Ornamental Plant Research

ARTICLE

Open Access

https://doi.org/10.48130/opr-0025-0035
Ornamental Plant Research **2025**, 5: e033

Aldehyde dehydrogenase superfamily of three *Cymbidium* orchids: genome-wide identification and response models to drought and salt stress

Linying Wang[1], Ruiyue Zheng[1], Cuili Zhang[1], Chen Chen[1], Jiayi Lu[1], Xi Zhao[1], Yang Tian[1], Siren Lan[1,2*] and Zhong-Jian Liu[1,2*]

[1] *Key Laboratory of National Forestry and Grassland Administration for Orchid Conservation and Utilization at the College of Landscape Architecture and Art, Fujian Agriculture and Forestry University, Fuzhou 350002, China*
[2] *College of Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, China*
* Corresponding authors, E-mail: lkzx@fafu.edu.cn; zjliu@fafu.edu.cn

Abstract

Aldehydes serve as intermediates in various metabolic processes, and their excessive accumulation can disrupt normal plant metabolic activities. Environmental stressors, including drought and salinity, can trigger the generation of reactive oxygen species (ROS), which results in elevated levels of aldehydes within plant cells. Aldehyde dehydrogenases (ALDHs) are recognized as key enzymes that act as 'aldehyde scavengers' capable of detoxifying the harmful aldehydes produced during oxidative stress conditions. Orchids constitute one of the largest families within angiosperms and possess significant economic value. However, genetic data for these non-model plants are limited, and the molecular mechanisms underlying orchid responses to abiotic stress remain poorly understood. In this study, we identified 17, 14, and 21 *ALDH* family members in three *Cymbidium* orchids: *C. sinense*, *C. ensifolium*, and *C. goeringii*, respectively. We conducted comprehensive analyses of their conserved domains, promoter elements, and collinearity relationships and predicted the interaction networks of ALDH proteins and microRNA binding sites in these three orchid species. Furthermore, we subjected *C. sinense* to drought and high salt-stress treatments and examined its expression patterns under these conditions using quantitative real-time PCR (qRT-PCR). Subcellular localization experiments were also performed on genes that exhibited significant responses to stress. Our findings provide valuable insights into the molecular mechanisms of stress resistance in orchids and contribute to a deeper understanding of their potential applications as ornamental plants.

Citation: Wang L, Zheng R, Zhang C, Chen C, Lu J, et al. 2025. Aldehyde dehydrogenase superfamily of three *Cymbidium* orchids: genome-wide identification and response models to drought and salt stress. *Ornamental Plant Research* 5: e033 https://doi.org/10.48130/opr-0025-0035

Introduction

Aldehydes serve as intermediates in numerous metabolic pathways within organisms; however, excessive aldehyde accumulation can disrupt normal metabolic processes and potentially become toxic[1]. Aldehyde dehydrogenases (ALDHs) constitute a superfamily of NAD(P)+-dependent enzymes that catalyze the oxidation of diverse aliphatic and aromatic aldehydes to their corresponding carboxylic acids[2]. This process produces NADPH and NADH, which play critical roles in maintaining redox balance and aldehyde homeostasis[3]. These enzymes are frequently termed 'aldehyde scavengers' due to their role in detoxifying harmful aldehydes[4]. In addition, *ALDH*s are involved in various metabolic pathways such as amino acid metabolism, retinoic acid metabolism, and osmoprotectant synthesis.

ALDHs are ubiquitous across all taxonomic groups, including animals, plants, algae, and others, and exhibit high conservation throughout evolution. According to the nomenclature guidelines established by the Aldehyde Dehydrogenase Gene Nomenclature Committee (AGNC), proteins are classified into 24 distinct families[5]. In plants, the *ALDH* superfamily encompasses 14 families (2, 3, 5, 6, 7, 10, 11, 12, 18, 19, 21, 22, 23, and 24), with seven families (*ALDH10*, *ALDH12*, *ALDH19*, *ALDH21*, *ALDH22*, *ALDH23*, and *ALDH24*) being plant-specific. These genes are expressed in multiple tissues and respond to various stressors. Extensive research indicates that upregulated *ALDH* expression is a common feature of activated stress response pathways.

Most studied plant *ALDH* exhibit upregulated expression in response to various abiotic stresses, including high salinity, dehydration, elevated temperatures, waterlogging, oxidative stress, and heavy metal exposure[6–8]. Environmental challenges such as drought and high salinity trigger rapid reactive oxygen species (ROS) production, leading to toxic aldehyde accumulation and disrupted metabolic processes. ALDH enzymes detoxify these harmful aldehydes, thereby enhancing plant survival under stress. Recent studies suggest that ectopic overexpression of *ALDH* enhances plant tolerance to abiotic stresses. Beyond *Arabidopsis thaliana*, *ALDH* families have also been characterized in diverse plant species, including moss, algae[9], and economically important crops such as sorghum, soybean, and cotton[10–12]. However, *ALDH* remain unstudied in orchids.

Cymbidium, a large genus within the orchid family (Orchidaceae), exhibits diverse growth habits, including terrestrial, epiphytic, and saprophytic lifestyles[13]. For a long time, orchids have fascinated people with their exquisite floral morphology[14]. Orchids have a long history of cultivation and ornamental appreciation globally[15]. However, global climate change, including global warming and rising sea levels, poses significant challenges for terrestrial orchids, particularly in terms of soil drought and salinization[16,17]. Notwithstanding China's vast geographical coverage and rich germplasm diversity, the country faces increasingly severe environmental challenges. To address this, we focused on three historically significant *Cymbidium* species in China—*C. sinense*, *C. ensifolium*, and *C. goeringii*—and conducted whole-genome identification of ALDH genes to study their evolutionary relationships and stress-responsive expression patterns. We identified and classified *ALDH*s from these three species into ten families and six protein subfamilies according to AGNC standards[5]. Further analysis included genomic

organization, sequence homology, and subcellular localization of the identified ALDH genes. We further analyzed ALDHs gene expression in *C. sinense* leaves under drought and salinity stress, and validated the subcellular localization of stress-responsive genes.

Materials and methods

Identification and characterization of ALDH genes in three *Cymbidium* orchids

The whole-genome sequencing data for *C. sinense, C. ensifolium*, and *C. goeringii* were obtained from their respective genomic resources[18–20]. Protein sequences of the *A. thaliana ALDH* family were sourced from TAIR (www.arabidopsis.org, accessed on August 10, 2024). For *Oryza sativa*, ALDH family protein sequences were retrieved from Phytozome (https://phytozome-next.jgi.doe.gov, accessed on August 12, 2024), and previous studies[21]. To conduct BLASTP comparisons using the ALDHs data from the genomes of *C. sinense, C. ensifolium*, and *C. goeringii*, we utilized TBtools (Version 2.4.1)[22]. The presence of the *ALDH* superfamily domain (Pfam00171) in the identified proteins was confirmed using HMMER (www.ebi.ac.uk/Tools/hmmer/search/phmmer, accessed on August 15, 2024). All identified *ALDH* family members were classified according to the criteria established by the AGNC[5]. The physicochemical properties of the proteins were analyzed using the parameter calculation function in TBtools (Version 2.4.1)[22]. Subcellular localization predictions were performed using Cell-PLoc (www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2, accessed on November 17, 2024).

Phylogenetic analysis and multiple sequence alignment

Phylogenetic trees for ALDH genes from *C. sinense*, *C. ensifolium*, *C. goeringii*, *A. thaliana*, and *O. sativa* were constructed using Phylosuite (version 1.2.3) with the Maximum Likelihood algorithm and 1,000 bootstrap replicates[23]. The resulting phylogenetic tree was visualized using iTOL (https://itol.embl.de/personal_page.cgi, accessed on August 20, 2024). Multiple sequence alignments of CsALDHs, CeALDHs, and CgALDHs were conducted using the MAFFT function in Phylosuite. ESPript 3.0 (https://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi, accessed on August 20, 2024) was used with default parameters to visualize amino acid sequences. Conserved motifs were analyzed using MEME (version 5.5.7, https://meme-suite.org/meme/tools/meme, accessed on August 15, 2024). The NCBI's Conserved Domain Database was utilized to examine conserved domain structures (www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi, accessed on August 15, 2024). TBtools was then used to visualize the results obtained from both MEME and NCBI analyses.

Analysis of *cis*-acting elements and gene structure

The upstream 2000 bp regions of the CsALDH, CeALDH, and CgALDH gene coding sequences were captured using TBtools. The PlantPAN (version 4.0, http://plantpan.itps.ncku.edu.tw/plantpan4/index.html, accessed on October 4, 2024) was utilized to predict and analyze the promoter regions within these 2,000-bp upstream sequences[24]. The numbers of exons, introns, coding sequences (CDSs) and untranslated regions (UTRs) were calculated using GXF stat. Data collection, analysis, and visualization were all performed using Microsoft Excel.

Chromosome localization and collinearity analysis

TBtools was employed to map ALDH genes onto the chromosomes of *C. sinense*, *C. ensifolium*, and *C. goeringii*. One-Step MCScanX was utilized for collinearity analysis. Following pairwise collinearity analysis among the three species and *A. thaliana*, FileMerge MCScanX was used to integrate the results. Finally, Dual Synteny Plot was applied for visual presentation and refinement.

Protein-protein interaction (PPI) network, and microRNA binding site prediction

All CsALDH protein sequences were uploaded to the STRING database (https://cn.string-db.org, accessed on September 27, 2024) to predict protein interactions. The resulting interaction data from STRING were imported into Cytoscape, where the Betweenness Centrality (BC) method was used to identify highly interacting proteins. The protein interaction network was then optimized and visualized. Additionally, psRNATarget (www.zhaolab.org/psRNATarget, accessed on November 16, 2024) was used to predict miRNA binding sites[25], with default parameters.

Material processing and transcriptome data analysis

Healthy *C. sinense* 'Jinhuashan' plants, cultivated in the nursery of Fujian Agriculture and Forestry University (Fujian, China), were selected and acclimated in an artificial climate chamber for 10 d under the following conditions: photoperiod of 12 h light/12 h dark, light intensity of 8,000 lux, day/night temperature of 28/25 °C, and humidity of 75%. For drought stress treatment, roots of intact plants were immersed in 20% PEG6000 solution for 15 min and then removed[26]. For high salt stress treatment, roots of intact plants were immersed in 300 mmol/L NaCl solution for 15 min and then removed[27]. The first mature leaf from plants was sampled at five time points: 0 h (control, untreated), 6, 12, 24, and 48 h post-treatment. Each time point included three biological replicates (independent plants). Transcriptome data processing followed the method of Yang et al.[18] (NCBI accession: SRR15174943–SRR15174947). Raw reads were quality-filtered using fastp[28], and TPM (Transcripts Per Million) values were calculated with Kallisto (version 0.48.0) with parameters set as: quant -l -b 100. Data were visualized using TBtools.

Fluorescence quantitative test and data analysis

Total RNA was isolated from leaves at various time points after drought and salt stress treatments using FastPure Plant Total RNA Isolation Kit (Vazyme, Nanjing, China; Cat. No. RC401-01). Hifair III One-Step RT-qPCR SuperMix with gDNA Digestion (Yeasen, Shanghai, China; Cat. No. 11151ES) was employed to synthesize cDNA. The Actin gene (*Mol022529*) was chosen as the internal control for its stability[29]. Primers for the Actin gene and 14 candidate genes were designed using Primer Premier 5.0.

Real-time PCR was conducted on the Applied Biosystems QuantStudio system using cDNA labeled with SYBR Green fluorescent dye (Yeasen, Shanghai, China; Cat. No. 10222ES60). The thermal cycling protocol was as follows: an initial denaturation at 95 °C for 1 min, followed by 40 cycles of 95 °C for 10 s and 60 °C for 20 s, with fluorescence signal acquisition during the annealing/extension phase. Dissociation curves were generated using the instrument's default settings, and CT values were recorded for each sample. Gene expression levels were normalized to the Actin internal control gene (*Mol022529*), and relative expression was calculated using the $2^{-\Delta\Delta C_T}$ method[30]. Expression level correlation analysis was conducted using OmicStudio (www.omicstudio.cn/home, accessed on September 25, 2024)[31].

Cloning and subcellular localization of *CsALDH5F1* and *CsALDH11A3*

We selected cDNA samples exhibiting the highest expression levels of target genes in qRT-PCR as templates for cloning. The templates used for cloning *CsALDH5F1* and *CsALDH11A3* were cDNAs collected at 24 and 48 h after salt stress treatment, respectively. The Super1300 vector was linearized using XbaI and KpnI

restriction enzymes. Primers with homologous arms were designed based on the coding sequences (excluding their terminators) of *CsALDH5F1* and *CsALDH11A3*, with reference to the Super1300 vector. Target gene fragments were amplified using homology-arm-containing primers, gel-purified, and ligated with the linearized Super1300 vector (Vazyme, Nanjing, China; Cat. No. DC301-01). The recombinant plasmids were then transformed into *Escherichia coli* for propagation. Positive clones were selected for plasmid amplification, followed by plasmid extraction (Vazyme, Nanjing, China; Cat. No. DC201-01) and subsequent transformation into *Agrobacterium*. The working suspension was prepared in deionized water with final concentrations of 1 M $MgCl_2$, 1 M MES buffer, and 20 mg/mL Acetosyringone. The bacterial suspension was adjusted to an OD_{600} value of 0.8–1.0. Bacterial suspension was injected into interveinal regions of the abaxial surface in newly matured leaves of tobacco (*Nicotiana benthamiana*). Following injection, plants were maintained in darkness for 24 h and subsequently cultured under normal conditions for 48 h. GFP fluorescence signals on tobacco leaves were observed using confocal laser microscopy (Leica STELLARIS) in complete darkness.

Results

Identification, nomenclature, and characteristics of *ALDH*s in three *Cymbidium* orchids

We identified 14, 17, and 21 ALDH genes in *C. sinense*, *C. ensifolium*, and *C. goeringii*, respectively. To elucidate their evolutionary relationships and functional correlations within the ALDH gene family, we constructed a phylogenetic tree using ALDH sequences from these three orchids, *A. thaliana*, and *O. sativa*. An unrooted phylogenetic tree was generated based on CeALDH, CsALDH, CgALDH, AtALDH, and OsALDH sequences, followed by cluster analysis.

As shown in Fig. 1a, according to the AGNC guidelines, protein sequences showing > 40% similarity to known ALDH sequences were classified into the same family, while sequences with > 60% similarity were assigned to the same subfamily. Sequences showing < 60% similarity (i.e., > 40% divergence) to known ALDH sequences were classified as new gene families. Based on this classification, all identified *ALDH* family members in the three orchids were grouped into ten main families (Families 2, 3, 5, 6, 7, 10, 11, 12, 18, and 22) and six subfamilies. Subfamilies A, B, C, F, and H correspond to those previously identified in *A. thaliana* and *O. sativa*, while subfamily L, newly identified in this study, exhibits 40%–60% amino acid sequence similarity to other known ALDH subfamilies. Among the main *ALDH* families, *ALDH3* contained the most members across the three orchid species, followed by *ALDH2*. Among subfamilies, the newly identified subfamily L has the most members, followed by subfamily B.

The physicochemical properties of the identified ALDH proteins in *C. sinense*, *C. ensifolium*, and *C. goeringii* are summarized as follows. Protein lengths ranged from 117 to 653 amino acids (aa), isoelectric points (pI) ranged between 4.71 and 9.75, and molecular weights (MW) ranged from 13.29 to 71.19 kDa. Additionally, subcellular localization prediction indicated that cytoplasm was the primary localization site for most ALDH proteins, followed by mitochondria, chloroplasts, peroxisomes, and the endoplasmic reticulum (Supplementary Table S1).

Multiple sequence alignment and phylogenetic analysis of ALDH proteins

We conducted a conservation analysis on the identified members of the *ALDH* family. During the identification process, we observed that some genes were relatively short in length. However, searches in the Pfam and NCBI Conserved Domain Databases revealed that all identified ALDH members possessed the ALDH or ALDH-SF superfamily domain (PF00171) (Fig. 2a). Therefore, we concluded that these shorter genes likely represent valid *ALDH* family members despite potential errors in genome annotation, and thus retained them for further analysis.

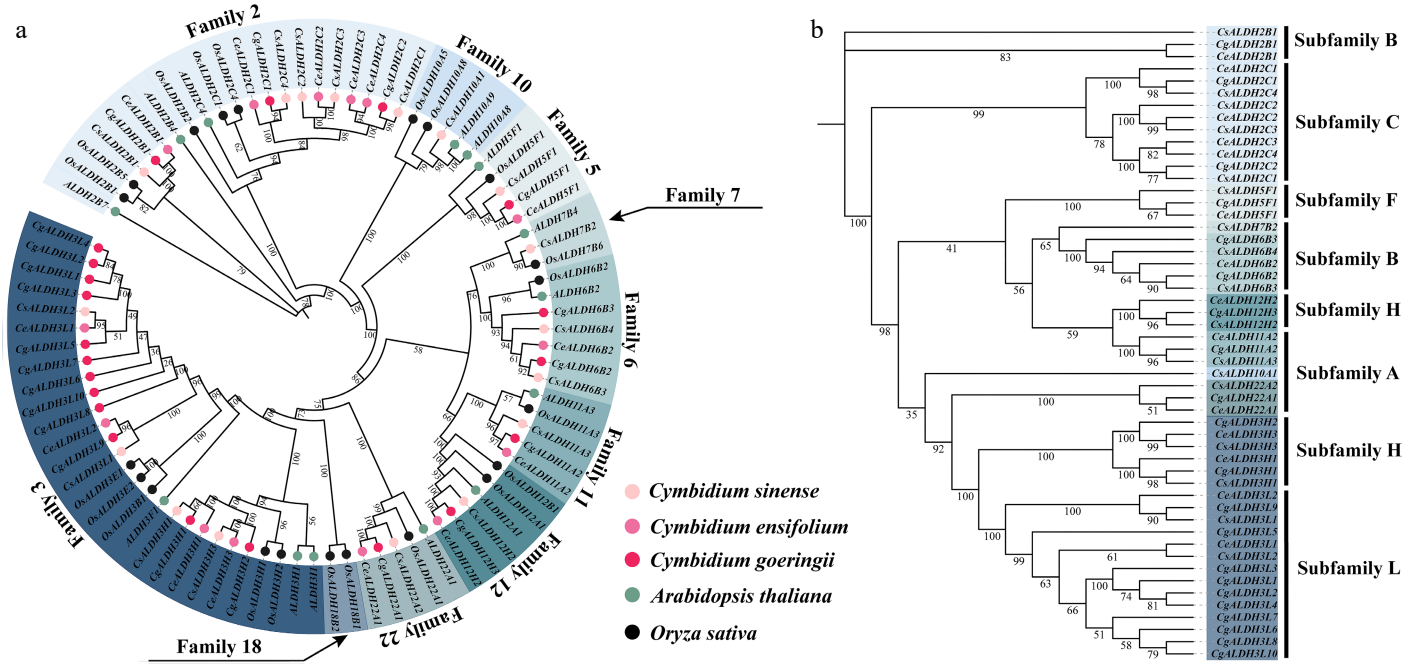

Fig. 1 Phylogenetic analysis of *ALDH* family members. (a) Phylogenetic tree constructed using ALDH family members from *C. sinense*, *C. ensifolium*, *C. goeringii*, *A. thaliana*, and *O. sativa*. (b) In the phylogenetic tree derived from *C. sinense*, *C. ensifolium*, and *C. goeringii*, the text on the right indicates the classification of protein subfamilies. These subfamilies are categorized into seven groups (A, B, C, F, H, and L), with subfamily L being a newly identified group in this study.

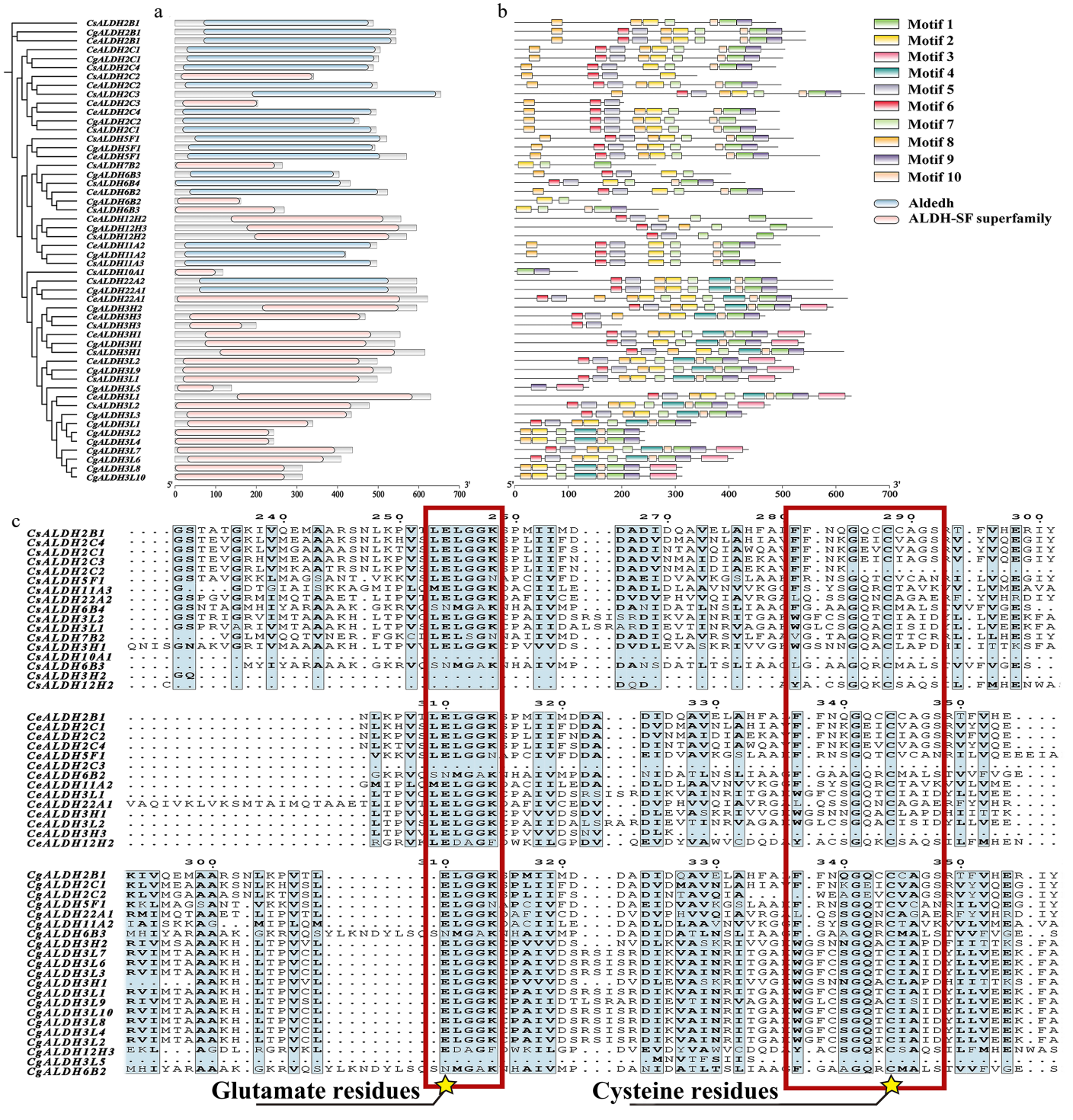

Fig. 2 Conserved domain analysis and multiple sequence alignment of the *CsALDH*, *CeALDH*, and *CgALDH* gene families. (a) Conserved domains, where blue represents the Aldedh domain (PF00171) and pink represents the ALDH-SF superfamily domain (PF00171). (b) Motifs of proteins, with different colors representing distinct motifs numbered from 1 to 10. The scale at the bottom allows for estimation of protein lengths. (c) Multiple sequence alignment of all identified proteins. Conserved motifs and active sites of glutamic acid residues and conserved motifs and active sites of cysteine residues are marked with red squares and yellow star symbols.

Motif analysis (Fig. 2b) indicated no significant differences in motifs across various families. Notably, most members of the *ALDH3* family exhibited a unique motif 3, while most members of the *ALDH3* and *ALDH22* families possessed motif 4, characteristics not observed in other families.

ScanProsite and multiple sequence alignment analyses (Supplementary Table S2; Fig. 2c) revealed that most ALDH proteins contained both glutamic acid (PS00687) and cysteine (PS00070) active sites. Among the three orchid species, all members of the *ALDH3* family lacked both PS00687 and PS00070. In *C. sinense*, 41.18% genes (7/17) contained both active sites; in *C. ensifolium*, 50.00% genes (7/14) contained both active sites; and in *C. goeringii*, 23.81% genes (5/21) contained both active sites.

Analysis of *cis*-elements and gene structure

Analysis of the 2,000-bp upstream region of the CsALDH, CeALDH, and CgALDH gene coding sequences (CDS) revealed numerous transcription factor binding sites. Specifically, *C. sinense*, *C. ensifolium*, and *C. goeringii* were found to contain 47, 47, and 43 types of transcription factor binding sites, respectively, with the number of sites ranging from 29 (e.g., *CsALDH3L2*, *CsALDH6B4*, *CeALDH2C3*, *CeALDH2C4*) to 37 (*CsALDH3H3*, *CsALDH10A1*, *CgALDH3L9*) (Fig. 3).

We further analyzed the regulatory functions associated with these transcription factors and found that *ALDHs* can respond to various plant hormones and abiotic stress signals. For instance, plant hormone-responsive transcription factors (e.g., *Aux/IAA*, *BES1*, and *EIN3*), cold stress-responsive factors (e.g., *AP2/ERF* and *HSF*), heat stress-responsive factors (e.g., *AT-Hook*), drought stress-responsive factors (e.g., *NF-Y* and *LEA*), and salt stress-responsive factors (e.g., *SRS* and *C2H2*).

Furthermore, we examined the structure and quantity of introns, exons, CDSs, and UTRs (Fig. 3). Most members contain at least one exon and intron, except for *CsALDH22A2* and *CsALDH2B1*, which lack introns. Most members of the *CsALDH* and *CeALDH* families do not contain UTRs. Among the *CgALDHs*, 33.3% members (7/21) lack UTRs, while the remaining members have one to four UTRs.

Chromosomal localization and collinearity analysis of *ALDH* in three *Cymbidium* orchids

The chromosomal locations of ALDH genes in the three *Cymbidium* species are shown in Supplementary Fig. S1. The distribution of ALDH genes on the chromosomes of *C. sinense* and *C. ensifolium* was relatively uniform, whereas the distribution in *C. goeringii* is more concentrated. Specifically, both *C. sinense* and *C. goeringii* exhibited a similar distribution pattern, with a maximum of three *ALDHs* per chromosome (e.g., Chr08 and Chr17 in *C. sinense*; Chr09 and Chr14 in *C. ensifolium*). In contrast, in *C. goeringii*, ALDH genes are predominantly located on chromosome Chr07, with no more than three *ALDHs* present on any single chromosome.

We conducted synteny analysis of *ALDHs* among the three orchid species and *A. thaliana* (Fig. 4). The results demonstrated that the three orchids exhibited stronger syntenic relationships among themselves than with the model plant *A. thaliana*. Through homology analysis of *ALDHs* across the four species, we identified 13 homologous gene pairs between *C. sinense* and *C. goeringii*, 10 between *C. goeringii* and *C. ensifolium*, but only one between *A. thaliana* and *C. ensifolium*. These findings indicate that *ALDHs* were more conserved within orchids than between orchids and *A. thaliana*.

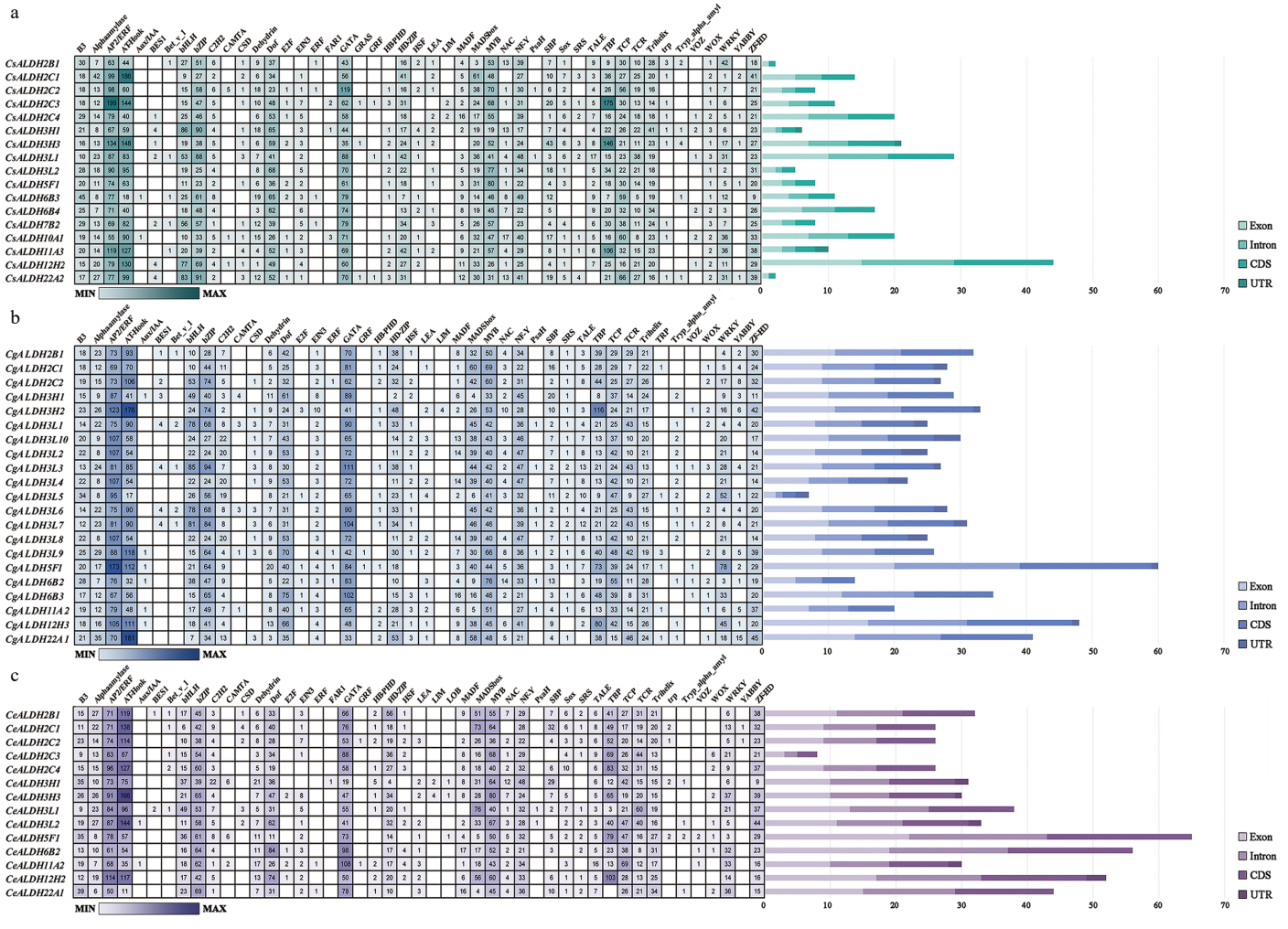

Fig. 3 *Cis*-element prediction and gene structure analysis. The gene structure of each *ALDH* family member is visualized, depicting the counts of exons, introns, CDSs, and UTRs. Distinct colors are used to differentiate these structural components. (a) *C. sinense*. (b) *C. goeringii*. (c) *C. ensifolium*. The quantity can be estimated based on the scale at the bottom.

Protein-protein interaction (PPI) network analysis

We constructed a PPI regulatory network based on the 17 identified CsALDH proteins (Fig. 5a). Analysis of the proteins interacting with CsALDHs revealed that the majority were associated with abiotic stress responses. The most strongly interacting protein was POP (Polyamine Oxidase). Cytochrome P450 (CYP), polyamine oxidase (PAO), and Alanine-Glyoxylate Transaminase (AGT) also exhibited strong interactions with CsALDH proteins.

MicroRNA binding sites and stress response mechanism prediction

MicroRNAs (miRNAs) were widely present in eukaryotic organisms as endogenous non-coding single-stranded RNA molecules approximately 20–24 nucleotides in length. By analyzing the *CsALDH* sequences using psRNATarget (Supplementary Table S3), we found that *CsALDH3H3* had the highest number of miRNA target sites (30). Among the predicted miRNAs, previous studies had validated that *miR414* responded to abiotic stresses in plants, particularly salt stress[32]. We speculate that *miR414* may inhibit the translation of its target genes post-transcriptionally, thereby reducing their protein level. This mechanism could enhance the conversion of toxic aldehydes into carboxylic acids as a salt stress response (Fig. 5b).

Expression patterns by transcriptome data and qRT-PCR analysis of *CsALDHs*

As shown in Fig. 6a, a comparative analysis of Transcripts Per Million (TPM) values was conducted across different tissues (Supplementary Table S4). The transcript abundance of each unigene in the *C. sinense* floral transcriptome was represented by log2(TPM) values. The expression levels of *CsALDHs* in the root were generally significantly lower than in other tissues. Notably, *CsALDH2B1*, *CsALDH6B3*, and *CsALDH7B2* showed no expression in roots, stems, leaves, flowers, or fruits. Additionally, *CsALDH2C1* was not expressed in the root. In the stem, *CsALDH3L1*, *CsALDH10A1*, and *CsALDH1A3* showed prominent expression. In leaves, most genes were highly expressed, particularly *CsALDH10A1*, which had a TPM value as high as 153.32. In flower, the overall expression of *CsALDHs* was lower than in stem, leaf, and fruit but slightly higher than in root. In fruit, *CsALDH3L1* and *CsALDH10A1* exhibited high expression levels.

Fourteen differentially expressed *CsALDHs* were screened, and their expression patterns at various time points after stress were analyzed by qRT-PCR. Figure 6b displayed 13 genes with significant expression changes, and the qRT-PCR data are presented in Supplementary Table S5. The primer sequences used for qRT-PCR are

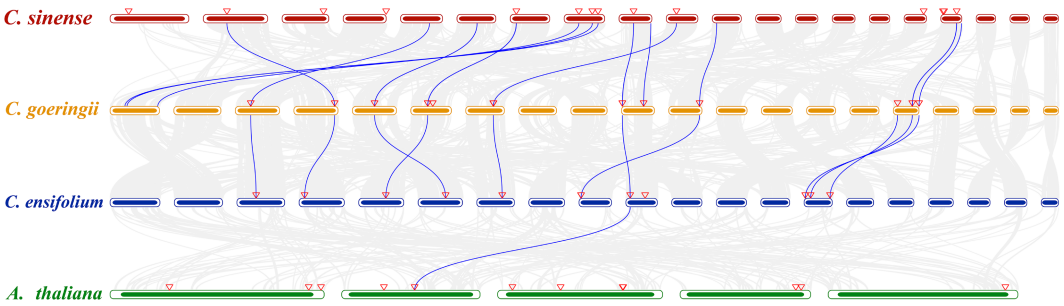

Fig. 4 Collinearity analysis between *C. sinense*, *C. goeringii*, *C. ensifolium*, and *A. thaliana*. The chromosomes of orchids are arranged from left to right as Chr01 to Chr20. The chromosomes of *A. thaliana* are arranged from left to right as Chr01 to Chr05. The gray lines in the background indicate collinear blocks within the three genomes. The blue lines highlight syntenic *ALDH* pairs. Red triangles indicate *ALDH* localization.

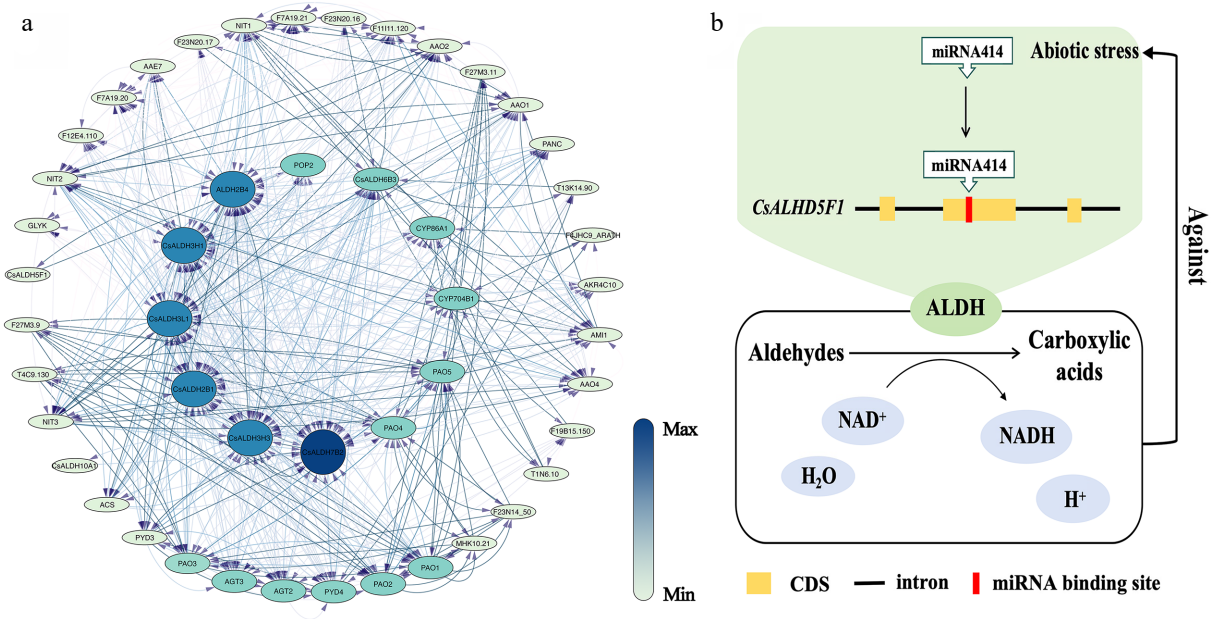

Fig. 5 (a) The PPI network of CsALDH. Each ellipse represents a protein, and lines connect pairs of interacting proteins. The strength of the interactions can be estimated using the color bar on the right. (b) Response pattern diagram of *CsALDH5F1* under abiotic stress.

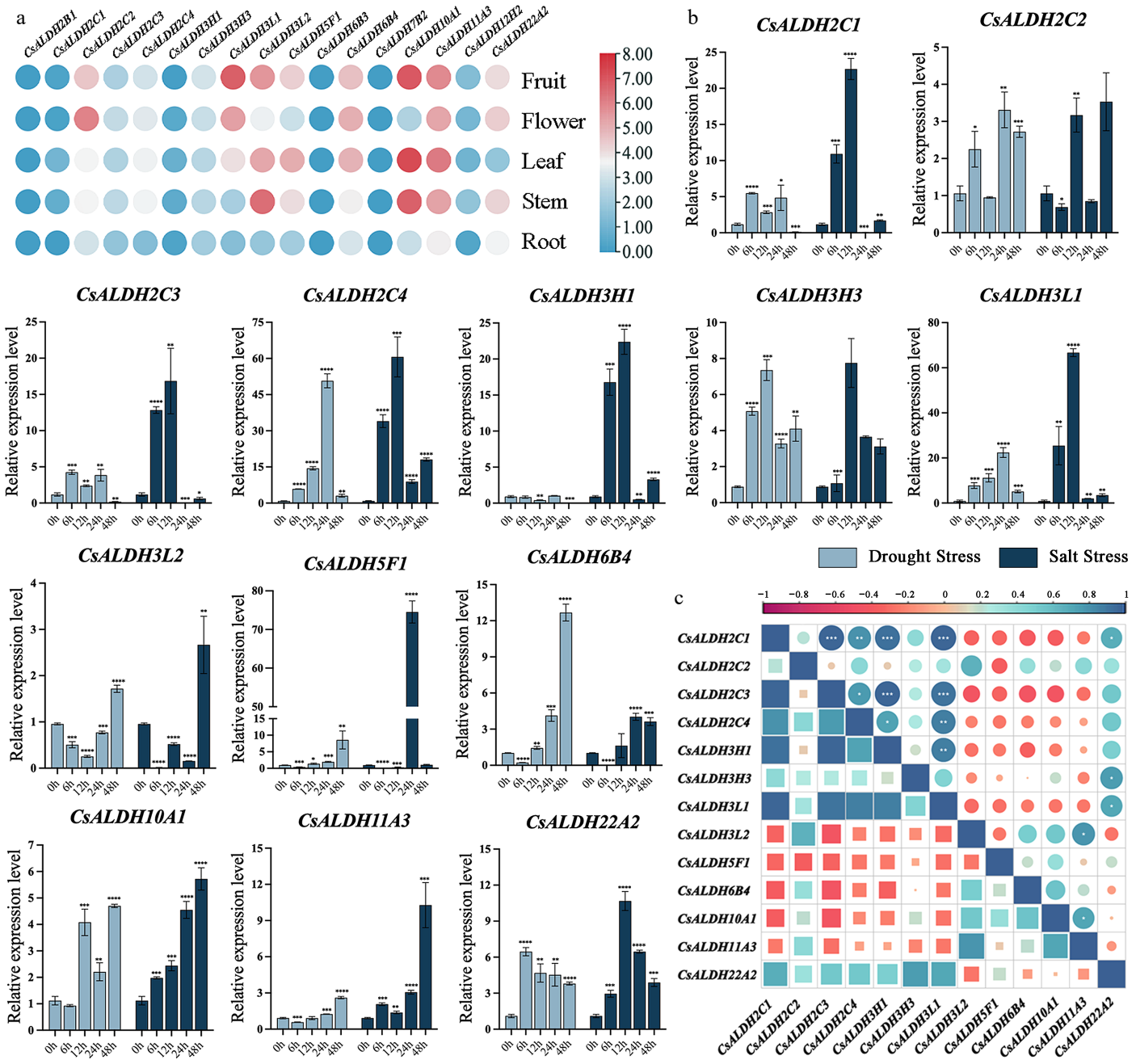

Fig. 6 (a) The heatmap of *ALDHs* expression in different tissues of *C. sinense*. (b) Expression patterns of *ALDHs* in *C. sinense* leaves at different time points after drought and salt stress. The x-axis represents time points (0 h as the control group, 6, 12, 24, and 48 h as experimental groups). The y-axis indicates the relative expression levels after exposure to stress for different durations. Light blue bars represent drought stress, while dark blue bars represent salt stress. The symbols *, **, ***, and **** denote significant differences at $p < 0.05$, $p < 0.01$, $p < 0.001$, and $p < 0.0001$, respectively. (c) Correlation analysis of qRT-PCR results. Correlation strengths are indicated by the graduated scale above.

provided in Supplementary Table S6. The qRT-PCR results indicated that the *ALDH* family actively responded to drought and salt stress. Overall, the response of *CsALDHs* to salt stress was more pronounced than to drought stress. These results suggest that the ALDH gene family is indeed involved in the response to both drought and salt stress, but different genes exhibit distinct response mechanisms to these stresses. Under drought stress, the expression of most genes significantly increased before or at 12 h, while under salt stress, the expression of all *CsALDHs* significantly increased starting from 12 h. Some genes also showed significant stress-specific correlations (Fig. 6c).

Cloning and subcellular localization of *CsALDH5F1* and *CsALDH11A3*

ALDH was a key element in oxygen-dependent metabolic transformation in cells. *CsALDH5F1* and *CsALDH11A3* were significantly upregulated at 24 and 48 h under salt stress. In phylogenetic trees, *CsALDH5F1* clustered with *AtALDH5F1*, while *CsALDH11A3* clustered with *AtALDH11A3*. Previous studies have shown that *AtALDH5F1* is highly responsive to hypoxia stress, and *AtALDH11A3* catalyzes the production of 3-phosphoglycerate and NADPH, which are essential for mannitol biosynthesis in many plant species[33]. Mannitol enhanced plants' ability to withstand stressful conditions, including drought and elevated soil salinity[34].

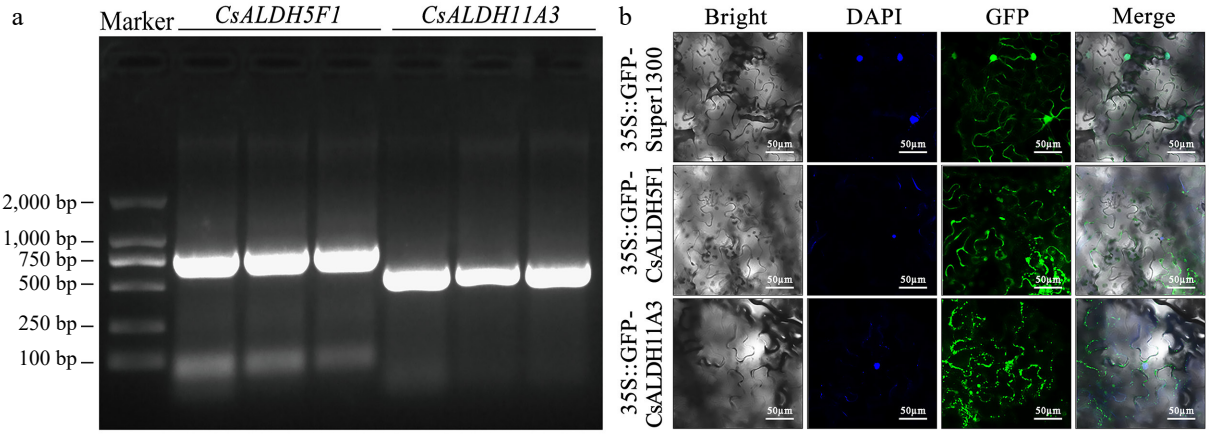

Fig. 7 (a) Agarose gel electrophoresis of PCR-amplified DNA full-length sequences. All samples exhibited a single band at the expected size, confirming the successful amplification of the target genes. (b) Subcellular localization of *CsALDH5F1* and *CsALDH11A3* in *N. benthamiana* epidermal cells.

Therefore, we hypothesized that *CsALDH5F1* and *CsALDH11A3* also play roles in *C. sinense*'s response to stress. We performed PCR amplification using *C. sinense* cDNA as a template with specific primers for *CsALDH5F1* and *CsALDH11A3*. Agarose gel electrophoresis showed bands corresponding to the expected sizes of target fragments, and the full-length sequences of *CsALDH5F1* and *CsALDH11A3* were 780 and 531 bp, respectively (Fig. 7a). Subcellular localization prediction indicated that CsALDH5F1 localized in mitochondria, while *CsALDH11A3* is located in the cytoplasm. To verify these predictions, we expressed the fusion proteins 35S::GFP-CsALDH5F1 and 35S::GFP-CsALDH11A3 in *N. benthamiana* leaves, with 35S::GFP-Super1300 as the control, and DAPI as the nuclear marker. The results confirmed that *CsALDH5F1* was indeed localized in the mitochondria, while *CsALDH11A3* was localized in the cytoplasm (Fig. 7b).

Discussion

ALDH genes are widely present in animals and plants and are involved in various life processes[35]. The genome-wide identification of *CsALDHs*, *CeALDHs*, and *CgALDHs* provides a foundation for further research on the functions of the *ALDH* superfamily in orchids. In this study, we identified a total of 17 ALDH genes in *C. sinense*, 14 in *C. ensifolium*, and 20 in *C. goeringii*. The numbers are comparable to those in *A. thaliana* (14)[35], rice (21)[36], pepper (27)[37], and melon (17)[38]. These results suggest that the *ALDH* family undergoes quantitative changes during long-term environmental adaptation, thereby enhancing plant resistance to environmental stress. As one of the most ecologically diverse plant families, orchids (Orchidaceae) are widely distributed across diverse habitats and exhibit unique life strategies[39]. The differential expansion of ALDH genes across orchid species likely reflects lineage-specific evolutionary responses to habitat-specific stressors. These findings offer new insights into the role of *ALDH* superfamily genetic adaptations in shaping the ecological success and niche-specific survival strategies of orchids. Multiple sequence alignment confirmed the conservation of key amino acid active sites in the *ALDH* family, specifically Glutamic (PS00687), and Cysteine (PS00070), findings that align with previous research[40]. Phylogenetic analysis provides a valuable theoretical framework for predicting functions of homologous genes across species. Genes clustered within the same phylogenetic branch exhibited conserved structural and functional characteristics, allowing for the inference of gene function based on prior studies. For instance, *AtALDH5F1* and *AtALDH11A3* in *A. thaliana* have been shown to respond to abiotic stress, with their expression levels decreasing under hypoxic conditions and being sensitive to temperature changes[1,33]. Therefore, it is reasonable to hypothesize that genes clustered with these *ALDHs* may possess similar or identical functions. Studies on *OsALDHs* suggest that certain ALDH genes involved in abiotic stress responses could contribute to the breeding improvement of rice[36]. Identifying stress-responsive ALDH genes in orchids will facilitate genetic improvement of orchid cultivars, enhancing their adaptability to extreme habitats.

The structural variations observed in ALDH genes are likely associated with the evolution and functional diversification of this gene family across different species. Structural analysis of the *CsALDH*, *CeALDH*, and *CgALDH* superfamilies revealed that while conserved motifs differ among families, they tend to be more similar within the same family. Analysis of intron and exon numbers indicated that most ALDH genes contain multiple introns and exons. Examination of *cis*-acting elements in the promoter regions showed that the upstream regions of *ALDHs* in the three orchid species contain numerous binding sites for transcription factors involved in light response, plant hormone signaling, cell development, and abiotic stress responses, such as *B3*, *E2F*, *SBP*, and *bZIP*[41]. In wheat, *TaNF-YA7-5B* has been shown to regulate stomatal closure under drought conditions, promoting leaf water retention and maintaining ROS homeostasis[42]. Heterologous expression of LEA proteins from *Pinus tabuliformis* in Escherichia coli demonstrated that most *PtaLEAs* enhanced the bacterium's tolerance to salt and heat stress[43].

Prediction of the CsALDHs PPI network revealed its interactions with numerous proteins associated with abiotic stress responses. For instance, Polyamine Oxidase (POP) plays a crucial role in polyamine metabolism, an essential process for plant growth, development, and stress tolerance[44]. Cytochrome P450 (CYP) is involved in suberin synthesis, significantly enhancing plant resistance to salt stress[45]. Alanine-Glyoxylate Transaminase (AGT) participates in glyoxylate metabolism, which is vital for maintaining metabolic homeostasis in plants[46]. Collectively, these proteins play pivotal roles in plant growth, development, stress responses, and metabolic regulation. As important regulatory factors, miRNAs participated in processes including plant growth, development, and stress responses[47]. The analysis of miRNA binding sites indicated that most miRNAs in *C. sinense* responded to various abiotic stresses; for example, *miR169* was responsive to salinity, temperature, and drought stresses[48], while *miR399a* responded to phosphorus deficiency[49]. In the present study, *CsALDH5F1* showed the most significant response under salt stress at the 24 h mark, which may be related to its targeting by *miRNA414*. Based on these predicted miRNA sites, we constructed a potential stress response pattern

diagram. Consequently, we investigated the responses of *C. sinense* to drought and high salt stress.

The *ALDH* superfamily represents a large and complex group of enzymes. In this study, we analyzed the expression profiles of the *CsALDH* family in roots, stems, leaves, flowers, and fruits of *C. sinense* using transcriptome data. Most *CsALDH*s (14 out of 17) were expressed in these tissues. The tissue-specific expression patterns provide foundational data for exploring their potential functions. Following drought and high salt stress treatments on *C. sinense*, qRT-PCR analysis was conducted at different time points post-stress to determine the expression patterns of each gene under abiotic stress. *ALDH*s generally showed positive responses to both drought and salt stress. Notably, the expression levels of *CsALDH5F1* and *CsALDH11A3* increased significantly during specific periods following stress exposure. Subcellular localization confirmed that these genes are not localized in the nucleus, as predicted, but function in mitochondria and cytoplasm, consistent with previous studies in *A. thaliana*[33]. Our findings provided valuable insights and served as a reference for future research on abiotic stress responses in orchid plants.

Conclusions

Through genome-wide identification in *C. sinense*, *C. ensifolium*, and *C. goeringii*, we detected 17, 14, and 21 ALDH genes respectively. Gene structure and homology analyses demonstrated that the *ALDH* family exhibited high evolutionary conservation. Furthermore, prediction of PPI networks and miRNA binding sites indicated close associations between *ALDH*s and abiotic stress responses. Transcriptome and qRT-PCR analyses revealed that 13 differentially expressed *CsALDH*s played crucial roles in abiotic stress responses. Subcellular localization experiments in tobacco with *CsALDH5F1* and *CsALDH11A3* validated prediction accuracy and confirmed these *CsALDH*s' cellular locations.

Author contributions

The authors confirm contributions to the paper as follows: study conception and design: Wang L, Liu ZJ, Lan S; data collection: Wang L, Zheng R; experimental operation: Wang L, Lu J, Tian Y; analysis and interpretation of results: Wang L, Zhang C, Chen C; draft manuscript preparation: Wang L, Zhao X. All authors reviewed the results and approved the final version of the manuscript.

Data availability

The genomic data of *C. sinense* are derived from NCBI: PRJNA743748. The genomic data of *C. ensifolium* are derived from NCBI: PRJNA694815. The genomic data of *C. goeringii* are derived from NCBI: PRJNA668552. All data generated or analyzed during this study are included in this published article and its supplementary information files, and also available from the corresponding author on reasonable request.

Acknowledgments

The research was funded by the National Key Research and Development Program of China (2023YFD1600504), and the Research Cooperation Program of Fujian Agriculture and Forestry University (KH240047A).

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary information accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/opr-0025-0035)

Dates

Received 31 December 2024; Revised 26 June 2025; Accepted 3 July 2025; Published online 28 August 2025

References

1. Stiti N, Missihoun TD, Kotchoni SO, Kirch HH, Bartels D. 2011. Aldehyde dehydrogenases in *Arabidopsis thaliana*: biochemical requirements, metabolic pathways, and functional analysis. *Frontiers in Plant Science* 2:65
2. Lindahl R. 1992. Aldehyde dehydrogenases and their role in carcinogenesis. *Critical Reviews in Biochemistry and Molecular Biology* 27:283–335
3. Gao Z, Loescher WH. 2000. NADPH supply and mannitol biosynthesis. Characterization, cloning, and regulation of the non-reversible glyceraldehyde-3-phosphate dehydrogenase in celery leaves. *Plant Physiology* 124:321–30
4. Sunkar R, Bartels D, Kirch HH. 2003. Overexpression of a stress-inducible aldehyde dehydrogenase gene from *Arabidopsis thaliana* in transgenic plants improves stress tolerance. *The Plant Journal* 35:452–64
5. Vasiliou V, Bairoch A, Tipton KF, Nebert DW. 1999. Eukaryotic aldehyde dehydrogenase (ALDH) genes: human polymorphisms, and recommended nomenclature based on divergent evolution and chromosomal mapping. *Pharmacogenetics* 9:421–34
6. Guo J, Sun W, Liu H, Chi J, Odiba AS, et al. 2020. Aldehyde dehydrogenase plays crucial roles in response to lower temperature stress in *Solanum tuberosum* and *Nicotiana benthamiana*. *Plant Science* 297:110525
7. Islam MS, Hasan MS, Hasan MN, Prodhan SH, Islam T, et al. 2021. Genome-wide identification, evolution, and transcript profiling of Aldehyde dehydrogenase superfamily in potato during development stages and stress conditions. *Scientific Reports* 11:18284
8. Islam MS, Ghosh A. 2022. Evolution, family expansion, and functional diversification of plant aldehyde dehydrogenases. *Gene* 829:146522
9. Wood AJ, Duff RJ. 2009. The aldehyde dehydrogenase (ALDH) gene superfamily of the moss *Physcomitrella patens* and the algae *Chlamydomonas reinhardtii* and *Ostreococcus tauri*. *The Bryologist* 112:1–11
10. Islam MS, Mohtasim M, Islam T, Ghosh A. 2022. Aldehyde dehydrogenase superfamily in sorghum: genome-wide identification, evolution, and transcript profiling during development stages and stress conditions. *BMC Plant Biology* 22:316
11. Ren Z, Liu N, Jia H, Sun M, Ma S, et al. 2024. Discovery of aldehyde dehydrogenase as a potential fungicide target and screening of its natural inhibitors against *Fusarium verticillioides*. *Journal of Agricultural and Food Chemistry* 72:19424–35
12. Gu H, Pan Z, Jia M, Fang H, Li J, et al. 2024. Genome-wide identification and analysis of the cotton ALDH gene family. *BMC Genomics* 25:513
13. Ma S, Wang M, Li P, Guo L, Xiong L, et al. 2024. Transcriptome and metabolome analysis reveal the lip color variation in *Cymbidium floribundum*. *Ornamental Plant Research* 4:e019
14. Zhang D, Zhao XW, Li YY, Ke SJ, Yin WL, et al. 2022. Advances and prospects of orchid research and industrialization. *Horticulture Research* 9:uhac220
15. Yang F, Gao J, Li J, Wei Y, Xie Q, et al. 2024. The China orchid industry: past and future perspectives. *Ornamental Plant Research* 4:e002
16. Hassani A, Azapagic A, Shokri N. 2021. Global predictions of primary soil salinization under changing climate in the 21st century. *Nature Communications* 12:6663
17. Berg A, Sheffield J. 2018. Climate change and drought: the soil moisture perspective. *Current Climate Change Reports* 4:180–91
18. Yang FX, Gao J, Wei YL, Ren R, Zhang GQ, et al. 2021. The genome of *Cymbidium sinense* revealed the evolution of orchid traits. *Plant Biotechnology Journal* 19:2501–16
19. Ai Y, Li Z, Sun WH, Chen J, Zhang D, et al. 2021. The *Cymbidium* genome reveals the evolution of unique morphological traits. *Horticulture Research* 8:255

20. Sun Y, Chen GZ, Huang J, Liu DK, Xue F, et al. 2021. The *Cymbidium goeringii* genome provides insight into organ development and adaptive evolution in orchids. *Ornamental Plant Research* 1:10
21. Kotchoni SO, Jimenez-Lopez JC, Gao D, Edwards V, Gachomo EW, et al. 2010. Modeling-dependent protein characterization of the rice aldehyde dehydrogenase (ALDH) superfamily reveals distinct functional and structural features. *PLoS One* 5:e11516
22. Chen C, Wu Y, Li J, Wang X, Zeng Z, et al. 2023. TBtools-II: a "one for all, all for one" bioinformatics platform for biological big-data mining. *Molecular Plant* 16:1733–42
23. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, et al. 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources* 20:348–55
24. Chow CN, Yang CW, Wu NY, Wang HT, Tseng KC, et al. 2024. PlantPAN 4.0: updated database for identifying conserved non-coding sequences and exploring dynamic transcriptional regulation in plant promoters. *Nucleic Acids Research* 52:D1569–D1578
25. Dai X, Zhuang Z, Zhao PX. 2018. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Research* 46:W49–W54
26. Shaffique S, Ali Shah A, Kang SM, Injamum-Ul-Hoque M, Shahzad R, et al. 2024. Melatonin: dual players mitigating drought-induced stress in tomatoes via modulation of phytohormones and antioxidant signaling cascades. *BMC Plant Biology* 24:1101
27. Feng Y, Chen Z, Chen L, Han M, Liu J, et al. 2025. Comprehensive evaluation of physio-morphological traits of alfalfa (*Medicago sativa* L.) varieties under salt stress. *Physiologia Plantarum* 177:e70044
28. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890
29. Zhu M, Wang Q, Tu S, Ke S, Bi Y, et al. 2023. Genome-wide identification analysis of the R2R3-MYB transcription factor family in *Cymbidium sinense* for insights into drought stress responses. *International Journal of Molecular Sciences* 24:3235
30. Schmittgen TD, Zakrajsek BA, Mills AG, Gorn V, Singer MJ, et al. 2000. Quantitative reverse transcription–polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods. *Analytical Biochemistry* 285:194–204
31. Lyu F, Han F, Ge C, Mao W, Chen L, et al. 2023. OmicStudio: a composable bioinformatics cloud platform with real-time feedback that can generate high-quality graphs for publication. *iMeta* 2:e85
32. Xu T, Zhang L, Yang Z, Wei Y, Dong T. 2021. Identification and functional characterization of plant MiRNA under salt stress shed light on salinity resistance improvement through MiRNA manipulation in crops. *Frontiers in Plant Science* 12:665439
33. Guan Y, Tanwar UK, Sobieszczuk-Nowicka E, Floryszak-Wieczorek J, Arasimowicz-Jelonek M. 2022. Comparative genomic analysis of the aldehyde dehydrogenase gene superfamily in *Arabidopsis thaliana* - searching for the functional key to hypoxia tolerance. *Frontiers in Plant Science* 13:1000024
34. Elbakary M, Hammad SF, Youseif SH, Soliman HSM. 2024. Revealing the diversity of Jojoba-associated fungi using amplicon metagenome approach and assessing the in vitro biocontrol activity of its cultivable community. *World Journal of Microbiology and Biotechnology* 40:205
35. Kirch HH, Bartels D, Wei Y, Schnable PS, Wood AJ. 2004. The ALDH gene superfamily of *Arabidopsis*. *Trends in Plant Science* 9:371–77
36. Gao C, Han B. 2009. Evolutionary and expression study of the aldehyde dehydrogenase (ALDH) gene superfamily in rice (*Oryza sativa*). *Gene* 431:86–94
37. Bhuya AR, Shuvo MRK, Al Nahid A, Ghosh A. 2025. Genome-wide identification, classification, and expression profiling of the aldehyde dehydrogenase gene family in pepper. *Plant Physiology and Biochemistry* 219:109413
38. Yang D, Chen H, Zhang Y, Wang Y, Zhai Y, et al. 2024. Genome-wide identification and expression analysis of the melon aldehyde dehydrogenase (ALDH) gene family in response to abiotic and biotic stresses. *Plants* 13:2939
39. Li MH, Liu KW, Li Z, Lu HC, Ye QL, et al. 2022. Genomes of leafy and leafless *Platanthera* orchids illuminate the evolution of mycoheterotrophy. *Nature Plants* 8:373–88
40. Kamarajan P, Rajendiran TM, Kinchen J, Bermúdez M, Danciu T, et al. 2017. Head and Neck squamous cell carcinoma metabolism draws on glutaminolysis, and stemness is specifically regulated by glutaminolysis via aldehyde dehydrogenase. *Journal of Proteome Research* 16:1315–26
41. Cao S, Ye Y, Zheng Z, Zhong S, Wang Y, et al. 2024. *Aux/IAA* gene family identification and analysis reveals roles in flower opening and abiotic stress response in *Osmanthus fragrans*. *Ornamental Plant Research* 4:e027
42. Zhao Y, Zhang Y, Li T, Ni C, Bai X, et al. 2022. *TaNF-YA7-5B*, a gene encoding nuclear factor Y (NF–Y) subunit A in *Triticum aestivum*, confers plant tolerance to PEG-inducing dehydration simulating drought through modulating osmotic stress-associated physiological processes. *Plant Physiology and Biochemistry* 188:81–96
43. Gao J, Lan T. 2016. Functional characterization of the late embryogenesis abundant (LEA) protein gene family from *Pinus tabuliformis* (Pinaceae) in *Escherichia coli*. *Scientific Reports* 6:19467
44. Sinha S, Mishra M. 2022. Polyamines: metabolism, regulation, and functions in crop abiotic stress tolerance. In *Augmenting Crop Productivity in Stress Environment*, eds Ansari SA, Ansari MI, Husen A. Singapore: Springer. pp. 317–44 doi: 10.1007/978-981-16-6361-1_19
45. Wang P, Wang CM, Gao L, Cui YN, Yang HL, et al. 2020. Aliphatic suberin confers salt tolerance to *Arabidopsis* by limiting Na^+ influx, K^+ efflux and water backflow. *Plant and Soil* 448:603–20
46. Liepman AH, Olsen LJ. 2001. Peroxisomal alanine : glyoxylate aminotransferase (AGT1) is a photorespiratory enzyme with multiple substrates in *Arabidopsis thaliana*. *The Plant Journal* 25:487–98
47. Li M, Yu B. 2021. Recent advances in the regulation of plant miRNA biogenesis. *RNA Biology* 18:2087–96
48. Yu Y, Ni Z, Wang Y, Wan H, Hu Z, et al. 2019. Overexpression of soybean *miR169c* confers increased drought stress sensitivity in transgenic *Arabidopsis thaliana*. *Plant Science* 285:68–78
49. Wang Y, Zhang F, Cui W, Chen K, Zhao R, et al. 2019. The FvPHR1 transcription factor control phosphate homeostasis by transcriptionally regulating miR399a in woodland strawberry. *Plant Science* 280:258–68