

Article

Genome-Wide Analysis of the *Trihelix* Gene Family and Their Response to Cold Stress in *Dendrobium officinale*

Yan Tong [†], Hui Huang [†]  and YuHua Wang ^{*} 

Bio-Innovation Center of DR PLANT, Key Laboratory of Economic Plants and Biotechnology, Yunnan Key Laboratory for Wild Plant Resources, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China; tongyan@mail.kib.ac.cn (Y.T.); huanghui@mail.kib.ac.cn (H.H.)

* Correspondence: wangyuhua@mail.kib.ac.cn

† These authors contributed equally to this work.

Abstract: *Trihelix* transcription factors play important roles in plant growth, development and various stress responses. In this study, we identified 32 *trihelix* family genes (*DoGT*) in the important Chinese medicinal plant *Dendrobium officinale*. These *trihelix* genes could be classified into five different subgroups. The gene structure and conserved functional domain of these *trihelix* genes were similar in the same subfamily but diverged between different subfamilies. Various stresses responsive *cis*-elements presented in the promoters of *DoGT* genes, suggesting that the *trihelix* genes might respond to the environmental stresses. Expressional changes of *DoGT* genes in three tissues and under cold treatment suggested that *trihelix* genes were involved in diverse functions during *D. officinale* development and cold tolerance. This study provides novel insights into the phylogenetic relationships and functions of the *D. officinale trihelix* genes, which will aid future functional studies investigating the divergent roles of *trihelix* genes belonging to other species.



Citation: Tong, Y.; Huang, H.; Wang, Y. Genome-Wide Analysis of the *Trihelix* Gene Family and Their Response to Cold Stress in *Dendrobium officinale*. *Sustainability* **2021**, *13*, 2826. <https://doi.org/10.3390/su13052826>

Academic Editor: Zhongwei Zou

Received: 28 January 2021

Accepted: 24 February 2021

Published: 5 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: *Dendrobium officinale*; *trihelix*; phylogenetic analysis; cold stress response

1. Introduction

Transcription factors are necessary regulatory factors in growth, development and abiotic stress responses of plant [1]. TFs activate or inhibit transcription by specifically combined with gene promoter regions and *cis*-acting elements [2,3]. *Trihelix* TF family feature a classic *trihelix* (helix–loop–helix–loop–helix) domain which binds to GT elements required for the light response, similar to the Myb DNA-binding domains in sequence, also termed GT factors [4–6]. According to the structure of the trihelical domain, *trihelix* TFs were grouped into five clades named GT-1, GT-2, SH4, GT γ and SIP1 [7]. There are differences between clades, the internal hydrophobic region of each α helix of each *trihelix* domain in GT-1 and SH4 contains a tryptophan residue [8]. In GT-2 and GT γ , the third conserved tryptophan is replaced by phenylalanine and by isoleucine in SIP1 [9,10]. The most notable feature of GT-2 is that there is another *trihelix* domain at the C-terminal [11]. Moreover, GT-1, GT-2, GT γ and SIP1 also contain an amphoteric α helix structure (the fourth helix) [12].

Past research has made clear that *trihelix* genes play important roles in flower development, embryo maturation and seed growth. *PETAL LOSS* (*PTL*) gene of *Arabidopsis* belongs to GT-2 subfamily, can regulate the development of petals, sepals and morphogenesis of floral organ [13–15]. Some members of the SIP1 subgroup have been reported to be associated with the process of plant embryo development and cell proliferation [16]. *ASIL1* isolated from *Arabidopsis thaliana* acts as a temporal regulator of seed filling by repressing the expression of master regulatory genes *LEC2*, *FUS3*, *ABI3* and other genes [17]. In *Brassica napus*, overexpression of the *BnSIP1-1* gene which fell in the SIP1 clade can improve seed germination under osmotic pressure, salt and ABA treatments [18]. The seed shattering of rice is controlled by a single dominant gene *SHATTERING1* (*SHA1*) belonging

to the SH4 clade [19]. Except for plant development, *trihelix* genes also play key roles in plant biotic and abiotic stress responses including pathogen-induced defense programs and response to drought, salt, cold stress et al. [7]. One of the most important functions of *trihelix* genes seems to be regulation of the cold stress response. Overexpressing of the GT-1 *trihelix* gene, *ShCIGT*, could enhance cold and drought tolerance in tomato [20]. The transgenic *Arabidopsis* plants expressing *GmGT-2A* and *GMGT-2B* from soybean displayed strong resistance to freezing stress [10].

Dendrobium officinale Kimura et Migo is a perennial epiphytic herb of *Dendrobium* in Orchidaceae [21]. As an important traditional herbal plant, *D. officinale* has over hundreds years of history of medicine in many Asian countries. Under nature condition, *D. officinale* grows compatibly on damp rock of mountain climates at 500–1600 m or tree trunks in primeval forests in warm and humid environments, so it is quite easily disturbed by abiotic stress, such as high temperature, low temperature, drought and salinity [22,23]. The species is extremely hypersensitive to low temperature above the freezing point, resulting in major yield losses [24]. As a result of its habitat shrinking, human overexploitation and its low natural reproduction rate, slow growth, wild *D. officinale* were listed in the IUCN red list of threatened species (<http://www.iucnredlist.org/details/46665/0> (accessed on 30 April 2004)). Nowadays, *D. officinale* is commonly produced in green house for saving from cold environment. Although important role of the *trihelix* genes in plant development and stress resistance and has been investigated in *Arabidopsis*, rice and soybean, but not yet well studied in *D. officinale*. Therefore, it is important to reveal molecular mechanism in response to cold stress in *D. officinale*, not only for breeding cold-tolerance cultivars but also for lifting productivity. The generation of draft genome of *D. officinale* provides a first-time opportunity to perform a genome-wide identification of *trihelix* gene family. We comprehensively characterized the number, structure and phylogenetic relationships of the *trihelix* members throughout the *D. officinale* genome. We also examined the expression differentiation of the *trihelix* genes among distinct tissues and under cold stress.

2. Materials and Methods

2.1. Identification of the *Trihelix* Gene Family in *D. officinale*

In order to identify the *trihelix* gene family, firstly, we downloaded the genome sequences of *D. officinale* from NCBI (<http://www.ncbi.nlm.nih.gov/> (accession codes: PRJNA262478, accessed on 26 Feb 2016)). Then, the HMM profile (PF13837) were acquired from Pfam database and used to search trihelix domains through HMMER 3.0 software with an E-value < 0.00001 [25]. Finally, after removing the incorrect and redundant members, the candidate trihelix protein sequences were further verified by Pfam and SMART online software [26]. ProtParam (<http://web.expasy.org/protparam/> (accessed on 4 June 2020)) was used to compute the pI (isoelectric point), MW (molecular weight) and GRAVY (grand average of hydropathy) of trihelix proteins. The subcellular locations of trihelix members were analyzed using Plant-mPLOC online software [27].

As a control, the *trihelix* genes of *Oryza sativa* and *Brachypodium distachyon* were downloaded from the Rice Information Resource and Phytozome, respectively [28]. In addition, the *trihelix* genes of *B. distachyon* (*BdGT*) were named according to a previous study [29].

2.2. Phylogenetic Analysis

A multiple alignment analysis of *DoGT* genes was run with ClustalW 2.0 and manually corrected using BioEdit 7.1. The phylogenetic tree was builded using MEGA 5.0 based on the NJ and ML methods [30]. The bootstrap values were calculated for 1000 iterations.

2.3. Conserved Structures and Motifs of *Trihelix* Genes

The exon/intron structure of *trihelix* genes was generated using the GSDS program (version 2.0) with coding and genomic sequences [31]. The motifs of each deduced trihelix

protein were analyzed by MEME suit software (version 4.12.0) with parameters as follows: maximum number of motifs, 10 [32].

2.4. Promoter Analysis and Gene Ontology (GO) Annotation

For promoter analysis, the 1500-bp upstream sequences of genes were scanned for the *cis*-elements using PlantCARE web site [33]. Gene ontology (GO) annotation of trihelix proteins was performed by InterproScan and used to predict the functions of DoGT proteins [34]. The GO annotation were then plotted using the OmicShare tool (<http://www.omicshare.com/tools> (accessed on 9 June 2020)).

2.5. Expression Profile Analysis

We downloaded the RNA-seq data titled SRR2014396, SRR2014297 and SRR2014236 from the NCBI to analyze tissue-specific expression patterns of *D. officinale trihelix* genes. The Illumina data were mapped to the *D. officinale* genome with HISAT2 (V2.1.0) and then the transcripts were next determined by StringTie (V1.3.5) [35]. The gene expression levels were calculated as the number of fragments per kilobase of gene length per million mapped reads (FPKM). Using the DESeq2 [36] package of R, we analyzed the differential expression patterns of DoGT genes and the Gplots package of R program (<https://www.R-project.org/> (accessed on 16 June 2020)) was used to draw the heatmap.

2.6. Plant Materials and Experimental Treatments

D. officinale plants were artificially cultivated and collected from the cultivation base of Puer, Yunnan Province, China. They were maintained in a greenhouse in Kunming Institute of Botany, Chinese Academy of Sciences (Kunming, China), in 25 °C with a constant photoperiod (12 h light/12 h dark). The fresh and healthy tissues including leaves, stems and flowers were harvested from about 1-year-old *D. officinale* plant, immediately frozen in liquid nitrogen and stored at −80 °C for subsequent studies.

For cold stress, the plants were applied by placing them at 0 °C in the cold cabinet (Haier) under a 24 h dark. Leaf samples were harvested from treatment at 0, 4, 8, 12, 16, 20 and 24 h after initiation of the treatment. Samples were quickly frozen at −80 °C until further processing. Three independent biological replications were performed for each treatment.

2.7. RNA Isolation, Expression and Statistical Analysis

Total RNA of the sample was extracted using a modified cetyltrimethylammonium bromide (CTAB) method [37]. The total RNA was eluted in 30 µL of RNase-free water and stored at −80 °C. The RNA concentration was calculated using a Nanodrop ND2000 spectrophotometer (Nano-Drop Technologies, Wilmington, DE, USA) with RNase-free water as a blank. The cDNA was synthesized through the FastKing gDNA Dispelling RT SuperMix (Tiangen, China).

2.8. qRT-PCR Analysis of DoGT Genes

Quantitative real-time PCR primers were designed according to the CDS sequence of DoGT by Primer 6 software (Table S3). The primer mass was tested using PCR amplification, agarose gel electrophoresis and melting curve analysis. The qRT-PCR experiment was run on the Bio-Rad CFX96 Real Time PCR System and SuperReal PreMix Plus (SYBR Green, Tiangen, China). The 10 µL reaction volume contained 5 µL 2×SuperReal PreMix Plus, 1 µL of diluted cDNA, 0.5 µL of each primer (10 µM), 1 µL 50×ROX Reference Dye and the addition of ddH₂O to bring the total volume to 10 µL. The PCR parameters were as follows: 95 °C for 30 s; 40 cycles of 95 °C for 5 s, 60 °C for 30 s, 95 °C for 10 s. According to previous results, *GAPDH* gene as the internal reference. Relative expression level of the DoGT gene was analyzed by the $2^{-\Delta\Delta CT}$ method. Three technical replicates were done for each sample. Normalizing all of the data based on setting the expression level at 0 h as a value of 1 for cold stress (values above 1 and below 1 were considered as up- and

down-regulated, respectively). For tissue expression analysis we used leaves as control (expression = 1.0) to calculate the fold change in the expression level of the relevant genes.

3. Results

3.1. Identification of the DoGT Genes in *D. officinale*

Using the consensus sequence of the trihelix domain, a total of 32 non-redundant *D. officinale* trihelix genes (*DoGT*) were identified and the detail information are listed in Table 1. The 32 trihelix genes were subsequently renamed from *DoGT1* to *DoGT32* according to their gene IDs. All of the 32 trihelix proteins contained a typical feature of the trihelix domain confirmed by Pfam and Smart. The detail information of *DoGT* genes including length, pI, Mw, GRAVY and location are also presented in Table 1. Lengths of the open reading frames (ORFs) of *DoGT* proteins ranged from 203 to 789 aa with the pI varying from 4.58 to 10.19 and the Mw spanning from 22.01 kD to 85.07 kD. The GRAVY values of these deduced trihelix proteins ranged from -1.414 to -0.244 , indicating that they were hydrophilic. Furthermore, the Plant-mPloc analysis showed that nearly all the *DoGT* proteins were located in the nucleus except for several members, including *DoGT1*, *DoGT2*, *DoGT6*, *DoGT9*, *DoGT10*, *DoGT23* and *DoGT30* which located in the chloroplast, chloroplast/nucleus/peroxisome, cell membrane/chloroplast, chloroplast/nucleus, chloroplast/nucleus, chloroplast and chloroplast, respectively.

Table 1. Summary of *DoGT* genes in the *D. officinale* genome.

Gene Name	Sequence ID	CDS (bp)	ORF (aa)	MW (kD)	PI	GRAVY	Group	Location
<i>DoGT1</i>	PKU61201.1	1845	614	68.32	8.66	-0.244	GT-1	Chloroplast
<i>DoGT2</i>	PKU61538.1	1173	390	44.48	6.54	-0.734	GT-1	Chloroplast/Nucleus/ Peroxisome
<i>DoGT3</i>	PKU64395.1	1002	333	37.46	9.08	-0.633	SIP1	Nucleus
<i>DoGT4</i>	PKU64635.1	960	319	37.35	6.66	-0.846	GT-2	Nucleus
<i>DoGT5</i>	PKU65280.1	2370	789	85.07	5.86	-0.691	GT-2	Nucleus
<i>DoGT6</i>	PKU66139.1	756	251	28.23	10.05	-0.496	SIP1	Cell mem- brane/Chloroplast
<i>DoGT7</i>	PKU68542.1	1095	364	42.29	5.8	-0.778	GT γ	Nucleus
<i>DoGT8</i>	PKU70315.1	1335	444	51.02	6.56	-0.905	GT γ	Nucleus
<i>DoGT9</i>	PKU72438.1	789	262	30.19	5.23	-1.039	SIP1	Chloroplast/Nucleus
<i>DoGT10</i>	PKU72660.1	882	293	32.82	9.79	-0.73	SIP1	Chloroplast/Nucleus
<i>DoGT11</i>	PKU72926.1	1272	423	47.82	5.12	-0.864	SIP1	Nucleus
<i>DoGT12</i>	PKU73007.1	882	293	33.06	9.21	-0.936	SIP1	Nucleus
<i>DoGT13</i>	PKU73055.1	1125	374	42.48	6.01	-0.794	GT γ	Nucleus
<i>DoGT14</i>	PKU73826.1	963	320	35.29	10.06	-0.865	SIP1	Nucleus
<i>DoGT15</i>	PKU74793.1	612	203	22.01	7.04	-0.368	SIP1	Nucleus
<i>DoGT16</i>	PKU75532.1	1860	619	69.43	6.27	-0.849	GT-2	Nucleus
<i>DoGT17</i>	PKU75541.1	936	311	36.63	6.39	-1.139	GT-1	Nucleus
<i>DoGT18</i>	PKU76359.1	1164	387	43.90	6.36	-0.697	GT-1	Nucleus
<i>DoGT19</i>	PKU76537.1	792	263	31.69	6.34	-1.414	GT-1	Nucleus
<i>DoGT20</i>	PKU77186.1	834	277	30.78	9.45	-0.973	SIP1	Nucleus
<i>DoGT21</i>	PKU77634.1	2136	711	77.12	5.55	-0.778	GT-2	Nucleus
<i>DoGT22</i>	PKU78493.1	672	223	24.97	5.07	-0.709	SH4	Nucleus
<i>DoGT23</i>	PKU79539.1	1782	593	64.86	6.27	-0.465	SIP1	Chloroplast
<i>DoGT24</i>	PKU79543.1	1344	447	49.48	6.01	-0.808	GT γ	Nucleus
<i>DoGT25</i>	PKU81577.1	837	278	31.63	10.19	-0.832	SIP1	Nucleus
<i>DoGT26</i>	PKU81817.1	1209	402	45.93	4.58	-1.286	SIP1	Nucleus
<i>DoGT27</i>	PKU81902.1	870	289	32.33	10.11	-0.709	SIP1	Nucleus
<i>DoGT28</i>	PKU83780.1	810	269	30.75	6.13	-0.996	SIP1	Nucleus
<i>DoGT29</i>	PKU84102.1	948	315	36.80	6.65	-1.156	GT-1	Nucleus
<i>DoGT30</i>	PKU84591.1	1749	582	63.77	6.61	-0.576	SIP1	Chloroplast
<i>DoGT31</i>	PKU86134.1	1440	479	53.07	8.65	-0.605	GT-2	Nucleus
<i>DoGT32</i>	PKU87767.1	993	330	35.74	9.33	-0.737	SH4	Nucleus

3.2. Sequence Alignment and Phylogenetic Analyses of Trihelix TFs

To uncover the classifications of the DoGT proteins, un-rooted Neighbor-Joining (NJ) and Maximum Likelihood (ML) phylogenetic trees were constructed using *D. officinale* and other species (rice and *B. distachyon*). The results of these two trees were consistent (Figure 1 and Figure S2). As shown in Figure 1 and Figure S2, the DoGT genes were divided into five subgroups, namely GT-1, GT γ , GT-2, SH4 and SIP1, respectively. This classification was highly similar to previous study in *Arabidopsis*, rice and wheat [7,29,38]. The 32 DoGT genes were distributed over all of these subfamilies. The distribution trends were similar to those in *Arabidopsis* and *Citrus sinensis* [39]: the SIP1 clade was the largest subfamily, containing 15 trihelix genes, whereas the SH4 subgroup was the smallest, only containing two members. This result indicates that DoGT genes are unevenly distributed in the five subgroups. In addition, the phylogenetic tree also revealed paralogous and orthologous relationships among these three species (Figure 1). Seven pairs of paralogous proteins were identified in *D. officinale*, including DoGT2 and DoGT18, DoGT5 and DoGT21, DoGT7 and DoGT13, DoGT9 and DoGT11, DoGT17 and DoGT29, DoGT23 and DoGT30 and DoGT25 and DoGT27 with strong bootstrap support 60, 100, 99, 100, 97, 100 and 99, respectively. Furthermore, 28 pairs of orthologous genes were found between rice and *B. distachyon*. Moreover, group GT-1 and GT-2 were located at the bottom of the phylogenetic tree, which is consistent with the hypothesis that GT1 and GT2 diverged early in the existence of land plants [40].

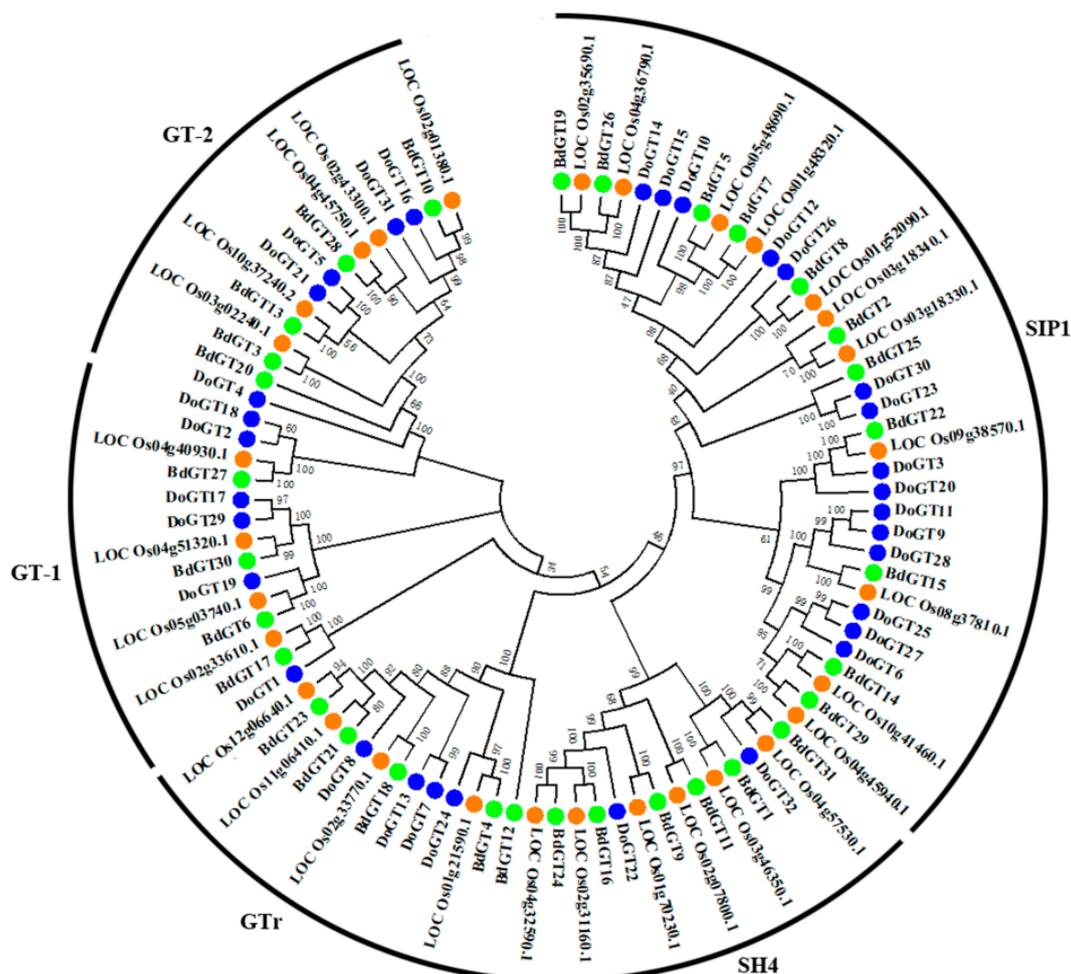


Figure 1. Phylogenetic tree of trihelix proteins in *D. officinale*, *B. distachyon* and rice which was constructed using NJ method. Reliability of the predicted trees were tested using bootstrapping with 1000 replicates. DoGT, BdGT and LOC indicated trihelix genes of *D. officinale*, *B. distachyon* and rice, respectively.

3.3. Protein Structure of the DoGT Gene Family

To reveal structural diversification of *DoGT* genes in *D. officinale*, we used MEME web site to predict the conserved motifs and a total of ten distinct motifs were analyzed [32,41]. We presented the schematic distribution of these motifs among different gene groups (Figure 2) to show their relative locations within proteins. The multi-level consensus sequences were produced among these motifs (Table 2). In total, the ten motifs were annotated by InterProScan. Almost all DoGT proteins had motif 1 and motif 4, except for DoGT18 (without motif 4, GT-1 subgroup), DoGT22 and DoGT32 (without motif 4, SH4 subgroup). Motif 3 and 7 were only found in clade members of GT-1 and GT-2 except for DoGT1, DoGT2 and DoGT18. Motif 9 was identified in all GT-2 members and some of GT-1 and SH4 subfamilies. Similarly, motif 6 and 8 were identified in all GT γ proteins, while motif 2 and 5 only appeared in SIP1 subgroup, respectively. These results showed that the gene structure and motifs of *DoGTs* were conserved. As a result, the majority of closely related members in the phylogenetic tree, as expected, had common motif compositions, suggesting their functional conservation. Our phylogenetic analysis results and previous studies clearly showed the reliability of this classification [41,42].

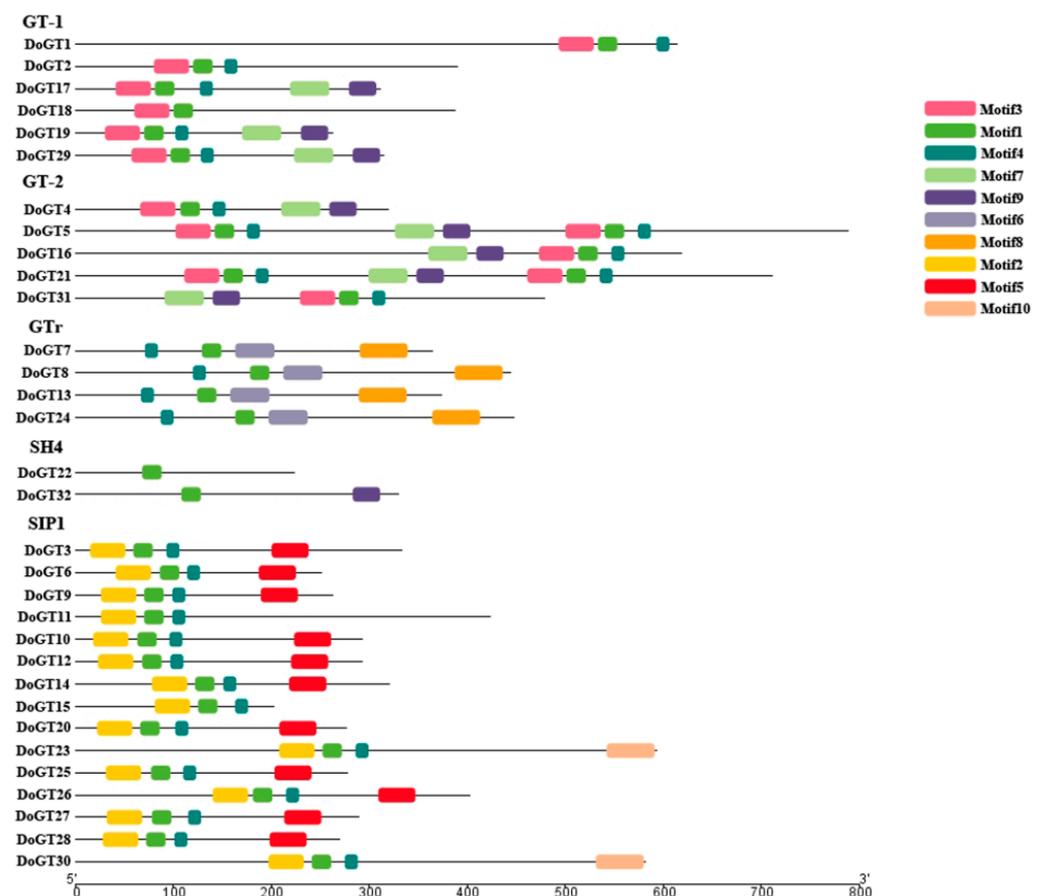


Figure 2. Conserved motif compositions of the *DoGT* genes in *D. officinale*. The conserved motifs were detected using MEME software and represented by colored boxes. The length of *DoGT* proteins can be estimated using the scale at the bottom and conserved motifs were shown in Table 2.

3.4. Exon-Intron Organization of DoGT Genes

To further investigate the evolution of *DoGT* genes, exon and intron structures of the *trihelix* gene family from *D. officinale* were visualized by CDS sequences with corresponding genomic sequences. As shown in Figure 3, the exon/intron structures were divergent among the *DoGT* genes. The number of exons spanned from 1 to 12. 14 *DoGT* genes (43.75%) contained 1 exon and 9 *DoGT* genes (28.13%) had 2 exons accounting for the

largest proportion, whereas only one *DoGT* gene harbored 12 exons, respectively. The results showed that the most closely related *trihelix* members in the same clade showed similar gene structures in intron numbers or exon lengths. The similarity in gene structures was consistent with the phylogenetic analysis. For example, all members of GT γ and SH4 subgroups contained one and two exons. In SIP1 subgroup, almost all *DoGTs* harbored one to two exons, with the exception of *DoGT11* (three exons), *DoGT23* (seven exons) and *DoGT30* (seven exons).

Table 2. Conserved motifs of *DoGT* proteins.

Motif ID	Conserved Motifs	Width	Sites
Motif 1	GYPRSPVQCKNKIENLKKRYK	21	32
Motif 2	WSEGETLALJDAYEEKWJSLNRGNLRAKDWEEVAATV	37	15
Motif 3	WPKQETRALIALRAELDRRFLESGPKPLWEEISARM	37	11
Motif 4	SSWPFKRLDALLR	14	29
Motif 5	GDVGEAEAJRKFGEGLRVERKKMEMMRELERERMEME	39	11
Motif 6	VVENPALLDSMSHVSHKAKDDVRKILSSKHLFYREMCA YHN	41	4
Motif 7	AAAFEGLLKZLMEQQEAMQQRFLETIERREQERMREEAWR	41	7
Motif 8	IQSEALELEKRRFKWQRFSSKKDRELEKMLKNERLRLNERMTLELRQK	50	4
Motif 9	REEERRAQERARAERDAAIISFLQKLTG	29	8
Motif 10	GFTAMDMTAISFCEENNIPVVLNLEPGNISRALCGDQVGTLLIDQSGRI	50	2

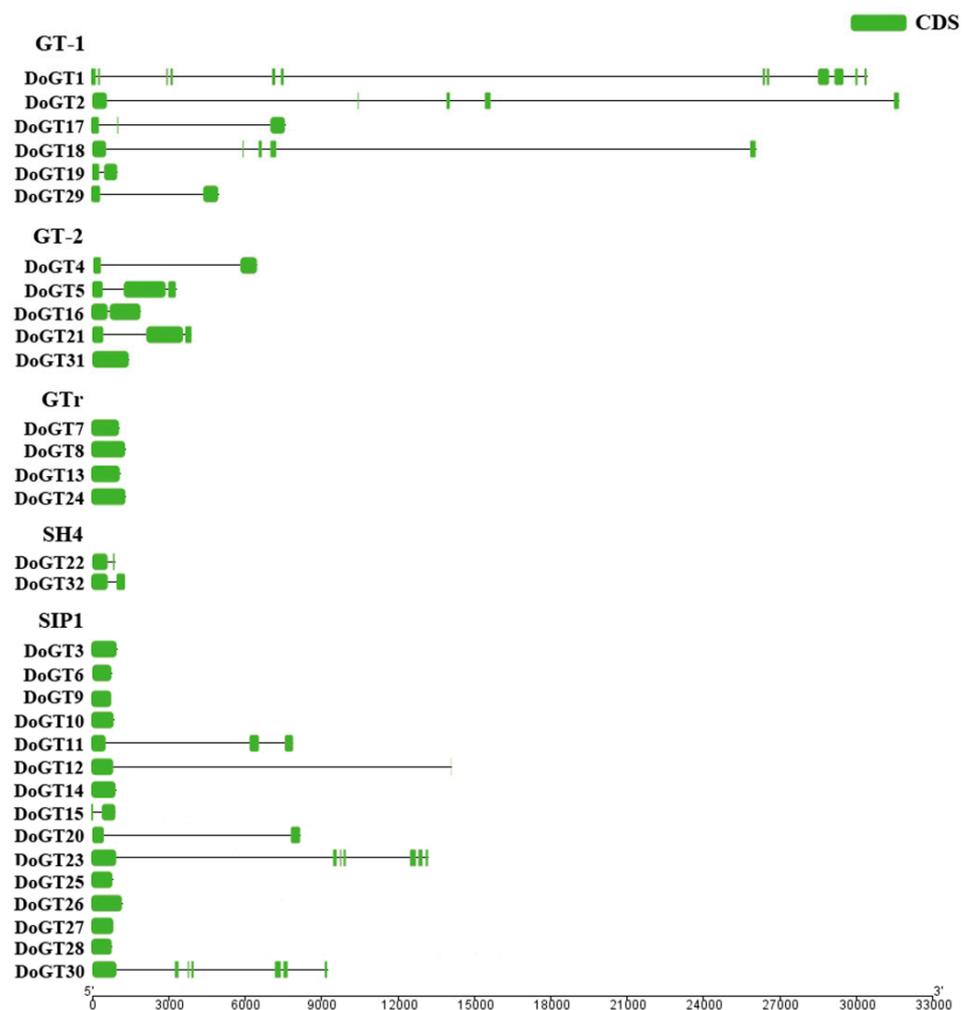


Figure 3. Gene structure compositions of the *DoGT* genes in *D. officinale*. Gene structure was analyzed through GSDS software. The green boxes indicate CDSs and black lines indicates introns.

3.5. Cis-Acting Elements Analysis and Gene Ontology (GO) Annotation

The *cis*-elements in promoter regions are associated with gene functions and expression patterns [43]. In order to investigate the evolution and functional divergence of *DoGT* genes, the upstream 1.5 Kb promoter regions of all *DoGT* genes in the *D. officinale* genome were analyzed using the PlantCARE online software. In this study, we examined the two types of *cis*-acting regulatory elements. One was related to plant development, including light responsive (box4, G-box, sp1 and MRE), endosperm expression (GCN4_motif), circadian (circadian control), meristem expression (CAT-box), meristem specific activation (CCGTCC-box) and zein metabolism regulation (O2-site); another one was associated with stress responses, including MeJA (methyl jasmonate) response elements (CGTCA-motif and TGACG-motif), ABA response (ABRE), heat and SA (salicylic acid) response (TCA element), anaerobic induction (ARE), drought stress (MBS), fungal elicitor response (Bow-W1), low-temperature stress (LTR) and so on. Our results showed that most of the *DoGT* genes contained the box4, MeJA response elements and MYB motifs, while only one *DoGT* gene (*DoGT5*) had circadian elements and two *DoGT* genes comprised MRE motif (Table S1).

In order to further predict the functions of *DoGT* proteins, gene ontology (GO) annotation analyses were performed. A total of 25 distinct functional groups were identified: 16 involved in biological processes, six involved in cellular components and three involved in molecular functions (Figure 4). In biological processes, GO classifications of “biological regulation”, “cellular process”, “metabolic process” and “regulation of biological process” were dominantly attributed. As for genes in the cellular component part, most *DoGT* genes were annotated with “cell”, “cell part” and “organelle”, while only three genes were associated with “macromolecular complex”. Under the molecular function term, 12 genes were annotated to have “nucleic acid binding transcription factor activity”, while only one gene was annotated to have “catalytic activity” (Figure 4 and Table S2). All these results indicate the multi-functions of *DoGT* genes.

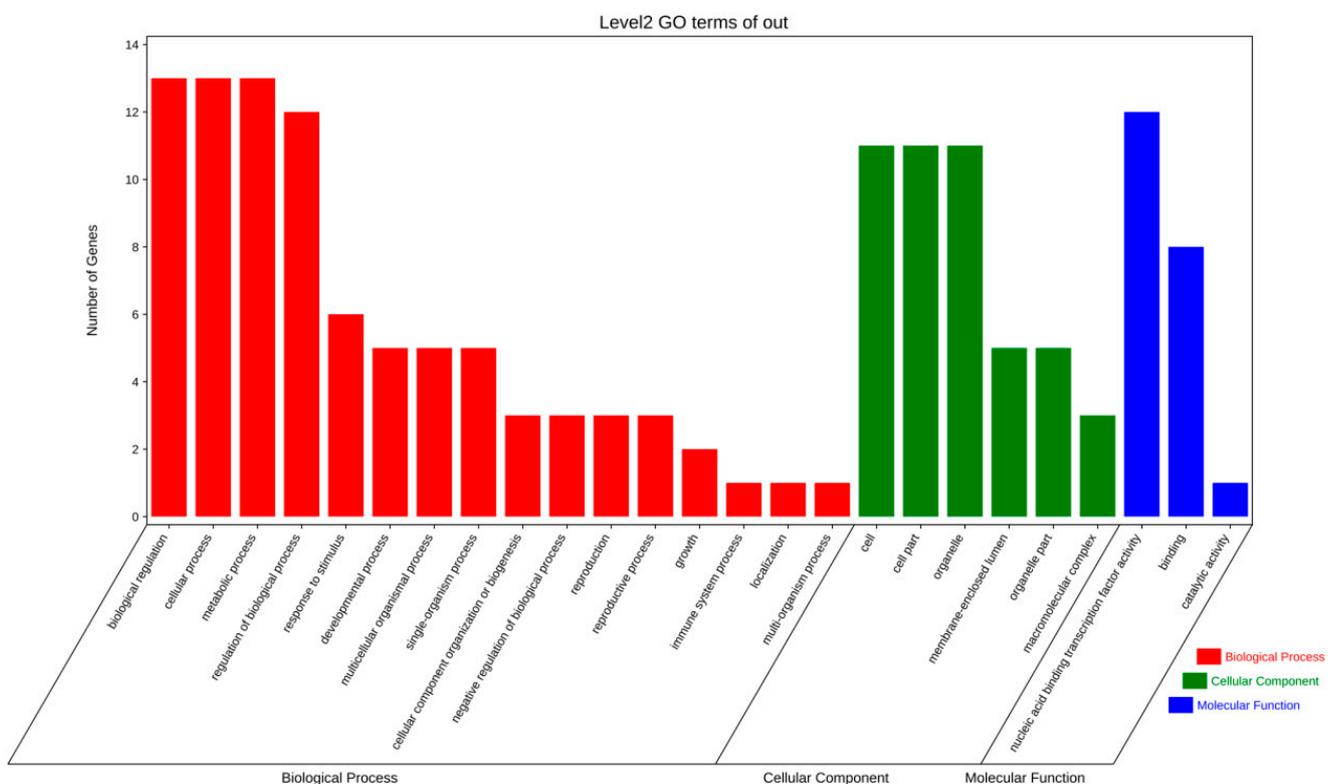


Figure 4. GO enrichment of *DoGT* transcription factors in *D. officinale*. According to the secondary terminology, the annotation results are divided into three ontology categories and distinguished by different colors.

3.6. DoGT Gene Expression Profiling

Previous studies reported that tissue-specific expression pattern of genes could reveal their potential biological functions [44,45]. Therefore, we downloaded the public Illumina RNA-seq data from three tissues (leaves, stems and flowers) to investigate the temporal and spatial expression levels in *D. officinale*. The gene expression profiles were calculated by FPKM value and the default empirical abundance threshold of 1 FPKM was used to evaluate whether a gene was positively expressed or not [46,47]. Of the 32 identified *DoGT* genes, only five members were actively expressed in all of the three tissues, while the remaining 27 (84.38%) genes were considered not expressed in at least one of the three tissues (Figure 5). All the top four highly expressed *DoGT* genes, including *DoGT5*, *DoGT27*, *DoGT28* and *DoGT22*, were observed in leaves, indicating their putative functions in the development and other physiological processes in leaves. Moreover, five genes (*DoGT4*, *DoGT19*, *DoGT24*, *DoGT29* and *DoGT32*) were found only up-regulated in flowers and two genes (*DoGT22* and *DoGT31*) shared similar high levels of gene expression in all of the three organs. Genes that are specific highly expressed in one tissue are often found to be able to regulate target genes involved in the processes of plant growth and development [47,48]. Therefore, tissue-specific *DoGT* genes reported herein might be valuable sources for further investigating their biological functions in the growth and development of *D. officinale*.

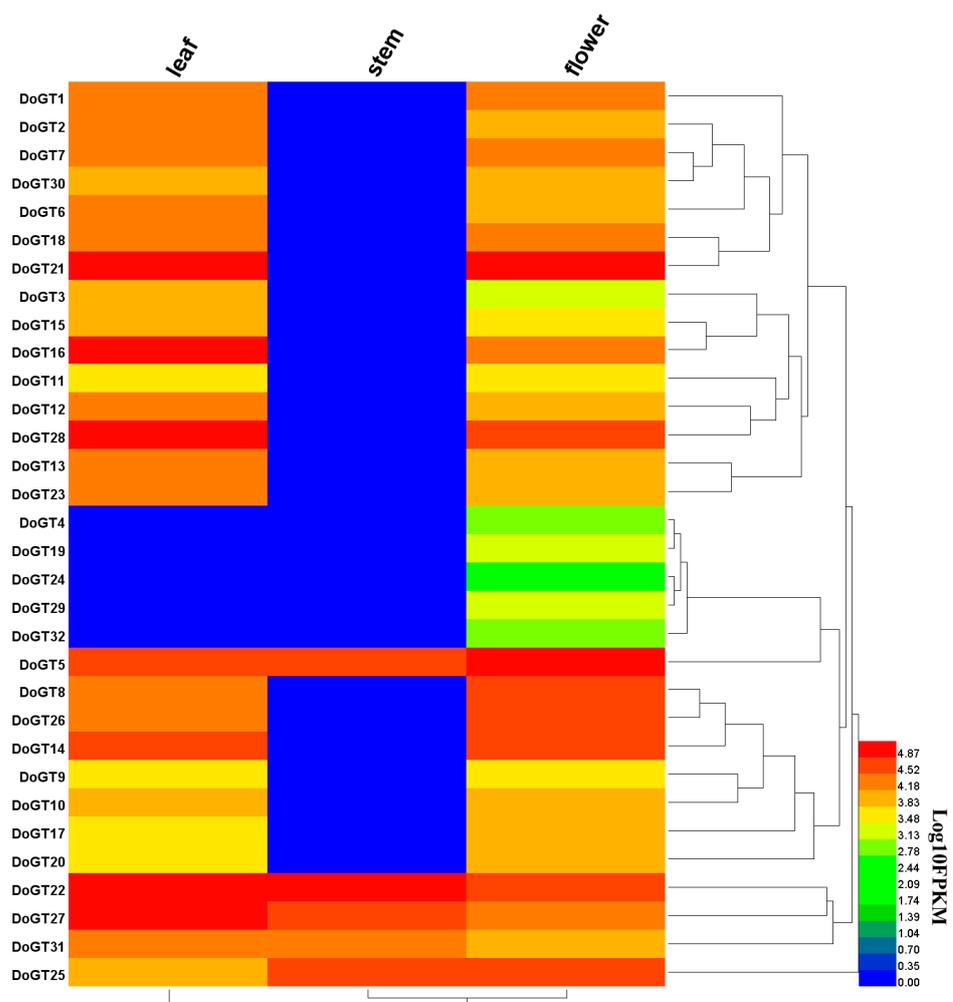


Figure 5. Tissue-specific expression profiles of *DoGT* genes among leaves, stems and flowers. The heatmap was drawn using R. Color bar represents log_{10} -transformed FPKM values. Red and blue colors indicate higher and lower expression levels, respectively.

3.7. Analysis of DoGT Genes Expression in Tissues

Ten *DoGT* genes were randomly selected from five subgroups to exhibit their expression profiling in *D. officinale*. The expression patterns of these 10 *DoGT* genes were examined in stems, leaves and flowers through qRT-PCR.

As illustrated in Figure 6, the expression patterns of those ten *DoGT* genes had a greater difference in these three organs. Meanwhile, the qRT-PCR analysis and transcriptome data of most genes were consistent. The expression levels of eight genes (*DoGT4*, *DoGT15*, *DoGT19*, *DoGT20*, *DoGT22*, *DoGT24*, *DoGT29* and *DoGT32*) were relatively higher in flowers. For example, *DoGT15* showed the greatest high expression in flower more than 57-fold in leaf. In contrast, *DoGT14* and *DoGT31* were relatively highly expressed in stems but lowly expressed in leaves.

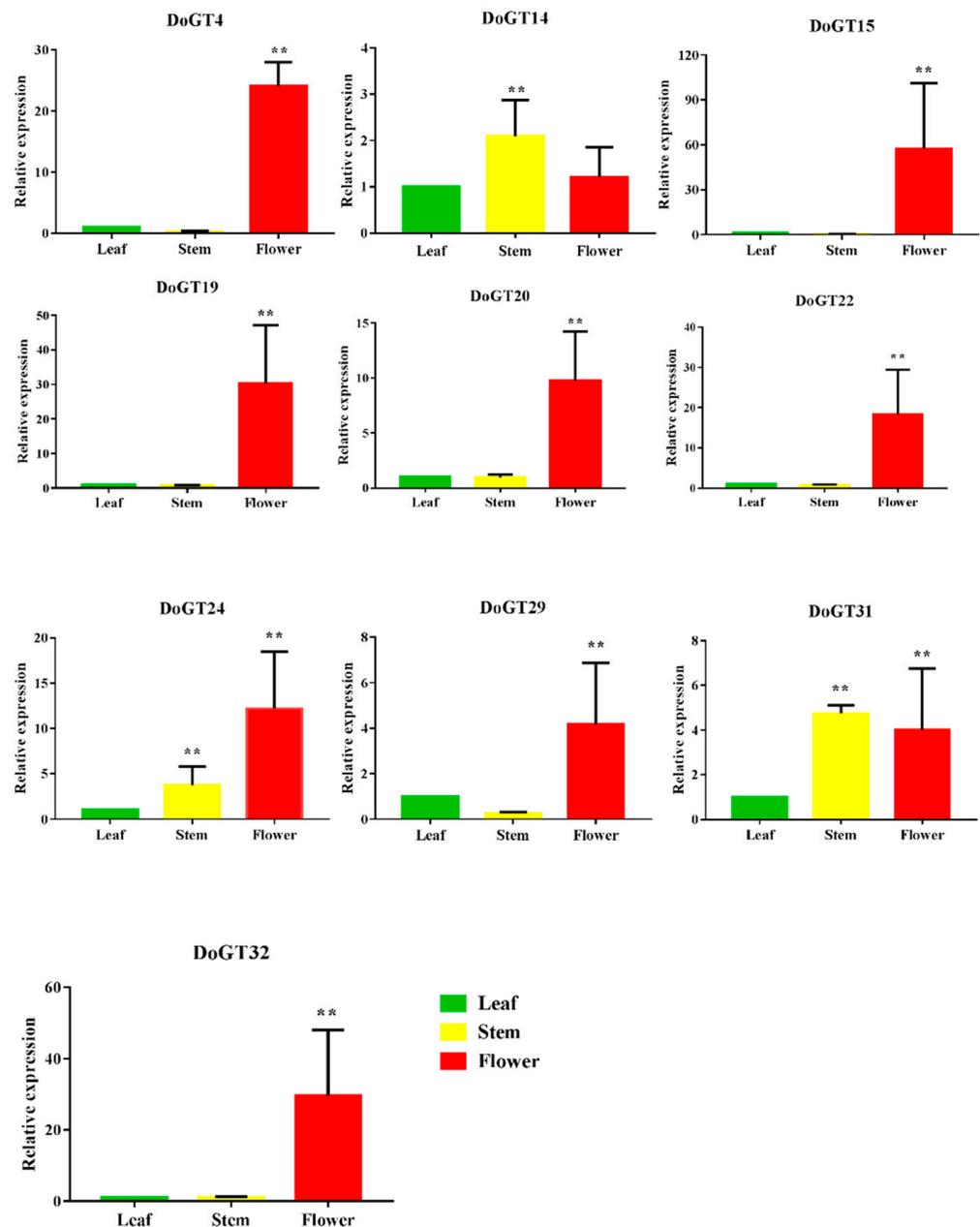


Figure 6. Verification of ten *GoGT* genes expression among leaves, stems and flowers in *D. officinale* by qRT-PCR. The expression of leaves was set to 1.0. Values are mean \pm SD of three replicates. **, very significant ($p < 0.01$) (Student's *t*-test).

3.8. Response of DoGT Genes to Cold Stress

Evidence has shown that *trihelix* genes important in response to cold stress [20,49,50]. To check the expression pattern of *DoGT* genes under cold stress, qRT-PCR was used to detect the expression levels of *DoGTs* at 0 °C. As showed in Figure 7, these selected ten *DoGT* genes showed different expression profiles under cold stress, suggesting that the *DoGT* genes were sensitive to cold stress. Some *DoGT* genes in leaves seemed to be more sensitive to cold stress. For example, *DoGT24* was significantly up-regulated by about 15-fold after 8 h. Our results also showed that the *DoGT* genes (*DoGT19*, *DoGT20* and *DoGT22*) belonging to different subgroups exhibited the same responses to cold stress: initially decreased, then increased at 8 h, and decreased again. Our results suggested that the 10 *DoGT* members were significantly affected under cold stress and these results indicated the involvement of *DoGT* gene family in their response against cold stress.

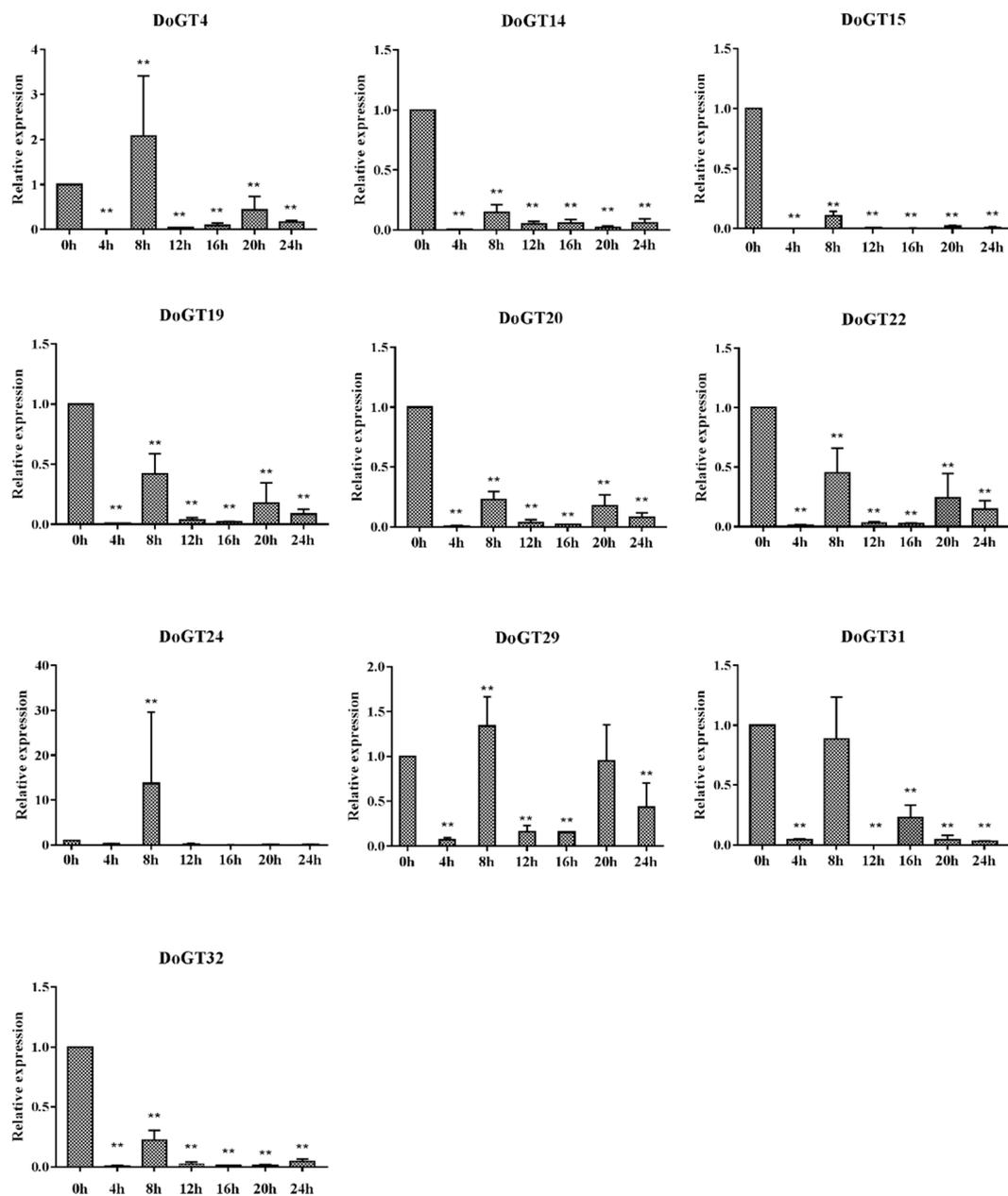


Figure 7. The real-time PCR analysis of ten *GoGT* genes in response to cold stress treatment in *D. officinale*. Statistically significant differences (Student's *t*-test) are indicated as followed: ** $p < 0.01$.

4. Discussion

Early studies suggested that the *trihelix* family genes contained three distinctive subfamilies (GT α , GT β and GT γ) [51]. Then, in 2012, Kaplan-Levy et al. proposed a new classification and divided the *trihelix* genes into five subgroups [7]. We compared the rice and *B. distachyon* sequences and constructed phylogenetic trees to classify the genes into five subgroups. The results supported previous findings. Members within same group had a similar gene structure, length and amino acid motif composition, indicating their close evolutionary relationship.

It was reported that genes within the same group might have similar functions because of sequence similarity. For example, *OsGT γ -1*, *OsGT γ -2* and *OsGT γ -3* in rice, both belonging to class GT γ , were involved in the stress responses, especially in salt stress [51].

In this current study, GO analyses showed that *DoGT* genes are grouped in 25 functional groups, including 16 involved in biological processes, six in cellular components and three in molecular functions, indicating the extensive functions of *trihelix* genes. Consistent with these results, many *cis*-acting elements were detected in promoters of *DoGT* genes and most of them were associated with plant growth/development and abiotic stress response.

We examined the expression patterns of *DoGT* genes in the flowers, stems and leaves of *D. officinale*. The expression profiles indicated spatial variations of *DoGT* expression in different tissues. In Figure 5, some *DoGTs* showed high expression levels in flowers, the same as function of previously report of *trihelix* genes in flower [51]. However, further research is necessary to make clear the functions of these *DoGT* genes.

Cold is a major abiotic stress that adversely influence plant growth and development. However, as described previously, *D. officinale* is often grown in the green house for saving from cold environment [52]. Therefore, understanding how respond and develops tolerance to cold stress is the first step towards improving the adaptation to cold environments. *DoGT* genes were significantly up-regulated/down-regulated after cold treatments, suggesting that these genes were associated with cold tolerance in leaves. It has been reported that analysis of *cis*-element distribution could provide insights into the functions of genes. Except for *DoGT20*, all of the 10 *DoGT* genes had MYB element which is related to stress response. In addition, *DoGT15* also contained a low-temperature response element (LTR). The obtained results may provide a number of the *DoGT* candidate genes for cold stress, which will facilitate the genetic improvement of cold tolerance in *D. officinale*.

5. Conclusions

Here, we first report a genome-wide analysis of *trihelix* genes in *D. officinale*. The classification and conserved domain of *DoGT* proteins, as well as stress-responsive elements in the promoters of *DoGT* genes were analyzed. Ten of the 32 *DoGT* genes were significant inducible, the results show that they might participate in response to cold stress. Our works will be helpful to increase knowledge about the molecular mechanism of these gene members in growth development and under cold stress in *D. officinale*. This research will expand our knowledge of the function of the *trihelix* family in cold stress regulation in plants.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2071-1050/13/5/2826/s1>, Figure S1. *DoGT* domain amino acid sequences. Figure S2. Maximum Likelihood (ML) phylogenetic tree for *trihelix* proteins in *D. officinale*, *B. distachyon* and rice. Table S1. Information of the promoter regions of *DoGT* genes. Table S2. GO enrichment of *DoGT* transcription factors in *D. officinale*. Table S3. Primers used in qRT-PCR analysis.

Author Contributions: Y.T. and Y.W. designed the project; Y.T. wrote the paper; Y.T. and H.H. performed the experiments and analyzed the data. Y.T. and H.H. contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (No. XDA20050204, XDA19050301 and XDA19050303) and grants from the Beijing DR PLANT Biotechnology Co., Ltd. (2019H078).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are available in this manuscript and supporting materials.

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

- Riechmann, J.; Heard, J.; Martin, G.; Reuber, L.; Jiang, C.-Z.; Keddie, J.; Adam, L.; Pineda, O.; Ratcliffe, O.; Samaha, R. *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes. *Science* **2000**, *290*, 2105–2110. [[CrossRef](#)] [[PubMed](#)]
- Lindemose, S.; O’Shea, C.; Jensen, M.K.; Skriver, K. Structure, Function and Networks of Transcription Factors Involved in Abiotic Stress Responses. *Int. J. Mol. Sci.* **2013**, *14*, 5842–5878. [[CrossRef](#)]
- Mizoi, J.; Shinozaki, K.; Yamaguchi-Shinozaki, K. AP2/ERF family transcription factors in plant abiotic stress responses. *BBA Gene Regul. Mech.* **2012**, *1819*, 86–96. [[CrossRef](#)]
- Nagano, Y. Several features of the GT-factor trihelix domain resemble those of the Myb DNA-binding domain. *Plant Physiol.* **2000**, *124*, 491–494. [[CrossRef](#)] [[PubMed](#)]
- Nagano, Y.; Inaba, T.; Furuhashi, H.; Sasaki, Y. Trihelix DNA-binding protein with specificities for two distinct *cis*-elements: Both important for light downregulated and dark-inducible gene expression in higher plants. *J. Biol. Chem.* **2001**, *276*, 22238–22243. [[CrossRef](#)] [[PubMed](#)]
- Zhou, D.X. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends Plant Sci.* **1999**, *4*, 210–214. [[CrossRef](#)]
- Kaplan-Levy, R.N.; Brewer, P.B.; Quon, T.; Smyth, D.R. The trihelix family of transcription factors—Light, stress and development. *Trends Plant Sci.* **2012**, *17*, 163–171. [[CrossRef](#)]
- Ayadi, M.; Delaporte, V.; Li, Y.F.; Zhou, D.X. Analysis of GT-3a identifies a distinct subgroup of trihelix DNA-binding transcription factors in *Arabidopsis*. *FEBS Lett.* **2004**, *562*, 147–154. [[CrossRef](#)]
- Kuhn, R.M.; Caspar, T.; Dehesh, K.; Quail, P.H. DNA binding factor GT-2 from *Arabidopsis*. *Plant Mol Biol.* **1993**, *23*, 337–348. [[CrossRef](#)]
- Xie, Z.M.; Zou, H.F.; Lei, G.; Wei, W.; Zhou, Q.Y.; Niu, C.F.; Liao, Y.; Tian, A.G.; Ma, B.; Zhang, W.K. Soybean Trihelix transcription factors GmGT-2A and GmGT-2B improve plant tolerance to abiotic stresses in transgenic *Arabidopsis*. *PLoS ONE* **2009**, *4*, e6898. [[CrossRef](#)]
- Ni, M.; Dehesh, K.; Tepperman, J.M.; Quail, P.H. GT-2: In vivo transcriptional activation activity and definition of novel twin DNA binding domains with reciprocal target sequence selectivity. *Plant Cell* **1996**, *8*, 1041–1059.
- Dehesh, K.; Smith, L.G.; Tepperman, J.M.; Quail, P.H. Twin autonomous bipartite nuclear localization signals direct nuclear import of GT-2. *Plant J. Cell Mol. Biol.* **2010**, *8*, 25–36. [[CrossRef](#)] [[PubMed](#)]
- Griffith, M.E.; Conceicao, A.D.; Smyth, D.R. *PETAL LOSS* gene regulates initiation and orientation of second whorl organs in the *Arabidopsis* flower. *Development* **1999**, *126*, 5635–5644. [[PubMed](#)]
- Brewer, P.B.; Howles, P.A.; Dorian, K.; Griffith, M.E.; Ishida, T.; Kaplan-Levy, R.N.; Kilinc, A.; Smyth, D.R. *PETAL LOSS*, a trihelix transcription factor gene, regulates perianth architecture in the *Arabidopsis* flower. *Development* **2004**, *131*, 4035–4045. [[CrossRef](#)] [[PubMed](#)]
- Lampugnani, E.R.; Kilinc, A.; Smyth, D.R. *PETAL LOSS* is a boundary gene that inhibits growth between developing sepals in *Arabidopsis thaliana*. *Plant J.* **2012**, *71*, 724–735. [[CrossRef](#)]
- Kitakura, S.; Fujita, T.; Ueno, Y.; Terakura, S.; Wabiko, H.; Machida, Y. The protein encoded by oncogene 6b from agrobacterium tumefaciens interacts with a nuclear protein of tobacco. *Plant Cell* **2002**, *14*, 451–463. [[CrossRef](#)] [[PubMed](#)]
- Ming-Jun, G.; Xiang, L.; Helen, L.; Gropp, G.M.; Lydiate, D.D.; Shu, W.; Hegedus, D.D. ASIL1 is required for proper timing of seed filling in *Arabidopsis*. *Plant Signal Behav.* **2011**, *6*, 1886–1888.
- Luo, J.; Tang, S.; Mei, F.; Peng, X.; Li, J.; Li, X.; Yan, X.; Zeng, X.; Liu, F.; Wu, Y. *BnSIP1-1*, a Trihelix Family Gene, Mediates Abiotic Stress Tolerance and ABA Signaling in *Brassica napus*. *Front. Plant Sci.* **2017**, *8*, e71136. [[CrossRef](#)]
- Lin, Z.W.; Griffith, M.E.; Li, X.R.; Zhu, Z.F.; Tan, L.B.; Fu, Y.C.; Zhang, W.X.; Wang, X.K.; Xie, D.X.; Sun, C.Q. Origin of seed shattering in rice (*Oryza sativa* L.). *Planta* **2007**, *226*, 11–20. [[CrossRef](#)]
- Yu, C.; Song, L.; Song, J.; Ouyang, B.; Guo, L.; Shang, L.; Wang, T.; Li, H.; Zhang, J.; Ye, Z. ShCIGT, a Trihelix family gene, mediates cold and drought tolerance by interacting with SnRK1 in tomato. *Plant Sci.* **2018**, *270*, 140–149. [[CrossRef](#)] [[PubMed](#)]
- Zhu, G.H.; Ji, Z.H.; Wood, J.J.; Wood, H.P. *Dendrobium*. In *Flora of China*; Wu, C.Y., Raven, P.H., Hong, D.Y., Eds.; Science Press: Beijing, China, 2009; pp. 367–397.
- Zhao, M.M.; Zhang, G.; Zhang, D.W.; Hsiao, Y.Y.; Guo, S.X. ESTs analysis reveals putative genes involved in symbiotic seed germination in *Dendrobium officinale*. *PLoS ONE* **2013**, *8*, e72705. [[CrossRef](#)]
- Ling, H.; Zeng, X.; Guo, S. Functional insights into the late embryogenesis abundant (LEA) protein family from *Dendrobium officinale* (Orchidaceae) using an *Escherichia coli* system. *Sci. Rep.* **2016**, *6*, 39693. [[CrossRef](#)]

24. Lavarack, B.; Harris, W.; Stocker, G. *Dendrobium and Its Relatives*; Timber Press: Portland, OR, USA, 2000.
25. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [[CrossRef](#)]
26. Letunic, I.; Doerks, T.; Bork, P. SMART 7: Recent updates to the protein domain annotation resource. *Nucleic Acids Res.* **2011**, *40*, D302–D305. [[CrossRef](#)]
27. Chou, K.C.; Shen, H.B. Plant-mPLOC: A top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS ONE* **2010**, *5*, e11335. [[CrossRef](#)] [[PubMed](#)]
28. Goodstein, D.M.; Shu, S.; Howson, R.; Neupane, R.; Hayes, R.D.; Fazo, J.; Mitros, T.; Dirks, W.; Hellsten, U.; Putnam, N.; et al. Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res.* **2012**, *40*, D1178–D1186. [[CrossRef](#)] [[PubMed](#)]
29. Xiao, J.; Hu, R.; Gu, T.; Han, J.; Qiu, D.; Su, P.; Feng, J.; Chang, J.; Yang, G.; He, G. Genome-wide identification and expression profiling of *trihelix* gene family under abiotic stresses in wheat. *BMC Genom.* **2019**, *20*, 287. [[CrossRef](#)]
30. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [[CrossRef](#)] [[PubMed](#)]
31. Hu, B.; Jin, J.; Guo, A.-Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* **2014**, *31*, 1296–1297. [[CrossRef](#)]
32. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208. [[CrossRef](#)] [[PubMed](#)]
33. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327. [[CrossRef](#)] [[PubMed](#)]
34. Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240. [[CrossRef](#)] [[PubMed](#)]
35. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protoc.* **2016**, *11*, 1650–1667. [[CrossRef](#)] [[PubMed](#)]
36. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550–558. [[CrossRef](#)]
37. Rong, T.J.; Wen, G.S.; Qian, X.; Zha, Y.H.; Xu, S.Z. Comparison on Two Modified Total RNA Isolation Methods from *Dendrobium officinale*. *J. Yunnan Agric. Univ.* **2012**, *27*, 703–707.
38. Gao, M.-J.; Lydiate, D.J.; Li, X.; Lui, H.; Gjetvaj, B.; Hegedus, D.D.; Rozwadowski, K. Repression of seed maturation genes by a trihelix transcriptional repressor in *Arabidopsis* seedlings. *Plant Cell* **2009**, *21*, 54. [[CrossRef](#)]
39. Wang, C.W.; Wang, Y.; Pan, Q.; Chen, S.K.; Feng, C.Z.; Hai, J.B.; Li, H.F. Comparison of Trihelix transcription factors between wheat and *Brachypodium distachyon* at genome-wide. *BMC Genom.* **2019**, *20*, 142. [[CrossRef](#)]
40. Qin, Y.; Ma, X.; Yu, G.; Wang, Q.; Wang, L.; Kong, L.; Kim, W.; Wang, H.W. Evolutionary history of trihelix family and their functional diversification. *DNA Res.* **2014**, *21*, 499–510. [[CrossRef](#)]
41. Song, A.; Wu, D.; Fan, Q.; Tian, C.; Chen, S.; Guan, Z.; Xin, J.; Zhao, K.; Chen, F. Transcriptome-wide identification and expression profiling analysis of chrysanthemum trihelix transcription factors. *Int. J. Mol. Sci.* **2016**, *17*, 198. [[CrossRef](#)]
42. Wang, Z.; Liu, Q.; Wang, H.; Zhang, H.; Xu, X.; Li, C.; Yang, C. Comprehensive analysis of *trihelix* genes and their expression under biotic and abiotic stresses in *Populus trichocarpa*. *Sci. Rep.* **2016**, *6*, 36274. [[CrossRef](#)]
43. Todeschini, A.L.; Georges, A.; Veitia, R.A. Transcription factors: Specific DNA binding and specific gene regulation. *Trends Genet.* **2014**, *30*, 211–219. [[CrossRef](#)] [[PubMed](#)]
44. Luo, M.; Dennis, E.S.; Berger, F.; Peacock, W.J.; Chaudhury, A. *MINISEED3 (MINI3)*, a *WRKY* family gene, and *HAIKU2 (IKU2)*, a leucine-rich repeat (*LRR*) *KINASE* gene, are regulators of seed size in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA.* **2005**, *102*, 17531–17536. [[CrossRef](#)]
45. Yu, F.; Huaxia, Y.; Lu, W.; Wu, C.; Cao, X.; Guo, X. *GhWRKY15*, a member of the *WRKY* transcription factor family identified from cotton (*Gossypium hirsutum* L.), is involved in disease resistance and plant development. *BMC Plant Biol.* **2012**, *12*, 144. [[CrossRef](#)]
46. González-Porta, M.; Frankish, A.; Rung, J.; Harrow, J.; Brazma, A. Transcriptome analysis of human tissues and cell lines reveals one dominant transcript per gene. *Genome Biol.* **2013**, *14*, R70. [[CrossRef](#)]
47. Yao, Q.Y.; Xia, E.H.; Liu, F.H.; Gao, L.Z. Genome-wide identification and comparative expression analysis reveal a rapid expansion and functional divergence of duplicated genes in the *WRKY* gene family of cabbage, *Brassica oleracea* var. capitata. *Gene* **2015**, *557*, 35–42. [[CrossRef](#)] [[PubMed](#)]
48. Nan, H.; Gao, L.Z. Genome-wide analysis of *WRKY* genes and their response to hormone and mechanic stresses in carrot. *Front. Genet.* **2019**, *10*, 363. [[CrossRef](#)]
49. Xi, J.; Qiu, Y.; Du, L.; Poovaiah, B.W. Plant-specific trihelix transcription factor *AtGT2L* interacts with calcium/calmodulin and responds to cold and salt stresses. *Plant Sci.* **2012**, *185–186*, 274–280. [[CrossRef](#)]
50. Xu, H.; Shi, X.; He, L.; Guo, Y.; Zang, D.; Li, H.; Zhang, W.; Wang, Y. *Arabidopsis thaliana* trihelix transcription factor *AST1* mediates salt and osmotic stress tolerance by binding to a novel AGAG-Box and some GT motifs. *Plant Cell Physiol.* **2018**, *59*, 946–965. [[CrossRef](#)] [[PubMed](#)]

-
51. Fang, Y.; Xie, K.; Hou, X.; Hu, H.; Xiong, L. Systematic analysis of GT factor family of rice reveals a novel subfamily involved in stress responses. *Mol. Genet. Genom.* **2010**, *283*, 157–169. [[CrossRef](#)]
 52. Wu, Z.G.; Jiang, W.; Chen, S.L.; Mantri, N.; Tao, Z.M.; Jiang, C.X. Insights from the cold transcriptome and metabolome of *Dendrobium officinale*: Global reprogramming of metabolic and gene regulation networks during cold acclimation. *Front. Plant Sci.* **2016**, *7*, 1653. [[CrossRef](#)]