

Genomic Identification and Expression Analysis of Acid Invertase (AINV) Gene Family in *Dendrobium officinale* Kimura et Migo

Yujia Liu

Shaoguan University

Boting Liu

Shaoguan University

Kefa Luo

Shaoguan University

Baiyin Yu

Shaoguan University

Xiang Li

Shaoguan University

Jian Zeng

Shaoguan University

Jie Chen

Shaoguan University

Rui Xia

South China Agricultural University

Jing Xu

South China Agricultural University

Yuanlong Liu (✉ liuyuanlong@scau.edu.cn)

South China Agricultural University

Research Article

Keywords: *Dendrobium officinale*, Acid invertase, Genomic identification, Expression analysis, Abiotic stress, protein-protein interaction

Posted Date: April 17th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-2780110/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background *Dendrobium officinale* Kimura et Migo, a traditional Chinese orchid herb famous for its great horticultural and medicinal value, grows in an adverse habitat and suffers from abiotic or biotic stresses. Acid invertases (AINVs) are widely considered enzymes involved in regulating sucrose metabolism and here revealed to participate in responses to environmental stress in plants. The AINV gene family members have been identified and characterized in multiple plant genomes. However, detailed information about this gene family and their expression patterns are unknown in *D. officinale* despite their significance in polysaccharide biosynthesis.

Results This study systematically analyzed the *D. officinale* genome and identified four *DoAINV* genes, classified into two subfamilies based on subcellular prediction and phylogenetic analysis. Comparison of gene structures and conserved motifs in *DoAINV* genes showed high conservation during their evolution. The conserved amino acids and domains of *DoAINV* proteins were characterized as pivotal for their function. Abiotic and biotic stress responses cis-elements were the most abundant motif in all *DoAINV* genes, indicating they were stress-responsive. Furthermore, bioinformatics analysis of transcriptome data, validated by quantitative real-time reverse transcription PCR (qRT-PCR), revealed distinct organ-specific expression patterns of *DoAINV* genes in eight tissues and in response to abiotic stresses. Soluble sugar content and interaction network were examined to provide information on stress release and sucrose metabolism.

Conclusions *DoAINV* genes are involved in various activities such as growth and development, stress response, and polysaccharide biosynthesis. The results provide valuable insights into the AINV family and will be helpful in further investigating *DoAINV* functions in *D. officinale*.

Background

Dendrobium officinale Kimura et Migo, also known as *Dendrobium catenatum*, is an endangered, perennial herbal plant of the Orchidaceae family endemic to China, with great ornamental and medicinal values [1–3]. As an epiphytic plant, wild *D. officinale* usually grows in adverse conditions, e.g., epiphytic on cliffs or tree trunks, and distributed at altitudes above 1,200 m. In nature, *D. officinale* prefers to grow in a warm and humid environment, while the reality is inevitably suffering abiotic stresses such as low temperature and drought, slowing growth, and lowering yield [4, 5]. However, even in such harsh habitats, *D. officinale* can grow well and accumulate important medicinal substances due to their evolved stress tolerance system. Moreover, *D. officinale* ranks first among the nine Chinese herbs for longevity and has had a documented use in folk medicine for over 1300 years [6]. In recent decades, bioactive ingredients, including polysaccharides, bibenzyls, alkaloids, and flavones, have been identified and isolated from *D. officinale* plants [7, 8]. Among them, polysaccharides in stems of *D. officinale* are the major and important active ingredient, with anti-oxidant [9], antitumor [1], and anti-inflammatory effects [10], supporting immune modulation [11], hepato protection [12], and hypoglycemia [13], resulting in their high commercial values as traditional medicines [14, 15].

Sucrose, the direct product of photosynthesis and the major form of carbohydrate transportation, is an essential element of the life cycle in higher plants [16]. Polysaccharides synthesis and metabolism involve multiple protein or enzyme genes. In stems of *D. officinale*, polysaccharides are produced by the hydrolysis of sucrose transported from leaves [17]. The sucrose hydrolysis depends on two critical gene families of sucrose synthase (SUS, EC 2.4.1.13) and invertase (INV, EC 3.2.1.26) [18]. SUS catalyzes the reversible reaction of sucrose and uridine diphosphate (UDP) into uridine diphosphate glucose (UDP-glucose) and fructose, while INV irreversibly hydrolyzes sucrose into glucose and fructose. Although SUS activity is pivotal in many important metabolic pathways, INV plays a more significant role than SUS in sucrose. Prior studies on *D. officinale* polysaccharides mainly concentrated on chemical composition and pharmacological activity. However, polysaccharides biosynthesis at the molecular level in *D. officinale* remains unclear, especially related to gene function and transcription regulation. As an essential enzyme for polysaccharides synthesis, INV promotes the accumulation of polysaccharides through sucrose hydrolysis, which moderates various biological processes such as growth and development [19, 20], stress resistance [21], signal transduction, and the biosynthesis of the secondary metabolite [22, 23].

Previous studies revealed that a multigene family in plants encoded INV proteins. According to the optimal pH value, solubility, and subcellular localization, INVs are mainly divided into three types: cytoplasmic invertase (CINV), insoluble cell wall invertase (CWINV), and soluble vacuolar invertase (VINV) [24]. CINV belongs to the neutral/alkaline Invertase (NINV). In contrast, CWINV and VINV pertain to AINV, also named β -fructofuranosidase, because of its ability to hydrolyze sucrose and other β -fructose oligosaccharides [25, 26]. Moreover, AINV can hydrolyze fructose-containing compounds aside from sucrose, such as raffinose and stachyose, which are strongly inhibited by heavy metals [24]. The first three AINV proteins isolated from carrot (*Daucus carota*) were encoded by different genes. In recent years, as more plant genomes have been assembled and reported, the AINV gene families were identified across many plant species, such as *Arabidopsis thaliana* [27], rice (*Oryza sativa*) [28], sweet sorghum (*Sorghum Bicolor*) [29], and poplar (*Populus trichocarpa*) [30].

Prior reports suggested that the number of genes encoding AINV proteins varies in different species, and they are expressed independently at specific stages and tissues during growth and development [31, 32]. For instance, two divergent VINV isoforms, named *At β fruct3* and *At β fruct4*, displayed distinct expression patterns during organ development in different organs of *A. thaliana* [27]. The two VINV genes also showed differential expression in carrot, *S1* was mainly expressed in primary roots, and *SII* was only expressed and had a central role in the development of root tips [33]. *GhVIN1* (*Gossypium hirsutum*) specific expression in the cotton seed coat is critical for pollination success and paired male and female fertility [32]. In other plant species, such as the tea plant (*Camellia sinensis*) [34], tomato (*Solanum lycopersicum*) [35], pepper (*Capsicum annuum*) [36], cassava (*Manihot esculenta*) [37], and potato (*Solanum tuberosum*) [38], AINV genes also have have also tissue-specific expression patterns, suggesting that they may have involved specialized physiological functions. Overall, AINV genes have been extensively explored in various species; however, there is little knowledge of these genes in *D. officinale*.

In addition, recent studies have uncovered the involvement of AINVs in responses to environmental stress [27, 34]. For example, *StVIN1* was induced by low temperatures in mature potato tubers and played an important role in saccharification [38]. The three *AINV* genes in poplar, *PtVINV1/2* and *PtVINV3*, were up-regulated under low-temperature treatment, while *PtVINV1/2* was up-regulated under NaCl treatment [30]. Under drought stress, the *AINV* gene expression in maize (*Zea mays*) ovary was decreased. In contrast, the expression could be restored when sucrose was supplemented, suggesting that the *AINV* played an important role in response to drought stress. Although there has been much research on *AINVs* in various plants, a comprehensive investigation of the *AINV* gene family in *D. officinale* has not yet been performed.

Identifying and characterizing the *AINV* genes is the first step toward detecting their specific roles involved in different metabolic pathways. In the present study, we employed bioinformatics methods to analyze the characters of *DoAINV* genes on a genome-wide scale in *D. officinale* based on publicly available data[6, 39]. Furthermore, we investigated the expression profiles of the *AINV* genes in different tissues and their response to several abiotic stresses. These comprehensive results will facilitate a better understanding of the potential functions of *AINV* enzymes in sucrose metabolism or polysaccharides accumulation of *Dendrobium* plants in further research.

Results

Identification of *AINV* gene family in *D. officinale*

To identify *DoAINV* genes, the Hidden Markov Model (HMM) was constructed by searching the sequences through glycosyl hydrolases family domains (PF00251 and PF08244), which was used to search against the *D. officinale* genome, whereas *A. thaliana* and rice *AINV* sequences were regarded as queries in the BLAST program. Furthermore, we confirmed the resulting sequences with an online CDD-search tool; four putative *AINV* members in the *D. officinale* genome were identified by removing redundant sequences. After manual reannotation and confirmation of the protein characteristic domain, the four *DoAINV* genes were designated *DoVIN1* and *DoVIN2*, belonging to the VINV sub-family, and *DoCWIN1* and *DoCWIN2*, belonging to CWINV sub-family, following the nomenclature proposed in a previous study [28]. The basic characteristics information of *DoAINV* members is shown in Table 1. The size of the *DoAINV* genes varied from 3,345 bp to 13,803 bp, but the coding DNA sequences (CDS) were similar, between 1,725 bp and 1,983 bp. Molecular analysis of the full-length deduced polypeptides indicated that the putative proteins of *DoAINVs* ranged from 574 to 660 amino acid residues with relative molecular weights and isoelectric points varied between 65.86 to 72.74 kDa, and 5.54 to 9.21 kDa, respectively. This range of variability implies that different *D. officinale* *AINV* proteins might operate in different microenvironments.

Table 1 Characteristics of *DoAINV* gene family members

Protein Type	Protein Symbol	Gene Name	Gene Symbol	Genomic DNA Size (bp)	CDS Size (bp)	Number of Amino Acids	Predicted Mw (kDa)	Theoretical pI	GRAVY	II	TMH	Localization	Functional Domains (Start-End)	
													a (N)	b (C)
Vacuolar Invertase	XP_020703196.1	DoVIN1	LOC110114607	5527	1887	628	69.84	6.09	-0.184	38.33	1	Vacuole	97-421	424-617
	XP_020677523.2	DoVIN2	LOC110096097	5604	1983	660	72.74	5.54	-0.165	35.55	1	Vacuole	123-448	451-643
Cell Wall Invertase	XP_020686802.1	DoCWIN1	LOC110102702	3345	1725	574	66.40	5.85	-0.339	41.59	0	Cell wall	50-369	372-567
	XP_020678379.2	DoCWIN2	LOC110096666	13803	1758	585	65.86	9.21	-0.507	35.81	0	Cell wall	59-377	380-576

Sequence comparison between DoAINV s revealed that the genes share a high sequence homology at the nucleotide level (47.09–60.93% identity) within the coding region and at the amino acid level (37.02–65.76% identity) (Table 2). Out of these, gene pairs *DoVIN1* and *DoVIN2*, and *DoCWIN1* and *DoCWIN2*, scored a higher sequence identity of nucleotide and amino acids level, respectively, indicating that DoAINV s were divided into two evolutionary sub-families. Furthermore, two conserved domains, glycosyl hydrolases family 32 N-terminal and 32 C-terminal domains, cataloged as typical plant AINV domains were also detected in all the DoAINV s and *A. thaliana* AINV s (Fig. 1). Twelve AINV s of *D. officinale* and *A. thaliana* contain the catalytic motifs NDPNG, RDP, and WECXDF. These findings above suggested that the four *DoAINV* s encode different AINV isozymes.

Table 2
Analysis of homology between DoAINV nucleotide and amino acid sequences

		Identity of amino acid sequences (%)			
		DoVIN1	DoVIN2	DoCWIN1	DoCWIN2
Identity of nucleotide sequences(%)	DoVIN1	-	65.76	39.48	42.14
	DoVIN2	60.93	-	37.02	40.75
	DoCWIN1	50.84	48.59	-	52.03
	DoCWIN2	50.94	47.09	58.90	-

To further illustrate the evolutionary relationship between AINV gene families from *D. officinale* and other species, we aligned a 103 full-length amino acid sequences from 14 plant species (comprising 2 early terrestrial plants, 2 base angiosperms, 5 eudicots, and 5 monocots, Table S1) and constructed an unrooted phylogenetic tree (Fig. 2A). All 103 AINV proteins were clustered into two major subfamilies according to the phylogenetic tree designated as VINV and CWINV, similar to the previous classification of the AINV family in rice and poplar [28, 30]. These groups could be further classified into early terraneous, base angiospermous, eudicotyledonous, and monocotyledonous sub-groups (Fig. 2B). This result

suggested that the evolutionary divergence of the *A/INV* genes could have occurred before the differentiation of the dicot and monocot ancestor. DoVIN1 and DoVIN2 proteins clustered in the VINV group, and DoCWIN1 and DoCWIN2 proteins clustered in the CWINV group.

Gene structure and conserved motifs of *D. officinale* AINVs

The distributions of protein-conserved motifs and gene intron-exons of *D. officinale*, *A. thaliana*, and rice genomes were comparatively analyzed to investigate the structural features and explore the evolutionary mechanisms of AINV gene families. Ten distinct motifs (motifs 1–10) were identified in three plants' AINV protein sequences, and the longest conserved motif was 50 amino acids (Table 3). Although most motifs classified as AINVs were similar, their distribution directions differed significantly between species. Moreover, the conserved motifs located within the N-terminal region in CWINVs were longer than VINVs (Fig. 3). The results suggested that the ancestral AINV genes are common to monocots and dicots in these positions.

Table 3 Putative conserved amino acid sequences of AINV motifs of *D. officinale* and *A. thaliana*

Motif	Conserved amino acid sequences with best possible match	width	E-value	Protein domain
1	GPMFYKGIYHLEFYQYNENGAVWGNIVW	27	9.7e-174	Glyco_hydro_32_N
2	LRFDYGNFYASKTFFDYKKNRRILWGWANESDTVEDDILKG WAGLQVI PR	50	6.4e-316	Glyco_hydro_32_N
3	GHAVSKOLINWLHLP IAMEPDRWYDINGVWSGSATLLHDGR LVMLYTGST	50	3.0e-284	Glyco_hydro_32_N
4	GEKLTMRILVDH SIVEAFGQGGRTCITSRVYPTTAIYGA	39	8.0e-232	Glyco_hydro_32_C
5	RRGIAYIYRSRDFKHVWKAKHPVH SKQSTGMWEC PDFFPVS	41	5.3e-204	Glyco_hydro_32_C
6	MSG LGPFGLMVLASNDME EYTSVYFRIFK	29	1.8e-134	Glyco_hydro_32_C
7	RQVQNVAF PKNLS DPYLR EWVKPDY N PVI	29	3.0e-133	None predicted
8	GERIKH I LKASLDDDRH DYAI GRYDANENKWPDDAEKDV	41	1.4e-136	Glyco_hydro_32_N
9	WQRTAFH FQPEKNWMNDPN	19	1.8e-124	None predicted
10	KQLVFWPIEEEIESLRGKNVQM	21	1.4e-085	None predicted

The exon-intron arrangement is regarded as an important parameter in gene phylogenies [40]. The *A/INV* gene structure in *D. officinale*, *A. thaliana*, and rice revealed that the number of introns varied from four to six. The number of exons ranged from five to seven, excluding a member of rice (Os02g01590.1, Fig. 3C). Most *A/INV* genes, except *DoCWINV2* and *Os02g01590.1*, included a mini exon composed of nine nucleotides. The mini exon encoding aspartic acid, proline, and asparagine, ranked second and was located in the β -fructosidase motif (NDPNG). Additionally, all *A/INV* genes contained a typical Glyco_32 domain belonging to the glycosyl hydrolases family 32. Only VINV genes held an extra domain of unknown function (DUF3357). The phylogenetic analysis between four *DoA/INVs* showed that they could

be classified into two distinct groups (Fig. 3A). This scenario was verified with the pattern observed on the phylogenetic tree between AINV sequences (Fig. 2A). These findings support the specific evolutionary features of AINV gene family between plant species.

Prediction of the protein structure of *D. officinale* AINV proteins

The secondary structure analysis showed that four DoAINV proteins comprised α -helices, extended β strands, β -turns, and random coils (Fig. S1, Table 4). The random coils were the major secondary structures among the DoAINV proteins, accounting for 51.62 to 55.41%, followed by extended β strands (23.76 to 26.88%). The α -helices were lower, ranging from 13.38 to 18.97%, implying that the DoAINV proteins were extremely unstable and easy to be degraded (Table 4). These secondary structure distributions were also highly conserved in four DoAINV polypeptide chains (Fig. S1).

Table 4
Secondary structural statistics of DoAINV proteins

Protein	Alpha helix (%)	Extended Beta strand (%)	Beta turn (%)	Random coil (%)
DoVIN1	13.38	25.8	5.41	55.41
DoVIN2	17.42	24.09	4.7	53.79
DoCWIN1	15.01	26.88	5.93	52.18
DoCWIN2	18.97	23.76	5.64	51.62

Structures of the four DoAINV proteins were predicted using the Swiss-model online software (Fig. 4), and the tertiary dimensional models (3D) of the DoAINVs were based on templates 3ugf.1.A (6-fructosyltransferase) and 2ac1.1.A (invertase). The tertiary structure of the four DoAINVs has only one monomer, consisting of a single main polypeptide chain. The 3D structures DoVIN1 and DoVIN2 were similar, and DoCWIN1 was closer to DoCWIN2.

Prediction of transmembrane structure (Fig. S2) showed that DoVIN1 and DoVIN2 protein sequences included a high-reliability hydrophobic region at the N-terminal; however, DoCWIN1 and DoCWIN2 did not (Fig. 1). However, the SignalP-5.0 prediction result showed that DoCWIN1 and DoCWIN2 contained one signal peptide composed of 19 ~ 27 amino acids and 34 amino acids, respectively, with a reliability of 0.5201 and 0.939, is no signal peptide was detected in DoVIN1 and DoVIN2 proteins (Fig. 1 and Fig. S3). This implies that DoVIN1 and DoVIN2 might be membrane proteins in the endoplasmic reticulum, while DoCWIN1 and DoCWIN2 might act as secreted proteins. Furthermore, phosphorylation is important in protein activity, function, and intracellular signal transduction. Prediction of the phosphorylation sites in DoAINV proteins showed that serine was the most common site for phosphorylation in all DoAINVs, followed by threonine and tryptophan phosphorylation sites (Fig. S4).

Promoter analysis of *D. officinale* AINV genes

The *cis*-acting elements play a vital role in regulatory networks controlling plant growth and development, and determining the spatial-temporal and tissue-specific expression of genes. Using the PlantCARE database, the *cis*-acting elements of *DoAINV* genes were classified into three categories, abiotic and biotic stress responses, phytohormone responses, plant growth and development in the promoter regions (Fig. 5). Notably, for all *cis*-acting elements, the abiotic and biotic stress classes were the most abundant, followed by the growth and development and phytohormone response classes, even in the *DoAINV* promoters (Fig. 5A, B). This implies that *AINV* genes might respond to abiotic stresses and have the potential to improve abiotic stress responses.

In the abiotic and biotic stress responses category, various stresses-related elements were observed, including ARE (anaerobic induction), MBS (drought induction), LTR (low-temperature responses), TC-rich repeats (defense and stress responses), WUN-motif (wound responses), and other *cis*-acting elements (Fig. 5C). In the phytohormone responsive category, the AuxRR-core and TGA-element required for auxin response, the ABRE and AAGAA-motif for the abscisic acid response, and ERE for ethylene response were identified. TCA-element associated with salicylic acid-responsive genes were identified in all *DoAINV* promoters. Notably, the most common motif was the TGACG-motif and CGTCA-motif *cis*-acting elements, related to MeJA (Methyl Jasmonate) responsiveness, accounting for 28% of the scanned hormone-responsive motifs. Unusually, no *cis*-acting element involved in gibberellin response was identified. In the last category, *cis*-acting elements were found sparsely in the promoter regions, including GCN4_motif required for endosperm expression, CAT-box for meristem expression, Circadian for circadian control, MRE and Box-4 for light responsiveness. Box-4 elements comprised the largest portion, accounting for 70% of all elements. These results imply that *DoAINV* genes may play multiple roles in plant growth and development.

The organ-specific expression patterns of *AINV* genes in *D. officinale*

To further understand the potential function of *AINV* gene family members in *D. officinale*, during development, the expression profiles of *DoAINV* genes were performed by reanalyzing the RNA-seq data from eight tissues: flower buds, sepal, gynostemium, labellum, leaf, stem, white root (equivalent to mature zone), and green root tip [39]. Based on the expression pattern, the four *DoAINV* genes exhibited distinct organ-specific expression (Fig. 6A). In the *VINV* group, the *DoVIN1* gene was highly expressed in sepal, gynostemium, labellum, and white root, while lower expressed in the green root tip, stem, leaf, and flower buds. *DoVIN2* was highly expressed in all examined reproductive organs, including the sepal, gynostemium, flower buds, and labellum, and in the stem and white root. The flower buds showed the highest expression of the *DoCWIN1* gene in the *CWINV* group, followed by white root. The *DoCWIN2* transcript had the highest abundance in leaves and gynostemium and was moderately expressed in flower buds and roots. These results suggested that the *DoAINV* genes might be indirectly or directly involved in developing and forming reproductive organs. In particular, each *DoAINV* family gene is highly expressed in reproductive organs and other tissues, such as leaves, stems, and roots. (Fig. 6B). The expression patterns for the *DoAINV* genes were also evaluated and confirmed with qRT-PCR in selected organs of the *D. officinale* DanXia cultivar, including flower, leaf, stem, and root (Fig. 6C). This implies a

rich diversity of tissue expression between genes of the same and different *DoAINV* clades, which may play redundant, complementary, or alternating roles in plant growth and development.

Regulation of the expression of *D. officinale* AINV genes by abiotic stresses

To investigate the expression patterns of *DoAINV* genes in response to abiotic stress-related stimuli, the transcript levels of *AINV* genes in the *D. officinale* 'DanXia' cultivar under cold, dehydration, and abscisic acid (ABA) treatments were analyzed. The four *DoAINV* genes exhibited different expression patterns in response to different treatments. Interestingly, under cold treatment, all of them were all up-regulated induced. Additionally, the *DoVIN1* gene was up-regulated in all stress treatments reaching the highest expression level at 6 h or 9 h. *DoVIN2* and *DoCWIN1* were almost unaffected during dehydration and ABA treatments, maintaining a high expression level relative to other two genes. At the same time, the *DoCWIN2* gene showed apparent changes in the patterns of expression levels (Fig. 7A). The data presented confirms the putative involvement of *DoAINV* genes in abiotic stress responses, especially in cold resistance.

To further determine whether the *AINV* genes were involved in abiotic stress resistance and the reliability of the transcriptome data, qRT-PCR was performed to measure the expression patterns and levels of *AINV* family genes in the *D. officinale* 'DanXia' cultivar, also treated with cold, dehydration, and ABA stresses. Figures 7B to 7D show that *DoVIN1* and *DoCWIN2* transcripts abundance increased in all three treatments. In contrast, *DoVIN2* and *DoCWIN1* significantly increase in cold treatment and are down-regulated or unaffected in other treatments. Under cold stress, the four *DoAINV* genes were up-regulated, peaking at 6 h, then gradually decreasing to untreated levels at 24 h of treatment (Fig. 7B). Among them, *DoCWIN1* responded particularly strong, with transcript levels increasing almost 600-fold at 6 h of cold stress, indicating that the *DoCWIN1* gene might play a vital role in cold resistance for *D. officinale*. In dehydration treatment, the *DoVIN1* and *DoCWIN2* expression levels were significantly up-regulated, 15-fold and 6-fold at 9 h, respectively (Fig. 7C). Although transcript levels varied among the family members, *DoAINV* genes exhibited a similar expression pattern among dehydration and ABA stress. For instance, *DoVIN1* and *DoCWIN2* transcript levels increased sharply at 9 h (over 200-fold and 1000-fold, respectively) after ABA treatment and then rapidly decreased (Fig. 7D). Similarly, *DoVIN2* and *DoCWIN1* shared similar expression patterns with dehydration stress and ABA treatment. This implies that *DoVIN1* and *DoCWIN2* genes, might affect the dehydration and ABA stress responses of *D. officinale*.

In addition, soluble sugar content in stems of *D. officinale* 'DanXia' cultivar under cold, dehydration, and ABA treatments revealed that the soluble sugar content was continuously increased during cold and dehydration treatments compared to untreated control (Fig. 8A and 8B). Interestingly, with ABA treatment, the dramatic expression of *DoVIN1* and *DoCWIN2* genes did not affect the soluble sugar content accumulation in *D. officinale* stems (Fig. 8C). In summary, these results illustrated that *DoAINV* genes expression varies in response to different abiotic stresses, indicating diverse roles of the *DoAINV* genes in response to abiotic stress.

Interaction network analysis of AINV proteins in *D. officinale*

To better understand the biological functions of DoAINVs in response to abiotic stress, possible protein-protein interaction (PPI) networks were further analyzed among DoAINV proteins and related proteins based on the potential interaction data for 1,363 differentially expressed genes (DEGs) with the new STRING 11.5 database. These DEGs were filtered by comparing RNA-seq data from cold stress treatment at 6 h and untreated control (0h). All genes harboring potential interaction relationships (Source-Target) are shown in Supplementary Table S2 in Supplementary Material. As shown in Fig. 9, a total of 147 nodes (genes) and 684 edges (interactions) were identified using Cytoscape software (Version 3.9.1).

The PPI network, involved in cold stress signal transduction, revealed that the DoAINV members had the strongest association with genes encoding calcium and calmodulin signaling network (CML14, CBP3, and A0A2I0W7M0) (Fig. 9). The second category of proteins that most closely interacted with DoAINVs were enzymes involved in carbohydrate synthesis and metabolism, including IRX9, Gluc18, XTH32, SPS1, SPS2, and UDP83A1. Furthermore, there were multiple receptor kinases in the network, including CDPK gene families related to Ser/Thr protein kinases (STN7, CIPK12, and ATR), heat shock proteins (HSP20 and HSP23), and signal recognition proteins (SRP14, CPRF2, and A0A2I0V753). Finally, also the most valuable, various kinds of transcription factors (TFs) were identified related to Adversity resistance and plant growth, which interact with DoAINV proteins, such as bHLH TFs (AIB, bHLH25, and bHLH35), MYB TFs (MYB46 and MYB86), ERF TFs (AP2/ERF, ERF1A, and ERF118) and NAC TFs (NAC10, NAC43, and NAC74). These results indicated that the AINV proteins of *D. officinale* have vital roles in multiple functions.

Discussion

Characteristic and evolutionary conservation among *D. officinale* AINV genes

Previous studies have shown that different species have different numbers of AINV family members, divided into two subgroups, namely CWINV and VINV. For instance, the *A. thaliana* AINV gene family consists of eight members, two VINVs and six CWINVs [41]; the rice AINV gene family contains nine members, two VINV and seven CWINV [28]; and the poplar AINV gene family has eight members, three VINVs and five CWINVs [30]. This study grouped all vacuolar INVs from base angiospermous, eudicotyledonous, and monocotyledonous in the VINV group based on the phylogenetic analysis of AINVs from fourteen plant species. The VINV group was distinctly separated from the cell-wall INV, grouped as the CWINV group; however, the AINVs from early terrestrial plants were not divided into VINV or CWINV subgroups, and remained clustered on the same branch (Fig. 2). The results suggested that the origin of VINVs from CWINVs probably occurred in the early period of angiosperm formation. This finding broadened the previous deduction that the separation of AINVs predated the last common ancestor of dicotyledons and monocotyledons [42]. In the phylogenetic tree of AINVs from fourteen plant species, the VINV groups were divided into four clear categories: early terraneanous, base angiospermous, eudicotyledonous, and monocotyledonous. Furthermore, the two VINVs of *D. officinale*, DoVIN1 and DoVIN2, were distributed in the monocotyledonous category. In the CWINV group, cell-wall INVs for fourteen plants also showed a similar evolutionary relationships, distributed in the four non-contiguous

categories, following different branches. These results indicated that forerunners of AINVs in the CWINV group were more divergent than those in the VINV group. The two CWINVs of *D. officinale* were distributed in two lineages of monocotyledonous plants, respectively, and the ancestor of DoCWIN1 originated older than DoCWIN2. Interestingly, only two CWINVs were identified from *D. officinale*, significantly different from other species in number, indicating no potential duplication events or gene functional redundancy in DoAINVs. This implies that the loss of CWINV genes occurred during the evolution of *D. officinale* to harsh environments such as cliffs and cold and dry conditions. In these cases, it may lead to the inability of *D. officinale* to form fruits, just like rice, potato, soybean (*Glycine max*) and other crops, with carbon sources and high sugars in the whole development process.

We also estimated the exon/intron structures of *DoAINV* genes to further understand the conservation and evolutionary relationship of *AINV* genes between dicot and monocot plants (Fig. 3). The integrated gene structure model in *DoAINVs* was very similar to that of *AINV* genes from *A. thaliana* and rice. The typical feature of the AINV gene family in plants is a mini second exon, located in the β -fructokinase motif (NDPNG), which only encoded three amino acids [43]. In *D. officinale*, other three DoAINVs contain the conserved mini exon, except *DoCWIN2*, suggesting *DoCWIN2* underwent exons-loss during evolution. This speculation, not a special case, was consistent with *Os02g01590.1* of CWINV in rice; *PtrCWINV1* and *PtrCWINV2* in poplar [30]; *ShCWINV3*, *ShCWINV8-1*, *ShCWINV9-1*, and *ShCWINV9-2* in sugarcane (*Saccharum officinarum*) [42]. Interestingly, the *DoCWIN2* with the longest DNA sequence, even longer than any *AINV* genes in *A. thaliana* and rice, contained only 5 exons. While motif comparison showed that the main motifs were kept in all the *DoAINV* genes, suggesting the gene structure variation was caused by exon splitting or intron length variation but not pseudo-exonization. These findings further supported the conclusions of Wang et al. [42]. Additionally, DoAINV proteins had very similar tertiary structural models, the four DoAINVs can form a β -propeller module at the N-terminal domain and a β -sandwich module at the C-terminal domain (Fig. 4), consistent with reported AINVs in other plants [44, 45]. Overall, although there were differences in genome length and exon/intron structure among *DoAINV* genes, they all had the typical conservative motifs and tertiary dimensional structure of AINVs, which confirmed the conclusion that the gene organization was highly conserved within the plant AINV gene family.

Sophisticated roles of AINV genes in *D. officinale*

The *AINV* gene family, as a key enzyme of sucrose metabolism regulation, contains varied members in different species, and their expression patterns were tissue-specific in different organs and developmental stages. Previous studies have shown that AINV proteins play an important role in life progression, such as plant morphogenesis, growth, and development [4]. It is widely accepted that analysis of gene expression patterns in different tissues is a reliable approach to studying the molecular functions of genes in various physiological processes. In the current study, our data revealed that *DoAINV* genes displayed tissue-specific expression patterns in *D. officinale*. For instance, several *cis*-acting elements associated with flowering were identified in *DoAINV* promoter regions, such as GCN4_motif required for endosperm expression, CAT-box for meristem expression, and circadian for circadian control elements (Fig. 5). The expression analysis revealed that all *DoAINV* genes were highly expressed in specific floral

organs of *D. officinalis*, indicating that *DoAINV* genes might have an important role in the formation and development of reproductive organs by regulating sucrose metabolism (Fig. 6A and 6B). *DoVIN1* was a highly expressed *AINV* gene in the white root, while *DoCWIN1* was the highest expressed *CWINV* gene in the tissue. This result suggested that these two genes may play an important physiological function in absorbing water and nutrients. By contrast, *DoCWIN2* had the highest transcript abundance of the *CWINV* gene in the leaf, indicating that the *DoCWIN2* gene plays a key role in phloem unloading of source organs; At the same time, *DoVIN2* was highly expressed *VINV* gene in stem, implying that the *DoVIN2* gene plays a vital role in vegetative organ to regulate sucrose metabolism and sugar accumulation. These results suggest that the functions of *DoAINV* genes are diversified almost without functional redundancy, which further brought us to hypothesize a functional collaboration between the different *DoAINV* proteins of *D. officinalis* in the polysaccharide biosynthesis and metabolic pathways.

Besides plant development, *AINV* proteins have been verified to be associated with plant signal transduction and various environmental stresses response [46, 47]. In the present study, a variety of frequently occurring *cis*-acting elements was identified in promoter regions of *DoAINVs*, such as LTR, MBS, MYB, ARE, STRE, and ABRE (Fig. 5). We also noted that all *DoAINV* genes contained at least ten abiotic stress *cis*-acting elements, suggesting that they might contribute to responding to biotic and abiotic stresses. Under cold, dehydration, and ABA treatments, the expression levels of *DoVIN1* and *DoCWIN2* were up-regulated. Thus, the key enzymes of *DoVIN1* and *DoCWIN2* encoded might be the predominant isoforms to respond to abiotic stress. Simultaneously, the transcription levels of *DoVIN2* and *DoCWIN1* were significantly increased in cold treatment, but down-regulated or unaffected after dehydration and ABA treatments. A similar expression response of *AINV* genes was also found in sugarcane [42]. Drought stress and cold treatment conspicuously induced *ShCWINV3*, *ShCWINV7*, and *ShCWINV9*, suggesting those genes are important in the stress responses of sugarcane.

In higher plants, the polysaccharide accumulation and sucrose metabolism are always disturbed by growth conditions and exogenous hormones, resulting in changes in sugar contents and various physiological functions, to resist environmental stresses [48]. Thus, we examined the soluble sugar content under cold, dehydration, and ABA treatments for 0 h, 12 h, 24 h, and 48 h, respectively. The soluble sugar content of *D. officinale* was affected by cold and dehydration treatment, but not altered in response to ABA treatment (Fig. 8). The observed increase of soluble sugar content in *D. officinale* stems in response to cold stress and dehydration treatment is likely due to up-regulation of expression of *AINV* genes to improve their protein activity and subsequently cleaved more sucrose into glucose and/or fructose to sustain osmotic pressure of cells and enhance resistance to environmental stresses [49, 50]. It is also possible that the raised *AINVs* levels are necessary to effectively increase the ability to stabilize the osmotic pressure of cells, whilst the down-regulated expression of *AINV* genes might be due to a requirement for receiving and interconnecting with sugar modulation signals blocking the downstream metabolism to adapt to the dehydration stress. Thus, we speculated there might be the difference in the molecular response mechanism between cold stress and dehydration stress (Fig. 7). Noteworthy, cold and dehydration may regulate gene expression, although they decrease the actual protein activity in the plant. In our study, the soluble sugar content of *D. officinale* during cold stress and dehydration treatment

was continuously increased compared with untreated control (Fig. 8). The increase in AINV protein may occur to enhance actual protein activity under cold and dehydration stresses.

In addition, ABA, an important phytohormone, *AINV* gene expression levels in green bamboo (*Bambusa oldhamii*) [51]. Similar results were revealed in our study; after ABA treatment, two of the DoAINV genes (*DoVIN1* and *DoCWIN2*) were significantly up-regulated, and the other two genes (*DoVIN2* and *DoCWIN1*) were down-regulated (Fig. 7D). Their expression patterns were consistent with dehydration stress (Fig. 7C), indicating that *DoVIN1* and *DoCWIN2* may provide the largest contribution to the activity of the entire AINV gene family in response to dehydration stress and ABA treatment. Interestingly, inconsistent with the result of soluble sugar content under dehydration treatment, the increased transcript abundance of *DoVIN1* and *DoCWIN2* genes did not particularly significantly alter the soluble sugar content of *D. officinale* during ABA treatment (Fig. 8C). The finding could explain this result that the AINVs may function as regulatory enzymes involved in sucrose degradation and maintenance of sugar concentration, in addition to acting as signaling molecules to modulate different signal transduction pathways, as well as protecting cellular structural integrity from damage.

These results indicated that *AINV* genes had various functions in *D. officinale*, and also provided evidence that plant AINV members could participate in abiotic stress responses. The gene expression patterns, soluble sugar content changes, and PPI network under abiotic stress could help further understand the interactive regulatory network of AINV genes and polysaccharide biosynthesis pathways.

Conclusions

In this study, four *D. officinale* AINV genes were bioinformatically identified and characterized. The *D. officinale* AINV family members shared two conserved domains and were classified into two subfamilies based on the cellular locations and phylogenetic analysis. Moreover, the DoAINV domains functional characteristics were identified. The expression of the four *DoAINV* genes was up-regulated under cold stress, in order to protect cellular structural integrity from damage. In comparison, under dehydration stress and ABA treatment, the expression of *DoVIN1* and *DoCWIN2* was also up-regulated to cope with the need to cleave more sucrose. Conversely, the expression of *DoVIN2* and *DoCWIN1* could be down-regulated to maintain sucrose homeostasis. Specifically, the expression of *DoAINV* genes enhanced the soluble sugar content of *D. officinale* in response to cold and dehydration stresses. Based on our experimental results, we speculated that DoAINVs play a significant role in the regulation of sucrose metabolism and polysaccharide accumulation, as well as serving as signaling molecules to modulate ABA signal transduction pathways. In total, our study provides informative insights on the *DoAINV* gene function of releasing stress and sucrose metabolism.

Methods

Plant materials, growth conditions and treatments

D. officinale used in this work were selected from cultivar 'DanXia', which were planted at the Engineering Technology Development Center of *Dendrobium* herb, Shaoguan, China. After surface disinfection of *D. officinale* capsules, the seeds were sowed in half-strength Murashige and Skoog (MS) medium (Sigma-Aldrich, St. Louis, MO, USA), with 20% potato flour, and sterile grown at a temperature of 25 ± 1 °C with a light/ dark cycle of 12/ 12 h and 60–70% relative humidity in a growth chamber at Shaoguan University (Shaoguan, China). The aseptic seedlings were subcultured every two months, until the stems were 6–8 cm in height, which were about six-month-old. Then young leaves were selected for different treatments. For each treatment, six leaves from six separate seedlings were mixed as one sample, and all the experiments were performed in triplicate.

For exogenous ABA treatment, seedlings were cultured in liquid half-strength MS supplemented with 100 μ M ABA for 2 days. For dehydration stress treatment, seedlings were cultured in liquid half-strength MS supplemented with 20% polyethylene glycol (PEG) 6000 for 2 days. Seedlings cultured in liquid half-strength MS were used as control. And for cold stress treatment, seedlings were moved to a growth chamber at a temperature of 4 ± 1 °C for 2 days. Untreated seedlings were used as controls. Then all the treated leaves were harvested at 0, 1, 3, 6, 9, 12, 24 and 48 h post-treatment, frozen in liquid nitrogen and stored at -80 °C for further expression analysis.

Moreover, the leaf, stem, root, and flower samples were collected from the *D. officinale* plants during the flowering stage which grown under normal conditions, and frozen in liquid nitrogen and stored at -80 °C for further expression analysis. Three biological replicates were collected for each sample.

Identification of the AINV family genes from *D. officinale* genome

The whole-genome sequence of *D. officinale* were downloaded from the National Centre for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/>). Then, the Hidden Markov Models (HMMs) profiles of the glycosyl hydrolases family 32 N-terminal domain (PF00251) and 32 C-terminal domain (PF08244), which were the conserved domain of AINV gene family, was downloaded from Pfam (<http://pfam.xfam.org/>, 35.0). And the HMMs were used to identify all AINV proteins from the *D. officinale* genome. Additionally, to avoid any nonspecific sequences outside the AINV cluster, eight AINVs of *A. thaliana*[26] and nine AINVs of rice [28] sequences were used as queries in the repeatedly BLASTs against to the *D. officinale* database via the default parameters. Finally, all putative DoAINV genes were manually verified with the Conserved Domains Database (<http://www.ncbi.nlm.nih.gov/cdd>) to confirm the completeness of all core domains.

Protein Properties And Sequence Analyses

The molecular weight and isoelectric points of the DoAINVs were predicted through the ExPASy database (<http://expasy.org/>). The motifs were identified using the MEME (Multiple Em for Motif Elicitation) software (<http://meme.nbcr.net/meme/cgi-bin/meme.cgi>) and were displayed through visualize motif pattern function of ToolKit Biologists Tools (TBtools) software (<https://github.com/CJ-Chen/TBtools>)

[52]. Motifs identified by MEME were further queried in the InterPro database (<http://www.ebi.ac.uk/interpro/>). The gene structure of *DoA/INV*s were carried out using Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/>, 2.0) by alignment of the complementary DNA (cDNA) sequences with their corresponding genomic DNA sequences. The protein domain was drafted using PROSITE biosoft (<https://prosite.expasy.org/cgi-bin/prosite/mydomains/>). Secondary and tertiary structures were predicted using SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) and the ExPaSy Swiss-Model online software (<http://swissmodel.expasy.org>), respectively. Transmembrane helix and signal peptide were predicted through TMHMM program (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>) and SignalP 5.0 Server (<https://services.healthtech.dtu.dk/service.php?SignalP-5.0>). Phosphorylation sites were predicted using NetPhos 3.1 Servers (<https://services.healthtech.dtu.dk/service.php?NetPhos-3.1>).

Cis -acting element and phylogenetic analyses

The promoter sequences, 2 kb upstream of the start codon, were acquired from *D. officinale* genomic sequences. The *cis*-acting elements of *DoA/INV* promoters were analyzed by PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). And the key promoter elements were categorized, which related to phytohormones, environmental response and adversity adaptation, were displayed through the Simple BioSequence Viewer function of TBtools. DNAMAN (version 6.0) were used to perform the multiple alignments of the identified *DoA/INV*s nucleotide and amino acid sequences. The phylogenetic tree of the different species *AINV* amino acids was constructed with MEGA 11 and Clustal X (version 2.0) using the neighbor-joining (NJ) algorithm and the bootstrap test carried out with 1 000 replicates [53, 54].

Transcriptomic Analysis

The sequencing procedures were completed by BGI Tech Co., Ltd. (Shenzhen, China), including RNA extraction, library construction, and sequencing. Illumina Hiseq™ 2500/Miseq™ (Illumina, USA) was used as the sequencing platform. Adapter sequences and low-quality sequences produced in sequencing were removed using the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) and FastQC program (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Expression levels were calculated and normalized as fragments per kilobase per million mapped fragments (FPKM). The DEGseq R package (version 1.12.0) was used to identify significantly DEGs (Log2-based fold-changes > 1; Log2-based fold-changes < - 1) in response to treatment.

In silico expression profiling of *D. officinale* *AINV* genes in different tissues

For the tissue expression profiling of *DoA/INV* genes, the raw RNA-seq reads of flower bud (SRR4431603), gynostemium (SRR4431596), labellum (SRR4431602), sepal (SRR4431597), leaf (SRR4431601), stem (SRR4431600), white part of root (SRR4431598) and green root tip (SRR4431599) were downloaded from the NCBI Sequence Read Archive (SRA) [39]. Reads of all the samples were aligned to the NCBI

Dendrobium reference genome through HISAT package [55]. FPKM for mRNA expression evaluation was calculated by StringTie [56]. The heat map was drawn through the eFP Browser function of TBtools.

Rna Extraction, Cdna Synthesis And Qrt-pcr

Total RNA was extracted from the leaves using MiniBEST Plant RNA Extraction Kit (9769, TaKaRa, Japan). First strand cDNA synthesis was performed using PrimeScript™ RT Master Mix (RR036A, TaKaRa, Japan). PCR reaction mixture (20 µL) was prepared according to the manual of TB Green® Premix Ex Taq™ (Tli RNaseH Plus) Kit (RR820A, TaKaRa, Japan) according to the manufacturer's instructions. The reaction was carried out using BIO-RAD CFX Connect™ Real-time System (BIORAD, USA). The *DoACTIN* gene was used as an internal control [57]. All reactions were performed in triplicate in each experiment and three biological repeats were conducted, and the primers used in the qRT-PCR experiments are listed in Table S3. The relative expression levels were calculated according to the Normalized Expression method ($2^{-\Delta\Delta CT}$ method) [58].

Measurement Of Total Soluble Sugar Content

Total soluble sugar was extracted from the *D. officinale* 'DanXia' cultivar seedlings by using water extraction and alcohol precipitation method. Then the soluble sugar content was measured by the phenol-sulfuric acid method, which using ultraviolet spectroscopy with anhydrous glucose as reference substance at 483nm wavelength. The content calculated through standard curve of anhydrous glucose, which with a good linearity ($r = 0.9993$), based on the regression equation was $y = 3.5696x + 0.0065$.

Protein-protein Interaction Network Construction And Visualization

The potential interaction data for DoAINVs and related proteins in cold stress treatment were extracted from the latest STRING v11.5 available database (<https://cn.string-db.org/>) [59]. The interactions between two genes was built using Cytoscape v3.9.1 software (<https://cytoscape.org/>) for visualization. The Cytoscape plugin CytoNCA was employed to screen the key targets [60].

Statistical Analysis

Differences between values were calculated using a one-way analysis of variance (ANOVA) with a student's t-test at a significance level of 0.05 in GraphPad Prism v9.5 (GraphPad Software, Inc., Chicago, USA, www.graphpad.com). All of the data analyses were performed for three biological repeats, and the values shown in the figures represent the average values of three repeats. Sample variability is given as the standard deviations (SDs) of the mean.

Declarations

Ethics approval and consent to participate

We declare that no wild plants were collected or used in this study. All experimental research was conducted in accordance with relevant institutional, national, and international guidelines and legislation. Permission was obtained for the collection of the *D. officinale* 'DanXia' cultivar from Prof. Baiyin Yu of the Engineering Technology Development Center of *Dendrobium* Herb in Shaoguan, China.

Consent for publication

Not applicable.

Availability of data and materials

The datasets supporting our conclusions of the current study are included in the manuscript and additional file. The genome sequence of *D. officinale* were downloaded from the NCBI database. The AINV protein sequences of *A. thaliana* and rice were acquired from The *Arabidopsis* Information Resource (TAIR, <http://www.arabidopsis.org/>) and the Rice Genome Annotation Project (RGAP, <http://rice.plantbiology.msu.edu/>) databases, respectively. The tissue expression profiling analysis used already published raw sequence data of *D. officinale* were obtained from NCBI-BioProject PRJNA348403.

The generated sequencing data in this study were deposited in NCBI's SRA database. The raw sequence data of *D. officinale* in response to abiotic stress are available in NCBI-BioProject database under the accession number PRJNA943399, PRJNA949800 and PRJNA949802, respectively.

Competing interests

The authors declare no conflict of interests.

Funding

This work was supported by grants from the National Science Foundation of China (No. 32000266); the Foundation for Basic and Applied Basic Research Project of Guangdong Province (No. 2020A1515011438); the Special Program of Key Fields for Ordinary University (No. 2022ZDZX4044); the Open Fund of the Guangdong Provincial Key Laboratory of Utilization and Conservation of Food and Medicinal Resources in Northern Region (No. FMR2022002Z) and the Key Projects of Shaoguan University (No. SZ2019ZK04).

Authors' contributions

YuL, BY and YL designed and coordinated the research. YuL, BL, KL and JX performed the main experiments. XL, JZ and JC participated in some of the experiments and data collection. YuL, BL, RX and YL analyzed all the data. YuL, BL and YL wrote and revised the manuscript. All authors reviewed and approved the manuscript.

Acknowledgements

The authors would like to thank all colleagues from our laboratory for providing the useful technical assistance. We are very grateful to “Professor Xia Rui's Research Group of South China Agricultural University” for assistance with the biological data analyses. We also sincerely thank the editor and reviewers for critically evaluating this manuscript and providing constructive comments for its improvement.

References

1. Ng TB, Liu J, Wong JH, Ye X, Wing Sze SC, Tong Y, et al. Review of research on *Dendrobium*, a prized folk medicine. *Appl Microbiol Biotechnol*. 2012;93:1795–803.
2. Zhu S, Niu Z, Xue Q, Wang H, Xie X, Ding X. Accurate authentication of *Dendrobium officinale* and its closely related species by comparative analysis of complete plastomes. *Acta Pharm Sin B*. 2018;8:969–80.
3. Li L, Jiang Y, Liu Y, Niu Z, Xue Q, Liu W, et al. The large single-copy (LSC) region functions as a highly effective and efficient molecular marker for accurate authentication of medicinal *Dendrobium* species. *Acta Pharm Sin B*. 2020;10:1989–2001.
4. Wang Y, Liu A. Genomic characterization and expression analysis of basic Helix-Loop-Helix (bHLH) family genes in traditional chinese herb *Dendrobium officinale*. *Plants*. 2020;9:1044.
5. Wang H, Dong Z, Chen J, Wang M, Ding Y, Xue Q, et al. Genome-wide identification and expression analysis of the Hsp20, Hsp70 and Hsp90 gene family in *Dendrobium officinale*. *Front Plant Sci*. 2022;13:979801.
6. Yan L, Wang X, Liu H, Tian Y, Lian J, Yang R, et al. The genome of *Dendrobium officinale* Illuminates the biology of the Important traditional chinese orchid herb. *Mol Plant*. 2015;8:922–34.
7. Tang H, Zhao T, Sheng Y, Zheng T, Fu L, Zhang Y. *Dendrobium officinale* Kimura et Migo: a review on its ethnopharmacology, phytochemistry, pharmacology, and industrialization. *Evid Based Complement Alternat Med*. 2017;2017:7436259.
8. Shen C, Guo H, Chen H, Shi Y, Meng Y, Lu J, et al. Identification and analysis of genes associated with the synthesis of bioactive constituents in *Dendrobium officinale* using RNA-Seq. *Sci Rep*. 2017;7:187.
9. Xing S, Zhang X, Ke H, Lin J, Huang Y, Wei G. Physicochemical properties of polysaccharides from *Dendrobium officinale* by fractional precipitation and their preliminary antioxidant and anti-HepG2 cells activities in vitro. *Chem Cent J*. 2018;12:100.
10. Liang J, Wu Y, Yuan H, Yang Y, Xiong Q, Liang C, et al. *Dendrobium officinale* polysaccharides attenuate learning and memory disabilities via anti-oxidant and anti-inflammatory actions. *Int J Biol Macromol*. 2019;126:414–26.
11. Zhong C, Tian W, Chen H, Yang Y, Xu Y, Chen Y, et al. Structural characterization and immunoregulatory activity of polysaccharides from *Dendrobium officinale* leaves. *J Food Biochem*. 2022;46:e14023.

12. Liu Y, Yang L, Zhang Y, Liu X, Wu Z, Gilbert RG, et al. *Dendrobium officinale* polysaccharide ameliorates diabetic hepatic glucose metabolism via glucagon-mediated signaling pathways and modifying liver-glycogen structure. *J Ethnopharmacol.* 2020;248:112308.
13. Kuang M, Li J, Yang X, Yang L, Xu J, Yan S, et al. Structural characterization and hypoglycemic effect via stimulating glucagon-like peptide-1 secretion of two polysaccharides from *Dendrobium officinale*. *Carbohydr Polym.* 2020;241:116326.
14. Li J, Li S, Huang D, Zhao X, Cai G. Advances in the of resources, constituents and pharmacological effects of *Dendrobium officinale*. *Sci Technol Rev.* 2011;29:74–9.
15. Huang K, Li Y, Tao S, Wei G, Huang Y, Chen D, et al. Purification, characterization and biological activity of polysaccharides from *Dendrobium officinale*. *Mol Basel Switz.* 2016;21:701.
16. Zhang B, Tolstikov V, Turnbull C, Hicks LM, Fiehn O. Divergent metabolome and proteome suggest functional independence of dual phloem transport systems in cucurbits. *Proc Natl Acad Sci U S A.* 2010;107:13532–7.
17. Li Q, Li B, Guo S. Advance in molecular biology of *Dendrobium* (Orchidaceae). *China J Chin Mater Medica.* 2016;41:2753–61.
18. Li M, Chen T, Gao T, Miao Z, Jiang A, Shi L, et al. UDP-glucose pyrophosphorylase influences polysaccharide synthesis, cell wall components, and hyphal branching in *Ganoderma lucidum* via regulation of the balance between glucose-1-phosphate and UDP-glucose. *Fungal Genet Biol FG B.* 2015;82:251–63.
19. Wang X, Wang S, Xue Y, Ren X, Xue J, Zhang X. Defoliation, not gibberellin, induces tree peony autumn reflowering regulated by carbon allocation and metabolism in buds and leaves. *Plant Physiol Biochem PPB.* 2020;151:545–55.
20. Zhou T, Hao G, Yang Y, Liu H, Yang M, Zhao Y. *Sicwinv1*, a Cell Wall Invertase from Sesame, Is Involved in Anther Development. *J Plant Growth Regul.* 2019;38:1274–86.
21. Chen B, Wang X, Lv J, Ge M, Qiao K, Chen Q, et al. GhN/AINV13 positively regulates cotton stress tolerance by interacting with the 14-3-3 protein. *Genomics.* 2021;113(1 Pt 1):44–56.
22. Wang X, Zheng L, Lin H, Yu F, Sun L, Li L. Grape hexokinases are involved in the expression regulation of sucrose synthase- and cell wall invertase-encoding genes by glucose and ABA. *Plant Mol Biol.* 2017;94:61–78.
23. Yan W, Wu X, Li Y, Liu G, Cui Z, Jiang T, et al. Cell wall invertase 3 affects cassava productivity via regulating sugar allocation from source to sink. *Front Plant Sci.* 2019;10:541.
24. Roitsch T, González M. Function and regulation of plant invertases: sweet sensations. *Trends Plant Sci.* 2004;9:606–13.
25. Ruan Y, Jin Y, Yang Y, Li G, Boyer JS. Sugar input, metabolism, and signaling mediated by invertase: roles in development, yield potential, and response to drought and heat. *Mol Plant.* 2010;3:942–55.
26. Sherson SM, Alford HL, Forbes SM, Wallace G, Smith SM. Roles of cell-wall invertases and monosaccharide transporters in the growth and development of *Arabidopsis*. *J Exp Bot.* 2003;54:525–31.

27. Haouazine-Takvorian N, Tymowska-Lalanne Z, Takvorian A, Tregear J, Lejeune B, Lechamy A, et al. Characterization of two members of the *Arabidopsis thaliana* gene family, *Atβfruct3* and *Atβfruct4*, coding for vacuolar invertases. *Gene*. 1997;197:239–51.
28. Ji X, Van den Ende W, Van Laere A, Cheng S, Bennett J. Structure, evolution, and expression of the two invertase gene families of rice. *J Mol Evol*. 2005;60:615–34.
29. Liu Y, Nie YD, Han FX, Zhao XN, Dun BQ, Lu M, et al. Allelic variation of a soluble acid invertase gene (*SAI-1*) and development of a functional marker in sweet sorghum [*Sorghum bicolor* (L.) Moench]. *Mol Breed*. 2014;33:721–30.
30. Chen Z, Gao K, Su X, Rao P, An X. Genome-wide identification of the invertase gene family in populus. *PLoS ONE*. 2015;10:e0138540.
31. Liu Y, Dun B, Zhao X, Yue M, Lu M, Li G. Correlation analysis between the key enzymes activities and sugar content in sweet sorghum (*Sorghum bicolor* L. Moench) stems at physiological maturity state. *Aust J Crop Sci*. 2013;7:84–92.
32. Wang L, Ruan Y. Critical roles of vacuolarinvertase in floral organ development and male and female fertilities are revealed through characterization of *GhVIN1*-RNAi cotton plants. *Plant Physiol*. 2016;171:405–23.
33. Sturm A. Molecular characterization and functional analysis of sucrose-cleaving enzymes in carrot (*Daucus carota* L.). *J Exp Bot*. 1996;47:1187–92. Spec No.
34. Qian W, Yue C, Wang Y, Cao H, Li N, Wang L, et al. Identification of the invertase gene family (INVs) in tea plant and their expression analysis under abiotic stress. *Plant Cell Rep*. 2016;35:2269–83.
35. Godt DE, Roitsch T. Regulation and tissue-specific distribution of mRNAs for three extracellular invertase isoenzymes of tomato suggests an important function in establishing and maintaining sink metabolism. *Plant Physiol*. 1997;115:273–82.
36. Shen L, Yao Y, He H, Qin Y, Liu Z, Liu W, et al. Genome-wide identification, expression, and functional analysis of the alkaline/neutral invertase gene family in pepper. *Int J Mol Sci*. 2018;19:224.
37. Yao Y, Geng M, Wu X, Liu J, Li R, Hu X, et al. Genome-wide identification, expression, and activity analysis of alkaline/neutral invertase gene family from cassava (*Manihot esculenta* Crantz). *Plant Mol Biol Report*. 2015;33:304–15.
38. Liu X, Zhang C, Ou Y, Lin Y, Song B, Xie C, et al. Systematic analysis of potato acid invertase genes reveals that a cold-responsive member, *StvacINV1*, regulates cold-induced sweetening of tubers. *Mol Genet Genomics MGG*. 2011;286:109–18.
39. Zhang G, Xu Q, Bian C, Tsai W, Yeh C, Liu K, et al. The *Dendrobium catenatum* Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. *Sci Rep*. 2016;6:19029.
40. Shao Z, Xue J, Wu P, Zhang Y, Wu Y, Hang Y, et al. Large-Scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiol*. 2016;170:2095–109.

41. Verhaest M, Lammens W, Le Roy K, De Coninck B, De Ranter CJ, Van Laere A, et al. X-ray diffraction structure of a cell-wall invertase from *Arabidopsis thaliana*. *Acta Crystallogr D Biol Crystallogr*. 2006;62(Pt 12):1555–63.
42. Wang L, Zheng Y, Ding S, Zhang Q, Chen Y, Zhang J. Molecular cloning, structure, phylogeny and expression analysis of the invertase gene family in sugarcane. *BMC Plant Biol*. 2017;17:109.
43. Bocock PN, Morse AM, Dervinis C, Davis JM. Evolution and diversity of invertase genes in *Populus trichocarpa*. *Planta*. 2008;227:565–76.
44. Yao Y, Geng M, Wu X, Liu J, Li R, Hu X, et al. Genome-wide identification, 3D modeling, expression and enzymatic activity analysis of cell wall invertase gene family from cassava (*Manihot esculenta* Crantz). *Int J Mol Sci*. 2014;15:7313–31.
45. Shen L, Qin Y, Qi Z, Niu Y, Liu Z, Liu W-X, et al. Genome-Wide analysis, expression profile, and characterization of the acid invertase gene family in pepper. *Int J Mol Sci*. 2018;20:15.
46. Jain R, Singh SP, Singh A, Singh S, Kishor R, Singh RK, et al. Soluble acid invertase (SAI) activity and gene expression controlling sugar composition in sugarcane. *Sugar Tech*. 2017;19:669–74.
47. Li R, Hou Z, Zou H, Wang Y, Liao X. Inactivation kinetics, structural, and morphological modification of mango soluble acid invertase by high pressure processing combined with mild temperatures. *Food Res Int Ott Ont*. 2018;105:845–52.
48. Vargas WA, Pontis HG, Salerno GL. Differential expression of alkaline and neutral invertases in response to environmental stresses: characterization of an alkaline isoform as a stress-response enzyme in wheat leaves. *Planta*. 2007;226:1535–45.
49. Kim JK, Bamba T, Harada K, Fukusaki E, Kobayashi A. Time-course metabolic profiling in *Arabidopsis thaliana* cell cultures after salt stress treatment. *J Exp Bot*. 2007;58:415–24.
50. Sturm A. Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. *Plant Physiol*. 1999;121:1–8.
51. Liao S, Lin C, Wang A, Sung H. Differential expression of genes encoding acid invertases in multiple shoots of bamboo in response to various phytohormones and environmental factors. *J Agric Food Chem*. 2013;61:4396–405.
52. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant*. 2020;13:1194–202.
53. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinforma Oxf Engl*. 2007;23:2947–8.
54. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol*. 2021;38:3022–7.
55. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12:357–60.
56. Perteira M, Perteira GM, Antonescu CM, Chang T, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*. 2015;33:290–5.

57. Yuan Y, Zhang J, Liu X, Meng M, Wang J, Lin J. Tissue-specific transcriptome for *Dendrobium officinale* reveals genes involved in flavonoid biosynthesis. *Genomics*. 2020;112:1781–94.
58. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods San Diego Calif*. 2001;25:402–8.
59. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015;43(Database issue):D447–452.
60. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498–504.

Figures

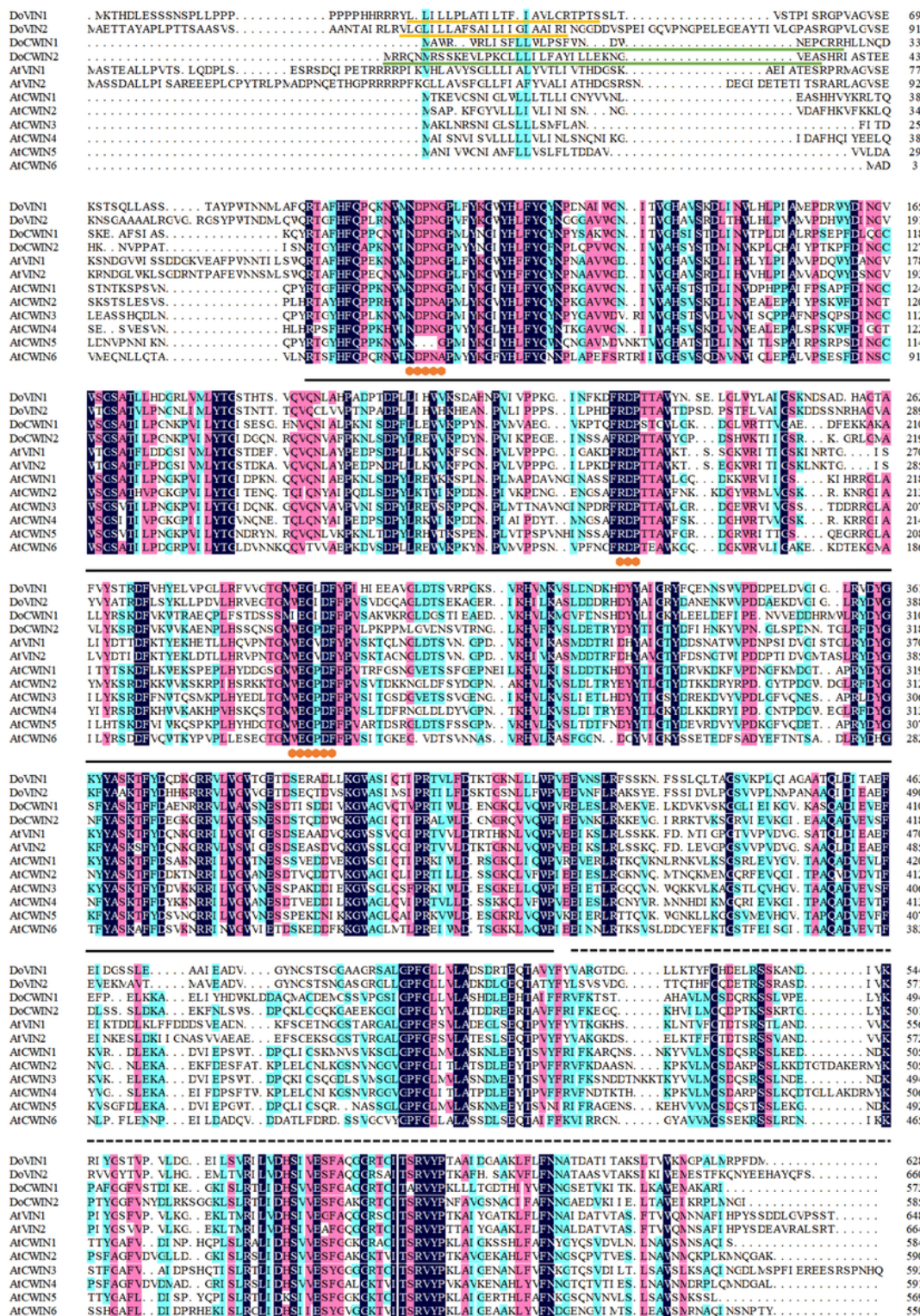


Figure 1

Sequence alignment of the A/NV genes from *A. thaliana* and *D. officinale*. Dark-blue, pink, and light blue shading reflects 100%, 75%, and 50% amino acid residue conservation, respectively. Yellow lines indicate the predicted transmembrane domain. Green lines indicate the predicted signal peptide. Orange dots indicate the conserved motifs of NDPNG (β-fructosidase), FRDP, and WECXDF. The solid black line indicates the conserved motif of NDPNG.

indicates the glycosyl hydrolases family GH32 N-terminal domain. The black dashed line indicates glycosyl hydrolases family GH32 C-terminal domain

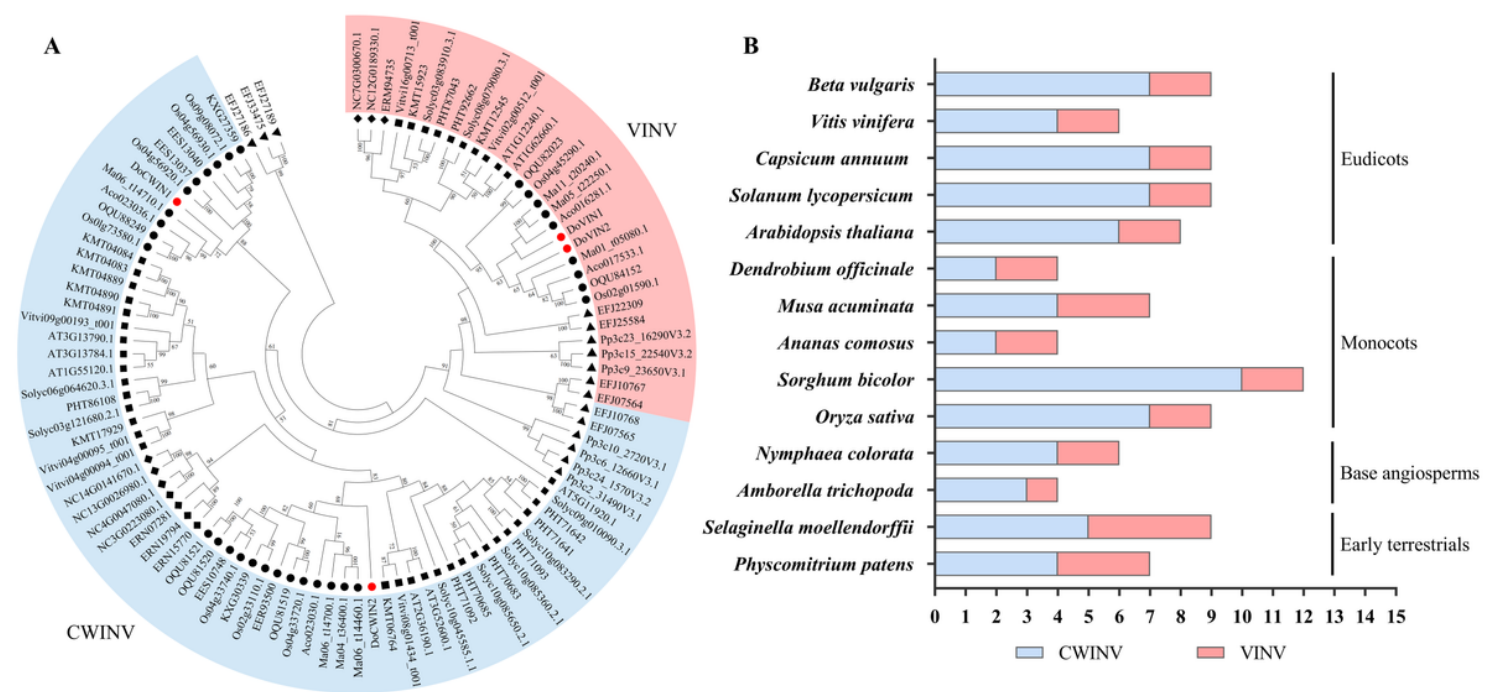


Figure 2

Phylogenetic analyses of AINVs from *D. officinale* and different plant species. (A) The phylogenetic tree was constructed using the NJ method with 1 000 bootstrap replicates. All *AINV* family genes were grouped into two subfamilies named VINV and CWINV. VINVs were colorized by pink zone. CWINVs were colorized by blue zone. Red points indicate *D. officinale* AINVs. The solid points indicate monocots, the solid square indicates eudicots, the solid rhombus indicates base angiosperms and the solid triangles indicate early terrestrial plants. (B) AINV family members of *B. vulgaris*, *V. vinifera*, *C. annuum*, *S. lycopersicum*, *A. thaliana*, *D. officinale*, *M. acuminata*, *A. comosus*, *S. bicolor*, *O. sativa*, *N. colorata*, *A. trichopoda*, *S. moellendorffii*, *P. patens*

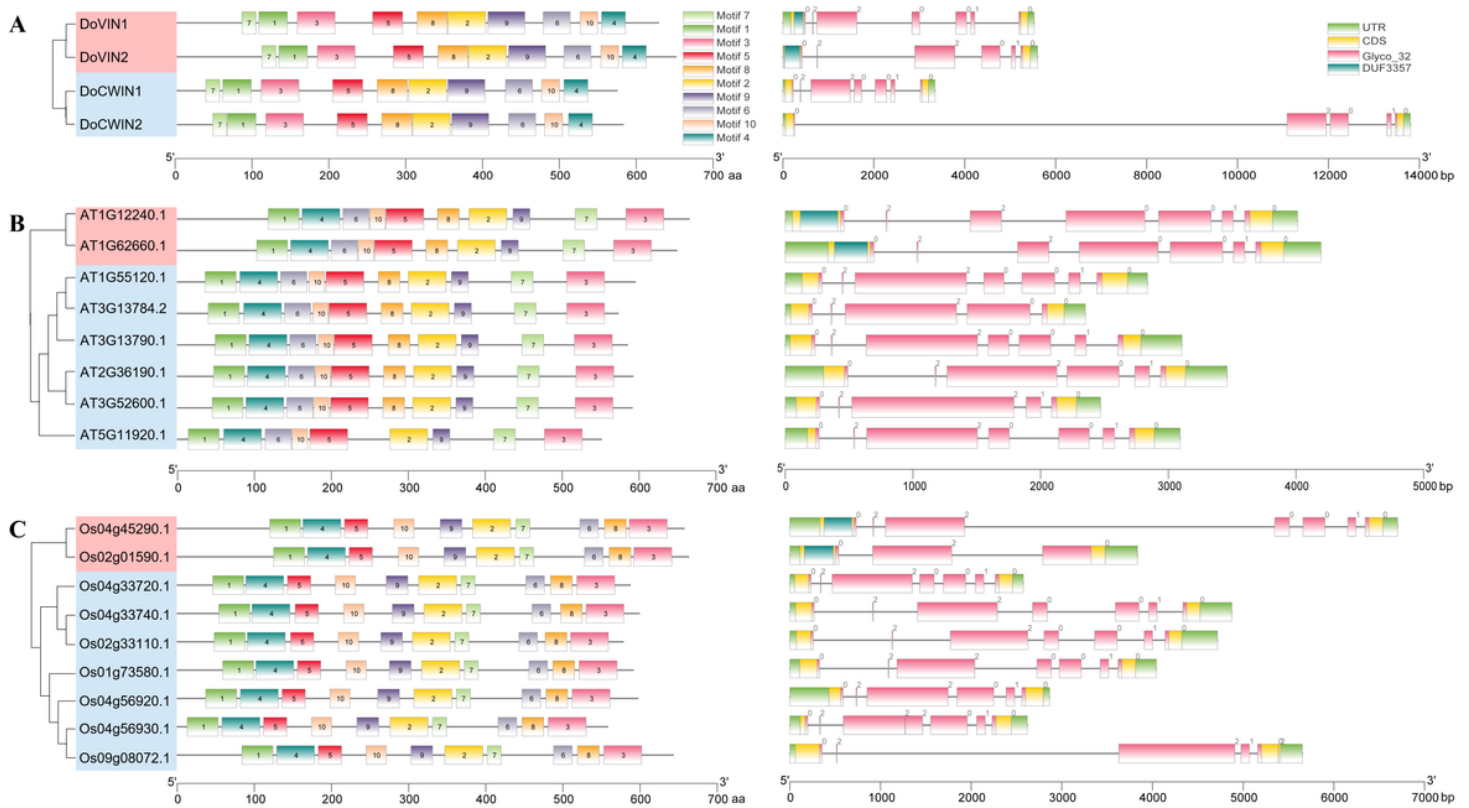


Figure 3

The conserved motifs and exon-intron organization analyses of plant *AINV* genes based on their evolutionary relationship. (A) Structural features of four *AINV* genes of *D. officinale*. (B) Structural features of eight *AINV* genes from the genome of *A. thaliana*. (C) Structural features of nine *AINV* genes from the rice genome (*Oryza sativa*). Different colors on the left mark conserved motifs. Different colors on the right also mark the exon-intron structures, and the numbers “0”, “1”, and “2” indicate the phases of introns. The phylogenetic tree between *AINVs* was constructed by MEGA 11

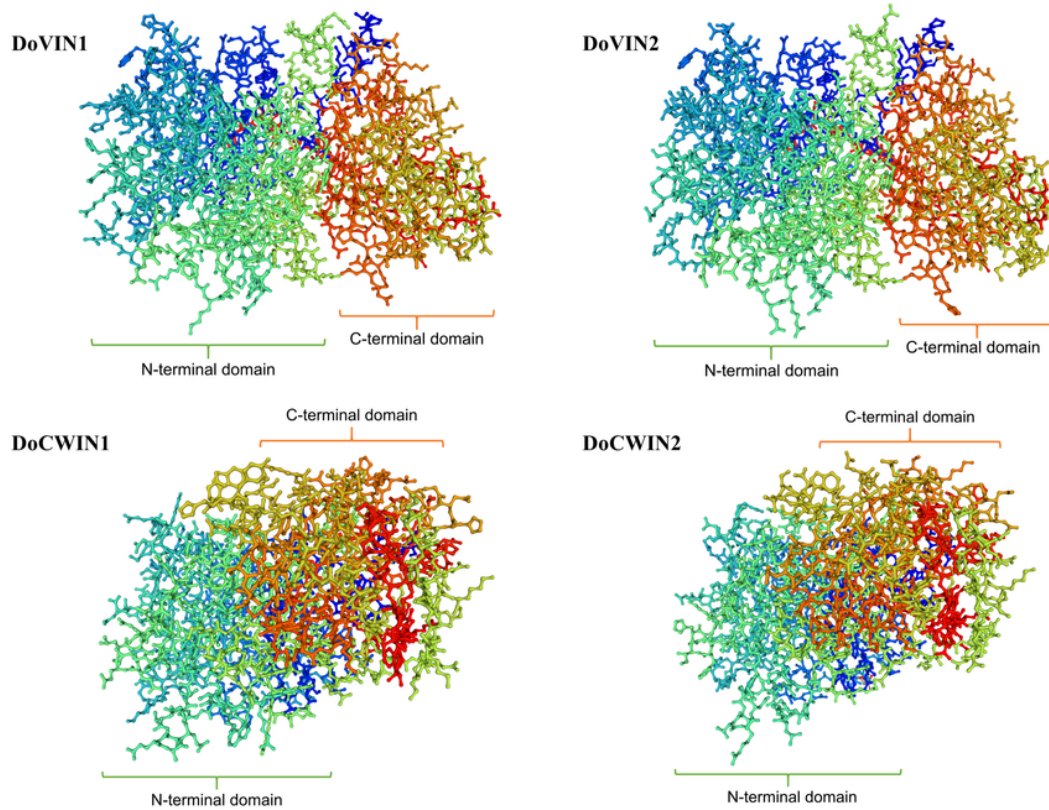


Figure 4

Predicted tertiary structure models of AINV proteins in *D. officinale*

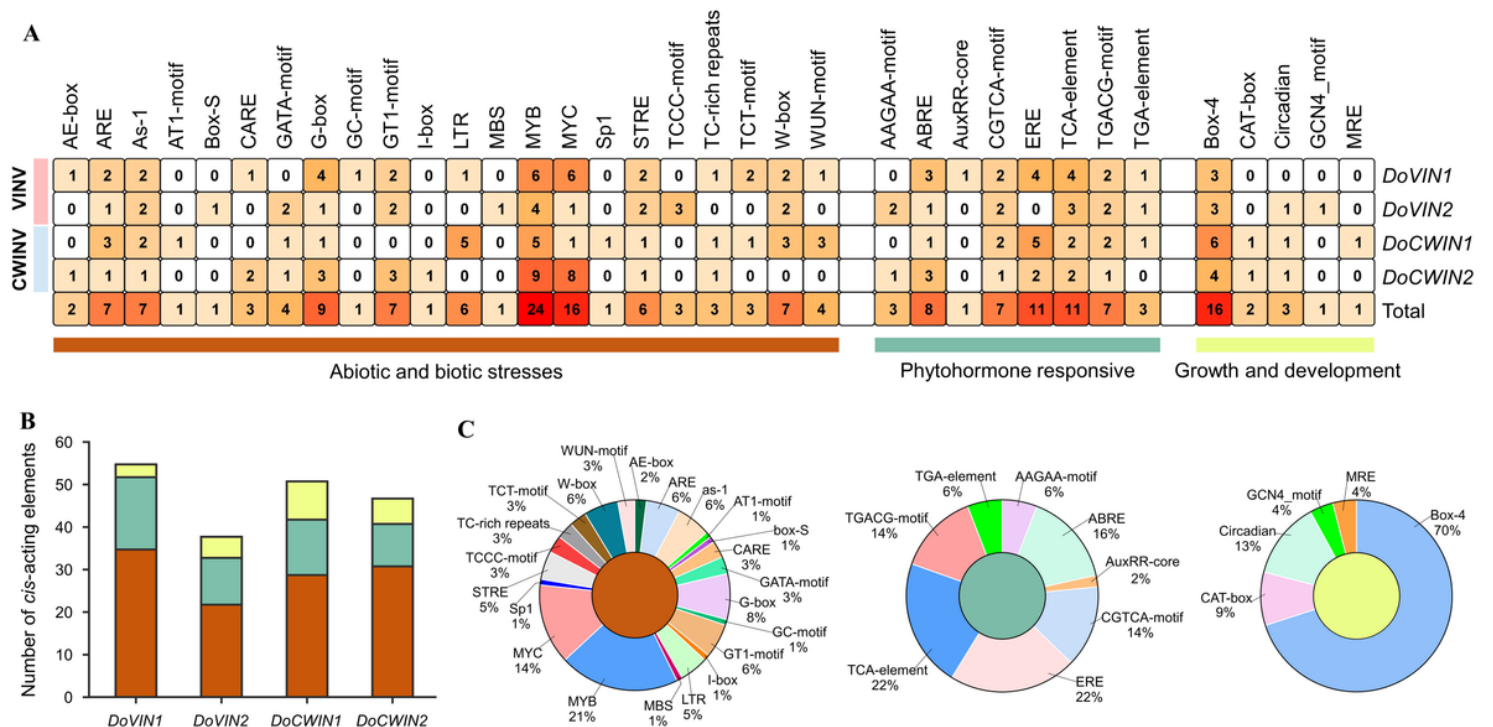


Figure 5

Investigation of *cis*-acting element numbers in *AINV* genes of *D. officinale*. (A) The different colors and numbers of the grid indicated the numbers of different promoter elements in four *DoAINV* genes. (B) The number of *cis*-acting elements belonging to each functional group in individual *DoAINV* promoter sequences. (C) Pie charts of different sizes indicated the ratio of each promoter element in each functional group, respectively

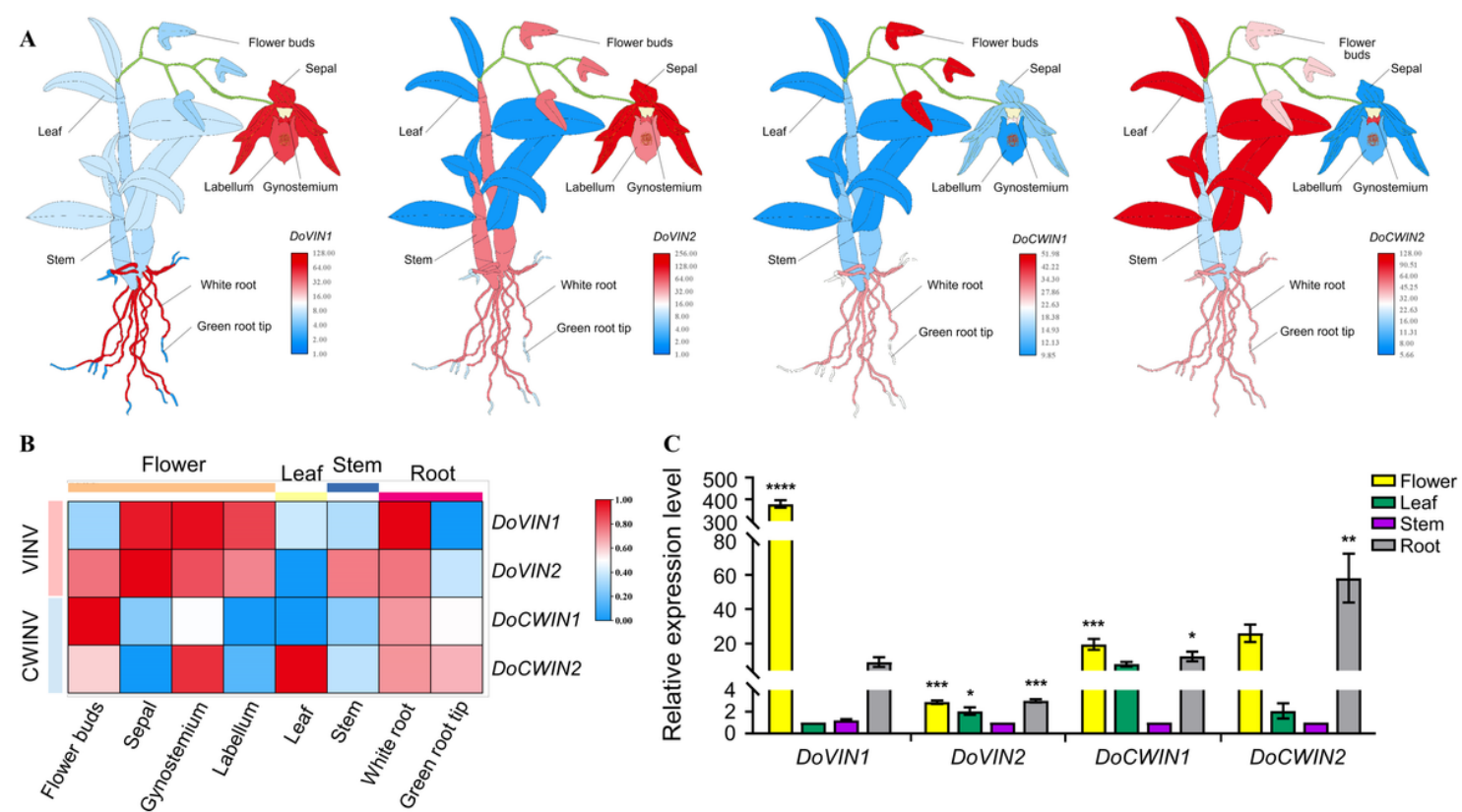


Figure 6

Tissue and organ expression pattern of *AINV* genes in *D. officinale*. (A) and (B) Organ-specific expression pattern of *DoAINV* genes in eight tissues: flower buds, sepal, gynostemium, labellum, leaf, stem, white root, and green root tip. Heatmap was generated using the TBtools software, based on $\log_2$ (FPKM). Blue and red indicated lower and higher levels of transcript abundance, respectively. (C) qRT-PCR validation of *DoAINVs* expression in different tissues was obtained from six-month-old aseptic seedlings of the *D. officinale* DanXia cultivar. The floral organs were collected from one-year-old plants at the flowering stage. Transcripts were normalized to the expression of the actin gene. The mean $\pm$ standard deviations (SDs) of three biological replicates are presented, and the significant differences were analyzed by t-test. **** and *** significant difference presented $P < 0.001$, ** $P < 0.01$, and * ($P < 0.05$)

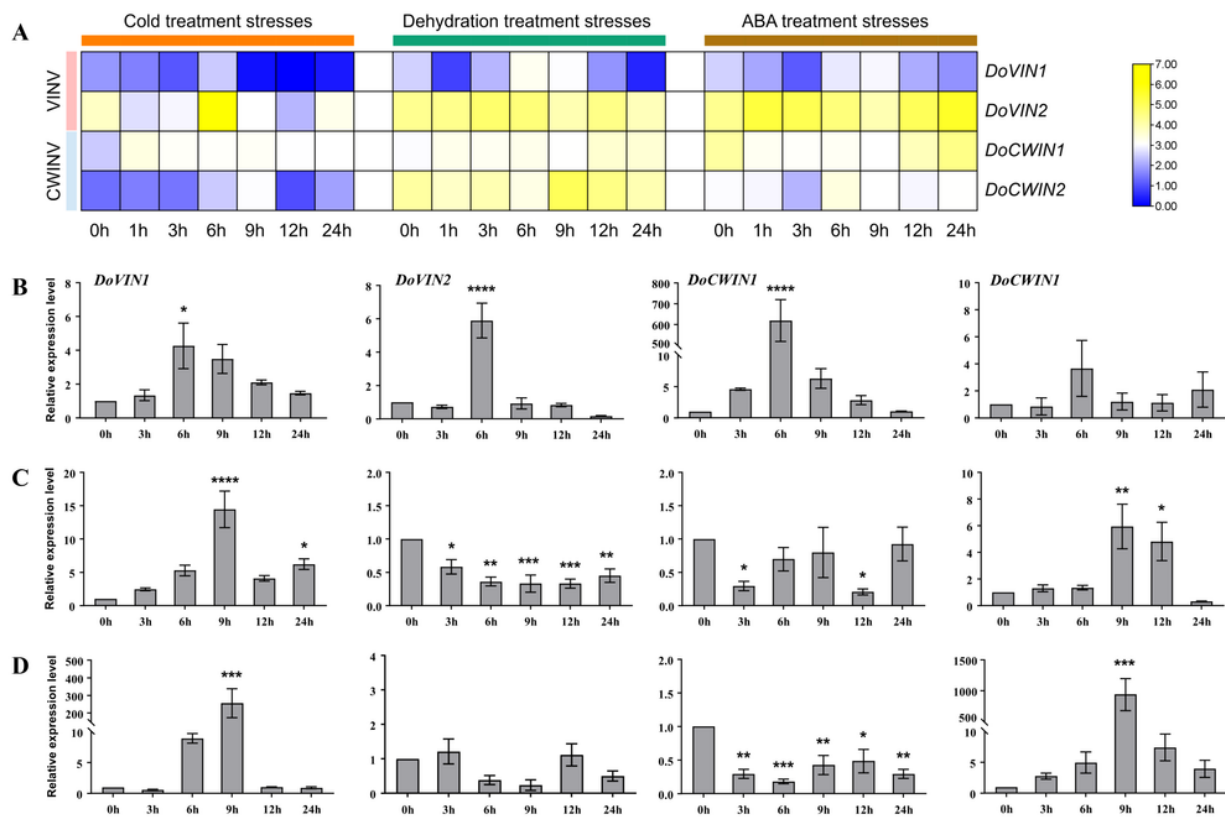


Figure 7

Expression patterns of *DoAINV* genes of *D. officinale* in response to abiotic stresses. (A) Expression profiles of *DoAINV* genes in *D. officinale* DanXia cultivar under cold, dehydration and ABA treatment stresses. Heatmap was generated using the TBtools software, based on the $\log_2$ (FPKM). Blue and red represent low and high levels of transcript abundance, respectively. (B), (C) and (D) Expression levels of *DoAINV* genes in *D. officinale* DanXia cultivar in response to cold, dehydration, and ABA treatment stresses as determined, respectively, by qRT-PCR experiment. The x-axis indicated the time course of each stress treatment, and the y-axis represented the relative expression level. The values are the mean of three biological replicates, SDs indicated by error bars. Significant differences were analyzed by t-test. **** and *** significant difference presented $P < 0.001$, ** Significant difference at $P < 0.01$, * significant difference ($P < 0.05$)

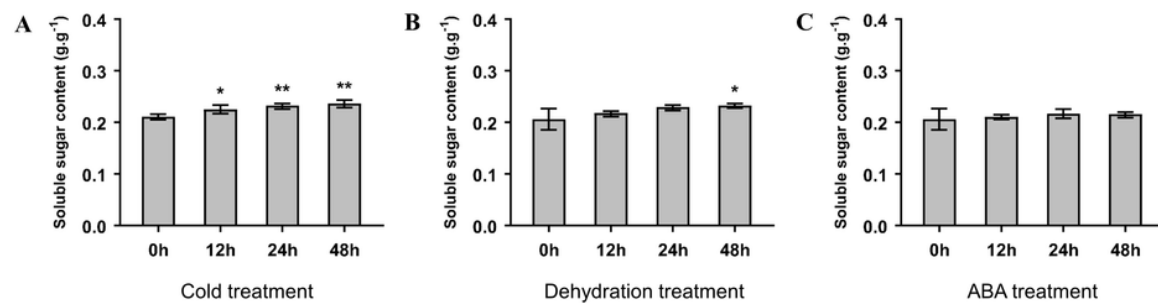


Figure 8

soluble sugar content of *D. officinale* stems at different times in the cold (A) dehydration (B) and ABA (C) treatments

