DOI: 10.1007/s10535-017-0767-y BIOLOGIA PLANTARUM 62 (2): 287-298, 2018

Identification of gene co-expression networks involved in cold resistance of *Lilium lancifolium*

Y.B. YONG, W.Q. LI, J.M. WANG, Y. ZHANG, and Y.M. LU*

Beijing Key Laboratory of Ornamental Germplasm Innovation and Molecular Breeding, China National Engineering Research Center for Floriculture, College of Landscape Architecture, Beijing Forestry University, Beijing, 100083, P.R. China

Abstract

Low temperature can affect the growth and development of lily, limiting the application of commercial cultivars in outdoor. *Lilium lancifolium* is an important cold-resistant wild lily, but little is known about how *L. lancifolium* tolerates cold stress at the molecular level. In this study, we identified and characterized genes and transcription factors associated with cold stress in control plants and plants treated by 4° C for 1 - 24 h. The construction of a highest reciprocal rank-based gene co-expression network along with its partition into defined functional modules using Markov cluster algorithm resulted in identification of 30 gene modules, and some of them were significantly enriched with various kinds of stress response under 4° C. These gene modules were associated with metabolic processes, cellular processes, regulation of biological processes, establishment of localization, and responses to stimuli. Moreover, three transcription factors that may regulate the downstream genes involved in response to stimuli were also found. We further studied the expression pattern and tissue specificity of these transcription factors. The functional evaluation of the various interesting genes in this study will probably provide novel discovery of pathway members and regulators associated with cold resistance in lily.

Additional key words: differentially expressed genes, Markov cluster algorithm, transcription factors.

Introduction

Lily is one of the most important flower crops. However, low temperature is the most important environmental constraint for lily outdoor planting. A wild relative *Lilium lancifolium* can acclimate to low temperatures up to -35 °C in winter and continuously germinate in the next spring (Wang *et al.* 2014b), suggesting that it may have distinctive molecular mechanisms of cold resistance. Cold acclimation is the main process involved in increasing cold resistance of some temperate plant species (Qu *et al.* 2015). Many of the biochemical and physiological changes occurs in plants during cold acclimation. These changes are mediated by the differential expression of certain genes in response to cold stress (for review see Chinnusamy *et al.* 2007). Moreover, induction of these cold responsive genes occurs mainly at the transcriptional level, and modification of the temporal and spatial expression patterns of specific cold responsive genes is an important part of cold acclimation (Bakshi and Oelmuller 2014). Numerous studies have demonstrated the vital roles of transcription factors (TFs) in the regulation of plant stress responses. Low temperatures trigger the expression of the cold responsive TFs, which in turn activate many downstream genes that confer cold tolerance to plants. Thus, detection and identification of cold responsive TFs contribute significantly to our understanding of the mechanism of cold tolerance in plants, and is beneficial for further research and in-depth exploration of cold resistance genes for breeding. Apart from the APETALA2/ethylene response factor (AP2/ERF) family,

Submitted 25 May 2017, *last revision* 16 July 2017, *accepted* 19 July 2017.

Abbreviations: CK - control; CT - cold treatment; DEG - differentially expressed gene; FDR - false discovery rate; GCN - gene co-expression network; GO - gene ontology; HRR - highest reciprocal rank; MCL - Markov cluster algorithm; PA - polyamine; RPKM - reads per kb of exon model per million mapped reads; RT-qPCR - reverse transcription quaintitative PCR; TF - transcription factor.

Acknowledgements: This study was supported by the China National Natural Science Foundation (Grant Nos. 31672190, 31272204, 31071815), the '863' research program (No. 2011AA10020804), and the D. Programs Foundation of the Ministry of Education of China (No. 20110014110006).

* Correspondence author; e-mail: bjluyingmin@163.com

which is known for playing an important regulatory role in signal transduction during various stresses (Du *et al.* 2016), a subset of Zinc finger family members, including WRKY and Zinc finger-homeodomain (ZFHD subfamilies, are also involved in responses to cold stress (Chen *et al.* 2012, Wang *et al.* 2016). As well as NAC (Tran *et al.* 2010), MYB (Agarwal *et al.* 2006) and basic leucine zipper (bZIP; Liu *et al.* 2012) family members are characterized as cold-inducible TFs. However, little is known about the role of these TFs in determining cold resistance in lily.

Next-generation sequencing technology (*e.g.*, *Illumina*/*Solexa*-based RNA-sequencing) has been successfully applied to the transcriptome sequencing of many plant species and has become an effective method for a comprehensive analysis of plant gene expression patterns under abiotic stresses. However, ordinary transcriptome analysis can only provide less information about gene-gene interrelationships (Bonthala *et al.* 2016). One promising approach to study the interactions between genes under biotic and/or abiotic stresses is network-based analyses like gene co-expression analysis (GCA) (Zhang *et al.* 2012, Wong *et al.* 2014). Gene co-expression networks (GCNs) are constructed from expression data generated by using either microarrays or deep-sequencing (RNA-seq) methods. Genes are usually represented as nodes, whereas edges represent pairwise relationships between nodes. A set of closely connected nodes represents a cluster, and the entire set of nodes, edges, and clusters forms the GCN (Wong *et al.* 2014). Usually, co-expressed genes within a cluster are expected to be functionally related to genes with a similar expression pattern. Furthermore, graph clustering algorithms such as Markov cluster algorithm (MCL; Enright *et al.* 2002), Heuristic cluster chiseling algorithm (HCCA; Mutwil *et al.* 2010) and weighted correlation network analysis (WCGNA; Langfelder and Horvath 2008) have been widely used to partition of the complex gene co-expression network to defined functional modules (Wong *et al.* 2014). This approach not only reveals gene-gene interrelationships but helps in the identification of candidate modules and genes under a specific stress conditions. Therefore, the goals of this study were: *1*) to identify the genes that are differentially expressed in 4° C for 0, 2, and 16 h using RNA-seq based transcriptome analysis; *2*) to construct a GCN from the differentially expressed genes; *3*) to divide GCN into individual modules and to identify functional modules that are enriched in cold stress; *4*) to detect the potential key functional genes and/or TFs and to analyze possible interrelationships between these genes during cold acclimation in *L. lancifolium*.

Materials and methods

Plants: Bulbs of *Lilium lancifolium* Thunb. purchased from a local nursery in Heilongjiang Province were cleaned, disinfected, and then stored in moist sawdust at 4° C for one month. Then they were cultivated at Beijing Forestry University greenhouse (116.3° E, 40.0° N) under day/night temperatures of 25/18 °C and a relative humidity of 70 % and irrigated every 3 d. After 4 - 8 weeks of cultivation, the seedlings were subjected to cold treatments [4° C for 0 (control), 1, 2, 3, 6, 12, 16, and 24 h, respectively]. At each time point, samples of leaves, bulbs, roots, and stems were collected, immediately frozen in liquid nitrogen, and stored at -80 °C before RNA extraction.

RNA extraction and RNA-seq library construction: Total RNA was extracted from the leaves of control and cold-treated (2 and 16 h) plants using an *RNAisomate RNA Easyspin* isolation system (*Aidlab Biotech*, Beijing, China) according to the manufacturer's instructions. A total of 3 RNA-seq libraries were constructed as described previously (Wang *et al.* 2014b) using the *Solexa* mRNA-seq platform (*Illumina*, San Diego, CA, USA). Sequencing using an *Illumina HiSeq™2000* platform was performed at the *ShoBiotechnology Corporation* (SBC), Shanghai, China, following the manufacturer's protocols. Three samples were sequenced independently as biological replicates.

RNA-seq data analysis: Using scaffolding contig methods with a minimum contig length of ≥ 400, *de novo* assembly was carried out with *CLC Genomics Workbench* (*v. 5.5*). Primary unigenes from the assembled *de novo* sequences of three samples were assembled using *CAP3 ES*, yielding final unigenes. Assembled final unigenes were used for *BLASTx* searches (E-value < 1e-5) against the *UniProt* database and the *Swiss-Prot* protein database (Wang *et al.* 2014a). *Blast2GO* program was used to assign Go Ontology (GO) terms. The false discovery rate (FDR) was used to determine the threshold of the *P*-value in multiple tests and analyses. Gene expression levels were calculated and normalized by reads per kb of exon model per million mapped reads (RPKM). A gene was defined as expressed if its RPKM value was ≥ 2 in at least one of the three transcriptomes (Wang *et al.* 2014a).

Identification of differentially expressed genes and construction of gene co-expression network: A differentially expressed gene (DEG) was declared if the associated FDR ≤ 0.05 and $|\log_2 (\text{ratio})| \geq 1.5$ were observed in three pairwise transcriptome comparisons (Choe *et al.* 2016). The heat-map of the total DEGs was generated using *MeV4.9* and clustered by hierarchical clustering (HCL) with default parameters. Highest-reciprocal rank (HRR) based co-expression network

methodology was employed to further investigate the interactions between the DEGs by constructing gene co-expression network (GCN). Correlation matrices were first calculated using all DEGs' RPKM with the Pearson's correlation coefficient (*r*) to define similarity of expression levels between DEGs. Next, raw *r* values for every relationship between DEGs were transformed into HRR (Mutwil *et al.* 2010). The HRR-based GCN was calculated using *R version 3.3.1*. (https://www.r-project.org/) with the parameter of HRR = 30, and then the GCN were visualized using *Cytoscape* (Smoot *et al.* 2011). A *Cytoscape Plugin Cluster Maker* (Morris *et al.* 2011), was used to detect the MCL modules using inflation score (*I*) parameter of 1.6. For network analysis, the *Cytoscape Plugin Network Analyzer* (Assenov *et al.* 2008) was used. Significantly enriched GO terms for each of the detected modules was carried out by *AgriGO* tool (Du *et al.* 2010). The *Arabidopsis* gene model (*TAIR9*) was chosen as the reference set, hypergeometric distribution adjusted by Hochberg FDR adjustment for the testing of multiple hypotheses with an adjusted threshold of $P < 0.05$ and a minimum number of mapping entries of 2 were used to evaluate the statistical significance of the functional enrichment in detected modules.

Reverse transcription quantitative PCR analysis: First-strand cDNA synthesis was performed using *Superscript II* reverse transcriptase (*Invitrogen*, Carlsbad, CA, USA) according to the manufacturer's instructions, using 1 µg total RNA and oligo (dT) primers. RT-qPCR was performed using a *Rotor-Gene 3000* real-time PCR detection system (*Qiagen*, Hilden, Germany) using *SYBR*® qPCR mix (*Toyobo*, Tokyo, Japan) according to the manufacturer's protocol. The primers used in this study were designed with *Primer Premier 5* (*Premier*, Palo Alto, CA, USA) and are listed in Table 1. All reactions were performed in three biological triplicates. Relative mRNA content was calculated using the $2^{-\Delta\Delta Ct}$ method against the internal reference gene encoding tonoplast intrinsic protein 1 (*TIP1*; Wang *et al.* 2014b), with expression at 0 h used as the internal control.

Table 1. Primers used for RT-qPCR.

Gene name	Forward primer sequence (5'- 3')	Reverse primer sequence (5'- 3')
MYB12	ATCTCATCATCAGGCTCCA	ACCTCAGACGAGTGTTCCA
ZFHD1	GTTTGCTGAAAGAGTTG	TTCTGCTTGTTGTTGTG
NAC	CGGAGGTCGGATTGA	TGGATGATTGGGTCTTG
TIP1 (reference gene)	CGAAGCCAGAAACGGAGAAGAAT	GGGTAGGGTGGATTGGGAAGA

Results and discussion

Transcriptome comparison was performed to investigate the transcript abundance changes in *L. lancifolium* subjected to cold stress [4 °C for 2 h (CT2h) or 16 h (CT16h)]. Total mapped reads were used for differential gene expression analysis and GCN construction. In total, approximately 74 million raw reads for the two cold-treated samples (CT2h and CT16h) and 41 million for the control sample (CK0h) were obtained after discarding low-quality reads (Wang *et al.* 2014b). After stringent quality checks and data cleaning, we obtained 115 421 520 raw reads containing a total of 11.6 Gb nucleotides (Wang *et al.* 2014a). The average read size was 90 bp, and the Q20 percentage (sequencing error rate < 1 %) and GC (guanine + cytosine) percentage were 98.1 and 43.7 %, respectively (Wang *et al.* 2014b). Based on the high quality reads, 46 516 contigs were assembled with an average length of 793 bp (Wang *et al.* 2014a). After local assembly with the unmapped ends to fill in the

Table 2. Simple parameters of constructed gene co-expression network (GCN).

Simple parameters			
Clustering coefficient	0.867	Avg. number of neighbours	12.793
Connected components	84	Number of nodes	1238
Network diameter	5	Network density	0.010
Network radius	1	Network heterogeneity	0.467
Network centralization	0.016	Isolated nodes	0
Shortest paths	26916 (1 %)	Number of self-loops	0
Characteristic path length	1.541	Multi-edge node pairs	15

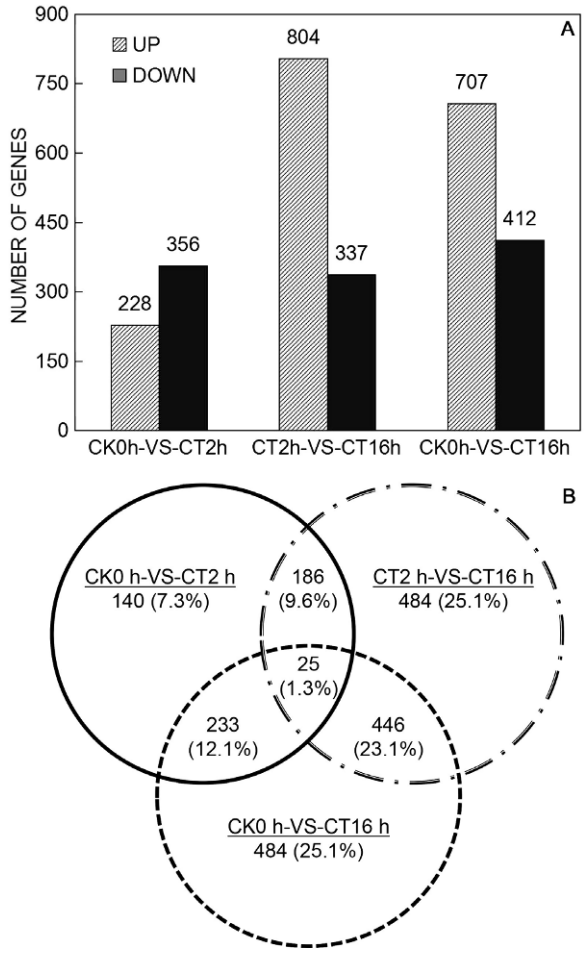

Fig. 1. Changes in gene expression profile among the different cold stress stages (*A*). The number of up-regulated and down-regulated genes between CK0h *vs*. CT2h, CK0h *vs*. CT16h, and CT2h *vs*. CT16h (*B*). Venn diagrams identify common DEGs from pairwise comparisons .

small gaps within the scaffolds, the *de novo* assembly yielded 37 843 unigenes with an average length of 971 bp, and 18 736 of them had significant similarity to known proteins in the *Swiss-Prot* database (Wang *et al.* 2014a).

Based on the transcriptome profiling results, 1 929 DEGs were identified from the 18 736 annotated genes (Fig 1 Suppl.). By comparing the RPKM of DEGs between samples, we found that the number of up-regulated genes during cold treatments was higher than the number of down-regulated genes (Fig 1*A*). Using the Venn diagram analysis, we identified 25 common genes in three pairwise transcriptome comparisons (Fig 1*B*), such as *LEA1* (encoding late embryogenesis abundant protein 1), *SDD* (sphingolipid delta-8 desaturase), *COX1&2* (cytochrome *c* oxidase subunit 1&2), and *MT-CYB* (cytochrome *b*) (Contig17189, Contig22718, Contig12966, Contig20456, and Contig9065). They were constantly and dramatically up/down-regulated under cold stress for 16 h, suggesting that these genes always play an important role in the cold acclimation. Furthermore, 186 common DEGs were found between in CK0h and CT2h as well as CT2h and CT16h transcriptomes, but not between CK0h and CT16h, such as *GER* (encoding GASA-like protein), *POD3* (pericarp peroxidase 3), *sHsp* (small heat shock protein), *CDPK* (calcium-dependent protein kinase), *MLP22* (major latex protein 22), and *POD2* (cationic peroxidase 2) (Contig13386, Contig15570, Contig17538, Contig22048, Contig2504, and Contig2543). The expressions of these genes were nearly decreased back to the original level (CK0h) after activation or repression by cold stimulus, and so the function of these genes in cold acclimation was probably instantaneous.

We used the MCL to detect biologically related modules in GCN, which resulted in 30 modules

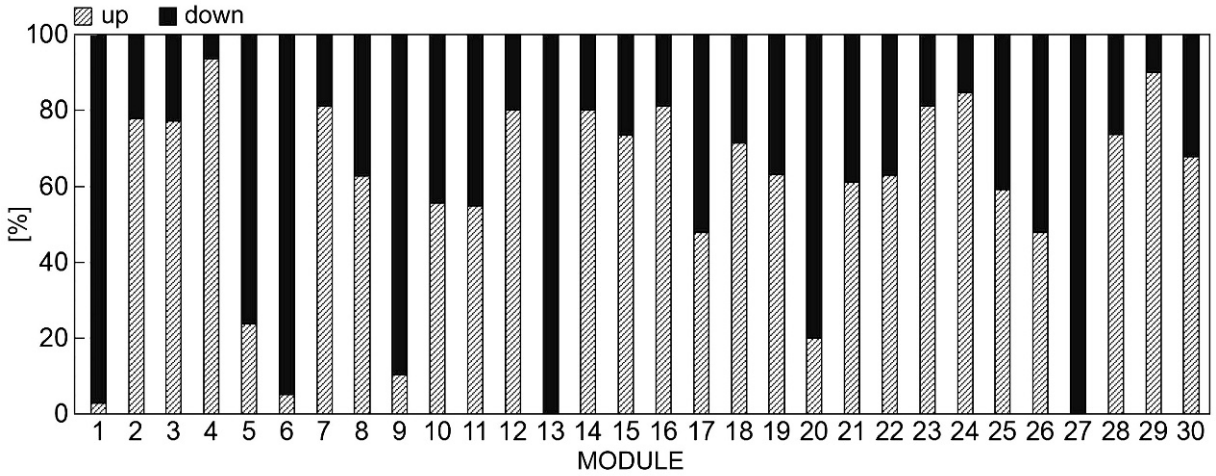

Fig. 2. Summary of expression of DEGs in the detected modules. The ratio of up- and down-regulated genes is represented by *dark* and *light grey* columns, respectively.

containing 1 238 DEGs with cluster size ranging from 15 to 65 nodes. The simple parameters of constructed GCN, and HRR values between all co-expressed genes are presented in Table 2 and Table 1 Suppl., respectively. Through gene expression analysis of DEGs in each detected module, we found that most modules were mainly composed of up-regulated genes (68 %), whereas Module 1, Module 5, Module 6, Module 9 Module 13, Module 26, and Module 27 were mainly composed of down-regulated genes (32 %) (Fig. 2).

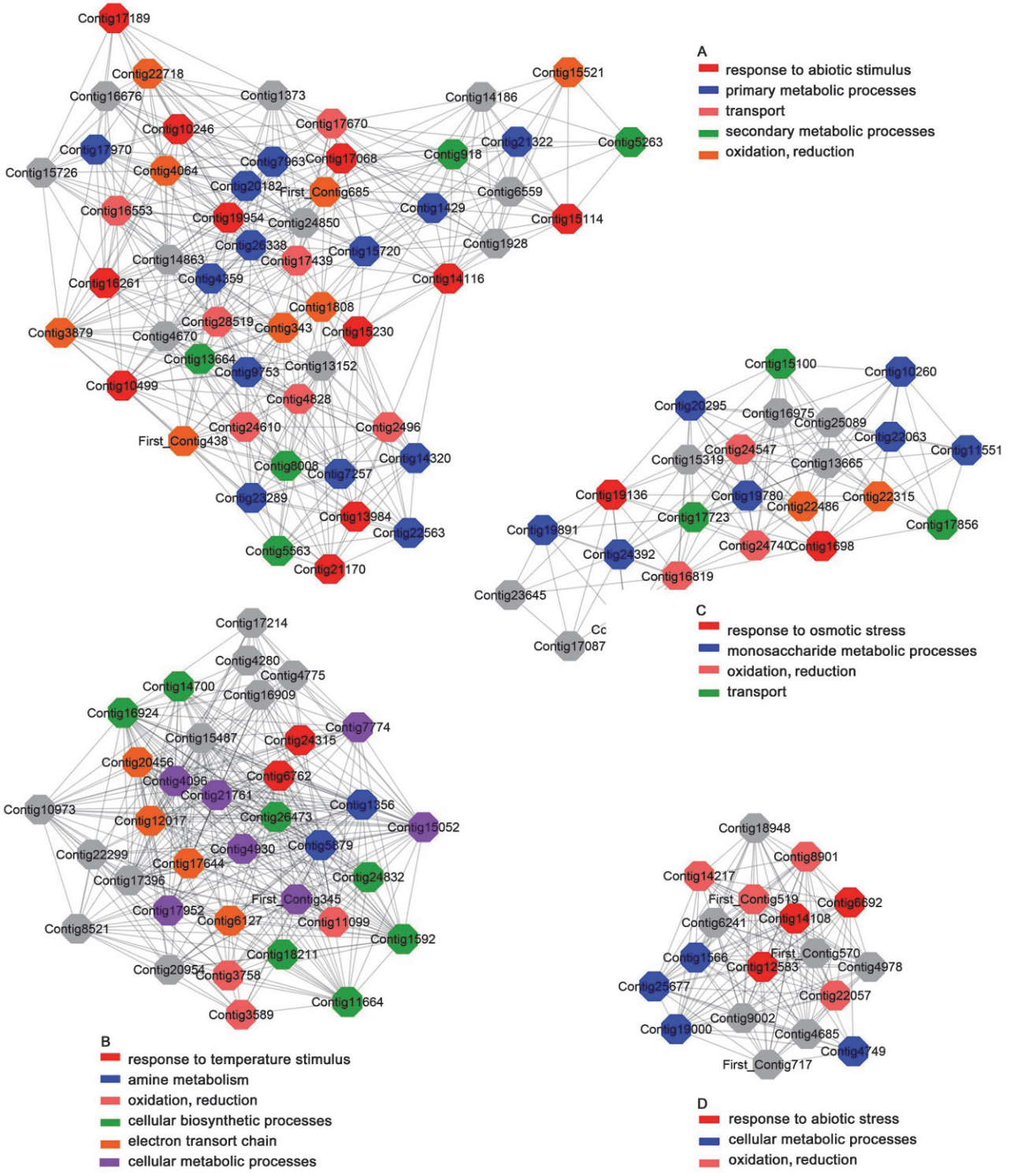

Fig. 3. Stressed related modules (Module 2, 5, 12, and 22) involved in response to stimuli and other biological processes.

GO annotation is one of important steps in the GCN analysis to understand the biological functions of each detected module (Bonthala *et al.* 2016). We carried out a GO enrichment analysis using *AgriGO* tool, which revealed that modules were mainly involved in metabolic processes (GO:0008152), including oxidation and reduction (GO:0055114), cellular processes (GO:0009987), regulation of biological processes (GO:0050789), biological regulation (GO:0065007), establishment of localization (GO:0051234), localization (GO:0051179), and response to stimulus (GO:0050896). The identification of over-represented and/or statistically significant biological processes can further reveal the functional features of each detected module. In this study, 13 modules were enriched to various kinds of stress response, and 4 of them (Module 2, 5, 12, and 22) were statistically significant, indicating that Module 2, Module 5, Module 12, and Module 22 are the stress related function modules.

In gene co-expression analysis, functional consistency among the highly inter connected genes present in the same modules at a given cut-off is very important for gene mining. In this study, functional enrichment analysis has resulted in significant enrichment of 4 modules (Module 2, 5, 12, and 22) containing genes principally involved in various kinds of stress response and metabolic processes.

Module 2 consisted of 54 gene nodes connected by 490 edges, and was highly enriched with cellular metabolic processes (GO:0044237, FDR = 0.00036), primary metabolism processes (GO:0044238, FDR = 0.0002), secondary metabolism process (GO:0019748, FDR = 0.0095), biosynthetic processes (GO:0009058, FDR = 0.024), nitrogen metabolism processes (GO:0006807, FDR = 0.02), and responses to stimuli (GO:0050896, FDR = 0.027) (Fig 3*A* and Table 2 Suppl.). In this module, 78 % of genes were up-regulated by cold stress, while rest of the genes was down regulated (Fig 2). A total of 11 genes associated with various kinds of stress responses, such as response to cold, osmotic stress, salinity, and excess irradiance, represented 44 % of annotated genes in this module (Fig 3*A* and Table 2 Suppl.). The late embryogenesis abundant (LEA) proteins are proved to be the common components associated with tolerance to various stresses (Tang *et al*. 2016). For example, the reduced expression of CeLEA1, a LEA-like protein from *Caenorhabditis elegans*, lead to reduced survival rate of plant under osmotic and heat stresses (Gal *et al*. 2004). Overexpression of *NtLEA7-3* in *Arabidopsis* transgenic plants enhances their resistance to cold, drought, and salt stresses (Gai *et al*. 2011). Moreover, novel *LEA* genes are constantly identified in various plant species and they are induced by multiple abiotic stresses. As nearly half of annotated genes in this module were related to stress response coupled with the presence of up-regulated *LEA1* genes (Contig19954 and Contig17189), we can conclude that the genes in Module 2 could be qualified as candidates for gene mining in stress response in *L. lancifolium*.

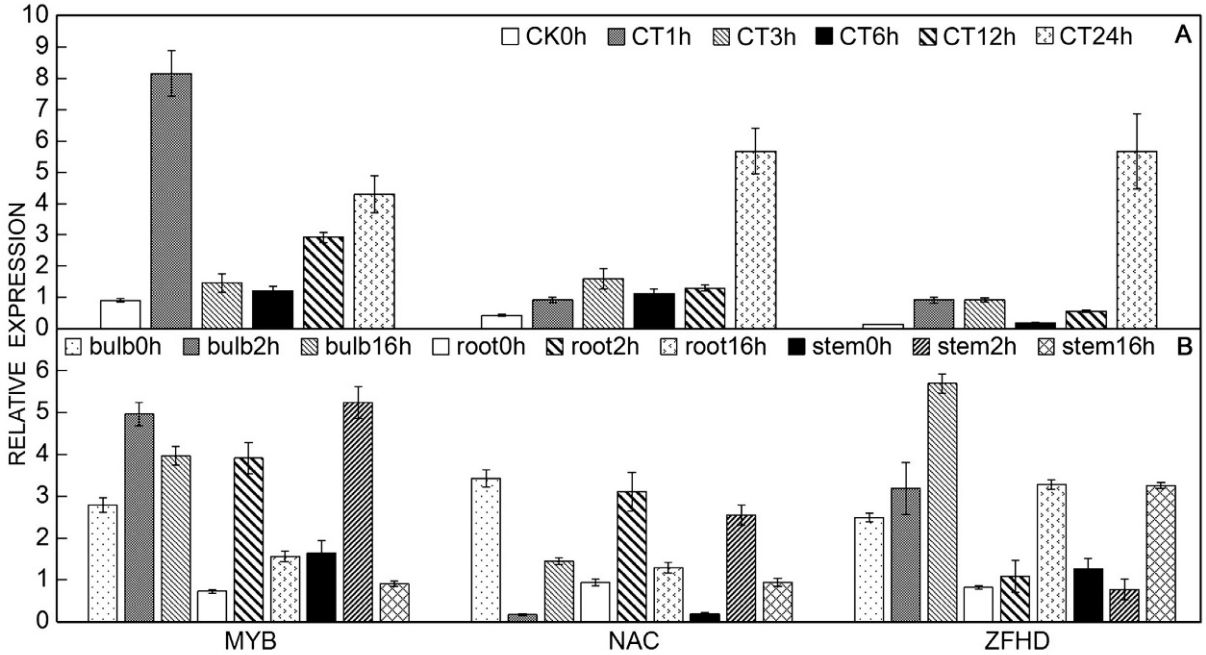

Fig. 4. The expression profiles of three candidate transcription factors gene, *MYB* (contig10499), *NAC* (contig16924), and *ZFHD* (contig11664) in *Lilium lancifolium* using RT-qPCR. *A* - Variation on gene expression in leaves under cold stress. Plants were treated with 4° C for 1, 3, 6, 12, and 24 h. *B* - Expressions of three candidate TFs in bulb, root, and stem under cold stress. Plants were treated with 4° C for 2 and 16 h.

Module 5 consisted of 36 gene nodes connected by 335 edges. Except for cellular metabolic processes (GO:0044237, FDR = 0.00036), cellular biosynthetic processes, oxidation and reduction (GO:0055114, FDR =

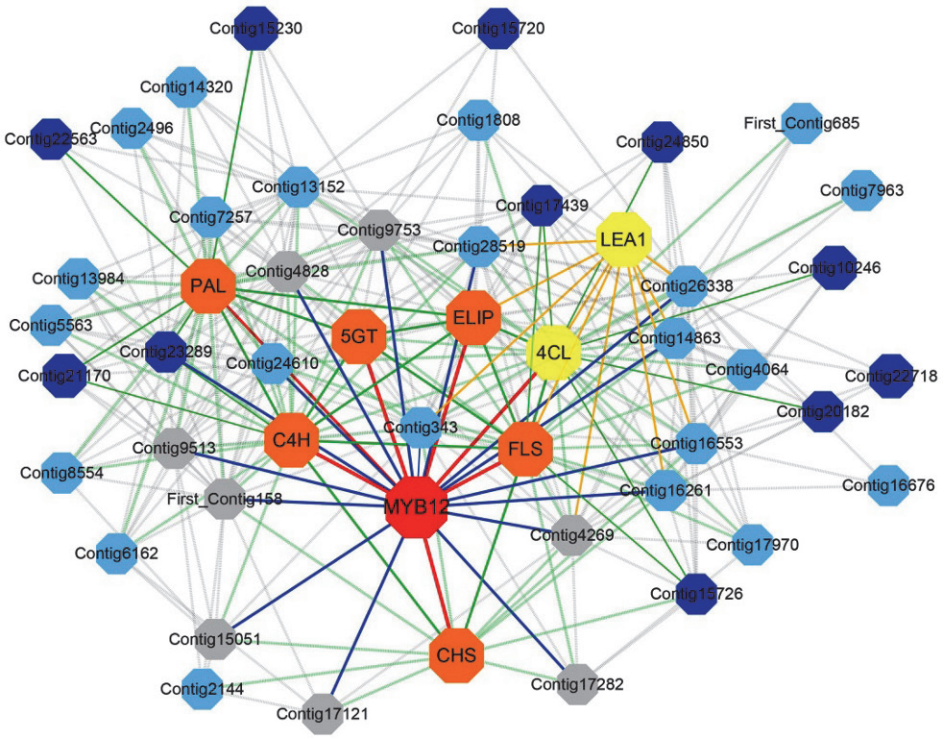

Fig. 5. *MYB12* (Contig10499) showing high node degree (*i.e.* marked connections) with many genes involved in phenylpropanoid and anthocyanin biosynthesis pathways at different highest reciprocal rank (HRR) values (Table 3 Suppl.).

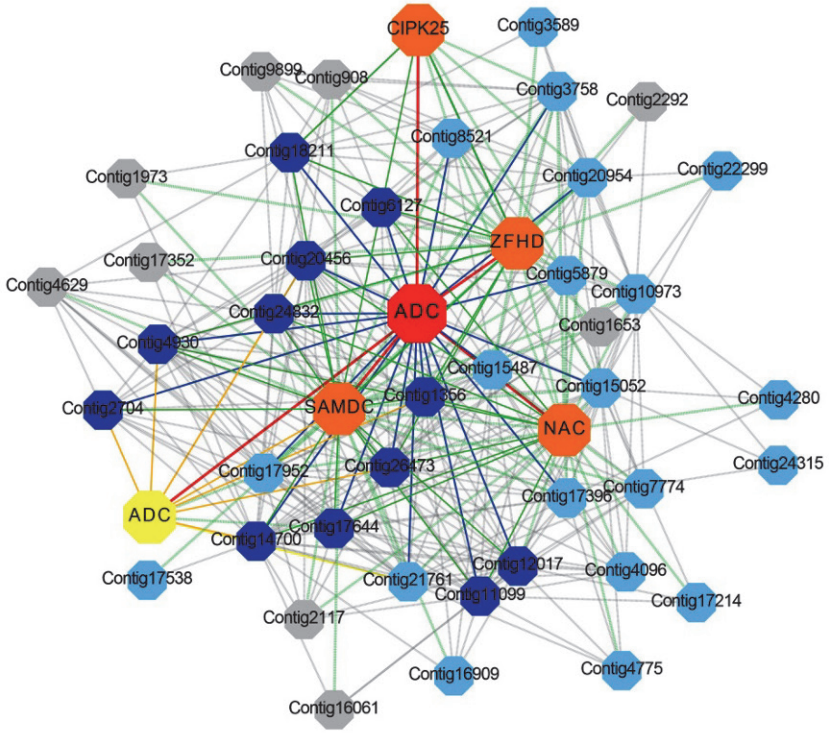

Fig. 6. *ADC* (Contig6762), *NAC* (Contig16924), and *ZFHD1* (Contig11664) showing high node degree (*i.e.*, marked connections) with many genes involved in polyamine biosynthesis pathway at different HRR values (Table 3 Suppl.).

4.1e-08) and response to temperature stimulus (GO:0009266, FDR = 0.018), the module was also highly enriched with electron transport chain (GO:0022900, FDR = 0.00048) and amine metabolism processes (GO:0009308, FDR = 0.02) (Fig 3*B* and Table 2 Suppl.). In contrast to Module 2, 76 % of genes in this module were down-regulated by cold stress (Fig 2), including an important gene called small heat shock protein (*sHsp*) gene (Contig17538). As molecular chaperones, sHsps can bind to partially folded or denatured proteins, and thereby prevent irreversible aggregation or promote correct protein folding (Sun *et al.* 2002). The expressions of some *sHsp* genes were regulated differentially by abiotic stresses including cold, implying that these *sHsp* genes may play a role in plant abiotic stress responses (Zou *et al.* 2009). In contrast to the induction of *sHsp* genes by multiple stresses (Chen *et al.* 2014, Muthusamy *et al.* 2017), our expression analysis showed that the *sHsp* gene (Contig17538) was significantly down-regulated during CT2h and CT16h. Similarly, some *ACD* (α-crystalline C-terminal domains) genes in rice (Sarkar *et al.* 2009) and some *Hsp20* genes in tomato (Yu *et al.* 2016) are also down-regulated under various abiotic stresses, although their functional roles under stress conditions have not been extensively worked out. The expression profile of *sHsp* genes differ depending on the exposure time and intensity of stress (Muthusamy *et al.* 2017). The functional role of the *sHsp* genes (Contig17538) in response to cold stress needs further study. Apart from sHsp, the up-regulated CBL-interacting protein kinase 25 (*CIPK25*; Contig345) was also important in Module 5. Recent studies in rice have found that some genes in the *CIPK* family are involved in the responses to multiple abiotic stresses, while some genes of this family are only responsive to specific stress (Xiang *et al.* 2007). In this study, we can suppose that the *CIPK25* gene in *L. lancifolium* can respond positively to cold stress.

Table 3. The list of four modules significantly enriched with various kinds of stress response after cold treatment. FDR - false discovery rate.

Module	Gene No.	Response to	FDR
2	GO:0009628	abiotic stimulus	0.0001
	GO:0006950	stress	0.0075
5	GO:0009611	wounding	0.0250
	GO:0009266	temperature stimulus	0.0180
12	GO:0009414	water deprivation	0.0300
	GO:0006970	osmotic stress	0.0500
	GO:0009415	water supply	0.0300
22	GO:0009409	cold	0.0310
	GO:0009628	abiotic stimulus	0.0380

Module 12 consisted of 24 gene nodes connected by 134 edges, and was highly enriched with monosaccharide metabolism processes (GO:0005996, FDR = 0.019), primary metabolism processes (GO:0044238, FDR = 0.009), oxidation and reduction (GO:0055114, FDR = 0.0028), transport (GO:0006810, FDR = 0.047), and response to osmotic stress (GO:0009415, FDR = 0.05) (Fig 3*C* and Table 2 Suppl.). In this module, 80 % of genes were up-regulated by cold stress, while the rest were down-regulated (Fig 2).

Module 22 consisted of 18 gene nodes connected by 117 edges. The highly enriched biological processes of the module were cellular metabolic process (GO:0044237, FDR = 0.0031), oxidation and reduction (GO:0055114, FDR = 4.3e-05), and response to abiotic stimuli (GO:0009628, FDR = 0.038) (Fig 3*D* and Table 2 Suppl.). In this module, 63 % of genes were up-regulated by cold stress, while the rest was down-regulated (Fig. 2). However, the role of genes in modules 12 and 22 in cold tolerance remains unclear.

TFs play important roles in the biological process involved in plant stress tolerance. In this study, we identified three TFs in the stress related modules which regulated downstream target genes involved in cold stress. In Module 2, one gene encoding TFs from MYB families called R2R3-MYB transcription factor *MYB12* (Contig10499) was identified. Module 5 contained two genes encoding TFs from two families: NAC domain protein (Contig16924) and Zinc finger-homeodomain protein 1 (*ZFHD1*; Contig11664). To further study the expression pattern and tissue specificity of these three TFs, we analyzed their expression profiles under cold stress in leaves, stems, roots, and bulbs using RT-qPCR (Fig. 4). The expression of *MYB12* gene increased rapidly after 1 h, and then decreased back to the original level after 3 and 6 h, but increased again after 12 and 24 h, suggesting that *MYB12* gene can be induced immediately by low temperature and may play different roles or participate in diverse signal pathways in different periods of cold acclimation. Furthermore, the expression patterns of *NAC* and *ZFHD1* genes were basically the same: both were up-regulated until 24 h, indicating that the response of *NAC* and *ZFDH1* genes to low temperature was coincident and much slower than that of *MYB12* gene. The expressions of *MYB12*, *NAC*, and *ZFDH1* genes in bulb, root, and stem were also analyzed (Fig 4*B*). The results showed that *MYB12* and *ZFDH1* genes were almost steadily up-regulated after 2 h, while *NAC* gene in bulb was down-regulated after 2 h and then up-regulated after 16 h. Given that the average expressions of these three genes were highest in bulb after 2 and 16 h, we can suppose that the bulb played an important role for cold acclimation in *L. lancifolium* and the *NAC* gene played a special role in cold resistance of bulb.

The whole phenylpropanoid pathway is known to be highly responsive to different abiotic stresses, such as wounding, drought, low, or high temperatures (Páldi *et al.* 2014). Cold stress has been shown to be an

important factor in increasing anthocyanin production in grape (Mori *et al.* 2005), maize (Christie *et al.* 1994), and apple (Ubi *et al.* 2006). Many cold-inducible genes involved in anthocyanin biosynthesis pathway are considered as cold-regulated (*cor*) genes (Christie *et al.* 1994). Thus, one important method to understand the role of anthocyanins in stress response is giving insight into the transcriptional regulation of its biosynthesis pathway. Anthocyanin biosynthesis is primarily controlled by functional and regulatory genes. The functional genes encode multiple enzymes including phenylalanine ammonia-lyase (PAL), cinnamic acid 4-hydroxylase (C4H), chalcone synthases (CHS), flavonol synthase (FLS), cinnamate-4-hydroxylase (4CL), flavanone-3-hydroxylase (F3H), anthocyanidin synthase (ANS) and UDP-glucose: flavonoid-3-*O*-glucosyltransferase (UFGT) (Winkel-Shirley 2001). The regulatory genes encode MYB-bHLH-WD40 repeat protein (MBW) complexes which control the expression of the functional genes (Broun 2005). However, in some cases, a single MYB TF can play an important negative regulatory role in the anthocyanin biosynthesis pathway (Cavallini *et al.* 2015). For example, the MYB4 homologs from *Arabidopsis* and petunia directly repress the expression of *C4H* or *4CL* (Colquhoun *et al.* 2010, Jin *et al.* 2000). The *ZmMYB42* and *ZmMYB31* can repress the expression of lignin genes affecting cell wall structure, composition, and degradability in maize (Sonbol *et al.* 2009, Fornale *et al.* 2010). Also, the expression of *BrMYB2-2* inhibits anthocyanin accumulation and cold tolerance in *Brassica rapa* (Ahmed *et al.* 2015). In this study, through the analysis of GCN, we found that the R2R3-MYB TF gene *MYB12* (Contig10499) in Module 2 was highly co-expressed with genes involved in phenylpropanoid and anthocyanin biosynthesis pathways, including *CHS* (Contig3366), *FLS* (Contig3879), *PAL* (Contig8008), *C4H* (First_Contig438), 4-hydroxycinnamoyl-CoA ligase 2 (*4CL2*, Contig4359), and *anthocyanin-5-O-glucosyltransferase* (Contig13664) (Fig 5 and Table 3 Suppl.). The similar mechanisms of anthocyanin biosynthesis pathway regulation can exist in many plants. Therefore, the MYB12 (Contig10499) in Module 2 might be a key TF for anthocyanin biosynthesis pathway in response to cold stress, or at least involved in the regulation of anthocyanin biosynthesis in cold acclimation in *L. lancifolium*.

The differential accumulation of polyamines (PAs) in different plant species in response to various environmental stresses such as oxygen deficiency, cold, high salinity, and low pH has been observed, suggesting that PAs play an essential role in response to adverse environments (Yamakawa *et al.* 1998). The PAs biosynthesis pathways have been well-established. In plants, putrescine can be produced directly from ornithine by ornithine decarboxylase (ODC) or indirectly from arginine by arginine decarboxylase (ADC), and can be subsequently converted into spermidine and spermine through the addition of propylamino groups from decarboxylated S-adenosylmethionine (dcSAM) (Walden *et al.* 1997). The dcSAM itself is produced from S-adenosylmethionine (SAM) catalyzed by S-adenosylmethionine dehydrogenase (SAMDC). Apart from the accumulation of putrescine, abiotic stresses increase ADC activity and induce the accumulation of *ADC* gene expression in plants (Liu *et al.* 2006). For instance, in apple callus, accumulation of putrescine was in accordance with the induction of ADC activity and the expression of *MdADC* under salt stress, while ODC activity showed an opposite pattern to ADC and does not parallel stress parameters (Liu *et al.* 2006). It indicates that ADC is tightly involved in salt stress response but ODC contributes little. Such results correspond with the work on *Arabidopsis thaliana* mutants with reduced ADC activity, where salt stress causes more damage, along with less putrescine accumulation (Kasinathan and Wingler 2004). In addition, Perez-Amador *et al*. (2002) also reported that *ADC2* is the only gene of PAs biosynthesis pathway involved in wounding response. These findings imply that *ADC* gene is responsive to stress conditions and important for putrescine biosynthesis under abiotic/biotic stresses. By the analysis of cold responsive genes in GCN in this study, two DEGs encoding ADC (Contig6762) and SAMDC (Contig1592) were found co-expressed in Module 5 (Fig. 6 and Table 3 Suppl.). Module 5 also contained two DEGs encoding Zinc finger-homeodomain protein 1 (ZFHD1, Contig11664) and NAC domain protein (Contig16924), which were highly co-expressed with both *SAMDC* and *ADC* genes and co-expressed with each other as well (Fig 6 and Table 3 Suppl.). In *Arabidopsis*, the co-expression of stress-inducible ZFHD1 and NAC TFs enhance the expression of *ERD1* (early responsive to dehydration stress 1) gene (Tran *et al.* 2007). Considering the results of RT-qPCR (Fig 4), the consistency of expression pattern of *NAC* and *ZFHD1* gene under continuous cold stress was confirmed. It can be assumed that ZFHD1 and NAC domain protein were involved in PAs biosynthesis pathway, and ZFHD1 might interact with NAC domain protein regulating the expression of *ADC* or/and *SAMDC* genes in *L. lancifolium* in response to cold stress.

The plant cytoskeleton can be readily remodelled in response to all sorts of intracellular and external stimuli. As an important component of cytoskeleton, microtubules (MTs) are essential for intra- and extra-cellular signalling (Zhu *et al.* 2013). The regulation of the dynamic instability of MTs has been shown to play a crucial role in plant resistance to salt and osmotic stresses (Sedbrook and Kaloriti 2008, Wang *et al.* 2011). The MT dynamics is precisely regulated by microtubule-associated proteins (MAPs) (Desai and Mitchison 1997). MAP65 is one of the most abundant plant MAPs, which was firstly isolated from tobacco BY2 cells as a group of 60 - 65 kDa proteins that co-purified with MTs (Chang and Sonobe 1993). Except for cell division and

elongation, root growth, and leaf senescence, several studies have shown that MAP65 proteins are also required for stabilizing MTs during low temperature and salt stress (Smertenko *et al.* 2004, Mao *et al.* 2005). In this study, a DEG (Contig16975) annotated as microtubule binding polypeptide TMBP200 was identified in Module 12. The TMBP200 is a 200 kDa protein isolated from tobacco BY2 cells, which shares homology with *Arabidopsis* MOR1, a member of the XMAP215 family of MAPs (Yasuhara *et al.* 2002). In view of the similarities between TMBP200 and MAP65, TMBP200 may participate in a response to low temperature in *L. lancifolium*.

In Module 22, a serine/threonine-protein kinase (STK, Contig14108) along with calcium-binding EF hand family protein (Contig6692) were found to be induced by cold treatment. They might be important factors in Ca^{2+} signalling pathway in response to cold stress. As a second messenger, Ca^{2+} plays an essential role in the signal transduction involved in various abiotic stresses. Cellular calcium signals are detected and transmitted by Ca^{2+} sensors such as calcium-binding proteins. Calcineurin B-like proteins (CBLs) represent a group of Ca^{2+} sensors and play a key role in decoding calcium transients by specifically interacting with and regulating a family of CBL-interacting protein kinases (CIPKs) (He *et al.* 2013). The plant CIPK contains a specific serine/ threonine protein kinase domain, which is activated by CBL. Activated CIPKs can transduce Ca^{2+} signals through phosphorylating downstream signalling components (He *et al.* 2013). The calcium-binding EF hand family protein belongs to a diverse superfamily of calcium sensors and calcium signal modulators. Differing from CBLs containing four EF hands for Ca^{2+} binding, most examples of this kind of protein contains only two active canonical EF hands. Thus, we can assume that calcium-binding EF hand family protein (Contig6692) in *L. lancifolium* might be a novel Ca^{2+} sensor in response to cold stress, which also interacted with a specific Ser/Thr protein kinase domain.

In conclusion, we identified four MCL modules which were significantly enriched under cold stress using a GCN analysis of *L. lancifolium* transcriptome profiles. Among the annotated genes presented in these four modules, some genes, such as *LEA1, sHsp*, and *CIPK25*, played crucial roles in plant response to abiotic stresses. Some genes were found to be key genes of anthocyanin and polyamines biosynthesis pathways which have also been considered as *cor* genes, including *CHS*, *FLS*, *4CL2*, *C4H,* and *ADC*. More importantly, three genes encoding TFs from MYB, NAC, and ZFHD families, which were highly co-expressed with *cor* genes within the same module, were likely to be important cold responsive transcriptional regulators in *L. lancifolium*. Apart from these previously characterized genes which are known to have roles in abiotic stresses, the genes encoding TMBP200 and calcium-binding EF hand family protein may be novel cold responsive genes firstly reported in the present study.

Reference

Agarwal, M., Hao, Y., Kapoor, A., Dong, C.H., Fujii, H., Zheng, X.. Zhu, J.K.: A R2R3 type MYB transcription factor is involved in the cold regulation of *CBF* genes and in acquired freezing tolerance. - J. biol. Chem. **281**: 37636-37645, 2006.

Ahmed, N.U., Park, J.I., Jung, H.J., Hur, Y., Nou, I.S.: Anthocyanin biosynthesis for cold and freezing stress tolerance and desirable color in *sBrassica rapa*. - Funct. integr. Genomics **15**: 383-394, 2015.

Assenov, Y., Ramirez, F., Schelhorn, S.E., Lengauer, T., Albrecht, M.: Computing topological parameters of biological networks. - Bioinformatics **24**: 282-284, 2008.

Bakshi, M., Oelmuller, R.: WRKY transcription factors:jack of many trades in plants. - Plant Signal. Behav. **9**: e27700, 2014.

Bonthala, V.S., Mayes, K., Moreton, J., Blythe, M., Wright, V., May, S.T., Massawe, F., Mayes, S., Twycross, J.: Identification of gene modules associated with low temperatures response in bambara groundnut by network-based analysis. - PLoS ONE **11**: e0148771, 2016.

Broun, P.: Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in *Arabidopsis*. - Curr. Opin. Plant Biol. **8**: 272-279, 2005.

Cavallini, E., Matus, J.T., Finezzo, L., Zenoni, S., Loyola, R., Guzzo, F., Schlechter, R., Ageorges, A., Arce-Johnson, P., Tornielli, G.B.: The phenylpropanoid pathway is controlled at different branches by a set of R2R3-MYB C2 repressors in grapevine. - Plant Physiol. **167**: 1448-1470, 2015.

Chang, J.J., Sonobe, S.: Identification and preliminary characterization of a 65 kDa higher-plant microtubule-associated protein. - J. cell. Sci. **105**: 891-901, 1993.

Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., Yu, D.: The role of WRKY transcription factors in plant abiotic stresses. - Biochim. biophys. Acta. **1819**: 120-128, 2012.

Chen, X., Lin, S., Liu, Q., Huang, J., Zhang, W., Lin, J., Wang, Y., Ke, Y., He, H.: Expression and interaction of small heat shock proteins (sHsps) in rice in response to heat stress. - Biochim. biophys. Acta. **1844**: 818-828, 2014.

Chinnusamy, V., Zhu, J., Zhu, J.K.: Cold stress regulation of gene expression in plants. - Trends Plant Sci. **12**: 444-451, 2007.

Choe, E., Drnevich, J., Williams, M.M.: Identification of crowding stress tolerance co-expression networks involved in sweet corn yield. - PLoS ONE **11**: e0147418, 2016.

Christie, P.J., Alfenito, M.R., Walbot, V.: Impact of low-temperature stress on general phenylpropanoid and anthocyanin pathways: enhancement of transcript abundance and anthocyanin pigmentation in maize seedlings. - Planta **194**: 541-549, 1994.

Colquhoun, T.A., Kim, J.Y., Wedde, A.E., Levin, L.A., Schmitt, K.C., Schuurink, R.C., Clark, D.G.: PhMYB4 fine-tunes the floral volatile signature of *Petunia×hybrida* through PhC4H. - J. exp. Bot. **62**: 1133-1143, 2010.

Desai, A., Mitchison, T.J.: Microtubule polymerization dynamics. - Annu. Rev. cell. dev. Biol. **13**: 83-117, 1997.

Du, C., Hu, K., Xian, S., Liu, C., Fan, J., Tu, J., Fu, T.: Dynamic transcriptome analysis reveals AP2/ERF transcription factors responsible for cold stress in rapeseed (*Brassica napus* L.). - Mol. Genet. Genomics **291**: 1053-1067, 2016.

Du, Z., Zhou, X., Ling, Y., Zhang, Z.H., Su, Z.: agriGO: a GO analysis toolkit for the agricultural community. - Nucl. Acids Res. **38**: W64-W70, 2010.

Enright, A.J., Van Dongen, S., Ouzounis, C.A.: An efficient algorithm for large scale detection of protein families. - Nucl. Acids Res. **30**: 157-1584, 2002.

Fornale, S., Shi, X.H., Chai, C.L., Encina, A., Irar, S., Capellades, M., Fuguet, E., Torres, J.L., Rovira, P., Puigdomenech, P., Rigau, J., Grotewold, E., Gray, J., Caparros-Ruiz, D.: ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. - Plant J. **64**: 633-644, 2010.

Gai, Y.P., Ji, X.L., Lu, W., Han, X.J., Yang, G.D., Zheng, C.C.: A novel late embryogenesis abundant like protein associated with chilling stress in *Nicotiana tabacum* cv. Bright Yellow-2 cell suspension culture. - Mol. Cell Proteomics **10**: M111 010363, 2011.

Gal, T.Z., Glazer, I., Koltai, H.: An LEA group 3 family member is involved in survival of *C. elegans* during exposure to stress. - FEBS Lett. **577**: 21-26, 2004.

He, L., Yang, X., Wang, L., Zhu, L., Zhou, T., Deng, J., Zhang, X.: Molecular cloning and functional characterization of a novel cotton CBL-interacting protein kinase gene (*GhCIPK6*) reveals its involvement in multiple abiotic stress tolerance in transgenic plants. - Biochem. biophys. Res. Commun. **435**: 209-215, 2013.

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B., Martin, C.: Transcriptional repression by *AtMYB4* controls production of UV-protecting sunscreens in *Arabidopsis*. - EMBO. J. **19**: 6150-6161, 2000.

Kasinathan, V., Wingler, A.: Effect of reduced arginine decarboxylase activity on salt tolerance and on polyamine formation during salt stress in *Arabidopsis thaliana*. - Physiol. Plant. **121**: 101-107, 2004.

Langfelder, P., Horvath, S.: WGCNA: an R package for weighted correlation network analysis. - BMC Bioinformatics **9**: 559, 2008.

Liu, C., Wu, Y., Wang, X.: bZIP transcription factor OsbZIP52/RISBZ5: a potential negative regulator of cold and drought stress response in rice. - Planta **235**: 1157-1169, 2012.

Liu, J.H., Nada, K., Honda, C., Kitashiba, H., Wen, X.P., Pang, X.M., Moriguchi, T.: Polyamine biosynthesis of apple callus under salt stress: importance of the arginine decarboxylase pathway in stress response. - J. exp. Bot. **57**: 2589-2599, 2006.

Mao, T.L., Jin, L.F., Li, H., Liu, B., Yuan, M.: Two microtubule-associated proteins of the *Arabidopsis* MAP65 family function differently on microtubules. - Plant Physiol. **138**: 654-662, 2005.

Mori, K., Sugaya, S., Gemma, H.: Decreased anthocyanin biosynthesis in grape berries grown under elevated night temperature condition. - Scientia Hort. **105**: 319-330, 2005.

Morris, J.H., Apeltsin, L., Newman, A.M., Baumbach, J., Wittkop, T., Su, G., Bader, G.D., Ferrin, T.E.: ClusterMaker: a multi-algorithm clustering plugin for Cytoscape. - BMC Bioinformatics **12**: 436, 2011.

Muthusamy, S.K., Dalal, M., Chinnusamy, V., Bansal, K.C.: Genome-wide identification and analysis of biotic and abiotic stress regulation of small heat shock protein (*HSP20*) family genes in bread wheat. - J. Plant Physiol. **211**: 100-113, 2017.

Mutwil, M., Usadel, B., Schutte, M., Loraine, A., Ebenhoh, O., Persson, S.: Assembly of an interactive correlation network for the *Arabidopsis* genome using a novel heuristic clustering algorithm. - Plant Physiol. **152**: 29-43, 2010.

Páldi, K., Rácz, I., Szigeti, Z., Rudnóy, S.: S-methylmethionine alleviates the cold stress by protection of the photosynthetic apparatus and stimulation of the phenylpropanoid pathway - Biol. Plant. **58**: 189-194, 2014.

Perez-Amador, M.A., Leon, J., Green, P.J., Carbonell, J.: Induction of the arginine decarboxylase *ADC2* gene provides evidence for the involvement of polyamines in the wound response in *Arabidopsis*. - Plant Physiol. **130**: 1454-1463, 2002.

Qu, Y., Zhou, A., Zhang, X., Tang, H., Liang, M., Han, H., Zuo, Y.: *De novo* transcriptome sequencing of low temperature-treated *Phlox subulata* and analysis of the genes involved in cold stress. - Int. J. mol. Sci. **16**: 9732-9748, 2015.

Sarkar, N.K., Kim, Y.K., Grover, A.: Rice *sHsp* genes: genomic organization and expression profiling under stress and development. - BMC Genomics **10**: 393, 2009.

Sedbrook, J.C., Kaloriti, D.: Microtubules, MAPs and plant directional cell expansion. - Trends Plant Sci. **13**: 303-310, 2008.

Smertenko, A.P., Chang, H.Y., Wagner, V., Kaloriti, D., Fenyk, S., Sonobe, S., Lloyd, C., Hauser, M.T., Hussey, P.J.: The *Arabidopsis* microtubule-associated protein AtMAP65-1: molecular analysis of its microtubule bindling activity. - Plant Cell **16**: 2035-2047, 2004.

Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L., Ideker, T.: Cytoscape2.8: newfeatures for data integration and network visualization. - Bioinformatics **27**: 431-432, 2011.

Sonbol, F.M., Fornale, S., Capellades, M., Encina, A., Tourino, S., Torres, J.L., Rovira, P., Ruel, K., Puigdomenech, P., Rigau, J., Caparros-Ruiz, D.: The maize ZmMYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in *Arabidopsis thaliana*. - Plant mol. Biol. **70**: 283-296, 2009.

Sun, W., Van Montagu, M., Verbruggen, N.: Small heat shock proteins and stress tolerance in plants. - Biochim. biophys. Acta **1577**: 1-9, 2002.

Tang, X., Wang, H., Chu, L., Shao, H.: *KvLEA*, a new isolated late embryogenesis abundant protein gene from *Kosteletzkya virginica* responding to multiabiotic stresses. - Biomed. Res. Int. **2016**: 9823697, 2016.

Tran, L.S., Nakashima, K., Sakuma, Y., Osakabe, Y., Qin, F., Simpson, S.D., Maruyama, K., Fujita, Y., Shinozaki, K., Yamaguchi-Shinozaki, K.: Co-expression of the stress-inducible zinc finger homeodomain ZFHD1 and NAC transcription factors enhances expression of the *ERD1* gene in *Arabidopsis*. - Plant J. **49**: 46-63, 2007.

Tran, L.S., Nishiyama, R., Yamaguchi-Shinozaki, K., Shinozaki, K.: Potential utilization of NAC transcription factors to enhance abiotic stress tolerance in plants by biotechno-

logical approach. - Genet. Modif. Crops **1**: 32-39, 2010.

Ubi, B., Honda, C., Bessho, H., Moriguchi, T.: Expression analysis of anthocyanin biosynthetic genes in apple skin: effect of UV-B and temperature. - Plant Sci. **170**: 571-578, 2006.

Walden, R., Cordeiro, A., Tiburcio, A.F.: Polyamines: small molecules triggering pathways in plant growth and development. - Plant Physiol. **113**: 1009-1013, 1997.

Wang, C., Zhang, L., Chen, W.: Plant cortical microtubules are putative sensors under abiotic stresses. - Biochemistry (Moscow) **76**: 320-326, 2011.

Wang, J., Wang, Q., Yang, Y., Liu, X., Gu, J., Li, W., Ma, S., Lu, Y.: *De novo* assembly and characterization of stress transcriptome and regulatory networks under temperature, salt and hormone stresses in *Lilium lancifolium*. - Mol. Biol. Rep. **41**: 8231-8245, 2014a.

Wang, J., Yang, Y., Liu, X., Huang, J., Wang, Q., Gu, J., Lu, Y.: Transcriptome profiling of the cold response and signaling pathways in *Lilium lancifolium*. - BMC Genomics **15**: 203, 2014b.

Wang, W., Wu, P., Li, Y., Hou, X.: Genome-wide analysis and expression patterns of ZF-HD transcription factors under different developmental tissues and abiotic stresses in Chinese cabbage. - Mol. Genet. Genomics **291**: 1451-1464, 2016.

Winkel-Shirley, B.: Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. - Plant Physiol. **126**: 485-493, 2001.

Wong, D.C., Sweetman, C., Ford, C.M.: Annotation of gene function in citrus using gene expression information and co-expression networks. - BMC Plant Biol. **14**: 186, 2014.

Xiang, Y., Huang, Y., Xiong, L.: Characterization of stress-responsive *CIPK* genes in rice for stress tolerance improvement. - Plant Physiol. **144**: 1416-1428, 2007.

Yamakawa, H., Kamada, H., Satoh, M., Ohashi, Y.: Spermine is a salicylate-independent endogenous inducer for both tobacco acidic pathogenesis-related proteins and resistance against tobacco mosaic virus infection. - Plant Physiol. **118**: 1213-1222, 1998.

Yasuhara, H., Muraoka, M., Shogaki, H., Sonobe, S.: TMBP200, a microtubule bundling polypeptide isolated from telophase tobacco BY-2 cells is a MOR1 homologue. - Plant Cell Physiol. **43**: 595-603, 2002.

Yu, J., Cheng, Y., Feng, K., Ruan, M., Ye, Q., Wang, R., Li, Z., Zhou, G., Yao, Z., Yang, Y., Wan, H.: Genome-wide identification and expression profiling of tomato *Hsp20* gene family in response to biotic and abiotic stresses. - Front. Plant Sci. **7**: 1215, 2016.

Zhang, L., Yu, S., Zuo, K., Luo, L., Tang, K.: Identification of gene modules associated with drought response in rice by network-based analysis. - PLoS ONE **7**: e33748, 2012.

Zhu, Y., Zuo, M., Liang, Y., Jiang, M., Zhang, J., Scheller, H.V., Tan, M., Zhang, A.: MAP65-1a positively regulates H2O2 amplification and enhances brassinosteroid-induced antioxidant defence in maize. - J. exp. Bot. **64**: 3787-3802, 2013.

Zou, J., Liu, A., Chen, X., Zhou, X., Gao, G., Wang, W., Zhang, X.: Expression analysis of nine rice heat shock protein genes under abiotic stresses and ABA treatment. - J. Plant Physiol. **166**: 851-861, 2009.