



# Genome-wide identification and expression analysis of the ERF transcription factor family in pineapple (*Ananas comosus* (L.) Merr.)

Youmei Huang<sup>1,\*</sup>, Yanhui Liu<sup>1,\*</sup>, Man Zhang<sup>1</sup>, Mengnan Chai<sup>1</sup>, Qing He<sup>1</sup>, Bello Hassan Jakada<sup>1</sup>, Fangqian Chen<sup>1</sup>, Huihuang Chen<sup>1</sup>, Xingyue Jin<sup>1</sup>, Hanyang Cai<sup>1</sup> and Yuan Qin<sup>1,2</sup>

<sup>1</sup>State Key Lab of Ecological Pest Control for Fujian and Taiwan Crops; Key Lab of Genetics, Breeding and Multiple Utilization of Crops, Ministry of Education; Fujian Provincial Key Lab of Haixia Applied Plant Systems Biology, College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou, Fujian Province, China

<sup>2</sup>State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, Guangxi Key Lab of Sugarcane Biology, College of Agriculture, Guangxi University, Nanning, Guangxi Province, China

\*These authors contributed equally to this work.

## ABSTRACT

Pineapple (*Ananas comosus* (L.) Merr.) is an important tropical fruit with high economic value. The quality and yield of pineapple will be affected by various environmental conditions. Under adverse conditions, plants can produce a complex reaction mechanism to enhance their resistance. It has been reported that the member of ethylene responsive transcription factors (ERFs) plays a crucial role in plant developmental process and stress response. However, the function of these proteins in pineapple remains limited. In this study, a total of 74 *ERF* genes (*AcoERFs*) were identified in pineapple genome, named from *AcoERF1* to *AcoERF74*, and divided into 13 groups based on phylogenetic analysis. We also analyzed gene structure, conserved motif and chromosomal location of *AcoERFs*, and the *AcoERFs* within the same group possess similar gene structures and motif compositions. Three genes (*AcoERF71*, *AcoERF73* and *AcoERF74*) were present on unanchored scaffolds, so they could not be conclusively mapped on chromosome. Synteny and *cis*-elements analysis of *ERF* genes provided deep insight into the evolution and function of pineapple *ERF* genes. Furthermore, we analyzed the expression profiling of *AcoERF* in different tissues and developmental stages, and 22 *AcoERF* genes were expressed in all examined tissues, in which five genes (*AcoERF13*, *AcoERF16*, *AcoERF31*, *AcoERF42*, and *AcoERF65*) had high expression levels. Additionally, nine *AcoERF* genes were selected for functional verification by qRT-PCR. These results provide useful information for further investigating the evolution and functions of ERF family in pineapple.

Submitted 21 April 2020  
Accepted 1 September 2020  
Published 22 September 2020

Corresponding authors  
Hanyang Cai, cai-hanyang123@163.com  
Yuan Qin, yuanqin@fafu.edu.cn

Academic editor  
Thiago Venancio

Additional Information and  
Declarations can be found on  
page 18

DOI 10.7717/peerj.10014

© Copyright  
2020 Huang et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Agricultural Science, Genomics, Plant Science

**Keywords** Pineapple, ERF transcription factors, Phylogenetic analysis, Synteny analysis, Expression profiles

## INTRODUCTION

Plant growth and development, yield and quality are frequently affected by unfavorable environmental factors such as drought, salinity, high temperature and cold. In order to survive these stress conditions, plants have developed complex reaction mechanisms at the molecular, cellular, and system levels (Mittler, 2006; Loudet & Hasegawa, 2017). Gene co-expression at the transcriptional level is one of the most important ways to regulate biological processes. Transcription factors (TFs) play important roles in regulating the expression of functional proteins when plants are exposed to unfavorable environmental conditions (Cui et al., 2016). Among these TFs, ethylene responsive transcription factors (ERFs) play important roles in various biological processes, such as defense responses and hormonal signal transduction (Agarwal et al., 2010; Sharma et al., 2010; Rashid et al., 2012). Therefore, it is important to study the evolution and function of these genes in order to improve yield and the ability to resist the adverse environmental conditions of plant.

The AP2/ERF super-family is one of the largest family in plant, and the member of this super-family characterized by one or two conserved AP2/ERF domains of approximately 60 to 70 amino acids (Sakuma et al., 2002). It has been reported that the AP2/ERF domain contains two conserved elements, YRG and RAYD element, and the core region of RAYD element has been predicted to form an amphipathic  $\alpha$ -helix (Okamoto et al., 1997). The three-dimensional (3D) structure of the AP2/ERF domain from AtERF1 shows that the domain is composed of a 3 anti-parallel  $\beta$ -sheets and an  $\alpha$ -helix (Allen et al., 1998). Furthermore, tryptophan (Trp) and arginine (Arg) residues of the  $\beta$ -sheet are in contact with DNA during transcription (Sharma et al., 2010). Based on the number of AP2/ERF domains and other DNA binding domains, the AP2/ERF super-family can be divided into three families: AP2, ERF and RAV (Sakuma et al., 2002; Nakano et al., 2006). Among the three families, only the AP2 family contains two conserved AP2/ERF domains, whereas ERF and RAV family contain just a single AP2/ERF domain (Riechmann & Meyerowitz, 1998). The members of the RAV family include an AP2/ERF domain and a B3 domain (Roman et al., 2009). Based on the ERF domain binding to DNA sequences, ERF family has been further classified into two major subfamilies: ERF (ethylene responsive transcription factors) subfamily and CBF/DREB (C-repeat binding factor/dehydration-responsive element binding factor) subfamily (Nakano et al., 2006). Proteins encoded by ERF subfamily genes bind to GCC box (AGCCGCC) (Ohmetakagi & Shinshi, 1995; Hao, Ohmetakagi & Sarai, 1998), whereas the CBF/DREB subfamily genes typically binds to the DRE and/or C-repeat (A/GCCGAC) (Stockinger, Gilmour & Thomashow, 1997; Riechmann & Meyerowitz, 1998). The residues at position 14 and 19 of ERF subfamily genes AP2/ERF domain is alanine (Ala) and aspartic acid (Asp), respectively, whereas DREB subfamily genes containing valine (Val) at position 14 and glutamine acid (Glu) at positions 19 (Sakuma et al., 2002).

ERFs are located downstream of the ethylene signaling pathway and regulate the transduction of ethylene as well as the trans-activation of certain transcription factors related to hormonal signal (Fujimoto et al., 2000; Sharma et al., 2010; Yu et al., 2012). In addition, ERF proteins are involved in various plant biological processes, such

as environmental stresses response (Liu, White & Macrae, 1999; Rashotte et al., 2006), beneficial symbiotic interaction (Vernie et al., 2008), as well as other developmental processes, such as leaf, flower and embryo development in some plants (Elliott et al., 1996; Moose & Sisco, 1996; Boutilier et al., 2002; Vernie et al., 2008). To date, ERF family has been reported in various plant species, such as *Arabidopsis thaliana* (Sakuma et al., 2002; Nakano et al., 2006), soybean (*Glycine max*) (Li et al., 2005; Zhang et al., 2009), rice (*Oryza sativa*) (Cao et al., 2006; Nakano et al., 2006; Sharoni et al., 2011; Rashid et al., 2012), cotton (*Gossypium barbadense* L.) (Jin & Liu, 2008; Meng et al., 2010), *Populus trichocarpa* (Zhuang et al., 2008), tomato (*Solanum tuberosum* L.) (Sharma et al., 2010; Pan et al., 2012), grape (*Vitis vinifera* L.) (Licausi et al., 2010), cucumber (*Cucumis sativus* L.) (Hu & Liu, 2011) and tartary buckwheat (*Fagopyrum Tataricum*) (Liu et al., 2019).

Pineapple (*Ananas comosus* (L.) Merr.) is one of the most important tropical fruit in the world (Su et al., 2017; Xie et al., 2018; He et al., 2019). The growth and development of pineapple is affected by various environmental conditions, such as drought, salt and cold stress (He et al., 2019). Additionally, pineapple is also an important monocotyledonous, and it can be considered as a proper model to study the monocot evolution (Ming et al., 2015; He et al., 2019). Recently, a comprehensive study of the pineapple genome provide a solid foundation for the study of pineapple gene functions (Ming et al., 2015; Fang et al., 2016; Su et al., 2017). Due to the great economic and research value of pineapple, it would be meaningful to make a further study ERF family in pineapple.

Previous study only proposed the existence of DREB subfamily in the ERF family of pineapple (Chai et al., 2020). However, detail information about the whole ERF family members in pineapple remains unexplored. In order to investigate the function and evolution of other ERF family members in pineapple, we conducted a detailed analysis of the entire ERF family. In this study, 74 pineapple *ERF* genes were identified and classified into 13 groups. Furthermore, we also conducted a systematic analysis including gene structure, motif compositions, chromosome distribution, phylogenetic and synteny analysis of each *AcoERF* gene. In order to further investigate the functions of *AcoERF* genes, we analyzed the expression profile of *AcoERF* genes in different tissues and stages. The data generated in this study will help to select the appropriate candidate genes for further functional studies during pineapple growth and development. This is of great significance to the investigation of pineapple stress response and variety improvement.

## MATERIALS & METHODS

### Genome-wide analysis of the ERF family in pineapple

The AP2/ERF amino acid sequences of pineapple and *Arabidopsis* were downloaded from Phytozome12 (<https://phytozome.jgi.doe.gov/pz/portal.html>) (Goodstein et al., 2012). The Hidden Markov Model (HMM) profiles of the AP2 domain (PF00847) and the B3 domain (PF02362) were obtained from PFAM database (<http://pfam.sanger.ac.uk/>) (Finn et al., 2008) as the queries for search predicted AP2/ERF proteins in the pineapple dataset using HMMER software 3.0 (<http://hmmer.wustl.edu/>) with a threshold of  $e\text{-value} < e^{-5}$  (Eddy, 2011). BLAST searches were used to determine the predicted AP2/ERFs

in pineapple database with all the *Arabidopsis* AP2/ERFs as queries. All of the candidate genes were further examined by Simple Modular Architecture Research Tool (SMART, <http://smart.embl-heidelberg.de/>) (Letunic & Bork, 2018) to confirm the protein sequences derived from the selected pineapple AP2/ERF domains. ExPasy proteomics server (<http://expasy.org/>) (Gasteiger et al., 2003) was used to predicate the basic information of *AcoERF* proteins, including molecular weight (MW) and isoelectric points (*pI*).

### Gene ontology (GO) annotation

To investigate the putative function of the candidate *AcoERF* genes, the gene ontology (GO) annotations of the pineapple ERF family members were downloaded from Phytozome12 (<https://phytozome.jgi.doe.gov/pz/portal.html>) (Goodstein et al., 2012).

### Sequence alignment and phylogenetic analysis

Multiple sequence alignment of the *AcoERF* proteins was performed using DNAMAN (version 9) with the default parameters (Wang, 2015), and the diagram was visualized using ESPript 3.0 (<http://esprict.ibcp.fr/ESPript/cgi-bin/ESPript.cgi>) (Gouet, Robert & Courcelle, 2003). To explore the evolution relationships of *ERF* gene family members in pineapple, the ERF amino acid sequences from pineapple and *Arabidopsis* were used. The multiple sequence alignments of all ERF proteins were performed by using MUSCLE (<http://www.ebi.ac.uk/Tools/msa/muscle/>) (Edgar, 2004), with the default parameters. Subsequently, the phylogenetic trees were constructed using MEGA6.0 software (<http://www.megasoftware.net>) (Tamura et al., 2013) via the Neighbor-joining (NJ) method with the following parameters: node robustness was detected using the bootstrap method, and the bootstrap was set to 1000 replications. Finally, the phylogenetic trees were generated by iTOL (<https://itol.embl.de>) (Letunic & Bork, 2016).

### Gene structure and conserved motif predictions

The exon/intron structures of *AcoERF* genes were obtained from the online program Gene Structure Display Server (GSDS, <http://gsds.gao-lab.org/>) (Hu et al., 2015). The conserved motifs in the *AcoERF* proteins sequences were investigated by Multiple EM for Motif Elicitation (MEME, <http://meme-suite.org/tools/meme>) (Bailey et al., 2015). The optimized parameters were as follows: maximum number of motifs was set to 20, the optimum width of motifs was set to 10-50 residues, and other options were set to default.

### Chromosomal localization and synteny analysis

The information of chromosome localization of all pineapple *ERF* genes was obtained from Phytozome12 (<https://phytozome.jgi.doe.gov/pz/portal.html>). And then, each genes was mapped on chromosomes using MapChart software based on positional information in the pineapple genome project (Voorrips, 2002). For synteny analysis, the sequence of rice and *Arabidopsis* ERF proteins were downloaded from Phytozome12. The potential anchors within pineapple, between pineapple and *Arabidopsis* or rice were searched by BLASTP with the score value of  $E < 1e^{-5}$  and top 5 matches (Zhang et al., 2018). The syntenic block of ERFs within pineapple was visualized using the Circos Program (version 0.69) (<http://circos.ca/>) (Krzywinski et al., 2009).

The syntenic blocks between pineapple and 2 representative plants were constructed using J. Craig Venter Institute (JCVI, <https://github.com/tanghaibao/jcvi>) (*Tanenbaum et al., 2010*). The Synonymous (*Ks*) and non-synonymous (*Ka*) values of *AcoERF* sequences were calculated using bio-pipeline-master (<https://github.com/tanghaibao/bio-pipeline>) (*Tang et al., 2008*).

### Expression analysis of *AcoERF* genes in different tissues

To investigate the expression patterns of the *AcoERF* genes in different development and growth processes of pineapple, the RNA-Seq data of pineapple roots, leaves, flowers, fruits (stage 1-6) were downloaded from the online website Pineapple Genomics Database (PGD, <http://pineapple.angiosperms.org/pineapple/html/download.html>) (*Ming et al., 2015*). The criterion of different stage of ovule, sepal, stamen and petal was referenced to the previously described methods (*Su et al., 2017*). Total RNA of these tissues was extracted using RNA extracted kit (Omega Bio-Tek, Shanghai, China) based on the supplier's instruction, and the samples were prepared based on the published method (*Chen et al., 2017*). The cDNA libraries for sequencing were constructed using the NEBNext Ultra<sup>TM</sup> RNA Library Prep Kit for Illumina (NEB). The reads were aligned to pineapple genome by the TopHat 2.0.0 software with default parameters (*Trapnell et al., 2012*). The Fragments Per Kilobase of transcript per Million fragments mapped (FPKM) value of each gene was calculated using Cuffdiff and Cufflinks software with default parameters. Finally, a heatmap of *AcoERF* expression profile was produced by the heatmaply R package (*Galili et al., 2018*).

### Analysis of *cis*-elements from *AcoERF* promoters

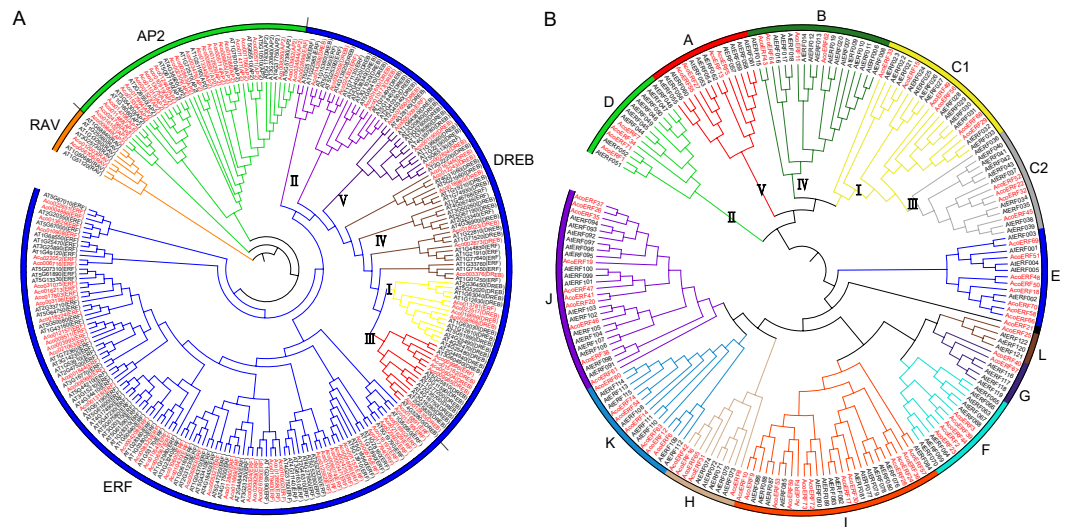
The promoter sequences of the *AcoERF* genes were downloaded from the Pineapple Genomics Database (<http://pineapple.angiosperms.org/pineapple/html/index.html>). And then, the stress-related *cis*-elements of *AcoERF* genes were predicted using the Plant *Cis*-Acting Regulatory Element (PlantCARE, <http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (*Lescot et al., 2002*).

### Plant material and abiotic treatments

The pineapple materials were provided by the Qin Lab (Center for Genomics and Biotechnology, Fujian Agriculture and Forestry University, Fujian, China). The seedlings were grown in a greenhouse at 25 °C, 16-h light/8-h dark photoperiod and a relative humidity of 70% for one month. And then the one-month-old pineapple plants were exposed to the following treatments: cold stress (4 °C), drought stress (350 mM Mannitol) and salt stress (150 mM NaCl). The samples of pineapple roots and leaves were collected after 6, 12, 24 and 48 h treatments. The one-month-old plants that without any stress treatment were used as controls in the experiment. The leaves and roots were harvested at the indicated time, and stored at -80 °C for subsequent analysis.

### RNA isolation and qRT-PCR analysis

Total RNA of the pineapple sample was extracted from roots and leaves using plant RNA extraction kit (OMEGA, Shanghai, China) according to the manufacturer's protocol. And then, total RNA was reverse-transcribed using AMV reverse transcriptase (Takara, Japan)



**Figure 1** The evolutionary relationship of the *ERFs*. (A) Unrooted phylogenetic tree representing relationship among *AP2/ERF* genes between pineapple and *Arabidopsis*; (B) Phylogenetic analysis of *ERF* proteins from pineapple and *Arabidopsis*. The genes in pineapple are marked in red, while those in *Arabidopsis* are marked in black. The pineapple *DREB* proteins are classified into five groups: I, II, III, IV and V.

Full-size [DOI: 10.7717/peerj.10014/fig-1](https://doi.org/10.7717/peerj.10014/fig-1)

based on the supplier's instruction. qRT-PCR analysis was conducted using SYBR Premix Ex Taq II system (Takara, Japan) and Bio-Rad Real-time PCR system (Foster city, CA, USA). The reactions were performed in a 20  $\mu$ l volume containing 10  $\mu$ l of 2 $\times$  SYBR Premix, 8.2  $\mu$ l of RNase free water, 1  $\mu$ l of template, 0.4  $\mu$ l of each specific primer. The primers used for RT-PCR were showed in Table S1. The RT-PCR reactions program was completed with the following conditions: 95  $^{\circ}$ C for 30 s; 40 cycles of 95  $^{\circ}$ C for 5 s and 60  $^{\circ}$ C for 34 s; 95  $^{\circ}$ C for 15 s (Cai et al., 2017; Cai et al., 2019). The analyses were confirmed in triplicate. The relative expression level of each *AcoERF* gene was calculated based on the comparison threshold period ( $2^{-\Delta\Delta C_t}$ ) method (Pan et al., 2012). The *protein phosphatase 2A* gene from pineapple was used as the reference gene (Sang et al., 2013). The final data were subjected to an analysis of variance test.

## RESULTS

### Genome-wide identification and characterization of pineapple *ERF* family

A total of 103 pineapple genes were identified as possible encoding proteins containing *AP2/ERF* domain (Fig. 1A). Among these, 27 genes containing two conserved *AP2/ERF* domains and 2 genes possessed a single *AP2/ERF* domain together with a *B3* domain were assigned to *AP2* and *RAV* family, respectively; Remaining 74 genes with a single *AP2/ERF* domain were grouped into *ERF* family, including 20 *DREB* and 54 *ERF* subfamily members. Earlier, the *DREB* subfamily members of pineapple had been divided into 5 groups: I, II, III, IV and V (Chai et al., 2020). To maintain uniformity, all the pineapple *ERF* genes were provisionally named as *AcoERF1*-*AcoERF74* according to the gene ID in ascending order

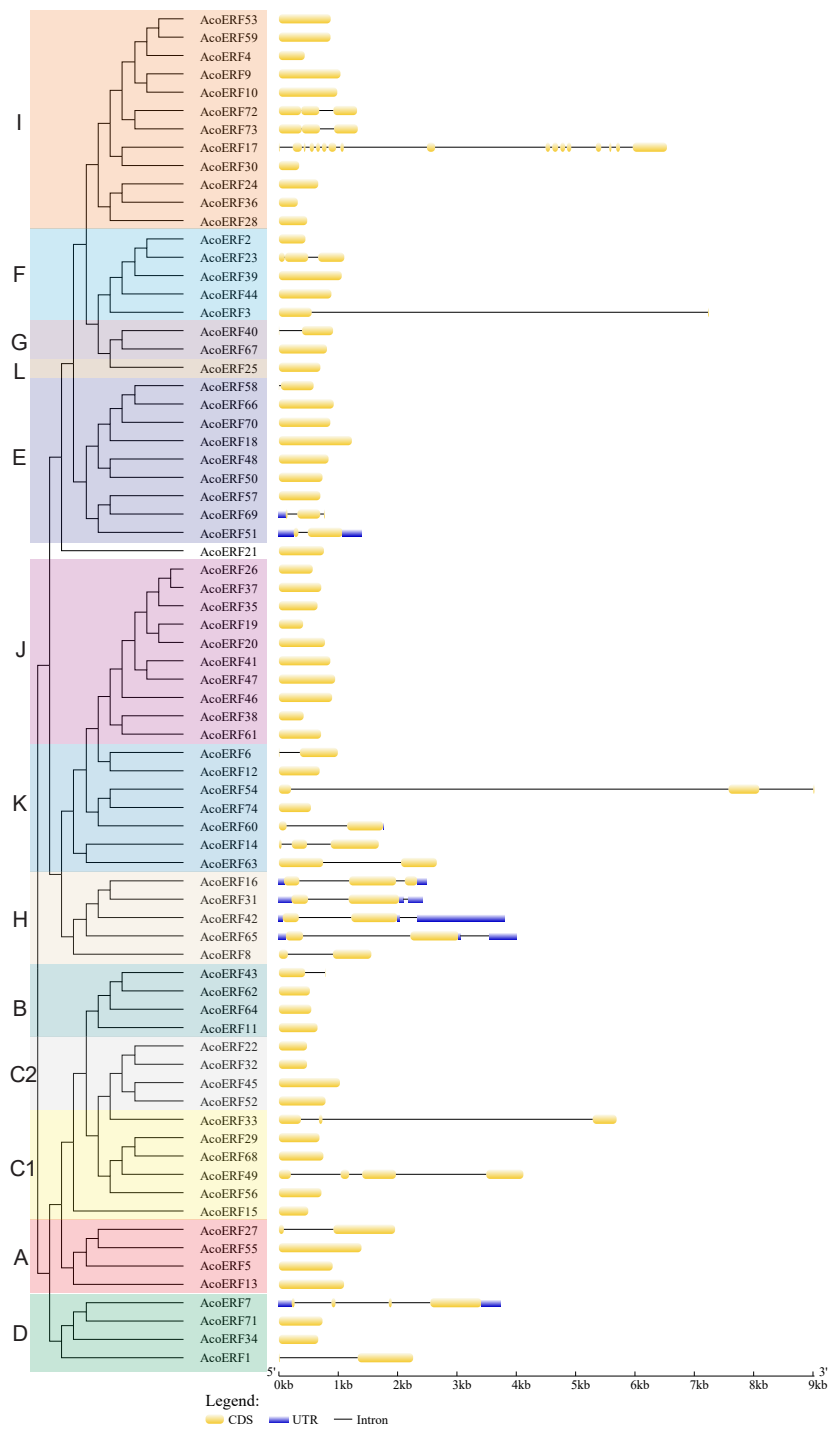
(Table S2). The full-length of the 74 *AcoERF* proteins range from 105 (*AcoERF36*) to 605 (*AcoERF17*) amino acid residues, while the coding sequence (CDS) length ranges from 318 to 1819 bp, and relative molecular weights (MW) range from 11.21 to 66.74 kDa. The predicted isoelectric points (*pI*) varied from 4.71 (*AcoERF18* and *AcoERF52*) to 9.63 (*AcoERF30*) (Table S2). Moreover, the Go annotation of *AcoERFs* included biological process (GO: 0006355) and molecular function (GO: 0003700) (Table S2). Within biological process, 72 *AcoERF* genes were associated with regulation of transcription and DNA-templated. For the molecular function categories, 73 *AcoERF* genes were involved in transcription factor activity and sequence-specific DNA binding.

### Multiple sequence alignment and phylogenetic analysis of pineapple *ERF* genes

Multiple sequence alignment of the 74 *AcoERFs* indicated that most *ERF* family members possessed conserved YRG and RAYD elements within the AP2/*ERF* domain region (Fig. S1). *ERF* subfamily proteins containing Ala at position 14 and Asp at position 19. However, the residues at position 14 and 19 of *DREB* subfamily proteins is Val and Glu, respectively. Evolution of *ERF* TFs was further explored based on the phylogenetic tree constructed with pineapple and *Arabidopsis ERF* genes (Fig. 1B). Earlier, seven *Arabidopsis ERF* proteins having low homology with other member of *ERF* family, were assigned to group VI-L and Xb-L according to the analysis of gene structure and conserved motif (Nakano et al., 2006). In this study, these seven *Arabidopsis ERF* proteins were also used for phylogenetic reconstruction and formed group G and L. The *Arabidopsis ERF* proteins of group I and III were from the same group in previous study (Nakano et al., 2006). Therefore, group I and III in this study were renamed as C1 and C2, respectively. The result showed that the *AcoERFs* were divided into 13 groups, except for *AcoERF21*. Group I contained the most *AcoERFs*, accounting for 16.22% of total *ERF* genes. Followed by group J, which has 10 *AcoERFs*. However, group L has only one *AcoERF*.

### Gene structure and motif composition of pineapple *ERF* family

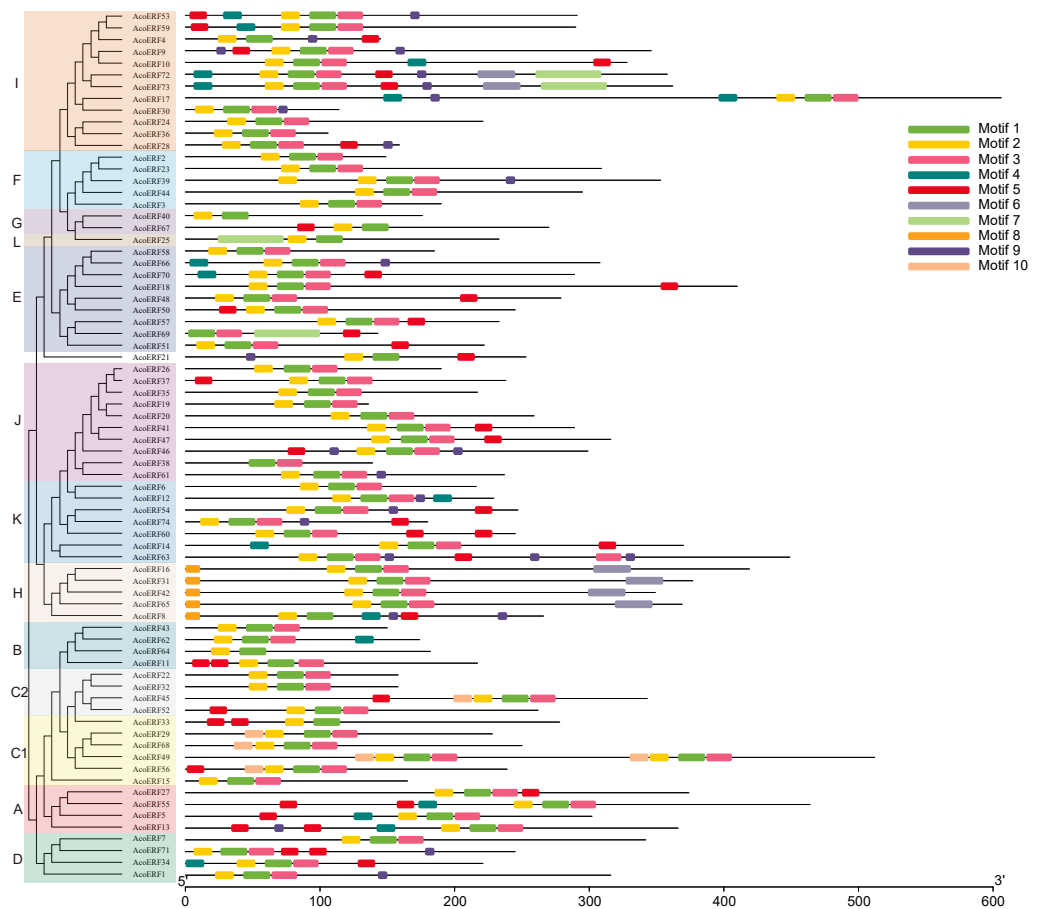
To gain more insight into the evolution and structural diversity of the *ERF* family in pineapple, we analyzed the gene structure and conserved motif of the *AcoERF* genes. As shown in Fig. 2, 49 *AcoERF* genes have no intron, accounting for 66.22% of the total number of *AcoERF* genes, and all genes of group J have no intron. Moreover, the number of introns in these genes varied from 1 to 16, and most *AcoERF* genes have 1 to 3 introns, whereas *AcoERF17* has the maximum number of introns. All genes of group H have 2 introns, except for *AcoERF8*. To further explore the divergence and functional relationship of *AcoERF* proteins, a total of 10 conserved motifs in the pineapple *ERFs* were identified by the MEME software, and the height of each letter in the logo was proportional to the conservation level of amino acid in all sequences analyzed (Fig. S2). As displayed in Fig. 3, motif 1 is present in every pineapple *ERF* protein, and almost all the proteins contain motif 2, except for *AcoERF38* and *AcoERF69*. In addition, the motifs in different groups indicated that varying degrees of divergence among them. For example, motif 8 is unique to group H. Motif 6 is only present in group H and I. Motif 10 is only present in the



**Figure 2** The exon-intron structure of *AcoERF* genes based on the evolutionary relationship. The yellow round-corner rectangle represents exons, the black shrunked line represents introns, and the blue round-corner rectangle represents UTR.

Full-size DOI: 10.7717/peerj.10014/fig-2





**Figure 3** Motif distribution of *AcoERF* proteins. The conserved motifs in the *AcoERF* proteins were identified with MEME software. Grey lines denote the non-conserved sequences, and each motif is indicated by a colored box numbered at the bottom. The length of motifs in each protein was presented proportionally.

Full-size DOI: 10.7717/peerj.10014/fig-3

members of group C1 and C2. In general, the pineapple ERF proteins in the same group usually contained similar motifs, which indicates that they may play similar roles in the development and growth of pineapple.

### Chromosomal distribution and synteny analysis of *AcoERF* genes

According to Fig. S3, 71 *AcoERF* genes were unevenly distributed on the 23 pineapple linkage groups (LG). Among them, LG02 contains the most diverse *AcoERF* genes, accounting for 12.16% of total genes, whereas LG04, LG14, LG19, and LG25 possessed only one *AcoERF* gene, respectively. Additionally, *AcoERF71*, *AcoERF73* and *AcoERF74* exist on unanchored scaffolds, so they could not be conclusively charted on any pineapple linkage groups (Table S2).

Segmental duplication provide an important mechanism for the expansion of gene families (Lynch & Conery, 2000; Vision, Brown & Tanksley, 2000; Cannon et al., 2004; Zhang et al., 2018). A total of 16 segmental duplication evens with 24 *AcoERF* genes

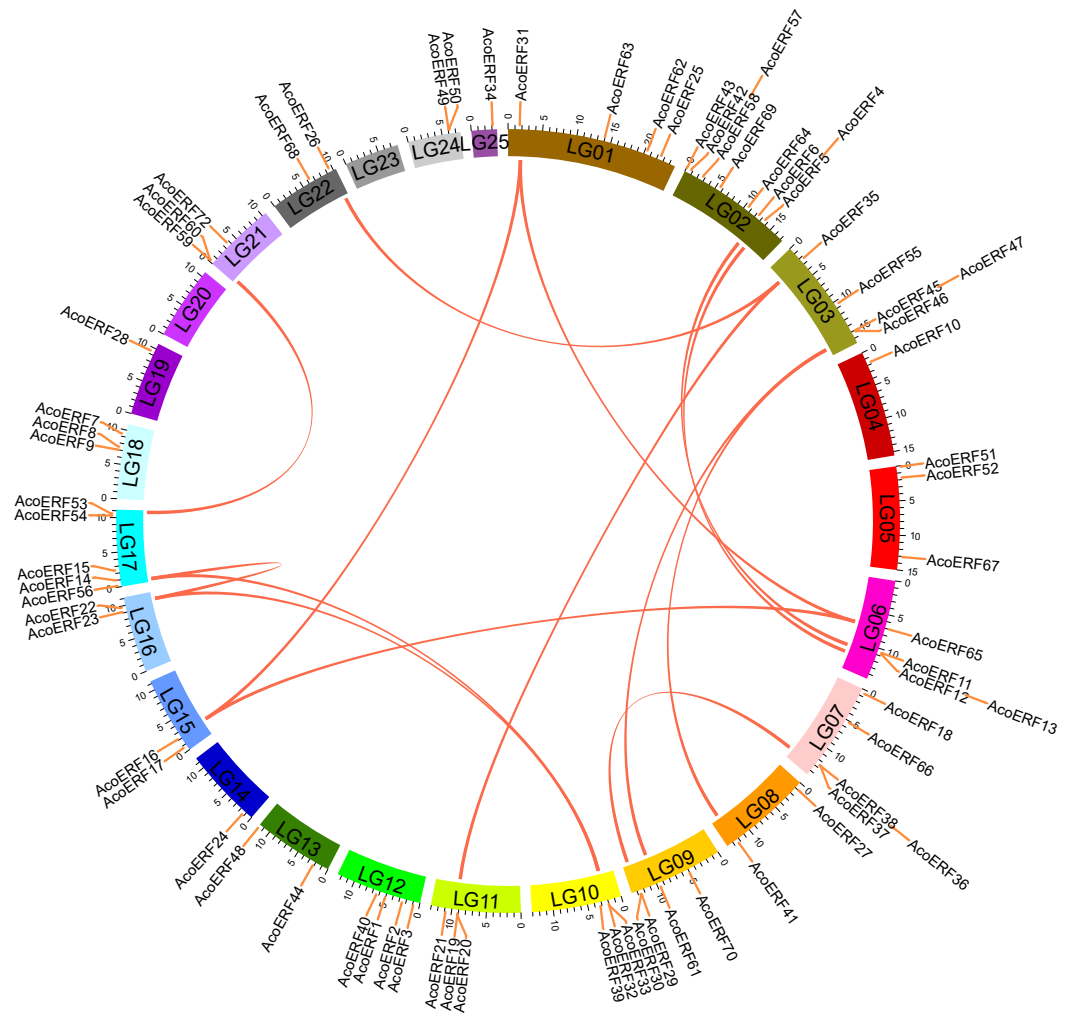
were found from pineapple genome, such as *AcoERF13/AcoERF5*, *AcoERF16/AcoERF31*, *AcoERF19/AcoERF35*, all of these were segmental duplicates (Fig. 4, Table S3), and these segmental duplication pairs play crucial roles in the expansion of ERF family in pineapple genome. In order to further study the phylogenetic and evolutionary trait of ERF family in pineapple, we set up 2 comparative syntenic maps of pineapple associated with 2 representative plant, including *Arabidopsis* and rice (Figs. 5A, 5B). According to the syntenic results, there are 37 and 44 orthologous gene pairs in *Arabidopsis* and rice, respectively (Tables S4 and S5). Some *AcoERF* genes had multiple orthologous gene pairs (one pineapple gene associated with multiple *Arabidopsis* or rice genes), such as *AcoERF62* associated with *AtERF008/019/020*, and *AcoERF29* associated with *OsERF025/026/027*, suggesting these genes might play crucial roles in the evolution of ERF family. Some *AcoERF* genes could be found orthologous gene pairs in *Arabidopsis* and rice, such as *AcoERF15*, *AcoERF22* and *AcoERF32*, suggesting that these genes might already exist before speciation. Some *AcoERF* genes only existed orthologous gene pairs between pineapple and rice, such as *AcoERF15*, *AcoERF22* and *AcoERF32*, suggesting that the appearance of these orthologous gene pairs before the divergence of monocotyledonous and dicotyledonous plants. In order to further understand the evolution of *ERF* gene family, the *Ka/Ks* ratios of the *ERF* gene pairs were calculated. The *Ka/Ks* ratio can represent different selection categories for duplication genes, such as the *Ka/Ks* >1 indicates positive selection, the *Ka/Ks* = 1 indicates neutral evolution, and the *Ka/Ks* <1 indicates negative selection (Li et al., 2019). According to the results, the *Ka/Ks* ratio of most orthologous *ERF* gene pairs are less than 1, indicating that they might have experienced strong negative selection pressure during the evolution of pineapple.

### Expression patterns of *AcoERF* genes in different tissues of pineapple

The expression profiles of 74 *AcoERF* genes in various pineapple tissues at different development stages was performed using RNA-seq expression data from MD2 pineapple plants recently published (Ming et al., 2015). According to the result, a total of 70 *AcoERF* genes were expressed in different tissues, whereas 4 *AcoERF* genes (*AcoERF22*, *AcoERF32*, *AcoERF43*, *AcoERF72*) were not found in the RNA-seq libraries (Fig. 6, Table S6). 22 genes were expressed in all tested tissues, in which 5 genes (*AcoERF13*, *AcoERF16*, *AcoERF31*, *AcoERF42*, *AcoERF65*) showed relative high expression levels (value > 10), suggesting these genes might played the indispensable roles in pineapple development. Moreover, 8, 16, 6, 9 genes showed high expression levels at all selected stages in ovule, sepal, stamen and petal, respectively; 16, 13, 12 genes showed high expression levels in leaf, flower and root, respectively; 10 genes showed high expression levels in fruit at all detected stages, indicating that these genes might be involved in the growth of relative tissues. Additionally, *AcoERF42* and *AcoERF65* showed the highest expression levels in fruit, suggesting that these genes might play positive roles in the growth and development of fruit.

### Identification of *cis*-elements in *AcoERF* promoters

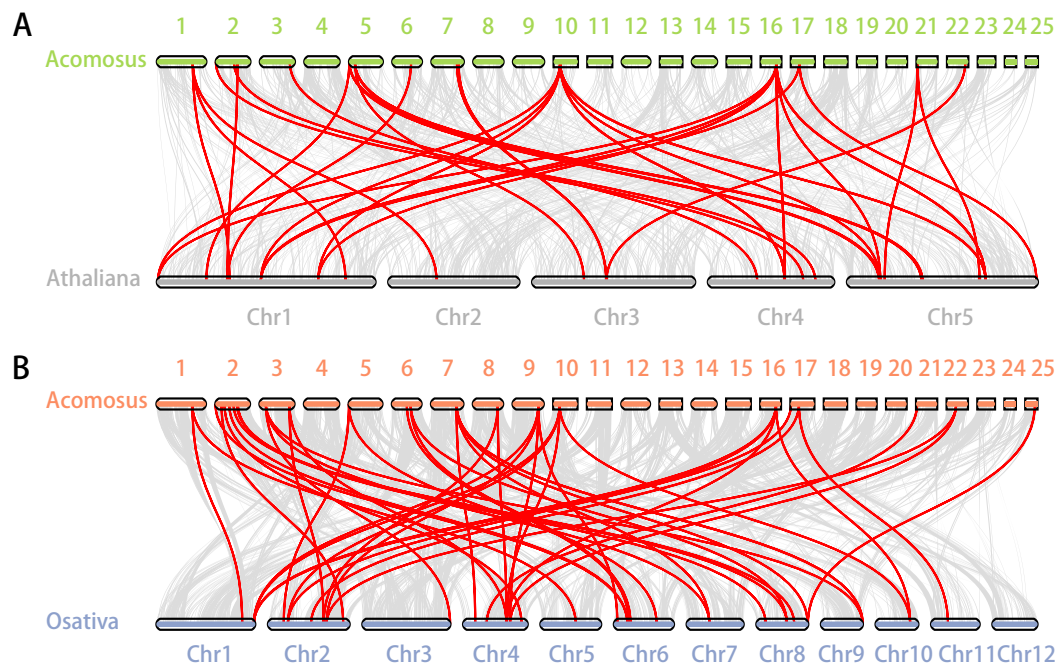
In order to determine the response mechanisms of *AcoERFs*, 13 stress-related *cis*-elements were found in pineapple ERF promoters, such as low temperature response elements



**Figure 4** Schematic representations for the chromosomal distribution and interchromosomal relationships of pineapple *ERF* genes. The red lines indicate duplicated *ERF* gene pairs, and the chromosome number is indicated at the bottom of each chromosome.

Full-size [DOI: 10.7717/peerj.10014/fig-4](https://doi.org/10.7717/peerj.10014/fig-4)

(LTR), dehydration responsive elements (DRE) and defense and stress responsive elements (TC-rich) (Table S7). LTRs are involved in low-temperature stress response (Choudhury *et al.*, 2008); DRE and TC-rich repeats play roles in dehydration, low temperature, and salt stress response (Diazdeleon, Klotz & Lagrimini, 1993; Germain *et al.*, 2012). According to the results, all *AcoERF* genes contain more than two *cis*-elements in their promoters. The promoter of *AcoERF23* only contained two *cis*-elements, whereas the most ten in the promoter of *AcoERF46*. In addition, at least one ABRE was present in 85.13% (63 out of 74) *AcoERF* gene promoters. The analysis of the 13 *cis*-elements suggested that most of *AcoERF* genes could response to different stress conditions.

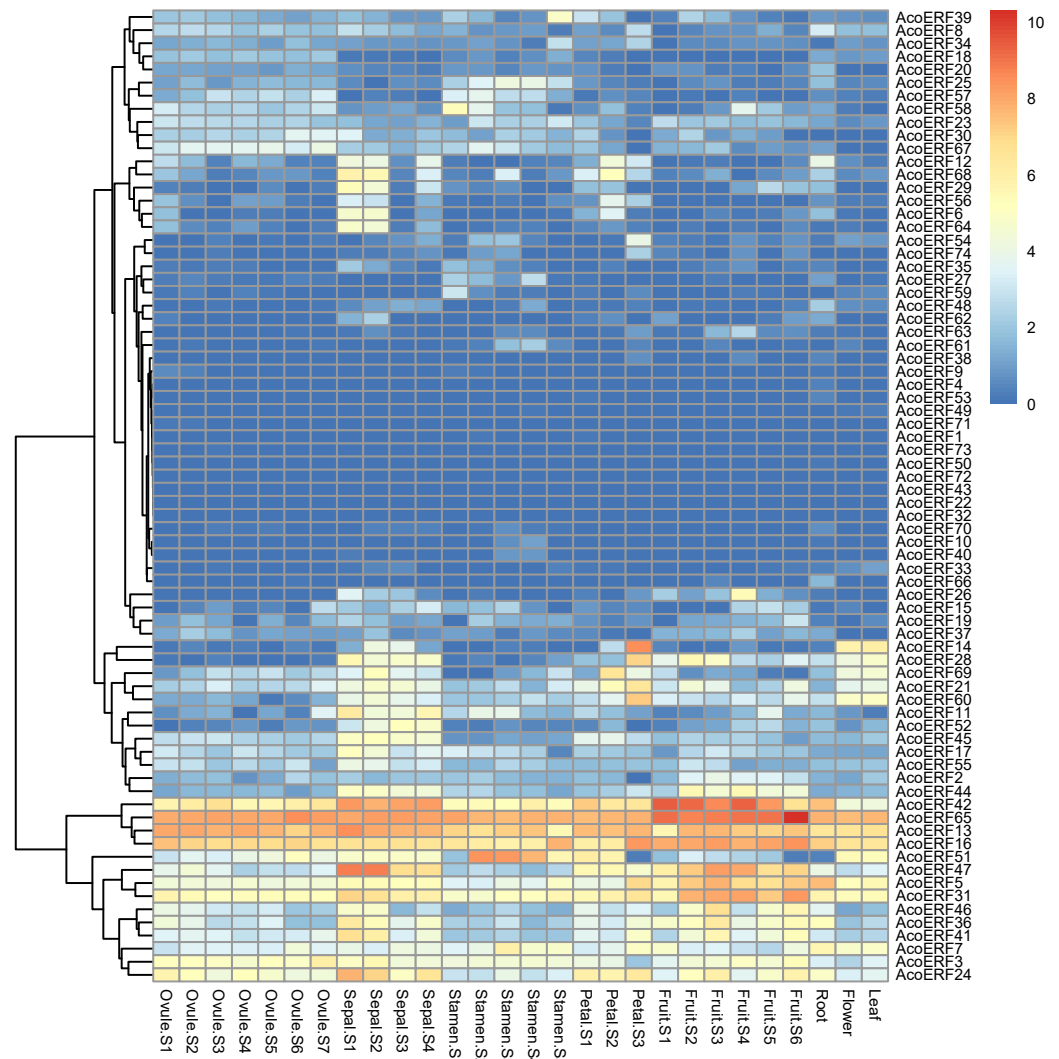


**Figure 5** Synteny analysis of *ERF* genes between pineapple and two representative plant species. (A) Synteny analysis of *ERF* genes between pineapple and *Arabidopsis*; (B) Synteny analysis of *ERF* genes between pineapple and rice. Gray lines in the background indicate the collinear blocks within pineapple and other plant genomes, while the red lines highlight the syntentic *ERF* gene pairs.

Full-size [DOI: 10.7717/peerj.10014/fig-5](https://doi.org/10.7717/peerj.10014/fig-5)

### Expression profiles of *AcoERF* genes under various abiotic stresses

According to previous studies, some *ERF* genes were involved in abiotic stress response in various plants, such as *Arabidopsis* (Sakuma et al., 2002), grape (Zhuang et al., 2009), tomato (Sharma et al., 2010), peach (*Amygdalus persica* L.) (Zhang et al., 2012). However, there was no report about the function of *ERF* genes in response to abiotic stress in pineapple. To investigate the roles of *AcoERF* genes in abiotic stress response, 9 *AcoERF* genes in roots and leaves were selected for functional verification. RT-PCR experiments were performed to analyze the expression patterns in response to different abiotic stress treatments, including cold, drought and salt stress (Figs. 7–9, Tables S8–S10). Overall, we found that the expression of these genes was influenced by the treatments. Among these treatments, some *AcoERF* genes in different pineapple tissues were significantly induced by the treatments, such as *AcoERF36*, *AcoERF45*, and *AcoERF67*. Interestingly, the expression patterns of *AcoERF* genes from root and leaf were different under the same treatments, such as *AcoERF13*, *AcoERF16* and *AcoERF42* under cold stress, *AcoERF7*, *AcoERF36* and *AcoERF60* in response to drought and salt stress, indicated that these genes may have different functions in various tissues of pineapple. Under cold and drought stress, the expression level of most *AcoERF* genes in leaves was higher than that in roots, while it was the opposite under salt stress, suggesting that *AcoERF* genes in root were sensitive to salt stress. Besides, most of *AcoERF* genes had maximal expression before 12 h in drought and



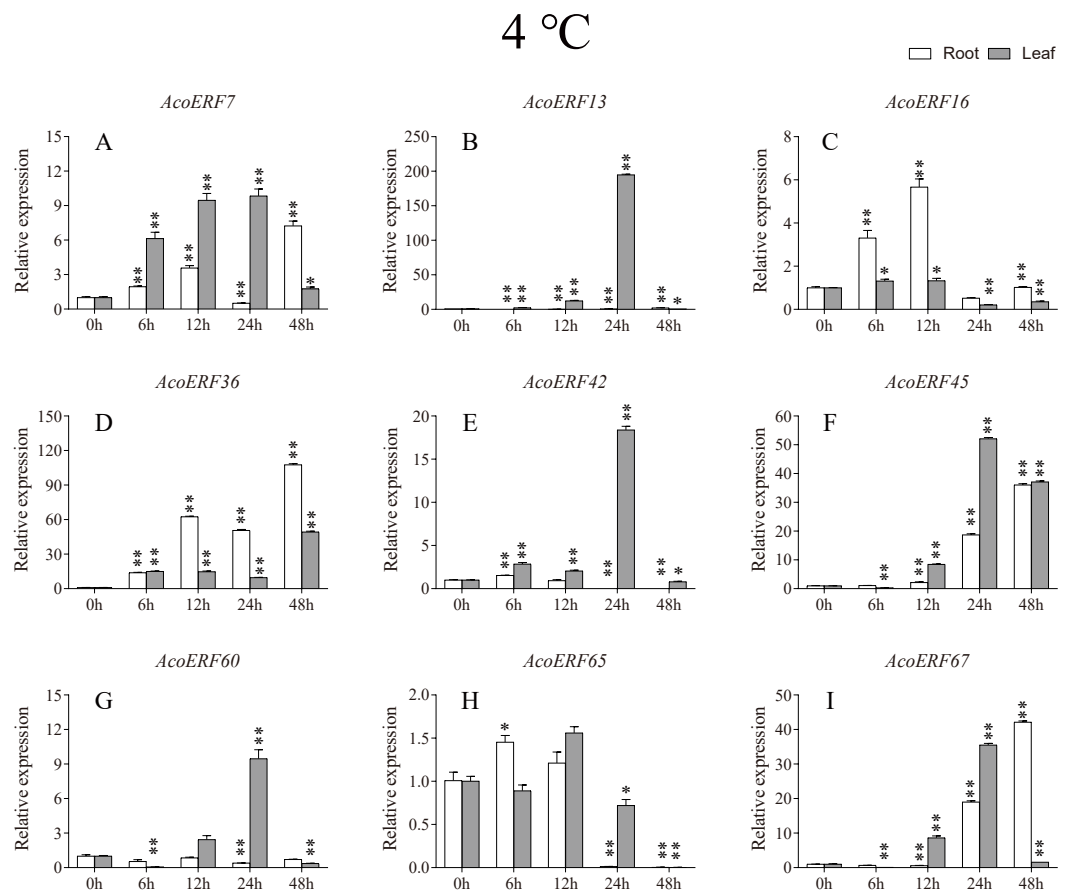
**Figure 6** Expression profiles of 74 *AcoERF* genes in different tissues and stages of pineapple. Different colors in map represent gene transcript abundance values as shown in bar by the side of the figure.

Full-size [DOI: 10.7717/peerj.10014/fig-6](https://doi.org/10.7717/peerj.10014/fig-6)

salt stress, indicating these genes might play key roles in the early stage of drought and salt stress responses.

## DISCUSSION

The ERF family is one of the most important transcription factor families, and it belongs to AP2/ERF super-family, which also contains AP2 and RAV family (Sakuma *et al.*, 2002; Nakano *et al.*, 2006). It has been reported that the member of ERF family plays a crucial role in the growth and development of various plants (Fits & Memelink, 2000; Banno *et al.*, 2001; Yu *et al.*, 2012; Muller & Munnebosch, 2015). Previous studies have identified 120 ERF family members in soybean (Li *et al.*, 2005), 103 in cucumber (Hu & Liu, 2011), 85 in tomato (Sharma *et al.*, 2010), 139 and 122 in rice and *Arabidopsis*, respectively



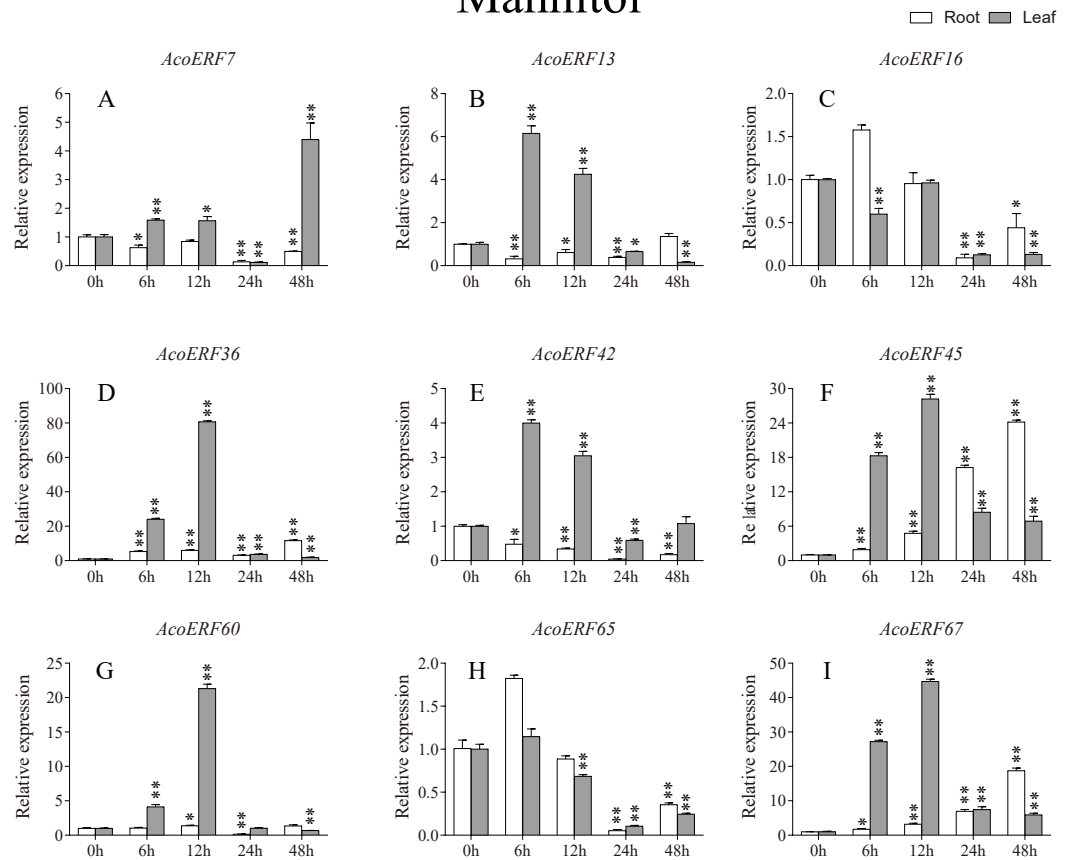
**Figure 7** Expression profile of nine selected *AcoERF* genes in response to cold stress treatment. The expression level of the 9 *AcoERF* genes (A–I) were gained by qRT-PCR. Error bars indicate the standard deviation. Asterisks on top of the bars indicating statistically significant differences between the stress and counterpart controls (\* $P < 0.05$ , \*\* $P < 0.01$ ).

Full-size DOI: 10.7717/peerj.10014/fig-7

(Nakano et al., 2006; Sharoni et al., 2011), but the identification of pineapple ERF gene family has not been reported. Pineapple has great economic and research value, so it is necessary to investigate the function of ERF family in pineapple. In this study, 103 candidate *AP2/ERF* genes were identified in pineapple genome, including 27 *AP2* family members, 2 *RAV* family members and 74 *ERF* family members (Fig. 1A). Compared with soybean, cucumber, tomato, rice and *Arabidopsis*, the pineapple *ERF* family is relatively small, indicating that some *ERF* members of pineapple may be lost during the evolution of species. To reveal the phylogenetic relationship of pineapple *ERF* family, a phylogenetic tree was constructed, 74 pineapple *ERF* family members were divided into 13 groups (Fig. 1B).

Gene structure analysis plays a crucial role in revealing the function of genes. Here, our results suggested that *AcoERFs* possess introns with number varying from 0 to 16, and the *ERF* members within the same group exhibited similar gene structure (Fig. 2). 66.22% *AcoERF* genes had no introns, which is similar to the status in *Arabidopsis*, cucumber and tartary buckwheat (Nakano et al., 2006; Hu & Liu, 2011; Liu et al., 2019). Presence of long,

## Mannitol

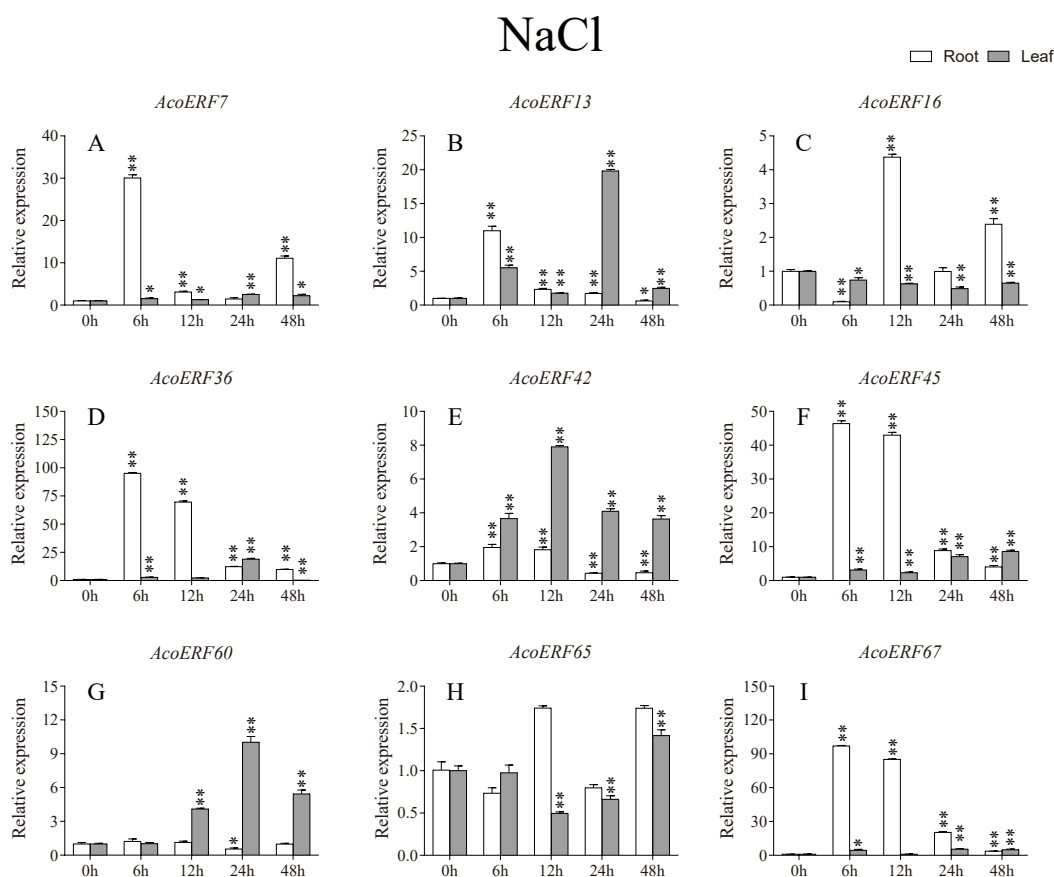


**Figure 8** Expression profile of nine selected *AcoERF* genes in response to drought stress treatment. The expression level of the 9 *AcoERF* genes (A–I) were gained by qRT-PCR. Error bars indicate the standard deviation. Asterisks on top of the bars indicating statistically significant differences between the stress and counterpart controls (\* $P < 0.05$ , \*\* $P < 0.01$ ).

Full-size DOI: 10.7717/peerj.10014/fig-8

multiple introns can delay transcriptional output, which may suppress the expression of genes under adverse conditions. Conversely, the genes with small or fewer introns may have efficient expression in response to stress environments (Jeffares, Penkett & Bahler, 2008; Heyn et al., 2015). Hence, a large number of intron-less *AcoERF* genes may react rapidly when the external environment changes. For instance, *AcoERF13*, *AcoERF36*, *AcoERF45* and *AcoERF67* had no introns, and they had efficient expression under various stresses within 48 h (Figs. 7–9, Tables S8–S10).

The domains and motifs of transcription factors play essential roles in proteins interaction, transcriptional activity and DNA binding (Liu, White & Macrae, 1999). Here, a total of 10 conserved motifs in the pineapple ERFs were identified (Fig. 3). Different types and numbers of motifs in the pineapple ERF proteins could affect the diversity of gene function. Motif 1, 2 and 3 correspond to the AP2/ERF domain region, and were highly conserved in pineapple ERF family members. Motif 4 is present in several AcoERFs and specifically rich in glutamine (Q). The presence of poly (Q) motif is related



**Figure 9** Expression profile of nine selected *AcoERF* genes in response to salt stress treatment. The expression level of the 9 *AcoERF* genes (A–I) were gained by qRT-PCR. Error bars indicate the standard deviation. Asterisks on top of the bars indicating statistically significant differences between the stress and counterpart controls (\* $P < 0.05$ , \*\* $P < 0.01$ ).

Full-size [DOI: 10.7717/peerj.10014/fig-9](https://doi.org/10.7717/peerj.10014/fig-9)

to protein aggregation and can stabilize protein-protein interactions (Liu *et al.*, 2011; Schaefer, Wanker & Andradenavarro, 2012). Motif 5 containing multiple serine residues, and the poly serine repeats role as flexible linker, as well as an important site of several posttranslational modifications (Uversky, 2015). Therefore, *AcoERFs* containing motif 5 may play crucial roles in various pathways. Motif 8 is specifically detected in group H, and the “MCGGAI” residues were highly conserved in the N-terminal region of the motif 8. It has been reported that the overexpression of the *ERF* genes with the “MCGGAI” motif can improve plant tolerance under hypoxia stress (Xu *et al.*, 2006). Besides, the genes containing “MCGGAI” motif is also involved in ethylene transcription activation (Buttner & Singh, 1997). Although the function of some conserved motifs is still unknown, they may also be involved in transcriptional regulation. In general, the *ERF* family members within the same group shared similar gene structures and motif compositions, which suggests that they may have similar roles in plant development and growth.

Segmental duplication occurs frequently in plants since most plants are diploidized polyploids (Zhu *et al.*, 2014). In this study, a total of 16 segmental duplication events with



24 *AcoERF* genes were found from pineapple genome (Fig. 4, Table S3). The diversity of gene combinations may be one of the factors that make the regulatory relationship complicated (Zhang et al., 2018; Zhang et al., 2020). Polyploidization, a widespread phenomenon among plant, is considered an important process of plant speciation and evolution, and the formation of polyploid includes hybridization and genomic doubling process before or after hybridization (Kohler, Scheid & Erilova, 2010; Zeng et al., 2020). Various *Arabidopsis* and rice *ERF* genes have been reported to be involved in the regulation of plant stress tolerance. For instance, *At5g52020* and *OsDREB1F* have been shown to play crucial roles in high-salt, low-temperature and drought stresses (Nakano et al., 2006). The synteny analysis could reveal the functional and evolutionary connections between two species. Here, 15 pineapple *ERF* genes and 25 *Arabidopsis* *ERF* genes, 23 pineapple and 32 rice *ERF* genes were identified as orthologous gene pairs, and different kinds of syntenic orthologous gene pairs could be investigated the evolutionary process of *ERF* family members in pineapple (Figs. 5A, 5B, Tables S4–S5). Some *ERF* genes did not find any orthologous gene pairs, which can be attributed to the rearrangement or fusion of chromosomes during their evolution (He et al., 2014; Zhang et al., 2018).

The analysis of gene expression patterns can be preliminarily predicted the function of genes (Peng et al., 2015; Su et al., 2017; Zhang et al., 2018). In this study, 5 genes (*AcoERF13*, *AcoERF16*, *AcoERF31*, *AcoERF42*, *AcoERF65*) showed high expression levels in all selected pineapple tissues (Fig. 6, Table S6), indicating that *ERF* genes may play a crucial role in pineapple development. Some *AcoERF* genes showed high expression levels in fruit, such as *AcoERF16*, *AcoERF45*, *AcoERF62*, suggesting that these genes play important roles in fruit ripening. Moreover, some *AcoERF* genes are expressed in various tissues or diverse stages, indicating that these genes could be more stable than those that only expressed in specific tissues or one stage of an organ (He et al., 2019).

When plants encounter stress conditions, a series of cell activities and molecules reaction mechanisms can improve the resistance of plants (Chinnusamy, Schumaker & Zhu, 2004; Mittler, 2006). According to previous studies, some *ERF* genes were involved in various stresses responses in plants, such as high-salt, low-temperature and drought stress (Ohmetakagi & Shinshi, 1995; Sakuma et al., 2002; Nakano et al., 2006). The expression profile of *AcoERF* genes suggested that several genes may be involved in the mechanism of abiotic stresses response (Figs. 7–9, Tables S8–S10). Besides, *AcoERF16*, *AcoERF42* and *AcoERF65* came from the same group H, and *AcoERF16* and *AcoERF65* was segmental duplication gene pair. Although these three genes had similar gene structure and motif compositions, the expression patterns of *AcoERF16* and *AcoERF65* were more similar under different treatments. The differential expression indicated that segmental duplication might influence the expression of gene. Overall, the above findings provide foundation to further investigate the potential function of pineapple *ERF* genes. These analyses are not only helpful in selecting valuable candidate *ERF* genes for further functional studies but also has important implication for genetic improvement for agricultural production and stress tolerance in pineapple crop.

## CONCLUSIONS

Based on genomic data of pineapple, a genome-wide identification of pineapple ERF family was performed, and 74 *AcoERFs* were identified. A comprehensive analysis of their phylogenetic relationships, gene structures and conserved motifs compositions showed high levels of similarity within the same group. Synteny analysis of ERF genes revealed the evolutionary characteristic of pineapple ERF family. The expression profile of *AcoERFs* verified their roles in responding to various abiotic stresses (cold, drought and salt stress). These results will help to further study the *ERF* genes and their role in abiotic and biotic stress tolerance for improve the agricultural productivity of pineapple crop.

## ACKNOWLEDGEMENTS

We would like to thank the reviewers for their helpful comments on the original manuscript.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This work was supported by the National Natural Science Foundation of China (U1605212, 31761130074 and 31970333 to Yuan Qin; 31700279 to Hanyang Cai) and a Guangxi Distinguished Experts Fellowship to Yuan Qin. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

National Natural Science Foundation of China: U1605212, 31761130074 and 31970333.  
Guangxi Distinguished Experts Fellowship.

### Competing Interests

The authors declare there are no competing interests.

### Author Contributions

- Youmei Huang, Yanhui Liu and Man Zhang performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Mengnan Chai performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Qing He, Bello Hassan Jakada, Fangqian Chen, Huihuang Chen and Xingyue Jin analyzed the data, prepared figures and/or tables, and approved the final draft.
- Hanyang Cai and Yuan Qin conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

### Data Availability

The following information was supplied regarding data availability:

The raw measurements are available as [Supplementary Files](#).

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.10014#supplemental-information>.

## REFERENCES

- Agarwal P, Agarwal PK, Joshi AJ, Sopory SK, Reddy MK. 2010. Overexpression of PgDREB2A transcription factor enhances abiotic stress tolerance and activates downstream stress-responsive genes. *Molecular Biology Reports* 37:1125–1135 DOI 10.1007/s11033-009-9885-8.
- Allen M, Yamasaki K, Ohmetakagi M, Tateno M, Suzuki M. 1998. A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO Journal* 17:5484–5496 DOI 10.1093/emboj/17.18.5484.
- Bailey TL, Johnson JR, Grant CE, Noble WS. 2015. The MEME suite. *Nucleic Acids Research* 43:39–49 DOI 10.1093/nar/gkv416.
- Banno H, Ikeda Y, Niu Q, Chua N. 2001. Overexpression of *Arabidopsis* ESR1 induces initiation of shoot regeneration. *The Plant Cell* 13:2609–2618 DOI 10.1105/tpc.010234.
- Boutillier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, Hattori J, Liu C, Lammeren AAMV, Miki B, Custers JBM, Campagne MMVL. 2002. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *The Plant Cell* 14:1737–1749 DOI 10.1105/tpc.001941.
- Buttner M, Singh K. 1997. *Arabidopsis thaliana* ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein. *Proceedings of the National Academy of Sciences of the United States of America* 94:5961–5966 DOI 10.1073/pnas.94.11.5961.
- Cai H, Zhang M, Chai M, He Q, Huang X, Zhao L, Qin Y. 2019. Epigenetic regulation of anthocyanin biosynthesis by an antagonistic interaction between H2A.Z and H3K4me3. *New Phytologist* 221:295–308 DOI 10.1111/nph.15306.
- Cai H, Zhao L, Wang L, Zhang M, Su Z, Cheng Y, Zhao H, Qin Y. 2017. ERECTA signaling controls *Arabidopsis* inflorescence architecture through chromatin-mediated activation of PRE1 expression. *New Phytologist* 214(4):1579–1596 DOI 10.1111/nph.14521.
- Cannon SB, Mitra A, Baumgarten A, Young ND, May G. 2004. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biology* 4:10 DOI 10.1186/1471-2229-4-10.
- Cao Y, Song F, Goodman R, Zheng Z. 2006. Molecular characterization of four rice genes encoding ethylene-responsive transcriptional factors and their expressions in response to biotic and abiotic stress. *Plant Physiology* 163:1167–1178 DOI 10.1016/j.jplph.2005.11.004.
- Chai M, Cheng H, Yan M, Priyadarshani S, Zhang M, He Q, Huang Y, Chen F, Liu L, Huang X, Lai L, Chen H, Cai H, Qin Y. 2020. Ananas comosus Identification

- and expression analysis of the DREB transcription factor family in pineapple ( (L.) Merr.). *PeerJ* **8**:e9006 DOI [10.7717/peerj.9006](https://doi.org/10.7717/peerj.9006).
- Chen P, Li Y, Zhao L, Hou Z, Yan M, Hu B, Liu Y, Azam S, Zhang Z, Rahman Z, Liu L, Qin Y. 2017.** Ananas comosus Genome-Wide Identification and Expression Profiling of ATP-Binding Cassette (ABC) Transporter Gene Family in Pineapple ( (L.) Merr.) Reveal the Role of in Pollen Development. *Frontiers in Plant Science* **8**:Article 2150 DOI [10.3389/fpls.2017.02150](https://doi.org/10.3389/fpls.2017.02150).
- Chinnusamy V, Schumaker K, Zhu J. 2004.** Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. *Journal of Experimental Botany* **55**:225–236 DOI [10.1093/jxb/erh005](https://doi.org/10.1093/jxb/erh005).
- Choudhury SR, Roy S, Das R, Sengupta D. 2008.** Differential transcriptional regulation of banana sucrose phosphate synthase gene in response to ethylene, auxin, wounding, low temperature and different photoperiods during fruit ripening and functional analysis of banana SPS gene promoter. *Planta* **229**:207–223 DOI [10.1007/s00425-008-0821-2](https://doi.org/10.1007/s00425-008-0821-2).
- Cui L, Feng K, Wang M, Wang M, Deng P, Song W, Nie X. 2016.** Genome-wide identification, phylogeny and expression analysis of AP2/ERF transcription factors family in *Brachypodium distachyon*. *BMC Genomics* **17**:1–19 DOI [10.1186/s12864-016-2968-8](https://doi.org/10.1186/s12864-016-2968-8).
- Diazdeleon F, Klotz KL, Lagrimini L. 1993.** Nucleotide sequence of the tobacco (*Nicotiana tabacum*) anionic peroxidase gene. *Plant Physiology* **101**:1117–1118 DOI [10.1104/pp.101.3.1117](https://doi.org/10.1104/pp.101.3.1117).
- Eddy SR. 2011.** Accelerated profile HMM searches. *PLOS Computational Biology* **7**:e1002195 DOI [10.1371/journal.pcbi.1002195](https://doi.org/10.1371/journal.pcbi.1002195).
- Edgar RC. 2004.** MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**:1792–1797 DOI [10.1093/nar/gkh340](https://doi.org/10.1093/nar/gkh340).
- Elliott R, Betzner A, Huttner E, Oakes M, Tucker W, Gerentes D, Perez P, Smyth D. 1996.** AINTEGUMENTA, an APETALA2-like gene of *Arabidopsis* with pleiotropic roles in ovule development and floral organ growth. *The Plant Cell* **8**:155–168 DOI [10.1105/tpc.8.2.155](https://doi.org/10.1105/tpc.8.2.155).
- Fang J, Miao C, Chen R, Ming R. 2016.** Genome-wide comparative analysis of microsatellites in pineapple. *Tropical Plant Biology* **9**:117–135 DOI [10.1007/s12042-016-9163-6](https://doi.org/10.1007/s12042-016-9163-6).
- Finn R, Tate J, Mistry J, Coggill P, Sammut S, Hotz H, Ceric G, Forslund K, Eddy S, Sonnhammer ELL, Bateman A. 2008.** The Pfam protein families database. *Nucleic Acids Research* **36**:D281–D288 DOI [10.1093/nar/gkm960](https://doi.org/10.1093/nar/gkm960).
- Fits LVD, Memelink J. 2000.** ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science* **289**:295–297 DOI [10.1126/science.289.5477.295](https://doi.org/10.1126/science.289.5477.295).
- Fujimoto S, Ohta M, Usui A, Shinshi H, Ohmetakagi M. 2000.** *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *The Plant Cell* **12**:393–404 DOI [10.1105/tpc.12.3.393](https://doi.org/10.1105/tpc.12.3.393).

- Galili T, Ocallaghan A, Sidi J, Sievert C. 2018.** heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* **34**(9):1600–1602 DOI [10.1093/bioinformatics/btx657](https://doi.org/10.1093/bioinformatics/btx657).
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003.** ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* **31**:3784–3788 DOI [10.1093/nar/gkg563](https://doi.org/10.1093/nar/gkg563).
- Germain H, Lachance D, Pelletier G, Fossdal CG, Solheim H, Séguin A. 2012.** The expression pattern of the *Picea glauca* Defensin 1 promoter is maintained in *Arabidopsis thaliana* indicating the conservation of signalling pathways between angiosperms, gymnosperms. *Journal of Experimental Botany* **63**:785–795 DOI [10.1093/jxb/err303](https://doi.org/10.1093/jxb/err303).
- Goodstein D, Shu S, Howson R, Neupane R, Hayes R, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar D. 2012.** Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* **40**:1178–1186 DOI [10.1093/nar/gkr944](https://doi.org/10.1093/nar/gkr944).
- Gouet P, Robert X, Courcelle E. 2003.** ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Research* **31**:3320–3323 DOI [10.1093/nar/gkg556](https://doi.org/10.1093/nar/gkg556).
- Hao D, Ohmetakagi M, Sarai A. 1998.** Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant. *Journal of Biological Chemistry* **273**:26857–26861 DOI [10.1074/jbc.273.41.26857](https://doi.org/10.1074/jbc.273.41.26857).
- He D, Lei Z, Xing H, Tang B. 2014.** Genome-wide identification and analysis of the aldehyde dehydrogenase (ALDH) gene superfamily of *Gossypium raimondii*. *Gene* **549**:123–133 DOI [10.1016/j.gene.2014.07.054](https://doi.org/10.1016/j.gene.2014.07.054).
- He Q, Liu Y, Zhang M, Bai M, Priyadarshani SVGN, Chai M, Chen F, Huang Y, Liu L, Cai H, Qin Y. 2019.** Genome-wide identification and expression analysis of the NAC transcription factor family in pineapple. *Tropical Plant Biology* **12**:255–267 DOI [10.1007/s12042-019-09233-3](https://doi.org/10.1007/s12042-019-09233-3).
- Heyn PC, Kalinka A, Tomancak P, Neugebauer K. 2015.** Introns and gene expression: cellular constraints, transcriptional, regulation and evolutionary consequences. *BioEssays* **37**:148–154 DOI [10.1002/bies.201400138](https://doi.org/10.1002/bies.201400138).
- Hu B, Jin J, Guo A, Zhang H, Luo J, Gao G. 2015.** GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* **31**:1296–1297 DOI [10.1093/bioinformatics/btu817](https://doi.org/10.1093/bioinformatics/btu817).
- Hu L, Liu S. 2011.** Genome-wide identification and phylogenetic analysis of the ERF gene family in cucumbers. *Genetics and Molecular Biology* **34**:624–634 DOI [10.1590/s1415-47572011005000054](https://doi.org/10.1590/s1415-47572011005000054).
- Jeffares D, Penkett C, Bahler J. 2008.** Rapidly regulated genes are intron poor. *Trends in Genetics* **24**:375–378 DOI [10.1016/j.tig.2008.05.006](https://doi.org/10.1016/j.tig.2008.05.006).
- Jin L, Liu J. 2008.** Molecular cloning, expression profile and promoter analysis of a novel ethylene responsive transcription factor gene GhERF4 from cotton (*Gossypium hirsutum*). *Plant Physiology and Biochemistry* **46**:46–53 DOI [10.1016/j.plaphy.2007.10.004](https://doi.org/10.1016/j.plaphy.2007.10.004).
- Kohler C, Scheid OM, Erilova A. 2010.** The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics* **26**:142–148 DOI [10.1016/j.tig.2009.12.006](https://doi.org/10.1016/j.tig.2009.12.006).

- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Research* 19:1639–1645 DOI 10.1101/gr.092759.109.
- Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Peer YVD, Rouze P, Rombauts S. 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research* 30:325–327 DOI 10.1093/nar/30.1.325.
- Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research* 44:W242–W245 DOI 10.1093/nar/gkw290.
- Letunic I, Bork P. 2018. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research* 46:D493–D496 DOI 10.1093/nar/gkx922.
- Li P, Zheng T, Li L, Zhuo X, Jiang L, Wang J, Cheng T, Zhang Q. 2019. Identification and comparative analysis of the CIPK gene family and characterization of the cold stress response in the woody plant *Prunus mume*. *PeerJ* 7:e6847 DOI 10.7717/peerj.6847.
- Li X, Tian A, Luo G, Gong Z, Zhang J, Chen S. 2005. Soybean DRE-binding transcription factors that are responsive to abiotic stresses. *Theoretical and Applied Genetics* 110:1355–1362 DOI 10.1007/s00122-004-1867-6.
- Licausi F, Giorgi F, Zenoni S, Osti F, Pezzotti M, Perata P. 2010. Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera*. *BMC Genomics* 11:719 DOI 10.1186/1471-2164-11-719.
- Liu Y, Chakrabortee S, Li R, Zheng Y, Tunnacliffe A. 2011. Both plant and animal LEA proteins act as kinetic stabilisers of polyglutamine-dependent protein aggregation. *FEBS Letters* 585:630–634 DOI 10.1016/j.febslet.2011.01.020.
- Liu M, Sun W, Ma Z, Zheng T, Huang L, Wu Q, Zhao G, Tang Z, Bu T, Li C, Chen H. 2019. Genome-wide investigation of the AP2/ERF gene family in tartary buckwheat (*Fagopyum Tataricum*). *BMC Plant Biology* 19:1–19 DOI 10.1186/s12870-019-1681-6.
- Liu L, White MJ, Macrae TH. 1999. Transcription factors and their genes in higher plants functional domains, evolution and regulation. *FEBS Journal* 262:247–257 DOI 10.1046/j.1432-1327.1999.00349.x.
- Loudet O, Hasegawa P. 2017. Abiotic stress, stress combinations and crop improvement potential. *Plant Journal* 90:837–838 DOI 10.1111/tpj.13604.
- Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290:1151–1155 DOI 10.1126/science.290.5494.1151.
- Meng X, Li F, Liu C, Zhang C, Wu Z, Chen Y. 2010. Isolation and characterization of an ERF transcription factor gene from cotton (*Gossypium barbadense* L.). *Plant Molecular Biology Reporter* 28:176–183 DOI 10.1109/TPWRS.2003.818738.
- Ming R, VanBuren R, Wai CM, Tang H, Schatz M, Bowers J, Lyons E, Wang M-L, Chen J, Biggers E, Zhang J, Huang L, Zhang L, Miao W, Zhang J, Ye Z, Miao C, Lin Z, Wang H, Zhou H, Yim W, Priest H, Zheng C, Woodhouse M, Edger PP, Guyot R, Guo H, Guo H, Zheng G, Singh R, Sharma A, Min X, Zheng Y, Lee H, Gurtowski

- J, Sedlazeck F, Harkes P. 2015.** The pineapple genome and the evolution of CAM photosynthesis. *Nature Genetics* **47**:1435–1442 DOI [10.1038/ng.3435](https://doi.org/10.1038/ng.3435).
- Mittler R. 2006.** Abiotic stress, the field environment and stress combination. *Trends in Plant Science* **11**:15–19 DOI [10.1016/j.tplants.2005.11.002](https://doi.org/10.1016/j.tplants.2005.11.002).
- Moose S, Sisco P. 1996.** Glossy15, an APETALA2-like gene from maize that regulates leaf epidermal cell identity. *Genes & Development* **10**:3018–3027 DOI [10.1101/gad.10.23.3018](https://doi.org/10.1101/gad.10.23.3018).
- Muller M, Munnebosch S. 2015.** Ethylene response factors: a key regulatory hub in hormone and stress signaling. *Plant Physiology* **169**:32–41 DOI [10.1104/pp.15.00677](https://doi.org/10.1104/pp.15.00677).
- Nakano T, Suzuki K, Fujimura T, Shinshi H. 2006.** Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiology* **140**:411–432 DOI [10.1104/pp.105.073783](https://doi.org/10.1104/pp.105.073783).
- Ohmetakagi M, Shinshi H. 1995.** Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *The Plant Cell* **7**:173–182 DOI [10.1105/tpc.7.2.173](https://doi.org/10.1105/tpc.7.2.173).
- Okamoto J, Caster B, Villarroel R, Montagu MV, Jofuku K. 1997.** The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **94**:7076–7081 DOI [10.1073/pnas.94.13.7076](https://doi.org/10.1073/pnas.94.13.7076).
- Pan Y, Seymour G, Lu C, Hu Z, Chen X, Chen G. 2012.** An ethylene response factor (ERF5) promoting adaptation to drought and salt tolerance in tomato. *Plant Cell Reports* **31**:349–360 DOI [10.1007/s00299-011-1170-3](https://doi.org/10.1007/s00299-011-1170-3).
- Peng X, Zhao Y, Li X, Wu M, Chai W, Sheng L, Wang Y, Dong Q, Jiang H, Cheng B. 2015.** Genomewide identification, classification and analysis of NAC type gene family in maize. *Journal of Genetics* **94**:377–390 DOI [10.1007/s12041-015-0526-9](https://doi.org/10.1007/s12041-015-0526-9).
- Rashid M, He G, Yang G, Hussain J, Xu Y. 2012.** AP2/ERF transcription factor in rice: genome-wide canvas and syntenic relationships between monocots and eudicots. *Evolutionary Bioinformatics* **8**:321–355 DOI [10.4137/EBO.S9369](https://doi.org/10.4137/EBO.S9369).
- Rashotte AM, Mason MG, Hutchison CE, Ferreira F, Schaller GE, Kieber JJ. 2006.** A subset of *Arabidopsis* AP2 transcription factors mediates cytokinin responses in concert with a two-component pathway. *Proceedings of the National Academy of Sciences of the United States of America* **103**:11081–11085 DOI [10.1073/pnas.0602038103](https://doi.org/10.1073/pnas.0602038103).
- Riechmann J, Meyerowitz E. 1998.** The AP2/EREBP family of plant transcription factors. *Biological Chemistry* **379**:633–646 DOI [10.1515/bchm.1998.379.6.633](https://doi.org/10.1515/bchm.1998.379.6.633).
- Romanel E, Schrago CG, Counago R, Russo CAM, Alvesferreira M. 2009.** Evolution of the B3 DNA binding superfamily: new insights into REM family gene diversification. *PLOS ONE* **4**:e5791 DOI [10.1371/journal.pone.0005791](https://doi.org/10.1371/journal.pone.0005791).
- Sakuma Y, Liu Q, Dubouzet J, Abe H, Shinozaki K, Yamaguchi-Shinozaki K. 2002.** DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochemical and Biophysical Research Communications* **290**:998–1009 DOI [10.1006/bbrc.2001.6299](https://doi.org/10.1006/bbrc.2001.6299).

- Sang J, Han X, Liu M, Qiao G, Jiang J, Zhuo R. 2013.** Selection and validation of reference genes for real-time quantitative PCR in hyperaccumulating ecotype of *Sedum alfredii* under different heavy metals stresses. *PLOS ONE* **8**:e82927 DOI [10.1371/journal.pone.0082927](https://doi.org/10.1371/journal.pone.0082927).
- Schaefer M, Wanker EE, Andradenavarro MA. 2012.** Evolution and function of CAG/polyglutamine repeats in protein-protein interaction networks. *Nucleic Acids Research* **40**:4273–4287 DOI [10.1093/nar/gks011](https://doi.org/10.1093/nar/gks011).
- Sharma MK, Kumar R, Solanke AU, Sharma R, Tyagi AK, Sharma AK. 2010.** Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato. *Molecular Genetics and Genomics* **284**:455–475 DOI [10.1007/s00438-010-0580-1](https://doi.org/10.1007/s00438-010-0580-1).
- Sharoni AM, Nuruzzaman M, Satoh K, Shimizu T, Kondoh H, Sasaya T, Choi I-R, Omura T, Kikuchi S. 2011.** Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant and Cell Physiology* **52**:344–360 DOI [10.1093/pcp/pcq196](https://doi.org/10.1093/pcp/pcq196).
- Stockinger E, Gilmour S, Thomashow M. 1997.** *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proceedings of the National Academy of Sciences of the United States of America* **94**:1035–1040 DOI [10.1073/pnas.94.3.1035](https://doi.org/10.1073/pnas.94.3.1035).
- Su Z, Wang L, Li W, Zhao L, Huang X, Azam SM, Qin Y. 2017.** Genome-wide identification of auxin response factor (ARF) genes family and its tissue-specific prominent expression in pineapple (*Ananas comosus*). *Tropical Plant Biology* **10**:86–96 DOI [10.1007/s12042-017-9187-6](https://doi.org/10.1007/s12042-017-9187-6).
- Tamura K, Stecher G, Peterson DS, Filipski A, Kumar S. 2013.** MEGA6: molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**:2725–2729 DOI [10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197).
- Tanenbaum D, Goll J, Murphy S, Kumar P, Zafar N, Thiagarajan M, Madupu R, Davidsen T, Kagan L, Kravitz S, Rusch D, Yooseph S. 2010.** The JCVI standard operating procedure for annotating prokaryotic metagenomic shotgun sequencing data. *Standards in Genomic Sciences* **2**:229–237 DOI [10.4056/sigs.651139](https://doi.org/10.4056/sigs.651139).
- Tang H, Bowers J, Wang X, Ming R, Alam M, Paterson A. 2008.** Synteny and collinearity in plant genomes. *Science* **320**:486–488 DOI [10.1126/science.1153917](https://doi.org/10.1126/science.1153917).
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012.** Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**:562–578 DOI [10.1038/nprot.2012.016](https://doi.org/10.1038/nprot.2012.016).
- Uversky VN. 2015.** The intrinsic disorder alphabet, III. Dual personality of serine. *Intrinsically Disordered Proteins* **3**:e1027032 DOI [10.1080/21690707.2015.1027032](https://doi.org/10.1080/21690707.2015.1027032).
- Vernie T, Moreau S, Billy FD, Plet J, Combiér J, Rogers C, Oldroyd GED, Frugier F, Niebel A, Gamas P. 2008.** EFD is an ERF transcription factor involved in the control of nodule number and differentiation in *Medicago truncatula*. *The Plant Cell* **20**:2696–2713 DOI [10.1105/tpc.108.059857](https://doi.org/10.1105/tpc.108.059857).



- Vision T, Brown DG, Tanksley SD. 2000.** The origins of genomic duplications in Arabidopsis. *Science* **290**:2114–2117 DOI [10.1126/science.290.5499.2114](https://doi.org/10.1126/science.290.5499.2114).
- Voorrips RE. 2002.** MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity* **93**:77–78 DOI [10.1093/jhered/93.1.77](https://doi.org/10.1093/jhered/93.1.77).
- Wang W. 2015.** The molecular detection of corynespora cassiicola on cucumber by PCR assay using DNA man software and NCBI. In: *International conference on computer and computing technologies in agriculture, vol. 479*. 248–258 DOI [10.1007/978-3-319-48354-2\\_26](https://doi.org/10.1007/978-3-319-48354-2_26).
- Xie T, Chen C, Li C, Liu J, Liu C, He Y. 2018.** Genome-wide investigation of WRKY gene family in pineapple: evolution and expression profiles during development and stress. *BMC Genomics* **19**:1–18 DOI [10.1186/s12864-018-4880-x](https://doi.org/10.1186/s12864-018-4880-x).
- Xu K, Xu X, Fukao T, Canlas PE, Maghirangrodriguez R, Heuer S, Ismail A, Baileyserres J, Ronald PC, Mackill DJ. 2006.** Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* **442**:705–708 DOI [10.1038/nature04920](https://doi.org/10.1038/nature04920).
- Yu Z, Li J, Yang C, Hu W, Wang L, Chen X. 2012.** The jasmonate-responsive AP2/ERF transcription factors AaERF1 and AaERF2 positively regulate artemisinin biosynthesis in Artemisia annua L. *Molecular Plant* **5**:353–365 DOI [10.1093/mp/ssr087](https://doi.org/10.1093/mp/ssr087).
- Zeng R, Zhu J, Xu S, Du G, Guo H, Chen J, Zhang Z, Xie L. 2020.** CymbidiumUnreduced male gamete formation in and its use for developing sexual polyploid cultivars. *Frontiers in Plant Science* **11**:Article 558 DOI [10.3389/fpls.2020.00558](https://doi.org/10.3389/fpls.2020.00558).
- Zhang M, Liu Y, He Q, Chai M, Huang Y, Chen F, Wang X, Liu Y, Cai H, Qin Y. 2020.** Genome-wide investigation of calcium-dependent protein kinase gene family in pineapple: evolution and expression profiles during development and stress. *BMC Genomics* **21**:72 DOI [10.1186/s12864-020-6501-8](https://doi.org/10.1186/s12864-020-6501-8).
- Zhang M, Liu Y, Shi H, Chai M, He Q, Yan M, Cao D, Zhao L, Cai H, Qin Y. 2018.** Evolutionary and expression analyses of soybean basic Leucine zipper transcription factor family. *Bmc Genomics* **19**:1–14 DOI [10.1186/s12864-018-4511-6](https://doi.org/10.1186/s12864-018-4511-6).
- Zhang G, Ming C, Li L, Xu Z, Chen X, Guo J, Ma Y. 2009.** Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. *Journal of Experimental Botany* **60**:3781–3796 DOI [10.1093/jxb/erp214](https://doi.org/10.1093/jxb/erp214).
- Zhang C, Shangguan L, Ma R, Sun X, Tao R, Guo L, Korir NK, Yu M. 2012.** Genome-wide analysis of the AP2/ERF superfamily in peach (Prunus persica). *Genetics and Molecular Research* **11**:Article 4789 DOI [10.4238/2012.October.17.6](https://doi.org/10.4238/2012.October.17.6).
- Zhu Y, Wu N, Song W, Yin G, Qin Y, Yan Y, Hu Y. 2014.** Soybean (Glycine max) expansin gene superfamily origins: segmental and tandem duplication events followed by divergent selection among subfamilies. *BMC Plant Biology* **14**:93 DOI [10.1186/1471-2229-14-93](https://doi.org/10.1186/1471-2229-14-93).
- Zhuang J, Cai B, Peng R, Zhu B, Jin X, Xue Y, Gao F, Fu X, Tian Y, Zhao W, Qiao Y, Zhang Z, Xiong A, Yao Q. 2008.** Genome-wide analysis of the AP2/ERF gene family in Populus trichocarpa. *Biochemical and Biophysical Research Communications* **371**:468–474 DOI [10.1016/j.bbrc.2008.04.087](https://doi.org/10.1016/j.bbrc.2008.04.087).

**Zhuang J, Peng R, Cheng Z, Zhang J, Cai B, Zhang Z, Gao F, Zhu B, Fu X, Jin X, Chen J, Qiao Y, Xiong A, Yao Q. 2009.** Genome-wide analysis of the putative AP2/ERF family genes in *Vitis vinifera*. *Scientia Horticulturae* **123**:73–81  
[DOI 10.1016/j.scienta.2009.08.002](https://doi.org/10.1016/j.scienta.2009.08.002).