20(1):238.
42. Tamura K, Stecher G, Kumar S, Battistuzzi FU. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38(7):3022–7.
43. Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–6.

44. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee Th, Jin H, Marler B, Guo H, et al. *MCScanX*: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49–49.
45. Rao X, Huang X, Zhou Z, Lin X. An improvement of the 2^(−delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. Biostat Bioinforma Biomath. 2013;3(3):71–85.
46. Zhang H, Zhang TT, Liu H, Shi DY, Wang M, Bie XM, Li XG, Zhang XS, Thioredoxin-Mediated ROS. Homeostasis explains natural variation in plant regeneration. Plant Physiol. 2018;176(3):2231–50.
47. Liu T, Liao J, Shi M, Li L, Liu Q, Cui X, et al. A jasmonate-responsive bHLH transcription factor TaMYC2 positively regulates triterpenes biosynthesis in Taraxacum antungense Kitag. Plant Sci. 2023;326:111506.
48. Salmon M, Thimmappa RB, Minto RE, Melton RE, Hughes RK, O'Maille PE, et al. A conserved amino acid residue critical for product and substrate specificity in plant triterpene synthases. Proc Natl Acad Sci. 2016;113(30):E4407–14.
49. Srisawat P, Fukushima EO, Yasumoto S, Robertlee J, Suzuki H, Seki H, Muranaka T. Identification of oxidosqualene cyclases from the medicinal legume tree *Bauhinia forficata*: a step toward discovering preponderant α-amyrin-producing activity. New Phytol. 2019;224(1):352–66.
50. Song A, Su J, Wang H, Zhang Z, Zhang X, Van de Peer Y, et al. Analyses of a chromosome-scale genome assembly reveal the origin and evolution of cultivated chrysanthemum. Nat Commun. 2023;14(1):2021.
51. Pryszcz LP, Huerta-Cepas J, Gabaldón T. MetaPhOrs: orthology and paralogy predictions from multiple phylogenetic evidence using a consistency-based confidence score. Nucleic Acids Res. 2011;39(5):e32–32.
52. Barb JG, Bowers JE, Renaut S, Rey JI, Knapp SJ, Rieseberg LH, Burke JM. Chromosomal evolution and patterns of introgression in *Helianthus*. Genet. 2014;197(3):969–79.
53. Kulkarni OS, Mazumder M, Kini S, Hill ED, Aow JSB, Phua SML, Elejalde U, Kjelleberg S, Swarup S. Volatile Methyl jasmonate from roots triggers host-beneficial soil Microbiome biofilms. Nat Chem Biol. 2023;20(4):473–83.
54. Nambara E, Van Wees SCM. Plant hormone functions and interactions in biological systems. Plant J. 2021;105(2):287–9.
55. Laloum T, Martín G, Duque P. Alternative splicing control of abiotic stress responses. Trends Plant Sci. 2018;23(2):140–50.
56. Kianersi F, Amin Azarm D, Pour-Aboughadareh A, Poczai P. Change in secondary metabolites and expression pattern of key rosmarinic acid related genes in Iranian lemon balm (Melissa officinalis L.) ecotypes using methyl jasmonate treatments. Molecules. 2022;27(5):1715.
57. Ali MB, Yu K-W, Hahn E-J, Paek K-Y. Methyl jasmonate and Salicylic acid elicitation induces ginsenosides accumulation, enzymatic and non-enzymatic antioxidant in suspension culture *Panax ginseng* roots in bioreactors. Plant Cell Rep. 2006;25(6):613–20.
58. Bertini L, Palazzi L, Proietti S, Pollastri S, Arrigoni G, de Polverino P, et al. Proteomic analysis of MeJa-Induced defense responses in rice against wounding. Int J Mol Sci. 2019;20(10):2525.
59. Dai Z, Yuan Y, Huang H, Hossain MM, Xiong S, Cao M, et al. Methyl jasmonate mitigates high selenium damage of rice via altering antioxidant capacity, selenium transportation and gene expression. Sci Total Environ. 2021;756:143848.
60. Mittler R, Zandalinas SI, Fichman Y, Van Breusegem F. Reactive oxygen species signalling in plant stress responses. Nat Rev Mol Cell Biol. 2022;23(10):663–79.
61. Han J-Y, Choi HS, Jo H-J, Jun KM, Choi Y-E. Heterologous production of multiple triterpenes, including taraxasterol, in Transgenic tobacco overexpressing *Taraxacum Coreanum* oxidosqualene cyclase (TcOSC1). Plant Biotechnol Rep. 2022;16(5):529–37.
62. Van Deenen N, Unland K, Prüfer D, Schulze Gronover C. Oxidosqualene cyclase knock-down in latex of *Taraxacum Koksaghyz* reduces triterpenes in roots and separated natural rubber. Molecules. 2019;24(15):2703.
63. Choi HS, Han JY, Cheong EJ, Choi YE. Characterization of a pentacyclic triterpene acetyltransferase involved in the biosynthesis of taraxasterol and ψ-taraxasterol acetates in lettuce. Front Plant Sci. 2022;12:788356.
64. Irmisch S, Krause ST, Kunert G, Gershenzon J, Degenhardt J, Köllner TG. The organ-specific expression of terpene synthase genes contributes to the terpene hydrocarbon composition of chamomile essential oils. BMC Plant Biol. 2012;12:84.

Publisher's note