Open access • Posted Content • DOI:10.1101/850164

Structural variation, functional differentiation and expression characteristics of the AP2/ERF gene family and its response to cold stress and methyl jasmonate in Panax ginseng C.A. Meyer — Source link

Jing Chen, Yuanhang Zhou, Qi Zhang, Li Li ...+11 more authors

Published on: 20 Nov 2019 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Gene family, Subfamily, Methyl jasmonate, Gene and Ginseng

Structural variation, functional differentiation and expression characteristics of the AP2/ERF gene family and its response to cold stress and methyl jasmonate in *Panax ginseng* C.A. Meyer

Short title: Structural and functional Analysis of AP2/ERF gene family in *Panax ginseng* C.A. Meyer

Jing Chen[1,2]☯,Yuanhang Zhou[1]☯, Qi Zhang[1], Li Li[1], Chunyu Sun[1,2], Kangyu Wang[1,2], Yanfang Wang[2,3], Mingzhu Zhao[1,2], Hongjie Li[1], Yilai Han[1], Ping Chen[1], Ruiqi Li[1], Jun Lei[1], Meiping Zhang [1,2*], and Yi Wang [1,2*]

[1] College of Life Science, Jilin Agricultural University, 2888 Xincheng Street, Changchun, Jilin 130118, China.

[2] Research Center for Ginseng Genetic Resources Development and Utilization, 2888 Xincheng Street, Changchun, Jilin 130118, China.

[3] College of Chinese Medicinal Materials, Jilin Agricultural University, 2888 Xincheng Street, Changchun 130118, Jilin, China.

☯These authors contributed equally to this work.

Correspondence should be addressed to meiping.zhang@jlau.edu.cn and wanglaoshi0606@163.com

Abstract:

The APETALA2/Ethylene Responsive Factor (AP2/ERF) gene family has been shown to play a crucial role in plant growth and development, stress responses and secondary metabolite biosynthesis. Nevertheless, little is known about the gene family in ginseng (*Panax ginseng*), an important traditional medicinal herb in Asia and North America. Here, we report the systematic analysis of the gene family present in ginseng using several transcriptomic databases. A total of 189 putative *AP2/ERF* genes, defined as *PgERF001* through *PgERF189*. The 93 *PgERF* genes that have the complete AP2 domain in their open reading frames were classified into five subfamilies, DREB, ERF, AP2, RAV and Soloist. The DREB subfamily and ERF subfamily were further clustered four and six groups, respectively, compared to the 12 groups of these subfamilies found in Arabidopsis. Gene ontology categorized these 397 transcripts of the 189 *PgERF* genes into eight functional subcategories, suggesting their functional differentiation and they have been especially enriched for the nucleic acid binding transcription factor activity subcategory. The expression activity and networks of the 397 *PgERF* transcripts have substantially diversified across tissues, developmental stages and genotypes. Then, the expression change of six *PgERF* genes randomly selected from DREB subfamily, i.e., *PgERF073*, *PgERF079*, *PgERF110*, *PgERF115*, *PgERF120* and *PgERF128* responding to cold stress suggesting that DREB subfamily genes played an important role in cold resistance of ginseng. Finally, we studied the responses of the *PgERF* genes to methyl jasmonate (MeJA). 288 (72.5%) of the 397 *PgERF* gene transcripts responded to the MeJA treatment, with 136 up-regulated and 152 down-regulated, indicating that most members of the *PgERF* gene family are responsive to MeJA. These results provide resources and knowledge necessary for family-wide functional analysis of the *PgERF* genes in ginseng and related species.

Keywords: Gene family, APETALA2/Ethylene Responsive Factor (AP2/ERF) genes, *Panax ginseng*, Phylogeny, Functional Differentiation, Co-expression network, cold stress, Methyl Jasmonate (MeJA)

1. Introduction

Plants are subjected to numerous biotic and abiotic stresses all time through their growth and development. Therefore, they have developed a variety of mechanisms by producing secondary signaling molecules (e.g., ethylene and jasmonic acid) and response networks at the molecular, biochemical and physiological levels to perceive the external signals from and response to the stresses [1]. It has been documented that a large number of genes are involved in these processes [2]. Therefore, it is important to decipher the regulatory mechanisms of the defense-related genes involved in the signal transduction pathways and the plant responses to these stresses for enhanced and efficient plant genetic improvement [3]. The APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors have been demonstrated to be one of the most important gene families actively functioning in plant response to biotic and abiotic stresses by binding to cis-acting elements of downstream target genes [4].

The AP2/ERF family has one or two conserved APETALA2 (AP2) domains (approximately 60 - 70 amino acids) [5]. Based on the number and amino acid sequence similarities of the AP2 domains, the AP2/ERF family is divided into the DREB (dehydration responsive element binding), ERF, AP2, RAV (Related to ABI3/VP1) and Soloist subfamilies [6],[7]. Both DREB and ERF subfamilies possess a single AP2 domain, with a specific WLG motif, and could be further subdivided into A1 to A6 and B1 to B6 groups, respectively [6]. Alternatively, the DREB and ERF subfamilies were also categorized into I to X, and VI-L and Xb-L groups, respectively [7]. The AP2 subfamily has two tandemly repeated AP2 domains, while the RAV subfamily has one AP2 domain and one B3 domain that are commonly found in other transcription factors [8]. The Soloist subfamily also has only one AP2 domain. It was classified into an independent subfamily due to its relatively low sequence homology with the DREB and ERF subfamilies [9]. Although the AP2 domain of the AP2/ERF family is highly conserved, its five

subfamilies, DREB, ERF, AP2, RAV and Soloist, recognize different DNA cis-acting elements and exhibit substantial functional diversity [10]. Specifically, the members of the AP2 subfamily bind to the GCAC(A/G)N(A/T)TCCC(A/G)ANG(C/T) element and regulate developmental processes of different plant tissues, e.g., embryo, flower, sepal and fruit [11],[12],[13]. The *RAV1* gene of the RAV subfamily was reported to bind to CAACA and CACCTG motifs in *Arabidopsis thaliana* [14]. The roles of the RAV subfamily in plant development and various biotic and abiotic stresses were investigated in several plant species [15],[16],[17]. The only gene of the Soloist subfamily in Arabidopsis, *APD1* (*At4g13040*), worked as a positive regulator of disease defense by up-regulating the accumulation of salicylic acid (SA) [18]. The members of the ERF subfamily typically bind to the cis-acting element GCC-box and are involved in the signaling pathways of plant hormone, e.g., ethylene (ET), SA, jasmonic acid (JA) and abscisic acid (ABA), which play an important role in both plant growth and development and response to stresses [19],[20],[21]. On the other hand, the DREB subfamily recognizes the conserved CCGAC motif of the dehydration-responsive element present in stress-responsive genes and is associated with the response of plants to abiotic stresses [22],[23],[24].

The AP2/ERF family has been well characterized in the model plants, *A. thanliana* [6],[7] and *Medicago truncatula* [25], several crops, such as rice [7], maize [26], soybean [27], Chinese cabbage [10] and grapevine [28], and *Populus trichocarpa* [29]. However, little is known about the AP2/ERF family in the medicinal herb, *Panax ginseng* (ginseng). *Ginseng* is a perennial of the *Araliaceae* family and has long been cultivated for human medicine in Asia, particularly in China, Korea, and Japan. Ginseng, known as the “king of all herbs” in China, is mainly cultivated in Jilin Province; therefore, it is often known as Jilin ginseng. Ginseng has been widely used as a medicinal herb due to its bioactive components, especially ginsenosides that have been shown to play significant roles in anti-inflammation

[30],[31], antitumor [32], and immunomodulation [33]. However, ginseng has been suffering from various biotic and abiotic stresses, which is greatly threating the ginseng production. Therefore, identification, characterization and utilization of the defense-related genes in ginseng are of significance for ginseng breeding and production. In the present study, we comprehensively studied the AP2/ERF family present in Jilin ginseng in several aspects, including gene identification, protein motif characterization, functional categorization and phylogenetic analysis. Moreover, the expression activities and patterns of AP2/ERF genes were also investigated at different developmental stages, in different tissues, different cultivars, under cold stress and under the methyl jasmonate (MeJA) treatment. The results of these studies have laid the foundation for deeply functional analysis and utilization of the genes of the AP2/ERF family and provided vital information on the molecular mechanism of plant response to biotic and abiotic stresses in ginseng and related plant species.

2. Materials and methods

2.1 Databases

We previously established a comprehensive transcriptome for Jilin ginseng from 14 tissues (fiber root, leg root, main root epiderm, main root cortex, rhizome, arm root, stem, leaf peduncle, leaflet pedicel, leaf blade, fruit peduncle, fruit pedicel, fruit flesh, and seed), from which 248,993 transcript unigenes (130,557 gene IDs) were assembled [34]. Moreover, we also sequenced and established the databases for the transcriptomes of the roots of 5-, 12-, 18- and 25-year-old plants [34] and the roots of four-year-old plants of 42 genotypes (named from S1 to S42) representing the diversity of Jilin ginseng [35]. In this study, a ginseng line IR826 genome sequence database [36] and another Ginseng Genome Database (http://ginsengdb.snu.ac.kr/index.php) reported by Kim *et al.* [37] were also used. In addition,

a transcriptome database of the adventitious roots of ginseng cv. Cheongsun treated with 200 μM MeJA for 0, 12, 24 and 48 h, respectively [38] was also consulted.

2.2 Identification of *PgERF* genes in ginseng

To identify the genes of the AP2/ERF family in ginseng, the Hidden Markov Model (HMM) profile of the AP2/ERF domain (Pfam: PF00847) and the protein sequences of the AP2/ERF genes downloaded from NCBI (http://blast.ncbi.nlm.nih.gov/Blast) were used to query the 248,993 Jilin ginseng transcript unigenes [34] by TBLASTN at E-value ≤ le-6. The obtained sequences were then used as a query to search for homologs in the ginseng line IR826 genome database [36]. Furthermore, TBLASTN were performed again to search the 248,993 transcript unigenes [34] using the homologs as query with E-value ≤ le-6 to maximize identification of the AP2/ERF family genes in ginseng. After merged all these aforementioned results, the identified genes were defined as *PgERF* for the AP2/ERF genes in ginseng and extracted by a Perl programming software. Finally, the predicted *PgERF* genes were analyzed by the conserved domain database (CDD) (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and the ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) at NCBI.

2.3 Multiple sequence alignment and phylogenetic analysis of *PgERF* genes

The encoded AP2 domain of the *PgERF* genes was aligned using the ClustalW program [39] and an unrooted phylogenetic tree was first constructed from the genes to have a preliminary phylogenetic tree of the *PgERF* genes. Then, another unrooted phylogenetic tree was constructed using 93 of the predicted *PgERF* genes representing the AP2/ERF family in ginseng and 147 *AtERF* genes previously identified and annotated in Arabidopsis [7]. These two phylogenetic trees were both constructed using

MEGA 5.0 by the Neighbor-Joining method with 1,000 bootstrap replications, using the Poisson correction model and the pairwise deletion [40].

2.4 Motif prediction of *PgERF* genes

The putative protein sequences of the above 93 *PgERF* genes used for construction of the AP2/ERF family phylogenetic tree were subjected to the online software, MEME (multiple EM for motif elicitation, V5.0.3) (http://meme-suite.org/tools/meme) [41] to identify the conserved motifs of these predicted *PgERF* genes in ginseng. Motif length was set to 6 - 50 amino acids and the maximum number of motifs was set to 25, while other parameters were set as default.

2.5 Expression and functional networks of *PgERF* genes

The expression profiles of all putative transcripts of *PgERF* genes identified above were extracted by a Perl programming software from the above four transcriptome databases: (1) the 14 tissues of a Jilin ginseng four-year-old plant, (2) the roots of Jilin ginseng 5-, 12-, 18- and 25-year-old plants, (3) the four-year-old roots in 42 Jilin ginseng genotypes and (4) the ginseng cv. Cheongsun adventitious roots treated with 200 μM MeJA for 0, 12, 24 and 48 h, respectively. The expression profiles of the putative transcripts of *PgERF* genes were measured as transcripts per million (TPM) and visualized by expression heatmap using the R programming language and software (http://www.r-project.org/, V3.3.3). Finally, the co-expression networks of these *PgERF* gene transcripts were constructed and analyzed among different tissues and different genotypes of Jilin ginseng using the BioLayout Express^{3D} software (Version 3.2) [42].

2.6 Expression activity of *PgERF* genes responding to cold stress

Equivalent ginseng hair roots (1 gram) were freshly cut from mature hair roots and cultured with

250 ml 1/2 Murashige and Skoog (MS) medium in dark culture at 22℃ for 30 days. Then, to simulate the cold stress treatments, the 30-days-old hair roots were placed in 4℃ for 6 h, 24 h, 48 h and 72 h, respectively. Afterwards, the ginseng hair roots were harvested and stored in -80℃ for the RNA isolation and further quantitative real-time PCR analysis. The total RNA of ginseng hair roots was extracted by TRIzol reagent (Bioteke, Beijing, China) according to the manufacturer's instructions, which was further reverse transcribed into cDNA using a PrimeScript™ RT reagent Kit with gDNA Eraser (TaKaRa, Tokyo, Japan), following the manufacturer's instructions. In this study, quantitative real-time PCR (qRT-PCR) of six *PgERF* genes, including *PgERF073*, *PgERF079*, *PgERF110*, *PgERF115*, *PgERF120* and *PgERF128,* was performed. The *PgGADPH* gene was used as the internal reference. The gene-specific primers used in qRT-PCR were designed by Primer Premier Software (version 5) and were listed in Table S1. The qRT-PCR was conducted by an Applied Biosystems 7500 Real Time PCR System (Thermo Fisher Scientific，Waltham, MA, USA) and SYBR Premix Ex Taq™ II (TaKaRa, Tokyo, Japan). The qRT-PCR was performed using the following steps: 30 seconds at 95℃; forty cycles of 5 seconds at 95℃ and 34 seconds at 60℃; one cycle of 15 seconds at 95 ℃ and 60 seconds at 60℃; 15 seconds at 95 ℃. The relative expression levels of these selected genes were calculated using formula $2^{-\Delta\Delta C_T}$ [43], and all the experiments were amplified in triplicate.

3. Results

3.1 Identification and classification of *PgERF* genes

A total of 397 transcripts that were derived from 189 predicted *PgERF* genes, including those containing a partial or complete AP2/ERF domain, were identified. These *PgERF* genes were defined *PgERF001* to *PgERF189*, with a suffix (e.g., -1) for different transcripts derived from the same gene (Table S2). Then, the sequences of 342 AP2/ERF gene transcripts downloaded from Ginseng Genome Database (http://ginsengdb.snu.ac.kr/index.php) were aligned with 397 transcripts identified in this study with identity≥95%, alignment length≥200 bp (about AP2 maximum domain length). As a result,

175 138 (73%) *PgERF* genes which contained 302 transcripts (88%) identified in this study were similar to
176 266 (78%) transcripts of the AP2/ERF gene from Ginseng Genome Database, which were supposed to
177 the same genes (Fig.S1 and Table S3). However, the other 51 (27%) *PgERF* genes, whose sequences
178 were quite different from Korean ginseng AP2/ERF genes, were assigned as newly discovered
179 AP2/ERF genes in ginseng (Fig.S1 and Table S3). The *PgERF* gene transcripts identified in this study
180 had nucleotide sequences ranging from 203 bp to 2,897 bp, with an average length of 1,216 bp. Of
181 these 397 *PgERF* gene transcripts, 176, derived from 96 *PgERF* genes, had partial AP2 domains, or
182 complete AP2 domains but being outside of open reading frames (ORFs). The remaining 221 *PgERF*
183 transcripts, derived from 93 *PgERF* genes, had complete AP2 domains within ORFs. Therefore, these
184 221 *PgERF* transcripts were further analyzed. The 221 *PgERF* gene transcripts encode putative
185 proteins with a length varying from 96 to 561 amino acids, with an average length of 256 amino acids
186 (Table S4). Analysis using the ExPASy Server showed that these putative proteins had an isoelectric
187 point between 4.43 (*PgERF025*) and 11.12 (*PgERF180*) and a molecular mass ranging from 11.00 kDa
188 (*PgERF140*) to 62.83 kDa (*PgERF159*) (Table S4).

189 Sakuma et al. [6] classified that the Arabidopsis AP2/ERF family into five subfamily, ERF, DREB,
190 AP2, RAV and Soloist. We also classified the 93 predicted *PgERF* genes whose transcripts had
191 complete AP2 domains within ORFs into these five subfamilies, according to the structures and the
192 number of AP2/ERF domains. These five *PgERF* subfamilies, ERF, DREB, AP2, RAV and Soloist,
193 contained 27, 48, 14, 2 and 2 genes, respectively (Fig. 1). Specifically, two genes, *PgERF035* and
194 *PgERF084*, were classified into the Soloist subfamily due to their low homology with the remaining
195 AP2/ERF genes and their high homology with the Arabidopsis Soloist subfamily gene, *AT4G13040*.
196 The *PgERF112* gene was the only one that codes one AP2/ERF domain with one B3 domain; therefore,
197 it was classified into the RAV subfamily. Moreover, *PgERF171* was also classified into the RAV

198 subfamily, even though it does not contain the B3 domain, because it has a high homology with the
199 Arabidopsis RAV subfamily genes, *AT1G51120* and *AT1G50680*. As seven *PgERF* genes, *PgERF062*,
200 *PgERF089*, *PgERF134*, *PgERF135*, *PgERF142*, *PgERF148* and *PgERF159*, contain two repeated
201 AP2/ERF domains, they were classified into the AP2 subfamily. In addition, another seven *PgERF*
202 genes, including *PgERF020*, *PgERF045*, *PgERF048*, *PgERF076*, *PgERF101*, *PgERF132* and
203 *PgERF140*, were also classified into the AP2 subfamily as they have high sequence similarity with the
204 members of the AP2 subfamily, even though they do not contain two repeated AP2/ERF domains. Of
205 the 75 remaining *PgERF* genes, 48 and 27 were classified into the ERF and DREB subfamilies,
206 respectively (Fig. 1). Furthermore, the DREB and ERF subfamilies of ginseng were each further divided
207 into six groups, A1 through A6 and B1 through B6, respectively, or these two subfamilies were divided
208 into 12 groups, from I to X, VI-L and Xb-L, based on Nakano et al. [7] (Fig. 1).

209 Figure 1. Phylogenetic tree of the AP2/ERF gene family present in ginseng and Arabidopsis. The
210 amino acid sequences of the AP2 domain were aligned using Clustal W and the phylogenetic tree was
211 constructed using neighbor-joining method.

212 **3.2 Phylogenetic analysis of the *PgERF* gene family**

213 To determine the phylogeny of the *PgERF* gene family, the 93 predicted *PgERF* genes whose
214 transcripts had complete AP2 domains were also used. An unrooted phylogenetic tree was constructed
215 from the 93 predicted *PgERF* genes for the *PgERF* gene family. The result showed that the *PgERF*
216 gene family was apparently classified into five clades, corresponding to ERF, DREB, AP2, RAV and
217 Soloist subfamilies (Fig. S2). Then, we constructed an unrooted phylogenetic tree of the *PgERF* gene
218 family using 147 Arabidopsis annotated *AtERF* genes as controls based on their conserved AP2
219 domains. The resultant phylogenetic tree clustered the 93 *PgERF* genes and 147 *AtERF* genes into 15
220 distinct clades, of which 13 was corresponding to the I, II, III, IV, V, VI, VII, VIII, IX and X groups of

the DREB and ERF subfamilies, and AP2, RAV and Soloist subfamilies of the *PgERF* family, and two, corresponding to Xb-L and VI-L groups of the DREB and ERF subfamilies, had only *AtERF* genes (Fig. 2). This result was consistent with the classification of the AP2/ERF family.

3.3 Motif identification and multiple sequence alignment

Next, the 93 predicted *PgERF* genes were subjected to conservative analysis for the conserved motifs of their proteins (Fig. 2; Fig. S3). A total of 25 conserved motifs were identified for the putative proteins of the 93 predicted *PgERF* genes, which were herein designated as Motif 1 through Motif 25. Motif 1 to Motif 6 were located in the AP2/ERF domain and the remaining 19 motifs, including Motif 7 through Motif 25, were found outside the AP2/ERF domain. The *PgERF* proteins encoded by gene members of the same subfamily or group contained similar conserved motifs. For example, Motifs 4, 7, 15, 16, 17, 21 and 25 were specifically shared by gene members of the AP2 subfamily. Motifs 14 and 18 were specifically present within Group I, while Motifs 9 and 20 were only present in the Group III gene members of the ERF subfamily. Motifs 10, 11, 12, 13, and 19 were specific for the DREB subfamily and absent in all four other subfamilies (ERF, AP2, RAV and Soloist subfamilies). These results suggested that most of the 25 motifs were divergent among subfamilies or groups, which might play an important role in their functional divergence [25].

Figure 2. Distribution of conserved motifs among the gene members of the *PgERF* family. Each motif is represented by a colored box. Box length corresponds to the motif length.

As the featured sequences within the specific domains of transcription factors are critical to their functions [44], the conserved amino acid residues of the AP2 domains were identified for the genes of both DREB and ERF subfamilies. By aligning the amino acid sequences of the AP2 domains of these two subfamilies in ginseng and Arabidopsis, 14 conserved amino acid residues, including 4G, 6R, 8R,

15W, 16V, 18E, 20R, 22P, 39W, 40L, 49A, 52A, 54D and 72N, and 11 conserved amino acid residues, including 4G, 6R, 8R, 11G, 17I, 30R, 42A, 46Y, 47D, 55G and 63F, were identified for the DREB and ERF subfamilies, respectively (Fig. S4 and S5). Besides, all gene members of the DREB subfamily and ERF subfamily obviously contained the two featured conserved elements, including YRG and RAYD, in their AP2 domains.

3.4 Functional categorization of the *PgERF* genes

To estimate the functional differentiation of the *PgERF* family, all 397 transcripts of the 189 *PgERF* genes identified in this study were annotated and functionally categorized using the Blast2GO software (Version 4.1.9) [45]. Surprisingly, only 195 (49%) of the *PgERF* gene transcripts could be annotated, while the remaining 202 could not be annotated using the database of the Blast2GO software, suggesting the uniqueness of the *PgERF* genes in ginseng. The annotated *PgERF* gene transcripts were categorized into all three primary gene ontology (GO) categories, molecular function (MF), biological process (BP) and cellular component (CC) (Fig. 3A). Of the 195 *PgERF* transcripts, 186 (95%) were categorized into all the three primary categories, MF, BP and CC. Only one *PgERF* gene transcript, *PgERF069*, had functions in both BP and CC categories, while two *PgERF* transcripts, *PgERF135-1* and *PgERF135-3*, and six *PgERF* transcripts, *PgERF152-3*, *PgERF152-4*, *PgERF152-1*, *PgERF152-2*, *PgERF105-3* and *PgERF087*, were categorized into BP and CC, respectively. At Level 2, these 195 transcripts were further categorized into eight subcategories, including nucleic acid binding transcription factor activity, binding, metabolic process, cellular process, developmental process, organelle, cell part and cell (Fig. 3B). Of the eight subcategories, three subcategories, including nucleic acid binding transcription factor activity, binding and cell part subcategories, have been significantly enriched, which is consistent with the transcription regulation functions of the *ERF* genes, while the

abundances of the remaining five subcategories are either not changed or significantly reduced relative to the whole genome background control.

Figure 3. Functional categorization of the AP2/ERF gene family in ginseng. (A) Venn diagram of the *PgERF* gene transcripts categorized into three primary categories, biological process (BP), molecular function (MF) and cellular component (CC). (B) The subcategories of the *PgERF* gene transcripts (Level 2). The enrichment of the *PgERF* gene transcripts in each subcategory was calculated using all the gene transcripts of ginseng as the background control. A single asterisk "*" indicates the significant difference of the number of *PgERF* gene transcripts categorized into the subcategory from that of all the gene transcripts of ginseng at $P \leq 0.05$, while double asterisks "**" indicate the difference at a significance level of $P \leq 0.01$.

Furthermore, the *PgERF* transcripts expressed in the roots of 5-, 12-, 18- and 25-year-old plants, 14 tissues of the 4-year-old plant, and the roots of 4-year-old plants of 42 genotypes were further categorized (Fig. 4). The *PgERF* transcripts expressed in differently-aged plant roots, different tissues and the roots of different genotypes were all categorized into these eight subcategories, suggesting that the functions of *PgERF* transcripts were consistent among developmental stages, tissues or genotypes. Nevertheless, a substantial variation of the categorization in the numbers of the *PgERF* transcripts categorized into these eight categories (Level 2) was observed across developmental stages, tissues or genotypes.

Figure 4. Variation in functional categories of the *PgERF* gene transcripts. (A) Variation in functional categories among the roots of differently aged plants. (B) Variation in functional categories among 14 tissues of a 4-year-old plant. (C) Variation in functional categories among the 4-year-old

plant roots of 42 genotypes.

3.5 Expression profiles and networks of the *PgERF* genes

To profile the activation patterns of *PgERF* genes, the expressions of all 397 transcripts were quantified in 5-, 12-, 18- and 25-year-old plant roots, 14 4-year-old plant tissues and the roots of 4-year-old plants of 42 genotypes. The expressions of the transcripts varied dramatically across developmental stages, tissues and genotypes, from silenced (0.0 TPM) to 586.3, 666.0 and 1159.2 TPM, respectively. Of the 397 *PgERF* transcripts, 136 (34.3%), 98 (24.7%) and 83 (20.9%) expressed in all 5-, 12-, 18- and 25-year-old plant roots, all 14 4-year-old plant tissues and the roots of 4-year-old plants of all 42 genotypes, respectively (Tables S5-S7), while 53 (13.4%), 39 (9.8%) and 14 (3.5%) of the 397 transcripts were development-, tissue- and genotype-specific, respectively. Nevertheless, the expression of a transcript varied dramatically across developmental stages, tissues and genotypes.

Moreover, we constructed the heatmaps of the *PgERF* genes expressed at different developmental stages of roots, in different tissues, and across different genotypes to find out whether the expressions of the genes were co-regulated. The results showed that although the expression co-regulation was observed for some of the genes at a developmental stage, a single tissue or a genotype and across developmental stages, it was not apparent across tissues or genotypes (Fig. 5). For instance, *PgERF140-12*, *PgERF046*, *PgERF089-3*, *PgERF093-3*, *PgERF108-1*, *PgERF184*, *PgERF118-2* and *PgERF170* were apparently co-regulated at a developmental stage and across developmental stages of roots (Fig. 5A).

Figure 5. Expression heatmaps of the *PgERF* gene transcripts at different developmental stages, in different tissues and across genotypes. (A) In the roots of different year-old plants. (B) In the 14 tissues of a 4-year-old plant. (C) In the 4-year-old plant roots of 42 genotypes.

To determine the functional relationships of the *PgERF* genes, the co-expression network of the

309 *PgERF* transcripts were constructed for 14 tissues of a four-year-old plant and the four-year-old plant
310 roots of 42 genotypes, respectively. Of the 397 *PgERF* gene transcripts, 364 (91.7%) formed a
311 co-expression network ($P \leq 0.05$) in the 14 tissues of the four-year-old plant (Fig. S6A). The network
312 consisted of 364 gene transcript nodes, 5,303 co-expression edges and 17 closer co-expression clusters
313 (Fig. S6A and B). Nevertheless, the tendency of this network formation had no substantial difference
314 from that of the network formed from randomly selected ginseng gene transcripts (Fig. S6C and D). In
315 the four-year-old plant roots of different genotypes, 341 (85.9%) of the 397 *PgERF* gene transcripts
316 formed a co-expression network ($P \leq 0.05$), consisting of 341 gene transcript nodes, 5,606
317 co-expression edges and 24 clusters (Fig. 6A and B). The tendency of this network formation was
318 stronger in terms of number of nodes and number of edges than that of the network formed from
319 randomly selected ginseng gene transcripts (Fig. 6C and D). Together, analysis of these networks
320 revealed that the gene members of the *PgERF* gene family were functionally quite independent, even
321 though some of them formed a co-expression network, because the tendency of the network formation
322 was similar to that of randomly-selected unknown genes.

323 Figure 6. Co-expression network of the *PgERF* gene transcripts in the 4-year-old plant roots of 42
324 genotypes. (A) The co-expression network constructed from 342 of the 397 *PgERF* gene transcripts at P
325 ≤ 0.05. (B) 17 clusters of the network. (C) Variation in number of nodes in the network of *PgERF*
326 transcripts at different *P*-values. (D) Variation in number of edges in the network of *PgERF* transcripts
327 at different *P*-values.

328 **3.6 Expression profiles of the *PgERF* genes in responding to cold stress**

329 As a perennial herb, ginseng is frequently suffering from various environmental stresses.
330 However, to date, the molecular mechanisms of the stress tolerance in ginseng were not clearly

clarified. To discover the potential functions of *PgERF* genes in resistance cold stress, the expression patterns of six *PgERF* genes randomly selected from DREB subfamily, i.e., *PgERF073*, *PgERF079*, *PgERF110*, *PgERF115*, *PgERF120* and *PgERF128*, in cold-stressed ginseng hair roots were analyzed by qRT-PCR. As shown in Fig.7D and 7E, two members of I group of DREB subfamily, i.e., *PgERF115* and *PgERF120* were firstly up-regulated but then somewhat different from each other by cold stress. Expression of *PgERF115* gradually rose and the highest change showed at cold stress for 24 h (about 13.06 times higher than the untreated hair roots), and after that time point, the expression level of *PgERF115* declined regularly. *PgERF120* respond quickly to cold stress, reaching a 26.88 times higher than the untreated hair roots in cold-stressed ginseng hair roots for 6 h and regularly declined to normal level. Similarly, the expression of *PgERF073* and *PgERF110*, two members of II group, were also somewhat different from each other. The expression of *PgERF073* showed similar trends to *PgERF115* while the expression of *PgERF110* showed similar trends to *PgERF120* (Fig.7A and 7C). *PgERF079*, a member from III group of DREB subfamily, responded rapidly and drastically to the cold stress, whose expression were significantly up-regulated by 1057.05, 274.44, 290.81 and 173.03 times in cold-stressed for 6 h, 24 h, 48 h and 72 h comparing to the untreated hair roots ($P < 0.01$) (Fig.7B). *PgERF128*, which belonged to the IV group of DREB subfamily, exhibited particular trend comparing with the other 5 *PgERF* genes. As shown in Fig.7F, the expression levels of *PgERF128* were rising gradually in ginseng hair roots under cold stress for 6 h, 24 h, 48 h and 72 h. At 72 h, the expression of *PgERF128* in cold-stressed ginseng hair roots were significantly up-regulated by 5.68 times than the untreated hair roots ($P < 0.01$).

Figure 7. Expression levels of *PgERF* genes in ginseng hair roots after 6, 24, 48 h and 96 h of cold stress treatment. The values were given as mean ± SD of triplicate samples. Different letters represent significant differences between the treatment means ($p < 0.05$, LSD).

3.6 Expression profiles of the *PgERF* genes in responding to MeJA

MeJA is a plant hormone and a kind of elicitors and has been widely used in regulation of genes involved in ginsenoside biosynthesis in ginseng [38]. Therefore, we further analyzed the expressions of the *PgERF* genes in the adventitious roots of ginseng treated with MeJA for 0, 12, 24 and 48 h, respectively. The expressions of the *PgERF* gene transcripts in the control and MeJA-treated adventitious roots varied from silent (0 TPM) to 197.731 TPM (Table S8). Of the 397 *PgERF* gene transcripts profiled, 173 (43.6%) expressed and 109 (27.5%) silenced in the control and all treated adventitious roots, and the remaining 115 (29.0%) either expressed or silenced in these adventitious roots. The expressions of the 288 *PgERF* gene transcripts expressed the adventitious roots were visualized by the expression heatmap (Fig. 8). Overall, all the 288 *PgERF* gene transcripts responded to the MeJA treatment, with 136 of them up-regulated and 152 down-regulated by MeJA. Among the three treatment times, 12 h, 24 h and 48 h, the responses of these *PgERF* gene transcripts to MeJA varied from time to time.

Figure 8. Expression heatmap of the *PgERF* gene transcripts treated with 200 μM MeJA for 0, 12, 24 and 48 h, respectively.

4. Discussion

The AP2/ERF gene family has been broadly studied in several plant species of economical or biological importance due to its important roles in various biological processes, including growth and development, and responses to environmental stresses based on genome and transcriptome sequences [5]. These species include Arabidopsis [6], rice [7], wheat [46], maize [26], cotton [3], grapevine [28], cucumbers [47] and rubber tree [48]. We have, in this study, comprehensively investigated the AP2/ERF genes in ginseng using several transcriptome databases, including those developed from 14 tissues of a four-year-old ginseng plant, the roots of 5-, 8-, 12- and 25-year-old plants and the four-year-old plant roots of 42 diverse genotypes. The *PgERF* gene family in ginseng is also a large

gene family, consisting of 189 or more gene members. This result is in consistence with those identified in other plant species such as Arabidopsis [6], rice [7] and grapevine [28]. Although the family size is non-comparable with those identified in the other species listed above due to the dramatic variation of gene family size within a plant species [49] and the difference of the databases used for these analyses, the *PgERF* gene family is unambiguously classified into five subfamilies, ERF, DREB, AP2, RAV and Soloist, as were those identified in Arabidopsis [6], rice [7] and grapevine [28]. These results indicate the *PgERF* gene family has a similar functional differentiation pattern as those in the three latter species.

It has been consensus that the conserved motifs of the AP2/ERF transcription factor are crucial to the function of transcription factors, such as nuclear localization and transcriptional activity [7]. The DNA binding domain of AP2/ERF transcription factors, i.e., AP2 domain, was highly conserved in plant species [50],[51]. The AP2 domain of the *PgERF* genes was also found to be highly conserved. This study has identified 14 and 11 completely conserved amino acid residues through all gene members of the DREB and ERF subfamilies, respectively, in both ginseng and Arabidopsis (Figs. S3 and S4). In Arabidopsis, the two conservative elements, YRG and RAYD, were shown to be critical to the binding of AP2/ERF transcription factors to the promoter regions of the target genes and modulate their expression [29]. The conservative YRG and RAYD elements identified in the AP2 domain of the DREB and ERF subfamilies in ginseng may suggest their necessity for similar functions of the *PgERF* genes. Nevertheless, subtle variation exists among the amino acid sequences of *PgERF* transcription factors, which has led to the separation of the DREB subfamily from the ERF subfamily. The difference between the DREB and ERF subfamilies might result in their functional divergence in ginseng. Moreover, the “EIR” in the AP2 domain was found to be shared by all gene members of the ERF

400 subfamily and the vast majority of the gene members of the DREB subfamily in both ginseng and
401 Arabidopsis, while the "EVR" exists only in Group III of the DREB subfamily in both ginseng and
402 Arabidopsis and only in Group II (*PgERF061*) of the DREB subfamily in ginseng. It has been reported
403 that the sequence similarity of the conserved motifs that exist outside of the DNA binding domain was
404 low [6],[27],[43]. In ginseng, 19 conserved motifs, except for Motif 1 to Motif 6, were identified
405 outside the AP2 domain. The vast majority of these 19 motifs were found to be divergent across
406 subfamilies or even subfamily groups in ginseng. The subfamily/group-specific distribution pattern of
407 these motifs might have led to the functional divergence between subfamilies or groups of the *PgERF*
408 transcription factors.

409 Because different transcripts alternatively spliced from the same gene may have different
410 functions [52], the 397 *PgERF* transcripts, instead of the 189 *PgERF* genes, were annotated and
411 functionally categorized in this study. The *PgERF* transcripts were categorized into eight subcategories
412 at Level 2. Although this result suggested a substantial functional differentiation of the *PgERF* genes,
413 the differentiation was much smaller than those observed in the *PgNBS* gene family [35], *PgRLK* gene
414 family [53] and *PgCYP* gene family [54] in ginseng. Interestingly, of the eight Level 2 subcategories,
415 only two, especially those in the nucleic acid binding transcription factor activity subcategory, were
416 significantly up-enriched, which is consistent with the roles of the *PgERF* genes as transcription factors
417 by binding to the promoters of target genes. While the functions of the AP2/ERF genes have been
418 shown to play important roles in plant growth and development, response to stresses and signal
419 pathway in the model plants such as Arabidopsis and rice [5],[55], further research is needed to
420 determine the functions of the *PgERF* genes in ginseng.

421 Companioned with their functional differentiation, the expressions of 397 *PgERF* transcripts
422 dramatically varied in a tissue, at a development stage or in a genotype. Moreover, the type and number

of expressed *PgERF* transcripts also diversified tempo-spatially and across genotypes. The differential expressions of the AP2/ERF genes were previously reported in other plant species, but mainly among tissues [25],[56]. Furthermore, the numbers of the *PgERF* transcripts categorized into each subcategory varied across tissues, developmental stages or genotypes. These variations might be an indication of their functional differentiation. On the other hand, co-expression network analysis revealed that most (>86%) of these *PgERF* transcripts express correlatively and tend to form a co-expression network in different tissues or different genotypes. These results suggest that the *PgERF* genes have functionally differentiated, but they are still somehow functionally collaborative.

As a perennial herb, ginseng frequently suffers from different kinds of environmental stresses. It was reported that members of AP2/ERF superfamily, especially DREB subfamily, played an essential role in response to biotic and abiotic stresses [16],[17],[57]. To tap the potential AP2/ERF genes of DREB subfamily resisting to cold stress in ginseng, the expression of six genes randomly selected from DREB subfamily under cold stress were analyzed using qRT-PCR. The expression of *PgERF079*, one gene from III group or A1 group of DREB subfamily, was dramatically changed (up to 1057.05 times with a brief period of cold-stressed for 6 h), suggesting it played an extremely important role in freezing tolerance. In fact, A-1 group was considered to be the major regulator of cold-stress responses as overexpressing any one of the three cold-inducible DREB1s, DREB1A/CBF3 (AT4G25480), DREB1B/CBF1 (AT4G25490) and DREB1C/CBF2 (AT4G25470), significantly improved freezing tolerance in Arabidopsis [57],[58],[59]. Besides, the expression level of the other five genes from other groups of DREB subfamily, i.e., *PgERF073*, *PgERF110*, *PgERF115*, *PgERF120* and *PgERF128*, also showed significant changes ($p<0.01$) in cold-stressed ginseng hair roots. Therefore, it speculates that besides A1 group, the other groups of DREB subfamily may also be effective in freezing tolerance in

445 ginseng, either directly or indirectly. Herein, the results of these cold-inducible genes would provide
446 some valuable information for the functional studies of *PgERF* genes in ginseng in the future.

447 It has been reported that some genes of the AP2/ERF family are involved in response to hormone
448 signals in plants [60],[61],[62]. It was showed that MeJA, as one of the signaling molecules, was rapidly
449 synthesized in plants, when subjected to various biotic and abiotic stresses, and then, induced
450 defense-related responses to the stresses and regulate plant growth and development [63]. MeJA has
451 been also used as an effective elicitor, since it can stimulate the biosynthesis of plant secondary
452 metabolites [64],[65],[66]. The biosynthesis and accumulation of ginsenosides, a cluster of important
453 secondary metabolites and the most valuable bioactive components in ginseng, were also reported to be
454 induced by MeJA [67],[68]. This study showed that the addition of exogenous MeJA to adventitious
455 roots dramatically changed the expression of a majority of the *PgERF* gene transcripts. The expressions
456 of some of the transcripts were up-regulated while those of the other down-regulation or inhibited by
457 MeJA, relative to the control not treated by MeJA. Given the demonstrated functions of MeJA in plant
458 responses to biotic and abiotic stresses, growth and development and secondary metabolite biosynthesis
459 in other plant species [57],[60],[61],[62],[63],[64],[65],[66], the *PgERF* genes may also be involved in
460 these processes, including the biosynthesis of ginsenosides.

461 ## 5. Conclusions

462 The present study, for the first time, reports identification and systematic characterization of the
463 AP2/ERF family present in ginseng, i.e., the *PgERF* gene family. A total of 189 *PgERF* genes that were
464 actively expressed in 14 tissues of a four-year-old ginseng plant were identified and these genes were
465 alternatively sliced into 397 transcripts. These *PgERF* genes were also classified into five subfamilies

(DREB, ERF, AP2, RAV and Soloist) as those previously identified in Arabidopsis. As expected, the conserved motifs that characterize the AP2/ERF family and several conserved domains were identified among the members of the *PgERF* gene family. Nevertheless, the transcripts of the *PgERF* genes were apparently categorized into eight subcategories by GO, especially into the subcategory for nucleic acid binding transcription factor activity, which indicates their functional differentiation. Along with their functional differentiation, the expressions of the *PgERF* genes, including the type, number and expression level of their transcripts, have also substantially diversified tempo-spatially and across genotypes. In spite of these differentiations, most of the *PgERF* genes remain to co-express and form a co-expression network, suggesting that most of the genes in the *PgERF* gene family remain functionally correlated. These *PgERF* genes and findings provide resources and knowledge valuable for family-wide functional analysis of the *PgERF* genes and determination of their roles in plant responses to biotic and abiotic stresses, growth and development, and biosynthesis of secondary metabolites, especially ginsenosides, in *P. ginseng* and related species.

SUPPLEMENTARY MATERIAL

Supplemental information is available with the online version of this manuscript.

ACKNOWLEDGMENT

This research was supported by an award from China 863 Project (2013AA102604-3), the Bureau of Science and Technology of Jilin Province (20190201264JC, 20170101010JC，20180414077GH，20180101027JC), the Development and Reform Commission of Jilin Province (2016C064, 2018C047-3), and a startup fund from Jilin Agriculture University (201801, https://www.jlau.edu.cn/).

AUTHOR CONTRIBUTIONS

MPZ and YW planned and designed this study; JC and QZ performed the bioinformatic analysis; JC wrote the manuscript; LL, YZ, PC, HL, RL, YH, CS, KW, JL, MZ and YFW prepared the tables and figures. MPZ revised the manuscript. All the authors read and approved the final version of the manuscript.

REFERENCES

1. Mawlong I, Ali K, Srinivasan R, Rai RD, and Tyagi A (2015) Functional validation of a drought-responsive AP2/ERF family transcription factor-encoding gene from rice in *Arabidopsis*. *Mol. Breeding* 35, 163.

2. Chinnusamy V, Zhu J, and Zhu JK (2007) Cold stress regulation of gene expression in plants. *Trends Plant Sci.* 12, 444-451.

3. Champion A, Hebrard E, Parra B, Bournaud C, Marmey P, Tranchant C, et al. (2009) Molecular diversity and gene expression of cotton ERF transcription factors reveal that group IXa members are responsive to jasmonate, ethylene and Xanthomonas. *Mol. Plant Pathol.* 10, 471-485.

4. Gutterson N, and Reuber TL (2004) Regulation of disease resistance pathways by AP2/ERF transcription factors. *Curr. Opin. Plant Bio.* 7, 465-471.

5. Francesco L, Masaru OT, and Pierdomenico P (2013) APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. *New Phytol.* 199, 639-649.

6. Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, and Yamaguchishinozaki K (2002) DNA-Binding Specificity of the ERF/AP2 Domain of Arabidopsis DREBs, Transcription Factors Involved in Dehydration- and Cold-Inducible Gene Expression. *Biochem. Bioph. Res.*

Co. 290, 998-1009.

7. Nakano T, Suzuki K, Fujimura T, and Shinshi H (2006) Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiol.* 140, 411.

8. Li MY, Xu ZS, Huang Y, Tian C, Wang F, and Xiong AS (2015) Genome-wide analysis of *AP2/ERF* transcription factors in carrot (*Daucus carota* L.) reveals evolution and expression profiles under abiotic stress. *Mol. Genet. Genomics* 290, 2049.

9. Mrunmay Kumar G, Swadhin S, Janesh Kumar G, Subaran S, Nidhi S, Lipika B, et al. (2014) The *Arabidopsis thaliana At4g13040* gene, a unique member of the AP2/EREBP family, is a positive regulator for salicylic acid accumulation and basal defense against bacterial pathogens. *J. Plant Physiol.* 171, 860-867.

10. Song X, Li Y, and Hou X (2013) Genome-wide analysis of the AP2/ERF transcription factor superfamily in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Bmc Genomics* 14, 573.

11. Nole-Wilson S, and Krizek BA (2000) DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic Acids Res.* 28, 4076-4082.

12. Krizek BA (2009) *AINTEGUMENTA and AINTEGUMENTA-LIKE6* Act Redundantly to Regulate Arabidopsis Floral Growthand Patterning. *Plant Physiol.* 150, 1916-1929.

13. Mi-Young C, Julia V, Rob A, Jemin L, Ryan MQ, Jae-Dong C, et al. (2010) A tomato (*Solanum lycopersicum*) APETALA2/ERF gene, *SlAP2a*, is a negative regulator of fruit ripening. *Plant J.* 64, 936-947.

14. Kagaya Y, Ohmiya K, and Hattori T (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res.* 27(2), 470-478.

15. Feng CZ, Yun C, Wang C, Kong YH, Wu WH, and Chen YF (2014) The Arabidopsis RAV1 transcription factor, phosphorylated by SnRK2 kinases, regulates the expressions of *ABI3*, *ABI4*, and *ABI5* during seed germination and early seedling development. *Plant J.* 80, 654-668.

16. Min H, Zheng J, and Wang J (2014) Maize *ZmRAV1* contributes to salt and osmotic stress tolerance in transgenic *arabidopsis*. *J. Plant Biol.*

17. Xiao JL, Mo L, Ying Z, Shan H, Rong H, Yun C, et al. (2015) Overexpression of cotton *RAV1* gene in Arabidopsis confers transgenic plants high salinity and drought sensitivity. *Plos One* 10, e0118056.

18. Giri MK, Swain S, Gautam JK, Singh S, Singh N, Bhattacharjee L, et al. (2014) The Arabidopsis thaliana *At4g13040* gene, a unique member of the AP2/EREBP family, is a positive regulator for salicylic acid accumulation and basal defense against bacterial pathogens. *J. Plant Physiol.* 171, 860-867.

19. Fujimoto SM, Usui A, Shinshi H, and Ohme TM (2000) Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *Plant Cell* 12, 393-404.

20. Luis OAS, and Singh KB (2002) Identification of Arabidopsis ethylene-responsive element binding factors with distinct induction kinetics after pathogen infection. *Plant Physiol.* 128, 1313-1322.

21. Brown RL, Kemal K, Mcgrath KC, Maclean DJ, and Manners JM (2003) A role for the GCC-box in jasmonate-mediated activation of the *PDF1.2* gene of Arabidopsis. *Plant Physiol.* 132, 1020-1032.

22. Yamaguchi-Shinozaki K, and Shinozaki K (1994) A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell* 6,

251-264.

23. Jiang C, Iu B, and Singh J (1996) Requirement of a CCGAC *cis*-acting element for cold induction of the *BN115* gene from winter *Brassica napus*. *Plant Mol. Biol.* 30(3), 679-684.

24. Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, et al. (2002) Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant cell* 14, 1737-1749.

25. Shu Y, Liu Y, Zhang J, Song, L., and Guo, C (2016) Genome-Wide Analysis of the AP2/ERF Superfamily Genes and their Responses to Abiotic Stress in *Medicago truncatula*. *Front. Plant Sci.* 6, 1247.

26. Zhou ML, Tang YX, and Wu YM (2012) Genome-Wide Analysis of AP2/ERF Transcription Factor Family in *Zea Mays*. *Curr. Bioinform.* 7, 324-332. DOI: 10.2174/157489312802460776

27. Zhang GC, Ming Chen X, Xu Z, Guan S, Li LC, Li A, et al. (2008) Phylogeny, gene structures, and expression patterns of the *ERF* gene family in soybean (*Glycine max* L.). *J. Exp. Bot.* 59, 4095-4107.

28 . Licausi F, Giorgi FM, Zenoni S, Osti F, Pezzotti M, and Perata P (2010) Genomic and transcriptomic analysis of the *AP2/ERF* superfamily in *Vitis vinifera*. *Bmc Genomics* 11, 719.

29. Zhuang J, Cai B, Peng RH, Zhu B, Jin XF, Xue Y, et al. (2008) Genome-wide analysis of the *AP2/ERF* gene family in *Populus trichocarpa*. *Biochem. Bioph. Res. Co.* 371, 468-474.

30. Je-Hyuk L, Jung-Hun L, Yu-Mi L, Pit-Na K, and Choon-Sik J (2008) Potential analgesic and anti-inflammatory activities of *Panax ginseng* head butanolic fraction in animals. *Food Chem. Toxicol.* 46, 3749-3752.

31. Tao Y, Yang Y, Kwak YS, Song GG, Kim MY, Man HR, et al. (2017) Ginsenoside Rc from *Panax*

575 *ginseng* exerts anti-inflammatory activity by targeting TANK-binding kinase 1/interferon
576 regulatory factor-3 and p38/ATF-2. *J. Gins. Res.* 41, 127-133.
577 32. Shu G, Jiang S, Mu J, Yu H, Duan H, and Deng X (2018) Antitumor immunostimulatory activity of
578 polysaccharides from *Panax japonicus* C. A. Mey: Roles of their effects on CD4 + T cells and
579 tumor associated macrophages. *Int. J. Biol. Macromol.* 111, 430-439.
580 33. Scaglione F, Ferrara F, Dugnani S, Falchi M, Santoro G, and Fraschini F (1990)
581 Immunomodulatory effects of two extracts of *Panax ginseng* C.A. Meyer. *Drug. Exp. Clin. Res.*
582 16, 537.
583 34. Wang K, Jiang S, Sun C, Lin Y, Rui Y, Yi W, et al. (2015) The Spatial and Temporal
584 Transcriptomic Landscapes of Ginseng, *Panax ginseng* C. A. Meyer. *Sci. Rep.* 5, 18283.
585 35. Yin R, Zhao M, Wang K, Lin Y, Wang Y, Sun C, et al. (2017) Functional differentiation and
586 spatial-temporal co-expression networks of the *NBS*-encoding gene family in Jilin ginseng,
587 *Panax ginseng* C.A. Meyer. *Plos One* 12, e0181596.
588 36. Xu J, Chu Y, Liao B, Xiao S, Yin Q, Bai R, et al. (2017) *Panax ginseng* genome examination for
589 ginsenoside biosynthesis. *Gigascience* 6, 1-15.
590 37. Kim NH, Jayakodi M, Lee SC, Choi BS, Jang W, et al. (2018) Genome and evolution of the
591 shade-requiring medicinal herb Panax ginseng. Plant Biotechnol J 16: 1904-1917.
592 38. Kang KB, Jayakodi M, Yun SL, Nguyen VB, Park HS, Koo HJ, et al. (2018). Identification of
593 candidate UDP-glycosyltransferases involved in protopanaxadiol-type ginsenoside biosynthesis
594 in *Panax ginseng*. *Sci. Rep.* 8, 11744.
595 39. Thompson JD, Higgins DG, and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of
596 progressive multiple sequence alignment through sequence weighting, position-specific gap

penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680.

40. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* 28, 2731-2739.

41. Bailey TL, Nadya W, Chris M, and Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34, 369-373.

42. Athanasios T, Stjin VD, Enright AJ, and Freeman TC (2009) Network visualization and analysis of gene expression data using BioLayout Express^{3D}. *Nat. Protoc.* 4, 1535-1550.

43. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

44. Xu W, Li F, Ling L, and Liu A (2013) Genome-wide survey and expression profiles of the AP2/ERF family in castor bean (*Ricinus communis* L.). *BMC Genomics* 14, 785.

45. Ana C, Stefan GT, Juan Miguel GG, Javier T, Manuel T, and Montserrat R (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674-3676.

46. Zhuang J, Chen JM, Yao QH, Xiong F, Sun CC, Zhou XR, et al. (2011) Discovery and expression profile analysis of *AP2/ERF* family genes from Triticum aestivum. *Mol. Biol. Rep.* 38, 745-753.

47. Hu L, and Liu S (2011) Genome-wide identification and phylogenetic analysis of the *ERF* gene family in cucumbers. *Genet. Mol. Biol.* 34(4), 624-633.

48. Duan C, Argout X, Gébelin V, Summo M, Dufayard JF, Leclercq J, et al. (2013) Identification of the *Hevea brasiliensis* AP2/ERF superfamily by RNA sequencing. *BMC Genomics* 14, 30.

49. Zhang M, Wu YH, Lee MK, Liu YH, Rong Y, Santos TS, et al. (2010) Numbers of genes in the

NBS and RLK families vary by more than four-fold within a plant species and are regulated by multiple factors. *Nucleic Acids Res.* 38, 6513-6525.

50. Kizis D, Lumbreras V, and Pagès M (2001) Role of AP2/EREBP transcription factors in gene regulation during abiotic stress. *Febs Lett.* 498, 187-189.

51. Li X, Zhang D, Gao B, Liang Y, Yang H, Wang Y, et al. (2017) Transcriptome-Wide Identification, Classification, and Characterization of *AP2/ERF* Family Genes in the Desert Moss *Syntrichia caninervis*. *Front. Plant Sci.* 8, 262.

52. Carvalho RF, Feijão CV, and Duque P (2013) On the physiological significance of alternative splicing events in higher plants. *Protoplasma* 250, 639-650.

53. Lin, Y., Wang, K., Li, X., Sun, C., Yin, R., Wang, Y., and Zhang, M (2018) Evolution, functional differentiation, and co-expression of the *RLK* gene family revealed in Jilin ginseng, *Panax ginseng* C.A. Meyer. *Mol. Genet. Genomics* 293, 845-859.

54. Wang Y, Li X, Lin Y, Wang Y, Wang K, Sun C, et al. (2018) Structural Variation, Functional Differentiation, and Activity Correlation of the Cytochrome P450 Gene Superfamily Revealed in Ginseng. *The Plant Genome* 11, 1-11.

55. Mizoi J, Shinozaki K, and Yamaguchi-Shinozaki K (2012) AP2/ERF family transcription factors in plant abiotic stress responses. *BBA - Gene Regul. Mech.* 1819, 86-96.

56. Guodong R, Jinkai S, Yanfei Z, Caiyun H, and Jianguo Z (2015) Genome-wide analysis of the *AP2/ERF* gene family in *Salix arbutifolia*. *Febs Open Bio.* 5, 132-137.

57. Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, and Shinozaki K (1998) Two transcription factors, *DREB1* and *DREB2*, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive

gene expression, respectively, in Arabidopsis. *Plant Cell* 10, 1391-1406.

58. Park S, Lee C-M, Doherty CJ, Gilmour SJ, Kim Y, Thomashow MF (2015) Regulation of the Arabidopsis CBF regulon by a complex low-temperature regulatory network. Plant Journal 82: 193-207.

59. Jaglo-Ottosen RK (1998) Arabidopsis CBF1 Overexpression Induces COR Genes and Enhances Freezing Tolerance. Science 280: 104-106.

60. Oscar L, Raquel P, Sánchez-Serrano JJ, and Roberto S (2003) ETHYLENE RESPONSE FACTOR1 integrates signals from ethylene and jasmonate pathways in plant defense. *Plant Cell* 15, 165-178.

61. Zarei A, Körbes AP, Younessi P, Montiel G, Champion A, and Memelink J (2011) Two GCC boxes and AP2/ERF-domain transcription factor *ORA59* in jasmonate/ethylene-mediated activation of the *PDF1.2* promoter in Arabidopsis. *Plant Mol. Biol.* 75, 321-331.

62. Moffat CS, Ingle RA, Wathugala DL, Saunders NJ, Knight H, and Knight MR (2012) ERF5 and ERF6 Play Redundant Roles as Positive Regulators of JA/Et-Mediated Defense against *Botrytis cinerea* in Arabidopsis. *PLOS ONE*.

63. Benevenuto RF, Seldal T, Hegland SJ, Rodriguez-Saona C, Kawash J, and Polashock J (2019) Transcriptional profiling of methyl jasmonate-induced defense responses in bilberry (*Vaccinium myrtillus* L.). *BMC Plant Biol.* 19, 70.

64. Chen H, Jones AD, and Howe GA (2006) Constitutive activation of the jasmonate signaling pathway enhances the production of secondary metabolites in tomato. *Febs Lett.* 580, 2540-2546.

65. Hyun-Jin K, Feng C, Xi W, and Rajapakse NC (2006) Effect of methyl jasmonate on secondary

663 metabolites of sweet basil (*Ocimum basilicum* L.). *J. Agric. Food. Chem.* 54(6), 2327-2332.

664 66. Wang H, Ma C, Li Z, Ma L, Hong W, Ye H, et al. (2010) Effects of exogenous methyl jasmonate
665 on artemisinin biosynthesis and secondary metabolites in *Artemisia annua* L. *Ind. Crop. Prod.*
666 31, 214-218.

667 67. Kim YS, Hahn EJ, Murthy HN, and Paek KY (2004) Adventitious root growth and ginsenoside
668 accumulation in *Panax ginseng* cultures as affected by methyl jasmonate. *Biotechnol. lett.* 26,
669 1619-1622.

670 68. Kim YJ, Zhang D, and Yang DC (2015) Biosynthesis and biotechnological production of
671 ginsenosides. *Biotechnol. Adv.* 33, 717-735.

672

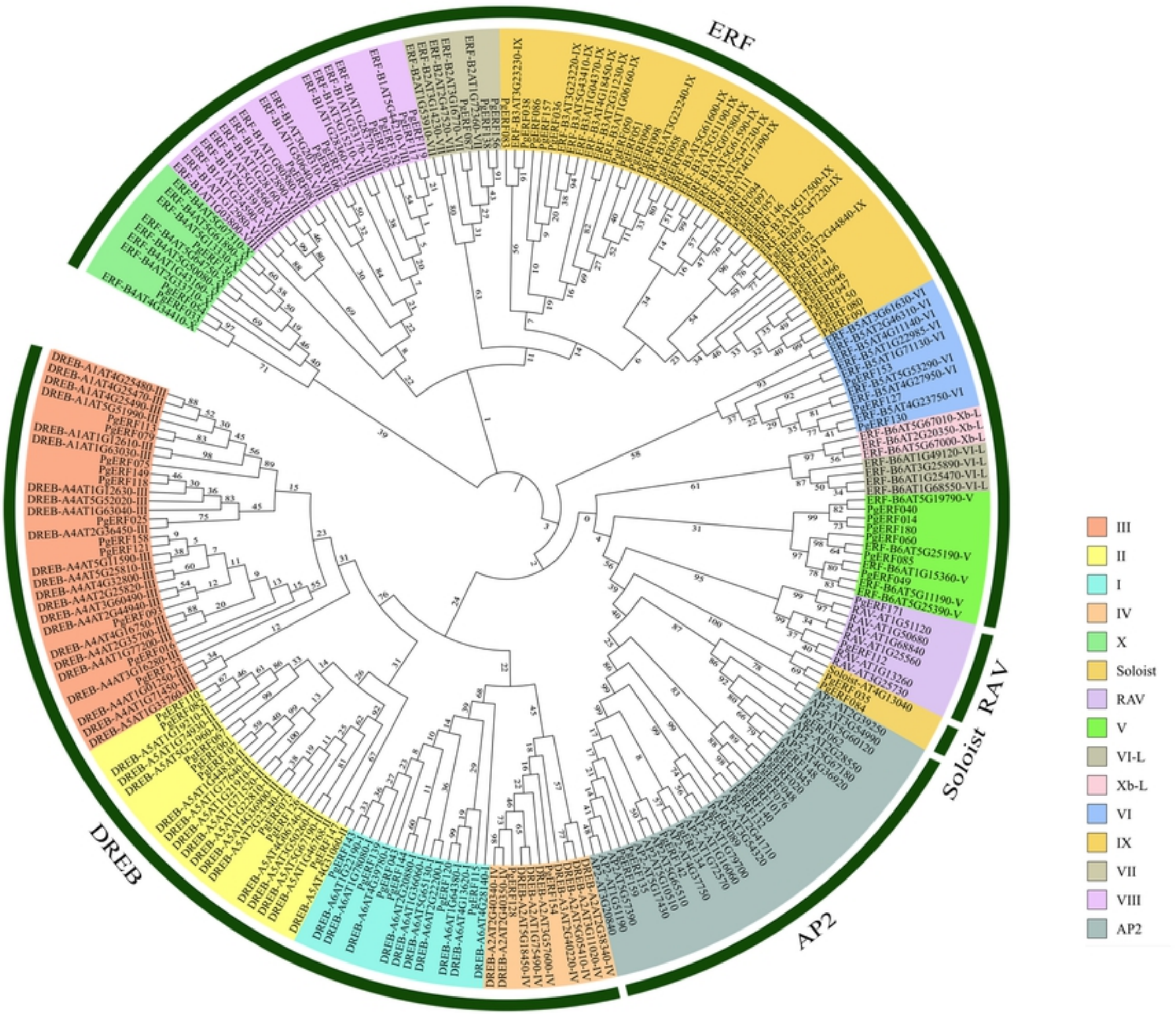


Figure 1

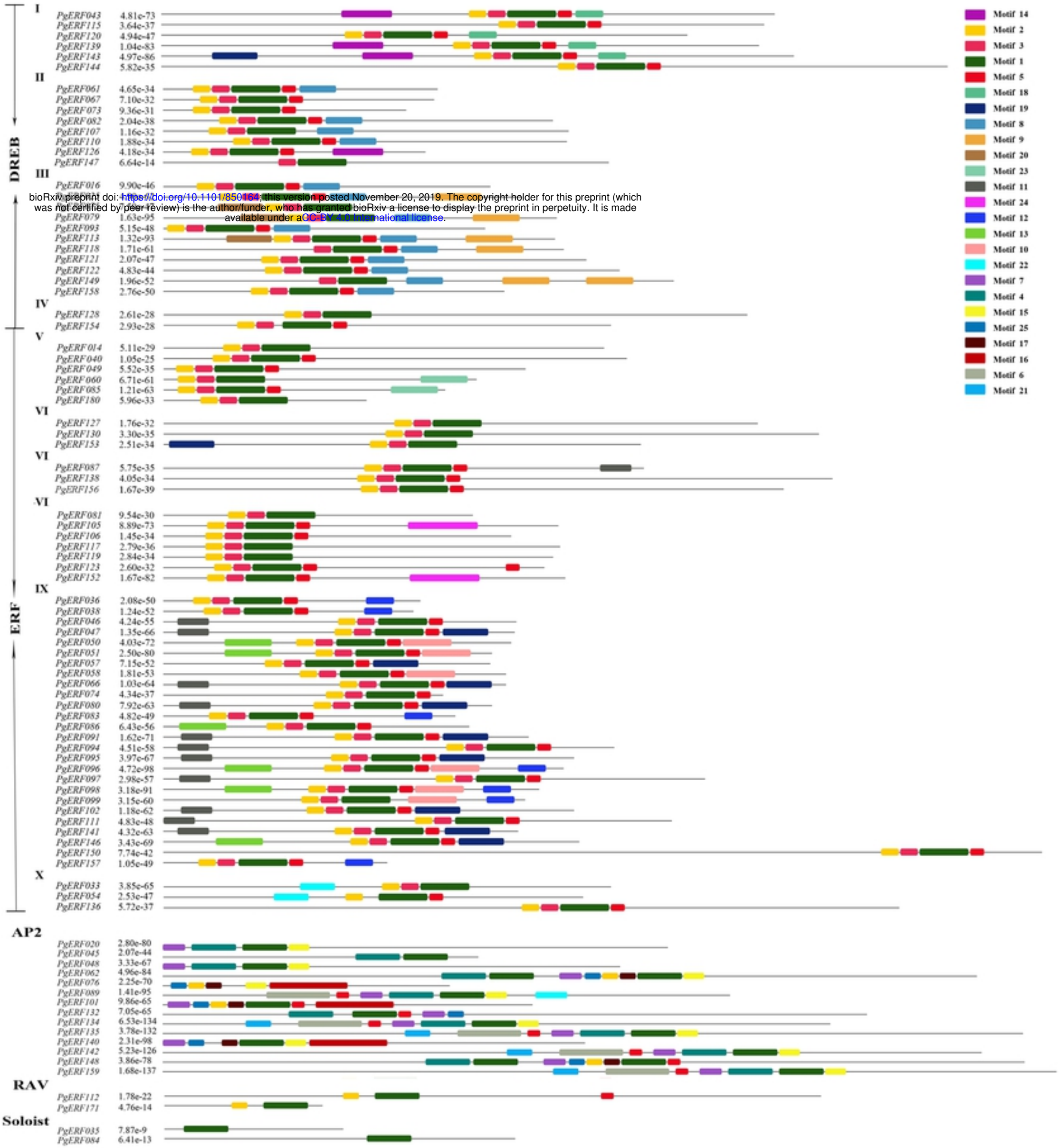


Figure 2

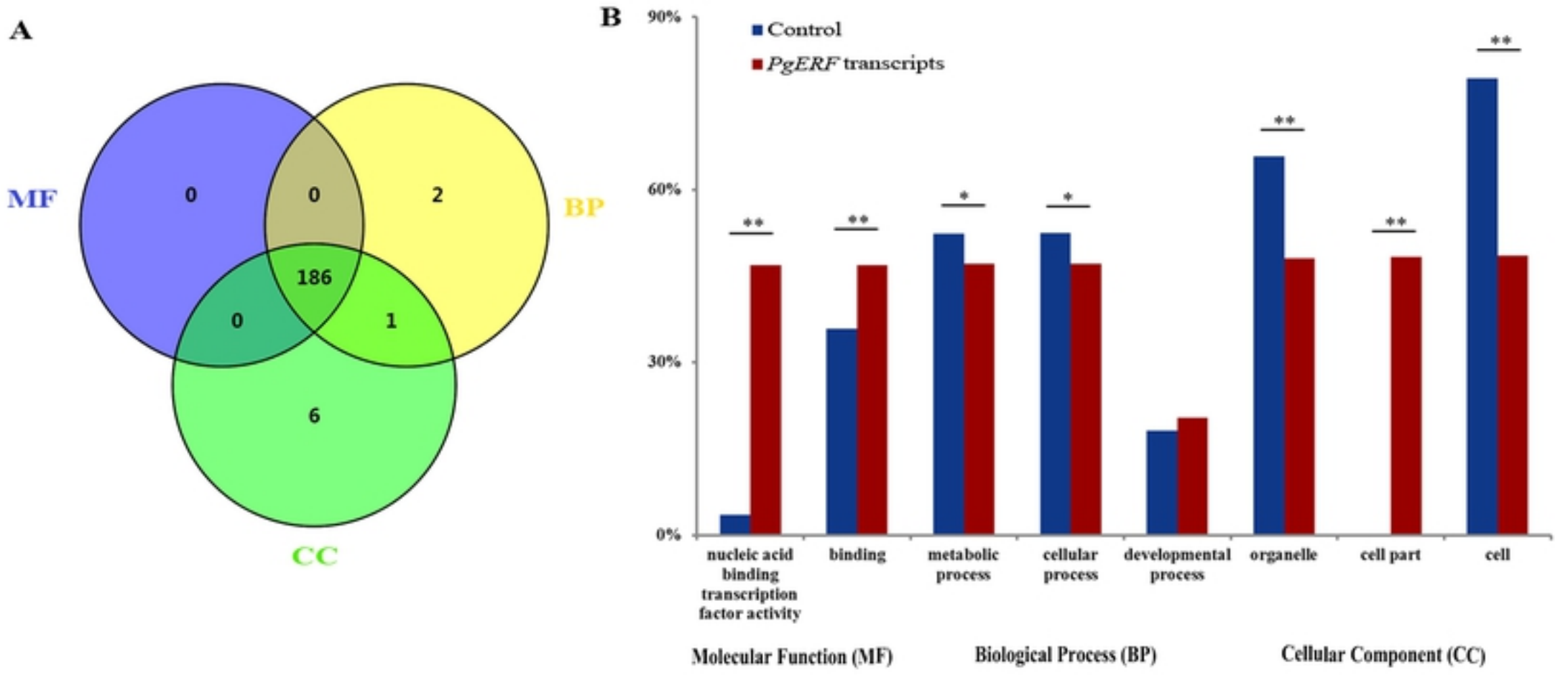

A
MF
BP
CC
0
0
2
186
0
1
6
B
90%
60%
30%
0%
Control
PgERF transcripts
**
**
*
*
**
**
**
nucleic acid binding transcription factor activity
binding
metabolic process
cellular process
developmental process
organelle
cell part
cell
Molecular Function (MF)
Biological Process (BP)
Cellular Component (CC)

Figure 3

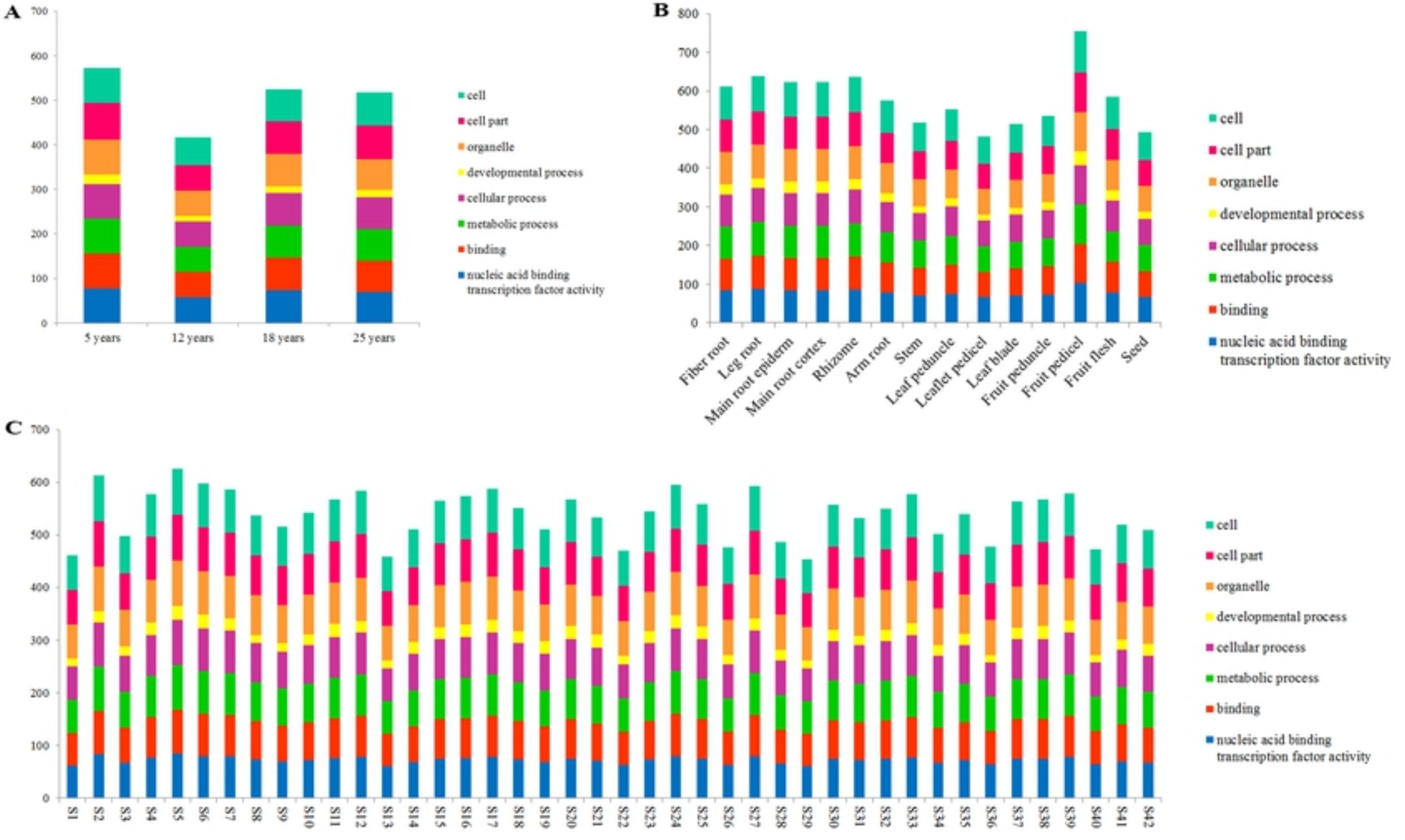

A
700
600
500
400
300
200
100
0
5 years
12 years
18 years
25 years
cell
cell part
organelle
developmental process
cellular process
metabolic process
binding
nucleic acid binding transcription factor activity
B
800
700
600
500
400
300
200
100
0
Fiber root
Leg root
Main root epiderm
Main root cortex
Rhizome
Arm root
Stem
Leaf peduncle
Leaflet pedicel
Leaf blade
Fruit peduncle
Fruit pedicel
Fruit flesh
Seed
cell
cell part
organelle
developmental process
cellular process
metabolic process
binding
nucleic acid binding transcription factor activity
C
700
600
500
400
300
200
100
0
S1 S2 S3 S4 S5 S6 S7 S8 S9 S10 S11 S12 S13 S14 S15 S16 S17 S18 S19 S20 S21 S22 S23 S24 S25 S26 S27 S28 S29 S30 S31 S32 S33 S34 S35 S36 S37 S38 S39 S40 S41 S42
cell
cell part
organelle
developmental process
cellular process
metabolic process
binding
nucleic acid binding transcription factor activity

Figure 4

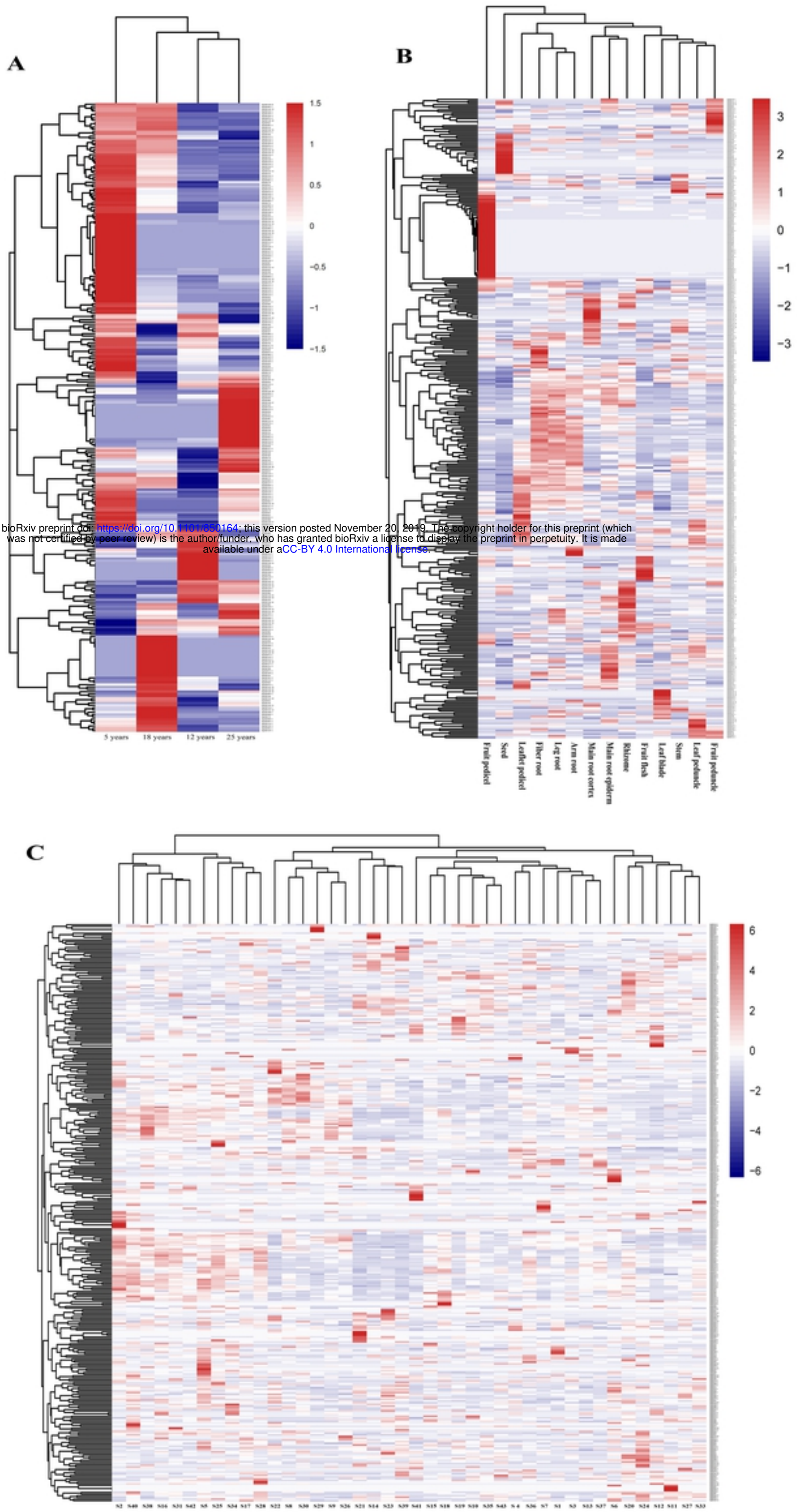
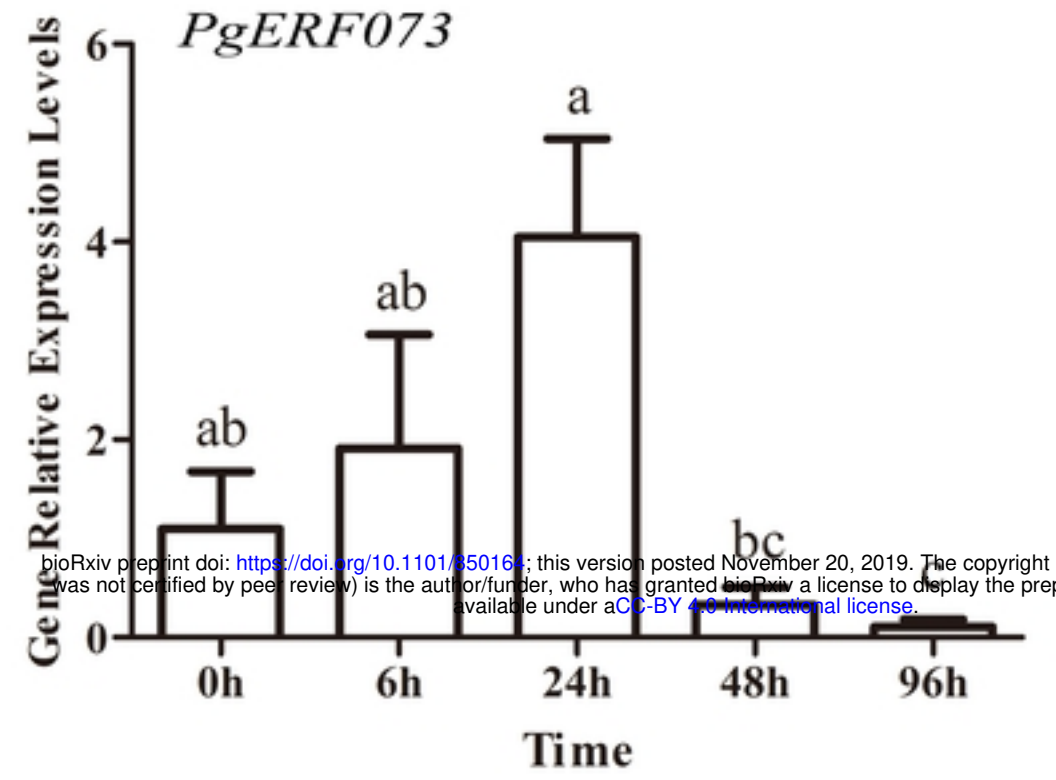
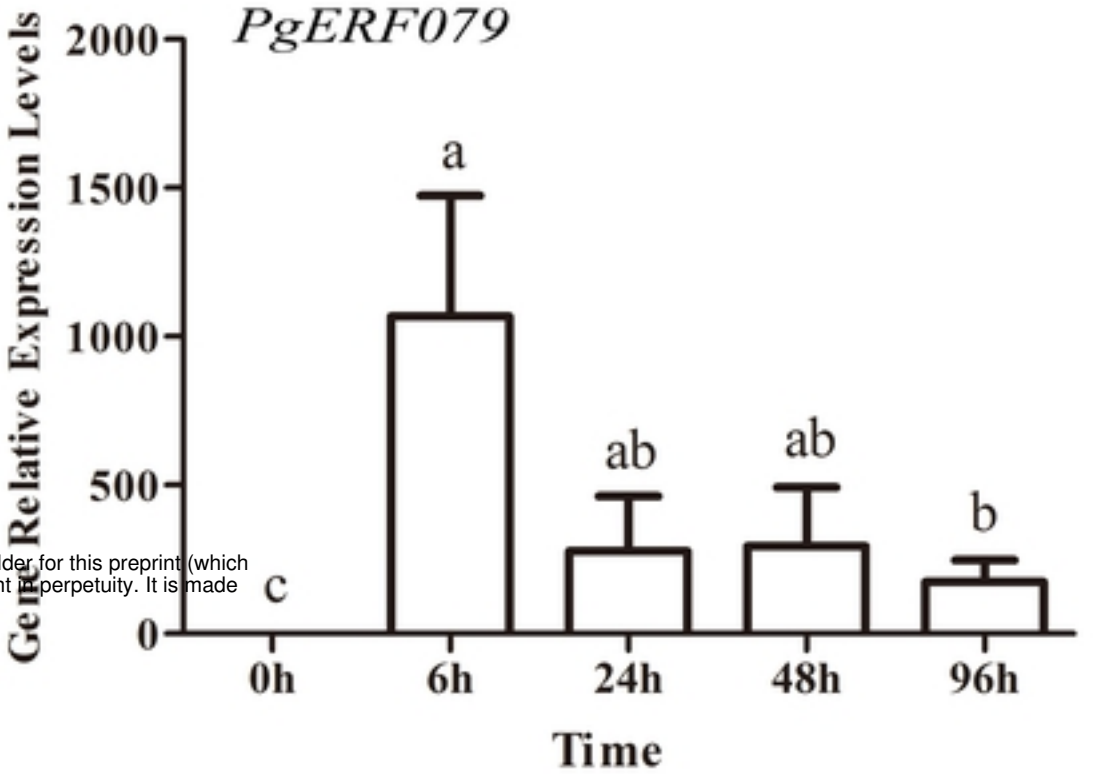
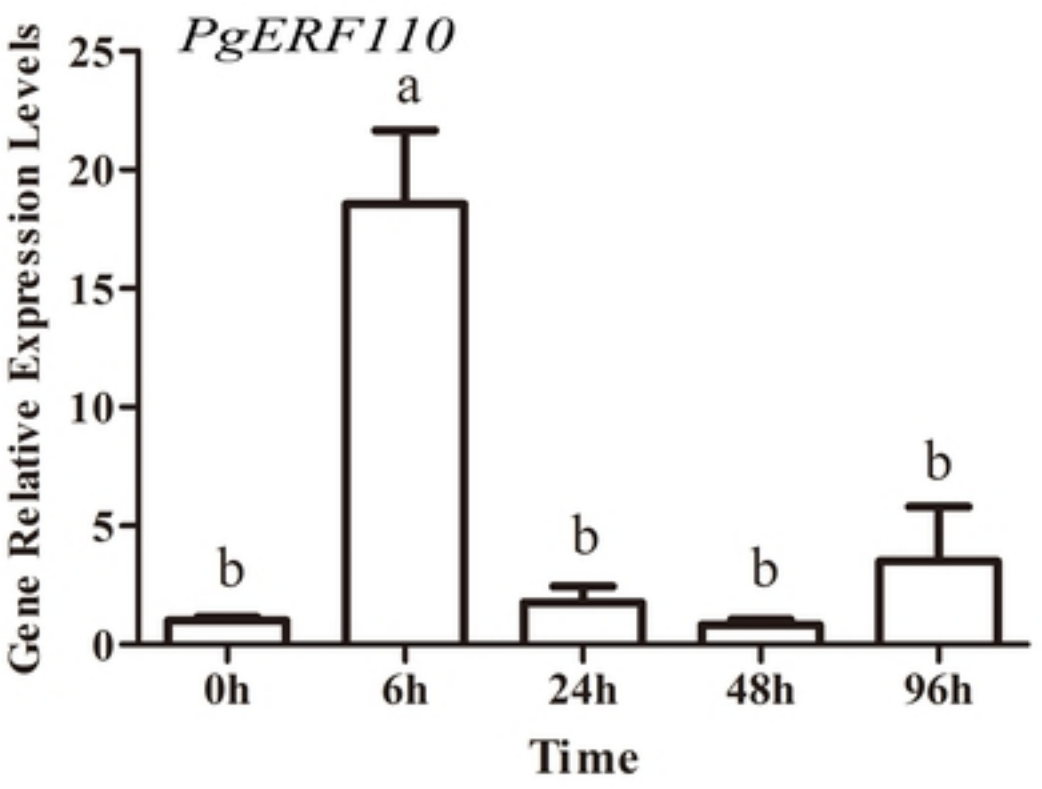
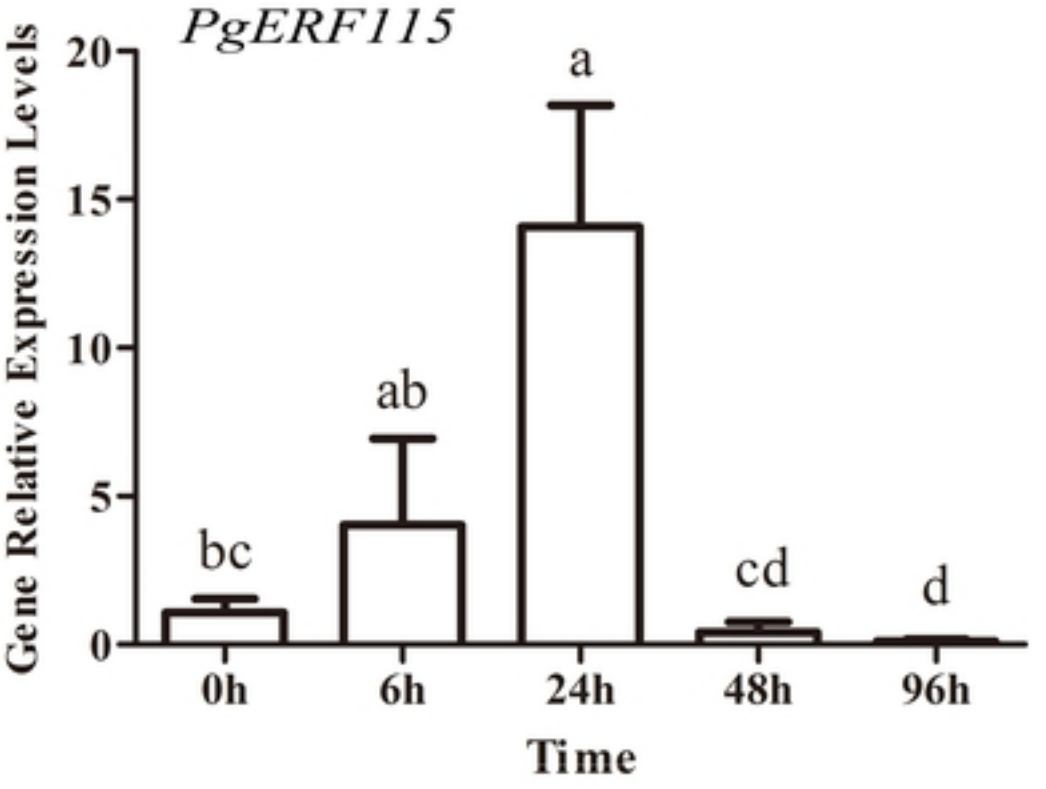


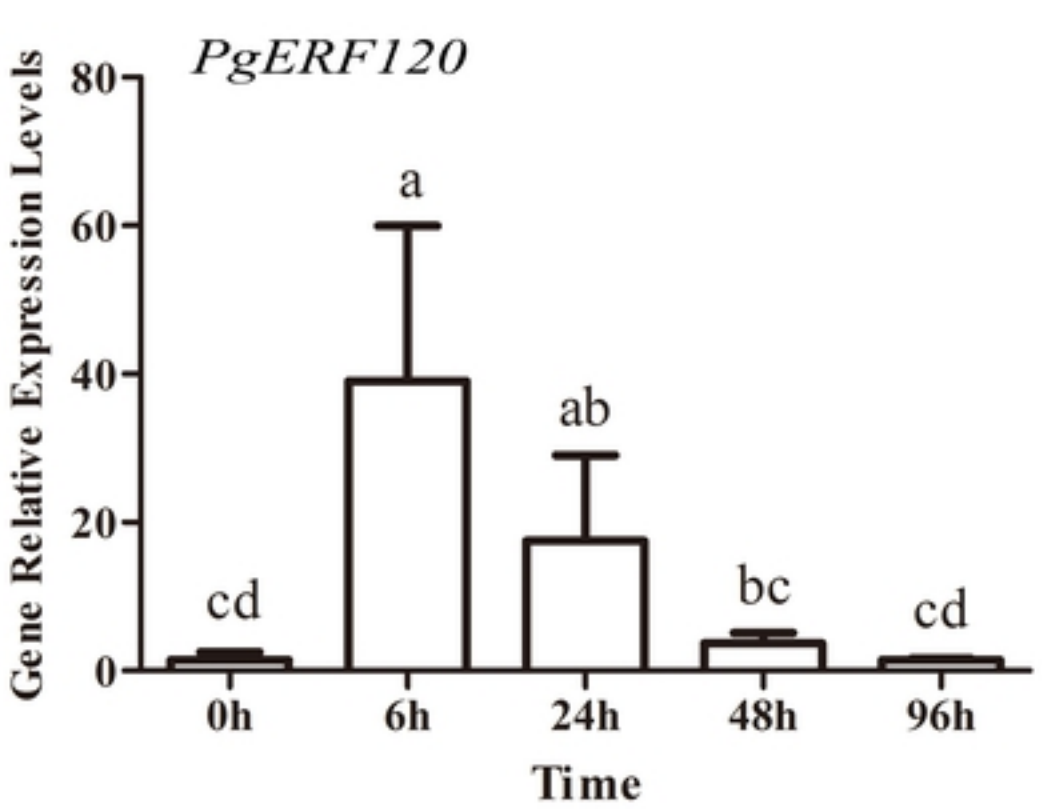
Figure 5











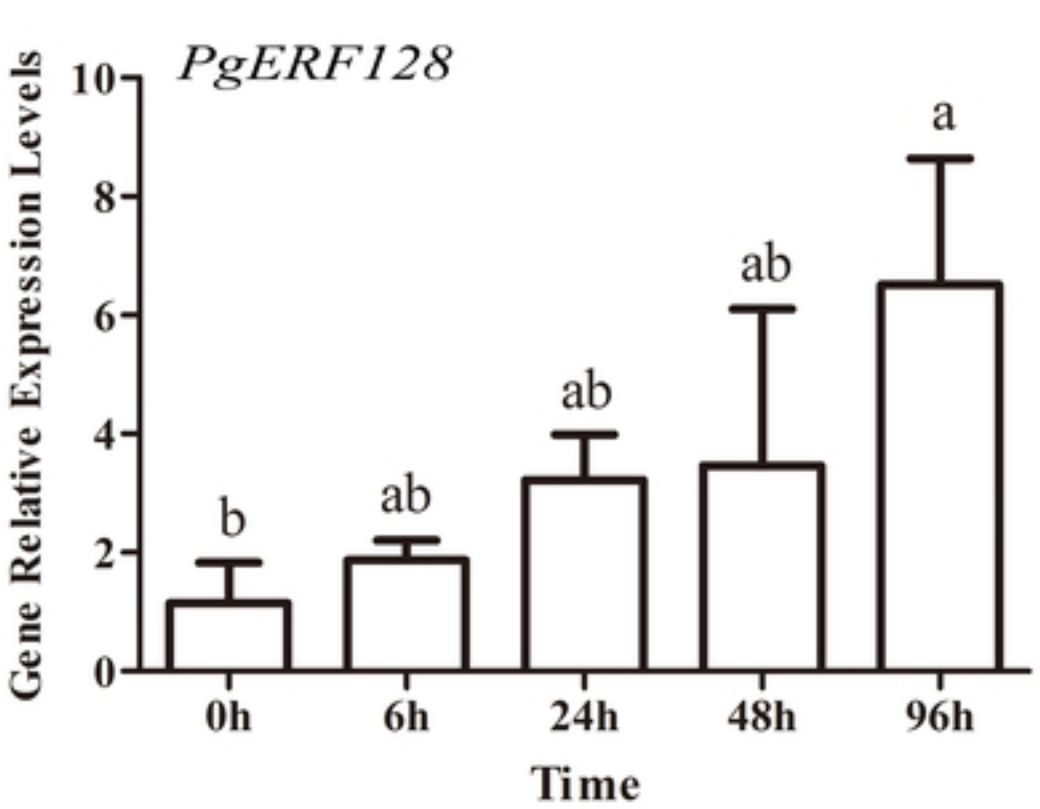
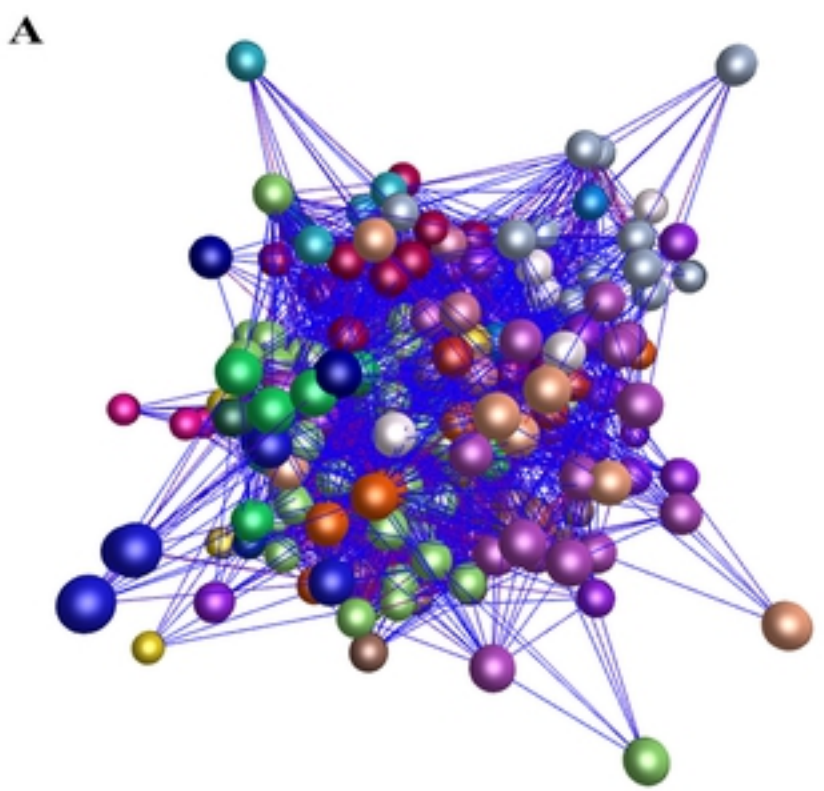
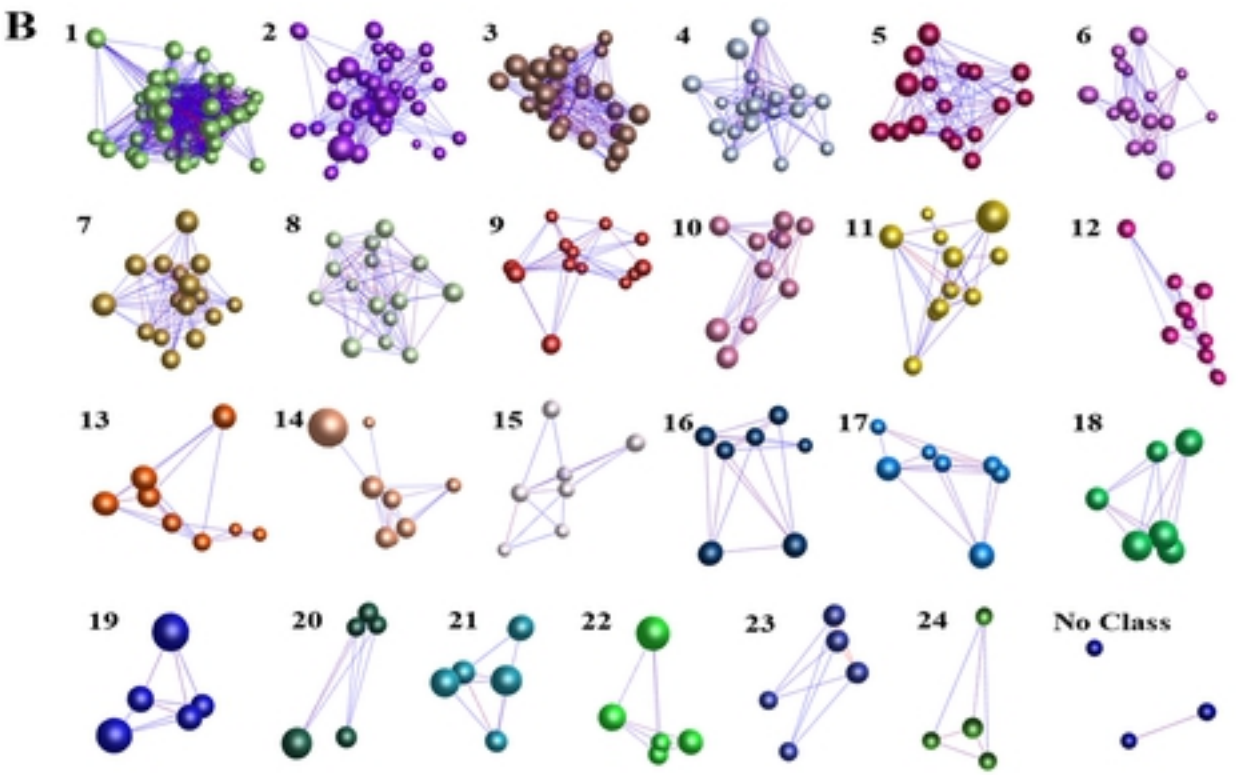
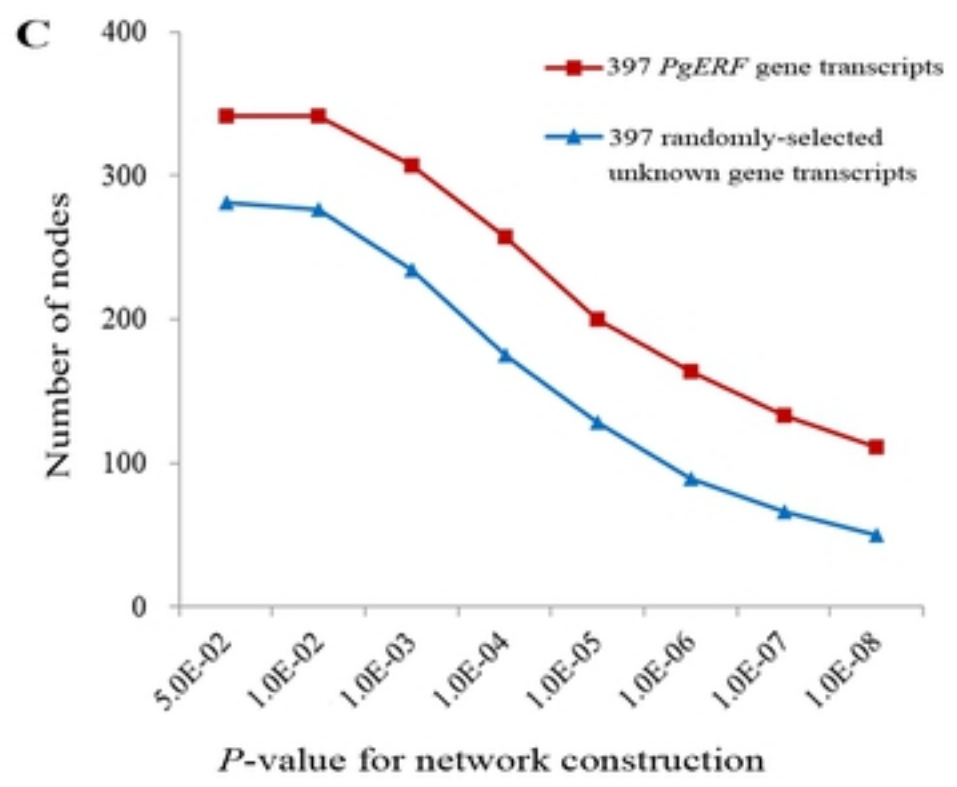


Figure 7


A


B
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
No Class


C
Number of nodes
400
300
200
100
0
397 PgERF gene transcripts
397 randomly-selected unknown gene transcripts
5.0E-02
1.0E-02
1.0E-03
1.0E-04
1.0E-05
1.0E-06
1.0E-07
1.0E-08
P-value for network construction

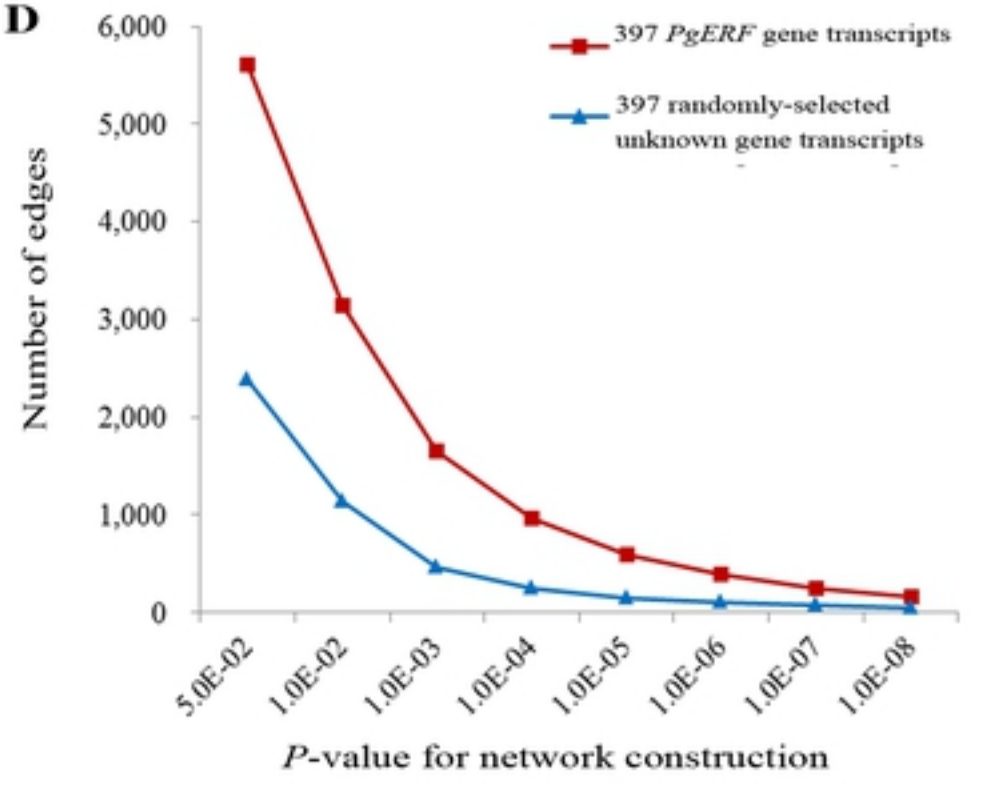

D
Number of edges
6,000
5,000
4,000
3,000
2,000
1,000
0
397 PgERF gene transcripts
397 randomly-selected unknown gene transcripts
5.0E-02
1.0E-02
1.0E-03
1.0E-04
1.0E-05
1.0E-06
1.0E-07
1.0E-08
P-value for network construction

Figure 6

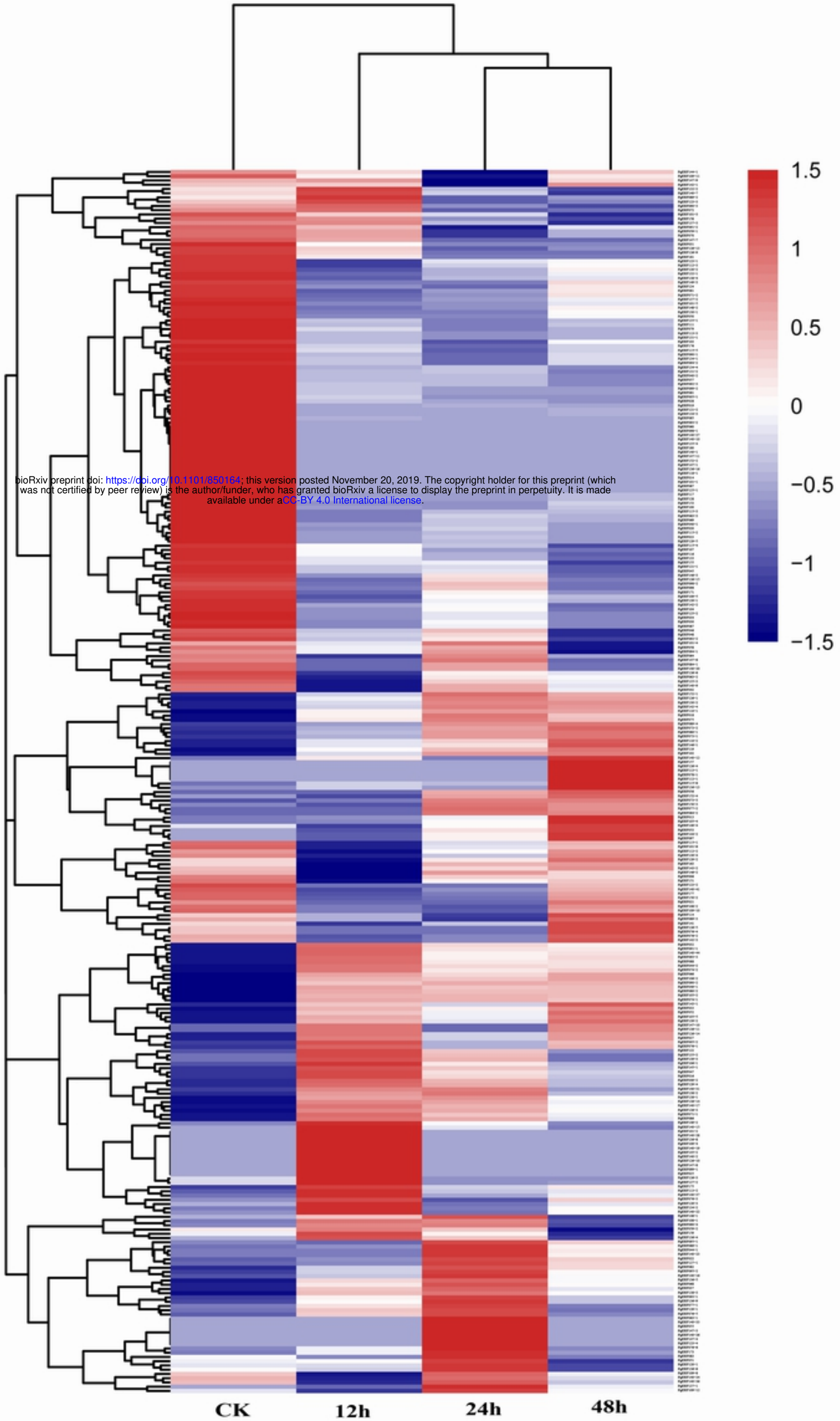


Figure 8