



# Genome-Wide Characterization of bHLH Genes in Grape and Analysis of their Potential Relevance to Abiotic Stress Tolerance and Secondary Metabolite Biosynthesis

## OPEN ACCESS

Pengfei Wang\*, Ling Su, Huanhuan Gao, Xilong Jiang, Xinying Wu, Yi Li, Qianqian Zhang, Yongmei Wang\* and Fengshan Ren\*

### Edited by:

Hemâni Gerós,  
University of Minho, Portugal

### Reviewed by:

Markus Rienth,  
Changins–School of viticulture and  
Enology, University of Applied  
Sciences and Arts of Western  
Switzerland, Switzerland  
Ana Margarida Fortes,  
Instituto de Biosistemas e Ciências  
Integrativas (BioISI), Portugal  
Andreia Figueiredo,  
Universidade de Lisboa, Portugal

### \*Correspondence:

Pengfei Wang  
fengqiaoyouzi@126.com  
Yongmei Wang  
wangym228@126.com  
Fengshan Ren  
rensd65@163.com

### Specialty section:

This article was submitted to  
Plant Breeding,  
a section of the journal  
Frontiers in Plant Science

Received: 20 July 2017

Accepted: 12 January 2018

Published: 01 February 2018

### Citation:

Wang P, Su L, Gao H, Jiang X, Wu X,  
Li Y, Zhang Q, Wang Y and Ren F  
(2018) Genome-Wide Characterization  
of bHLH Genes in Grape and Analysis  
of their Potential Relevance to Abiotic  
Stress Tolerance and Secondary  
Metabolite Biosynthesis.  
*Front. Plant Sci.* 9:64.  
doi: 10.3389/fpls.2018.00064

Shandong Academy of Grape, Jinan, China

Basic helix-loop-helix (bHLH) transcription factors are involved in many abiotic stress responses as well as flavonol and anthocyanin biosynthesis. In grapes (*Vitis vinifera* L.), flavonols including anthocyanins and condensed tannins are most abundant in the skins of the berries. Flavonols are important phytochemicals for viticulture and enology, but grape bHLH genes have rarely been examined. We identified 94 grape bHLH genes in a genome-wide analysis and performed Nr and GO function analyses for these genes. Phylogenetic analyses placed the genes into 15 clades, with some remaining orphans. 41 duplicate gene pairs were found in the grape bHLH gene family, and all of these duplicate gene pairs underwent purifying selection. Nine triplicate gene groups were found in the grape bHLH gene family and all of these triplicate gene groups underwent purifying selection. Twenty-two grape bHLH genes could be induced by PEG treatment and 17 grape bHLH genes could be induced by cold stress treatment including a homologous form of MYC2, VvbHLH007. Based on the GO or Nr function annotations, we found three other genes that are potentially related to anthocyanin or flavonol biosynthesis: VvbHLH003, VvbHLH007, and VvbHLH010. We also performed a cis-acting regulatory element analysis on some genes involved in flavonoid or anthocyanin biosynthesis and our results showed that most of these gene promoters contained G-box or E-box elements that could be recognized by bHLH family members.

**Keywords:** grape, bHLH transcription factor, abiotic tolerance, flavonol, anthocyanins, E-box, G-box

## INTRODUCTION

The bHLH protein is a kind of transcription factor, and it has DNA binding and dimerization capabilities as bHLH domain existed. bHLH domain contains approximately 60 amino acids with two functionally distinct regions of the HLH region and the basic region (Murre et al., 1989; Ledent and Vervoort, 2001). The HLH region contains two amphipathic  $\alpha$  helices separated by a loop region of variable length which could act as a dimerization domain that promote protein-protein interactions and form homo-dimers or hetero-dimers (Massari and Murre, 2000). The basic region

contains approximately 17 amino acids, and it is located at the N-terminus of the bHLH domain which allows bHLH transcription factors to bind to a consensus E-box (5'-CANNTG-3') or G-box (5'-CACGTG-3') cis elements to regulate gene expression (Atchley et al., 1999; Massari and Murre, 2000; Toledo-Ortiz et al., 2003; Li et al., 2006).

bHLH transcription factor family is very large in eukaryotes (Ledent and Vervoort, 2001; Toledo-Ortiz et al., 2003). There are 162 bHLH genes identified in *Arabidopsis* (Bailey et al., 2003), 167 bHLH genes in rice (Li et al., 2006), and 152 bHLH genes in tomato (Wang et al., 2015).

bHLH transcription factors could act as transcriptional activators or repressors and play important roles in metabolic and developmental processes (Feller et al., 2011). Most bHLH proteins identified have been functionally characterized in *Arabidopsis*. For example, PIF3 is a direct phytochrome reaction partner in the signaling network of photoreceptors, and it is involved in controlling the expression of light-regulated genes (Toledo-Ortiz et al., 2003). PIF4 plays a major role in growth regulation by integrating with multiple signals (Choi and Oh, 2016). bHLH transcript factor NAI1 is involved in the formation of endoplasmic reticulum bodies. RHD6 and RSL1 are involved in the formation of root hairs while LHW is involved in root development. PRE1, PRE2, PRE3, PRE4, and PRE5 are involved in gibberellin signaling transduction in plants (Feller et al., 2011).

Some bHLH family members have also been reported to be related to responses to abiotic stresses such as cold, drought, and salt stress. The tomato bHLH transcription factor MYC-type ICE, SLICE1a, could confer cold, osmotic, and salt tolerance of plants (Feng et al., 2013). A stress-responsive gene, bHLH transcript factor RsICE1 from *Raphanus sativus* could increase the cold tolerance of genetically modified rice (Man et al., 2017). bHLH transcript factor ICE1 could increase cold tolerance of *Pyrus ussuriensis* by enhancing PuDREBa transcriptional levels through interacting with PuHHP1 (Huang et al., 2015). *Arabidopsis* bHLH92 could function in plant's responses to osmotic stresses (Jiang et al., 2009). *Eleusine coracana* bHLH57 are related to drought, salt, and oxidative tolerance while bHLH122 plays an important role in drought and osmotic resistance in *Arabidopsis* (Liu et al., 2014). *Arabidopsis* bHLH112 was also reported to regulate the expression of genes that involved in abiotic stress tolerance (Liu et al., 2015). In addition to participate in adversity responses, bHLH genes in plants were also found to be involved in the biosynthesis pathways of anthocyanins (Mol et al., 1998; Nesi et al., 2000; Winkel-Shirley, 2001) and flavonols (Walker et al., 1999; Nesi et al., 2000; Baudry et al., 2004).

Grapes (*Vitis vinifera* L.) are one of the most widely cultivated fruit crops in the world, and they are of great economic importance. However, grape production is often severely limited by various abiotic stresses such as low temperature and winter droughts. In northern China of the main grape-producing areas, there have almost no grape cultivars survive under natural conditions of low temperatures and low air humidity in the winter because of the continental climate (Huang, 1980). Since the bHLH family has been proven to be associated with abiotic stress tolerance in many species, we want to identify

bHLH transcription factors in grapes to analysis their potential relevance to cold and drought tolerance.

Here, a total of 94 grape bHLH genes were identified in grape genome, and phylogenetic analyses were carried out to evaluate the relationships among these genes. The expression patterns of 94 grape bHLH genes in various tissues under cold and drought stress conditions were analyzed. Three grape bHLH genes that may have relevance to anthocyanin and flavonol biosynthesis were identified. Our purpose of this study was to find more candidate genes that may have relevance to cold and drought stresses tolerance, and to improve the current understanding of cold and drought tolerant mechanisms in *V. vinifera*. We also aimed to identify candidate genes that may participate in anthocyanin and flavonol biosynthesis for further breeding and cultivation of the new grape varieties.

## MATERIALS AND METHODS

### Data Collection and Identification of Grape bHLH Genes

The whole grape genome was obtained from the newest foxtail millet genome database (version v2.0; <http://genomes.cribi.unipd.it/DATA/V2/>). The conserved bHLH domain sequence (PF00010; <http://pfam.xfam.org/>) was used as a query to identify all possible bHLH protein candidates in the grape genome database using BLASTP ( $E < 0.001$ ). SMART online software (<http://smart.embl-heidelberg.de/>) was used to identify integrated bHLH domains in putative grape bHLH proteins. The date of the query was December 2016. The bHLH genes identified were named according to their order in the *V. vinifera* genomic sequence.

### Protein and GO Annotation

Nr annotations were based on the NCBI BLASTP and Nr databases ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)). GO function annotations were based on the GO database (<http://geneontology.org/page/go-database>).

### Phylogenetic Analysis of Grape bHLH Members

We performed a grape bHLH protein sequence alignment using Clustal Omega online software (<http://www.ebi.ac.uk/Tools/msa/clustalo>), and constructed Neighbor-Joining (NJ) trees using MEGA 6.0 with the aligned grape bHLH protein sequences. One-thousand bootstrap samples were generated to support the calculated relationships (Wang et al., 2016).

### Gene Duplication and Triplication Analyses

Grape bHLH duplicate gene pairs and triplicate gene groups were identified with two standards. Protein pairs with  $\geq 30\%$  identity, covering  $\geq 70\%$  of the protein length fell within the low-stringency standard. Protein pairs with  $\geq 50\%$  identity, covering  $\geq 90\%$  of the protein length were included in the high-stringency standard (Rizzon et al., 2006; Sun et al., 2015).

## Conserved Motif Analysis of Grape bHLH Protein Sequences

We analyzed conserved motifs in grape bHLH protein sequences using the Multiple Em for Motif Elicitation (MEME) suite 4.11.1 software (<http://meme-suite.org/tools/meme>) (Bailey et al., 2006) with the following parameter settings: output motifs, 20; minimum motif width, 6; and maximum motif width, 300 (Wang et al., 2016).

## Selective Pressure Analysis

The non-synonymous/synonymous ( $K_a/K_s$ ;  $\omega$ ) value of duplicate gene pairs or triplicate gene groups (between any two genes in one triplicate gene groups) were calculated using PAL2NAL (<http://www.bork.embl.de/pal2nal/>; Suyama et al., 2006; Yadav et al., 2016).

## Codon Usage Bias Analysis

Codon bias refers to the unequal use of synonymous codons for an amino acid (Hershberg and Petrov, 2008; Larracunte et al., 2008; Plotkin and Kudla, 2011; Guo et al., 2017). Coding sequences of 94 bHLH genes were used to calculate the frequency of optimal codons (FOP), GC content, GC content at the 3rd site of the synonymous codon (GC3s content), relative synonymous codon usage (RSCU), codon adaptation index (CAI), and codon bias index (CBI) with the online software CodonW 1.4 (<http://codonw.sourceforge.net>).

A plot of the effective number of codons (ENC) vs. GC3s (also called an ENC plot) is a strategic investigation into patterns of synonymous codon usage, providing a visual display of the main features of codon usage patterns for a number of genes. The effective number of codon values are always within the range of 20 (only one codon effectively used for each amino acid) to 61 (codons used randomly). The expected ENC values were calculated as follows (Wright, 1990):

$$ENC_{exp} = 2 + S + (29/(S^2 + (1 - S^2))) \quad (1)$$

Where, S is the frequency of G + C (i.e., GC3s).

## Promoter cis-Acting Regulatory Element Analysis

Plantcare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) was used to analyze the cis-acting regulatory elements of bHLH genes (Wang et al., 2016).

## In silico Gene Expression Analysis

The microarray expression profiles of 49 grape samples (GSE36128) (Fasoli et al., 2012), PEG (simulated drought), cold stress library, control library (GSE31594; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31594>), and CBF over-expression transgenic grape library (GSE29948; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29948>) were retrieved from the Plexdb (<http://www.plexdb.org/>) and GEO databases (<https://www.ncbi.nlm.nih.gov/gds/>). A heatmap was generated using MultiExperiment Viewer (MeV) v4.0.

## RESULTS

### Identification, Phylogenetic Analysis, and Annotation of Grape BHLH Genes

We identified 94 bHLH genes in *Vitis vinifera*. They were named from VvbHLH001 to VvbHLH094 according to their order in the *V. vinifera* genomic sequence. All bHLHs were annotated based on the best-hit proteins in the Nr database (Table S1). Our Nr-annotation showed that the grape bHLH family contained some homologous bHLH genes that had been previously reported in Arabidopsis (Feller et al., 2011). However, the annotation did not contain the homologs of FIT (involved in the regulation of iron uptake), AIB (involved in ABA signaling), RHD6 (involved in root hair formation), RSL1, RSL4, RSL3, RSL2, RSL5 (involved in root hair development), PRE1, PRE2, PRE3, PRE4, PRE5 (involved in gibberellin signaling), SISTYLE2.1 (involved in cell elongation in developing styles), KDR (involved in light signal transduction), and some orphans identified in Arabidopsis (Feller et al., 2011).

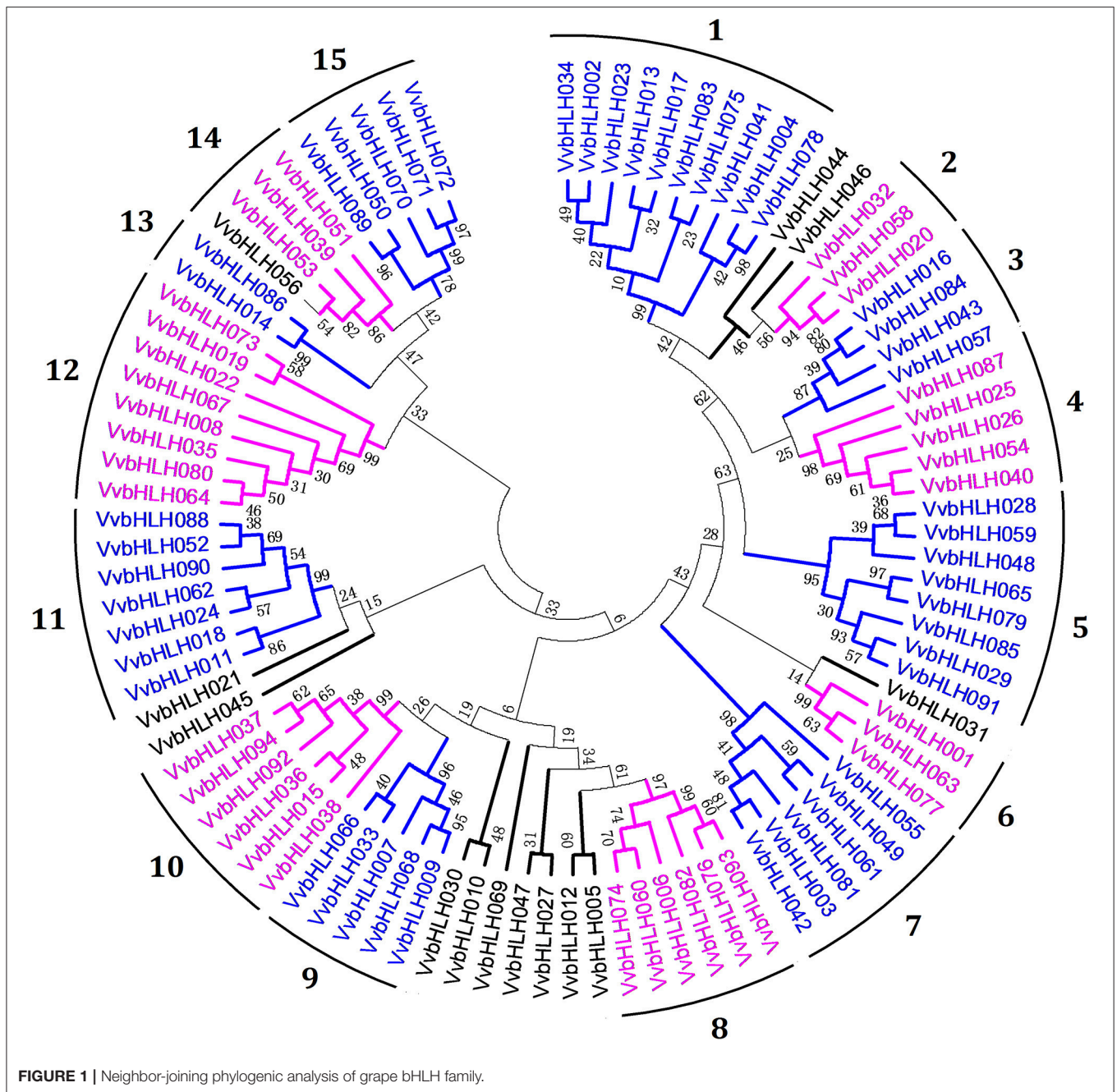
We used the full-length amino acid sequences of the 94 grape bHLHs for a phylogenetic analysis, in which subgroups with relatively high bootstrap support ( $\geq 70$ ) were identified. The phylogenetic tree revealed 15 clades (groups 1–15) in the group bHLH family and 12 orphans (Figure 1). A phylogenetic tree with Arabidopsis showed that 17 grape bHLHs were tightly grouped with the AtbHLHs (bootstrap support  $\geq 70$ ) (Figure S1). They may be orthologous to the 17 AtbHLHs and have similar functions.

GO functional annotation analyses were performed to survey the function of all grape bHLH transcription factors (Table S2). All grape bHLH transcription factors could be divided into three main terms: cellular component, molecular function, and biological process. Our results indicated that many grape bHLHs may be involved in the regulation of transcription, as well as many biological processes such as abiotic stress tolerance (Abe et al., 2003; Lee et al., 2005), organ development (Feller et al., 2011), anthocyanin biosynthesis (Mol et al., 1998; Nesi et al., 2000; Winkel-Shirley, 2001), and flavonol biosynthesis (Walker et al., 1999; Nesi et al., 2000; Baudry et al., 2004).

Our GO functional annotation showed that many grape bHLHs may respond to stimuli, including 10 different types of biotic stressors and 25 abiotic stressors (Table S2) such as response to cold, drought, UV-B, fungus, and bacteria. Many grape bHLHs were involved in secondary metabolite biosynthesis including anthocyanin-containing compound biosynthetic processes, xylan biosynthetic processes, and fatty acid beta-oxidation (Table S2). GO functional annotation also showed that grape bHLHs may be involved in the development of 28 types of tissues or cells (Table S2) and 10 types of cellular components (Table S2).

### Structure Analysis

Structure analyses of all of grape bHLH genes revealed that the number of exons varies from one to 11. Additionally, VvbHLH091, VvbHLH009, VvbHLH033, VvbHLH066, and VvbHLH068 are intron-less. The five intron-less genes are distributed across two clades, particularly clade nine (containing



**FIGURE 1** | Neighbor-joining phylogenetic analysis of grape bHLH family.

four intron-less genes) and clade five (containing the remaining one intron-less gene, VvbHLH091). The structural features of each grape bHLH gene are listed in **Figure S2**, while the others are listed in **Figure S1**. In *Salvia miltiorrhiza*, Zhang et al. (2015) also found intron-less genes mainly (four of six intron-less genes) distributed in one subfamily.

Exons with the same splicing phase at both ends are called symmetric exons, and an excess of symmetric exons and phase 0 introns are likely to facilitate exon shuffling, recombinational fusion, and protein domain exchange (Gilbert,

1987; Patthy, 1987; Zhang et al., 2015). According to the 464 exons analyzed herein, 202 exons are symmetric with phase 0 introns, two exons are symmetric with phase 1 introns, and 11 exons are symmetric with phase 2 introns. Among the 331 introns found in the bHLH genes, 276 are phase 0, 20 are phase 1, and 35 are phase 2. Therefore, our analysis of the bHLH gene structures in grapes strongly indicates a large diversity of bHLH transcription factors, which agrees with findings from *Salvia miltiorrhiza* (Zhang et al., 2015).

## Motif Analysis

A total of 30 conserved motifs were characterized (motifs 1–30; **Table S3**). The grape bHLHs VvbHLH043 (in clade three), VvbHLH057 (in clade three), and VvbHLH021 (orphan) contain only two types of motifs, whereas the grape bHLH VvbHLH027 (orphan) contains the highest number of motifs (10 types). In *Salvia miltiorrhiza*, three types of motifs are repeated twice (Zhang et al., 2015), and in grapes, only motif two is repeated twice in VvbHLH012. Furthermore, certain conserved motifs are nested in specific clades. For example, motif 15 only exists in clade 10, motif eight only exists in clade 14, motif 14 only exists in clade 12, and motif 13 exists in clade one and clade 14. Moreover, every grape bHLH gene contains motif one and motif two. Motif 14 existed in many clades (clades 6, 7, 8, 13, and 15; **Figure S3**).

## Gene Duplication and Triplication Analysis in the Grape bHLH Family

A previous study speculated that many bHLH genes may be generated by gene duplication in some plants (Sun et al., 2015). Therefore, we sought to detect duplicated gene pairs in the grape bHLH gene family. Forty-one duplicated gene pairs were found in grape genomes, including 26 high-stringency standard duplicated gene pairs and 15 low-stringency standard duplicated gene pairs (**Table S4**). Intraspecies synteny analysis showed that the many duplicated gene pair blocks were collinear, such as VvbHLH049–VvbHLH061 and VvbHLH058–VvbHLH032, which indicated that these gene duplications may derive from chromosome segmental duplication or a large-scale duplication event. Additionally, VvbHLH053–VvbHLH056 may be derived from a tandem duplication because they are located on the same chromosome.

We identified nine triplicated gene groups in the grape bHLH gene family, such as the VvbHLH058–VvbHLH032–VvbHLH020 group. Additionally, VvbHLH058, VvbHLH032, and VvbHLH020 were collinear, showing that the triplication may derive from chromosome segmental triplication or a large-scale triplication event.

All the triplicated gene groups in the grape bHLH family underwent purifying selection (**Table 1**). Purifying selection may generate genes with conserved functions or pseudogenization (Zhang, 2003). Based upon our Nr and GO function annotations and motif analysis, we found that the genes may be functionally conserved within a triplicated gene group such as the VvbHLH093–VvbHLH076–VvbHLH082 group (ICE), VvbHLH060–VvbHLH074–VvbHLH006 (ICE), VvbHLH007–VvbHLH068–VvbHLH009 (myc). The motifs of one bHLH family member are also similar to the other members in the same triplicated gene group (**Figure S3**).

## Expression Profile of Grape bHLH Genes

The expression profiles of 94 grape bHLHs in 49 tissues (tissue name details are listed in **Table S5**) were analyzed based on the microarray expression profile data collected from GSE36128 on the NCBI database (<https://www.ncbi.nlm.nih.gov/gquery/?term=GSE36128>) (Fasoli et al., 2012). We found all 94 grape bHLHs expressed in the 49 tissues (**Figure 4**). The log<sub>2</sub> values

(RMA-normalized signal intensities value of grape bHLH genes) were used to represent the expression level. The log<sub>2</sub> values vary from four to 15. Some genes were found to have a log<sub>2</sub> value of more than 14 in a certain tissue; therefore, they were considered to be highly expressed in that tissue. VvbHLH007 (GO function annotation: response to abscisic acid stimulus, chitin, and insects; jasmonic acid mediated signaling pathway) was highly expressed in stamens, post-fruit set berry pericarp, latent bud, mid-ripening berry flesh, ripening berry flesh, post-harvest withering I (1st month) berry flesh, post-harvest withering II (2nd month) berry flesh, post-harvest withering III (3rd month) berry flesh, carpels, petals, post-harvest withering II (2nd month) berry pericarp, fruit set rachis, post-fruit set rachis, véraison rachis, mid-ripening rachis, ripening rachis, post-harvest withering II (2nd month) berry skin, green stem, well developed tendril, and mature tendril (**Figure 4**; **Table S4**). A list of tissue abbreviations can be found in **Table S4** (Fasoli et al., 2012). VvbHLH005 (GO function annotation: pollen development; anther wall tapetum cell differentiation) was only highly expressed in stamens (**Figure 4**; **Table S4**), indicating that VvbHLH005 may play an important role in pollen development in grape stamens. VvbHLH048 (Nr annotation: pif4; GO function annotation: red or far-red light signaling pathway; regulation of biological quality) was

**TABLE 1** | Selection pressure of each grape bHLH duplication gene pairs in triplicate gene-group.

Each Gene Pairs in Triplicate Gene-Group	Non-Synonymous/Synonymous
<b>GROUP 1</b>	
VvbHLH025- VvbHLH040	ka/Ks = 0.1814
VvbHLH054- VvbHLH040	ka/Ks = 0.2561
<b>GROUP 2</b>	
VvbHLH058- VvbHLH020	ka/Ks = 0.1500
VvbHLH032- VvbHLH020	ka/Ks = 0.1488
<b>GROUP 3</b>	
VvbHLH063- VvbHLH001	ka/Ks = 0.1958
VvbHLH077- VvbHLH001	ka/Ks = 0.1533
<b>GROUP 4</b>	
VvbHLH093- VvbHLH082	ka/Ks = 0.2203
VvbHLH076- VvbHLH082	ka/Ks = 0.2293
<b>GROUP 5</b>	
VvbHLH060- VvbHLH006	ka/Ks = 0.1974
VvbHLH074- VvbHLH006	ka/Ks = 0.2618
<b>GROUP 6</b>	
VvbHLH068- VvbHLH007	ka/Ks = 0.0497
VvbHLH009- VvbHLH007	ka/Ks = 0.0113
<b>GROUP 7</b>	
VvbHLH088- VvbHLH090	ka/Ks = 0.2519
VvbHLH052- VvbHLH090	ka/Ks = 0.2819
<b>GROUP 8</b>	
VvbHLH053- VvbHLH039	ka/Ks = 0.2762
VvbHLH056- VvbHLH039	ka/Ks = 0.4034
<b>GROUP 9</b>	
VvbHLH072- VvbHLH070	ka/Ks = 0.1978
VvbHLH071- VvbHLH070	ka/Ks = 0.2605

highly expressed in post-fruit set berry skin, véraison berry skin, mid-ripening berry skin, young leaf, mature leaf, fruit set berry pericarp, post-fruit set berry pericarp, véraison berry pericarp, and post-fruit set berry flesh, showing that it may be involved in quantitative berry and leaf traits. VvbHLH075 was highly expressed in mid-ripening berry flesh and ripening berry flesh. VvbHLH088 was highly expressed in ripening berry flesh. VvbHLH048 was highly expressed in mid-ripening berry flesh and mid-ripening berry skin (**Figure 2; Table S5**). This indicates that they may play an important role in fruit ripening.

## Grape bHLH Gene Codon Usage Bias Analysis

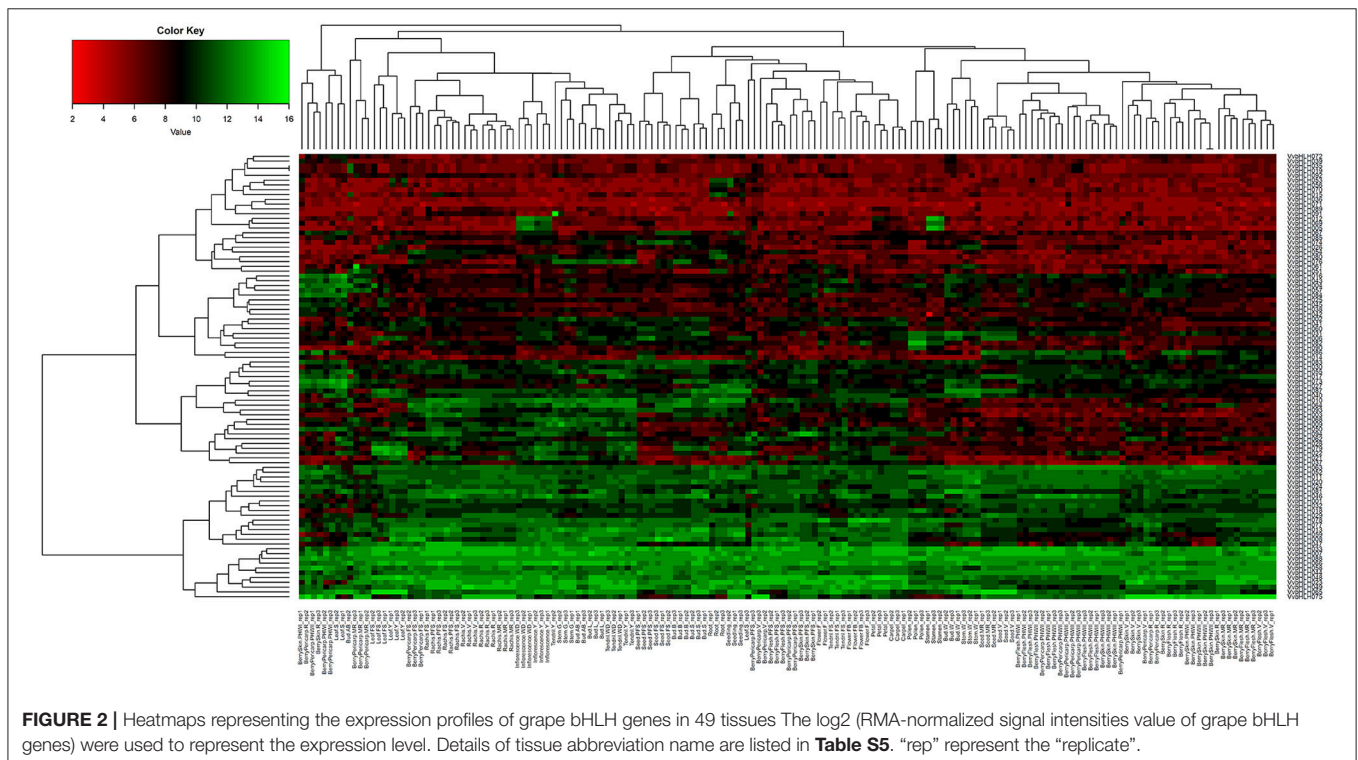
To understand whether grape bHLH genes have codon bias, the frequency of optimal codons (FOP), GC content, GC content on the 3rd site of a synonymous codon (GC3s content), the relative synonymous codon usage (RSCU), codon adaptation index (CAI), and codon bias index (CBI) were calculated. The average FOP was  $0.397892 \pm 0.040998$ , the average CAI was  $0.197194 \pm 0.023114$ , and the average CBI was  $-0.04456 \pm 0.071419$ . The GC content for the 94 grape bHLH genes varied from 41.5 to 56.5% and the average GC content was 47.1%. The GC3s content for grape bHLH genes varied from 31.6 to 63.4% and the average was 45.29%. To further understand the influences of codon bias on gene properties, a correlation analysis was performed. GC content was positively correlated with CBI and FOP ( $P < 0.01$ ), but is not strongly correlated with CAI; GC3s content is positively correlated with CBI and FOP ( $P < 0.01$ ), but is not correlated with CAI ( $P < 0.01$ ). Gene length and exon

length are not strongly correlated with CBI, CAI, or FOP (**Table 2**).

The ENC in grape bHLH genes ranged from 30.28 to 60.96, with an average of 53.25. Among the 94 grape bHLH genes, only one gene (VvbHLH091) exhibited high codon bias ( $ENC < 35$ ), indicating that in general, grape bHLH genes reflect random codon usage without strong codon bias (Li et al., 2016). Codon usage bias could be affected by many stressors including mutation pressure and natural selection (Li, 1987). If codon usage bias is corrected with GC content, we could deduce that the codon usage bias was affected by mutation pressure during its evolutionary history (Knight et al., 2001; Chen et al., 2004). The FOP and CBI of grape bHLH genes are positively correlated with GC and GC3s content, indicating that grape bHLH genes mainly evolved by mutation pressure.

An ENC plot was generated to explore the influence of GC3s on codon bias in grape bHLH genes. If a gene is located on the expected curve, the codons of that gene are not biased (Zhang et al., 2007). Here, all ENC values were lower or higher than expected and were located above or below the curve (**Figure 3**), indicating that other factors combined with mutation pressure affect codons. The distribution width of the GC3s might be related to the variation in the strength of directional selection against mutation pressure (Kawabe and Miyashita, 2003). In grape bHLH genes, the distribution of GC3s was between 0.314 and 0.651, indicating that grape bHLH genes evolved by mutation pressure.

In grape bHLH genes, many codons with larger RSCU values had larger corresponding tRNA copy numbers, such as the synonymous codons in glu, gln, cys, thr, val, leu, pro, ser, and lys (**Figure 4**).



**TABLE 2** | Correlation analysis.

	FOP	CBI	CAI
GC content	0.5*	0.5*	0.2*
GC3s content	0.6*	0.6*	0.2*
Gene length	0	0	0
Exon length	0	0.1*	0.1*

\* $p < 0.01$ .

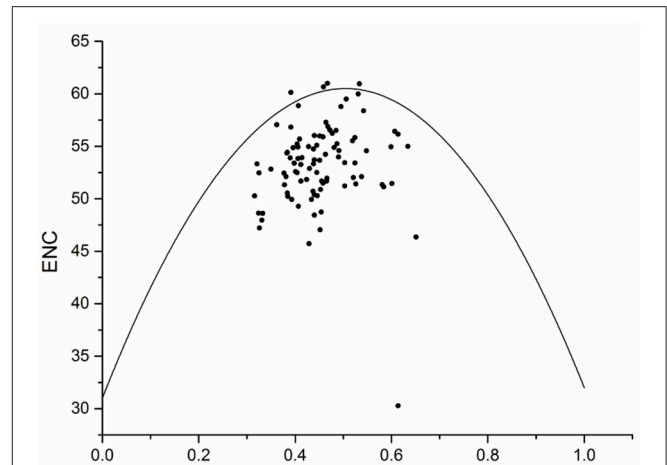
## The Expression Level of Grape bHLH Genes in Abiotic Stress Tolerance

Based on the microarray expression profile data of the PEG treatment library and the control library (GSE31594; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31594>), we found three bHLH genes were up-regulated 1.5-fold under a 1 h treatment with PEG. Four bHLH genes were up-regulated 1.5-fold under a 4 h PEG treatment. Sixteen bHLH genes were up-regulated 1.5-fold under an 8 h treatment with PEG (Figure S4; Table S6).

In total, 22 grape bHLH genes could be induced by PEG treatment. Among these 22 genes, 12 promoters contained ABRE elements (ABA responsive), 17 contained MBS or MER elements (Myb transcription factor binding; drought inducible), 15 contained HSE elements (Hsf transcript factor binding), and 17 contained G-box elements (bHLH transcript factor binding; Table S7).

Based on the microarray expression profile data of the cold stress library and control library (GSE31594; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31594>), we found seven bHLH genes were up-regulated 1.5-fold under 1 h of cold treatment (Figure 5; Table 3). Five bHLH genes were up-regulated 1.5-fold under 4 h of cold treatment. Thirteen bHLH genes were up-regulated 1.5-fold under 8 h of cold treatment (Figure 5; Table 3). In total, 17 grape bHLH genes could be induced by cold stress treatment including a homolog of MYC2 VvbHLH007 (Figure 5; Table 3). Among these 17 genes, four promoters contained ABRE elements, four contained LTR elements (low temperature responsive), 10 contained MBS or MER elements, 10 contained HSE elements, and 11 contained G-box elements (Table S7).

VvbHLH039 (org2), VvbHLH021 (bHLH92), and VvbHLH007 (MYC2) could be significantly induced by both PEG and cold treatment. The promoter of VvbHLH039 contained G-box elements, VvbHLH021 contained G-Box, LTR, and MBS elements, and VvbHLH007 contained ABRE, G-Box, HSE, and MBS elements. MBS elements could be responsive to drought and ABRE could be responsive to ABA. ABA-dependent pathways play a crucial role in the response of a plant to abiotic (drought, salinity, cold, and hypoxia) and biotic stressors. ABA-dependent transcription factors could be involved in cold, drought, and salinity stress gene expression such as MYC (Huang et al., 2012). The promoter of VvbHLH007 contained ABRE elements, so VvbHLH007 may be affected by ABA-dependent pathways and could be up-regulated by cold and drought stress. VvbHLH021 could be up-regulated both by PEG and cold stress, which may be due to

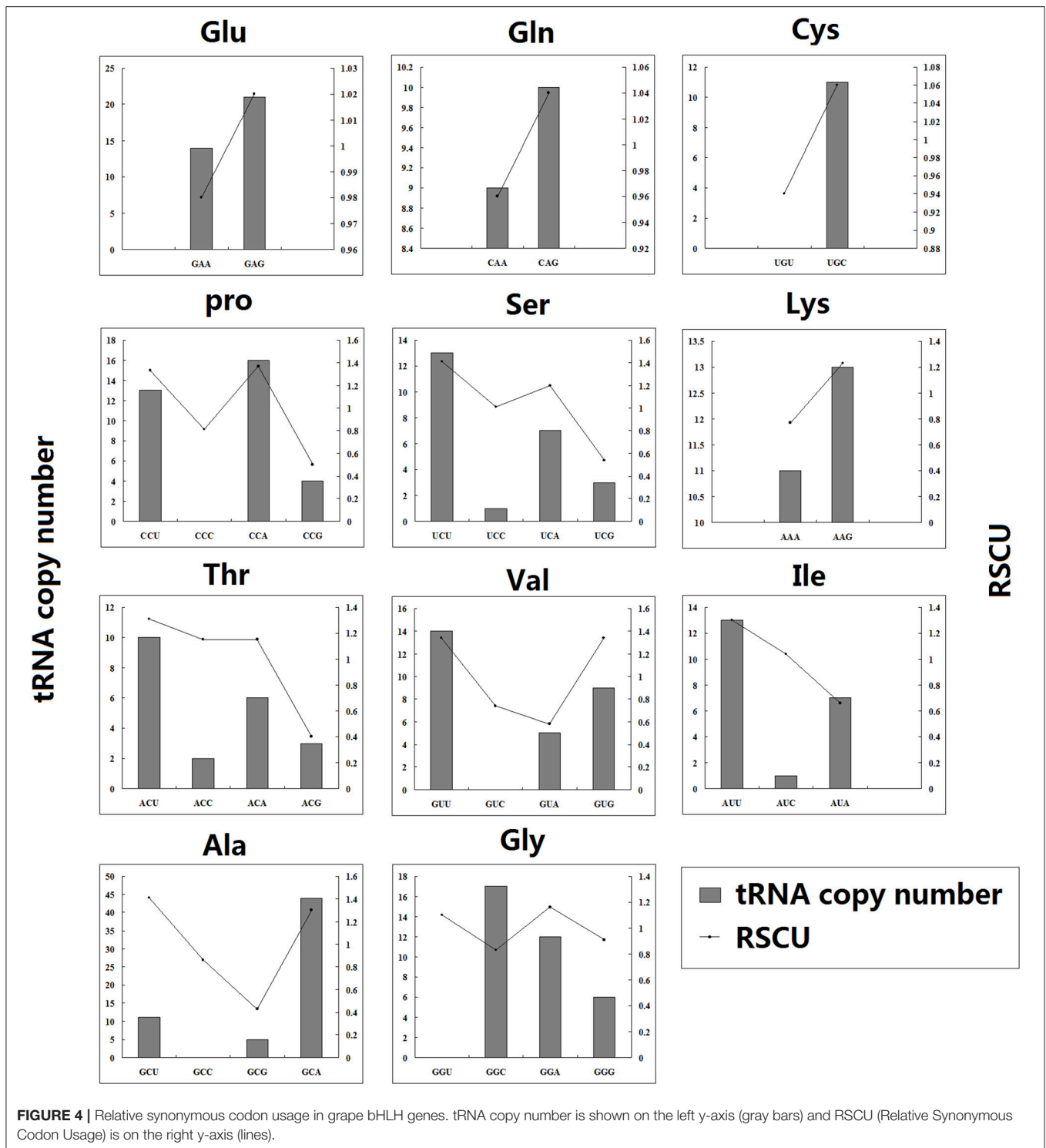
**FIGURE 3** | Effective number of codons ENC-plot showing relationship between ENC and GC3s.

the promoter of VvbHLH021 containing LTR elements and MBS elements. VvbHLH039 contained G-box elements, so it may be regulated by other grape bHLHs under cold or PEG treatments (Table S7).

Overexpression of CBF4 in Arabidopsis resulted in constitutive expressions of CRT/DRE containing stress responsive genes and enhanced tolerance to drought and freezing stressors (Haake et al., 2002; Chinnusamy et al., 2003). CBF could recognize the cold and dehydration responsive element CRT/DRE (Yang et al., 2005). Based on the microarray expression profile data of the CBF4 over-expression transgenic grape library (GSE29948; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29948>), we found seven bHLH genes that were differentially expressed between a CBF4 over-expression transgenic grape compared with the control (Figure S5; Table S8). The promoter of VvbHLH037 and VvbHLH073 contained the CRT/DRE (A/GCCGAC) element (Table S7).

## The Role of Grape bHLHs in Flavonoid and Anthocyanin Biosynthesis Pathways

Previous studies have identified that the grape bHLH genes VvMYCA1 (ABM92332; VIT\_215s0046g025601; Matus et al., 2010) and VvMYC1 (EU447172; VIT\_207s0104g000902; Nr annotation in this study: TT8; Hichri et al., 2010) are involved in flavonol biosynthesis. Based on our GO or Nr function annotations, we found three additional genes that are potentially related to anthocyanin or flavonol biosynthesis including VvbHLH003 (GO annotation: anthocyanin-containing compound biosynthetic process; Nr annotation: transcription factor bhlh30), VvbHLH007 (GO annotation: positive regulation of flavonoid biosynthetic process; Nr annotation: MYC2), and VvbHLH010 (Nr annotation: myc-like anthocyanin regulatory protein; Table S2). Two of these three bHLH genes also have other functions: VvbHLH003 also functions in cellular response to phosphate starvation and VvbHLH007 functions in a jasmonic acid mediated signaling pathway, regulation of defense response

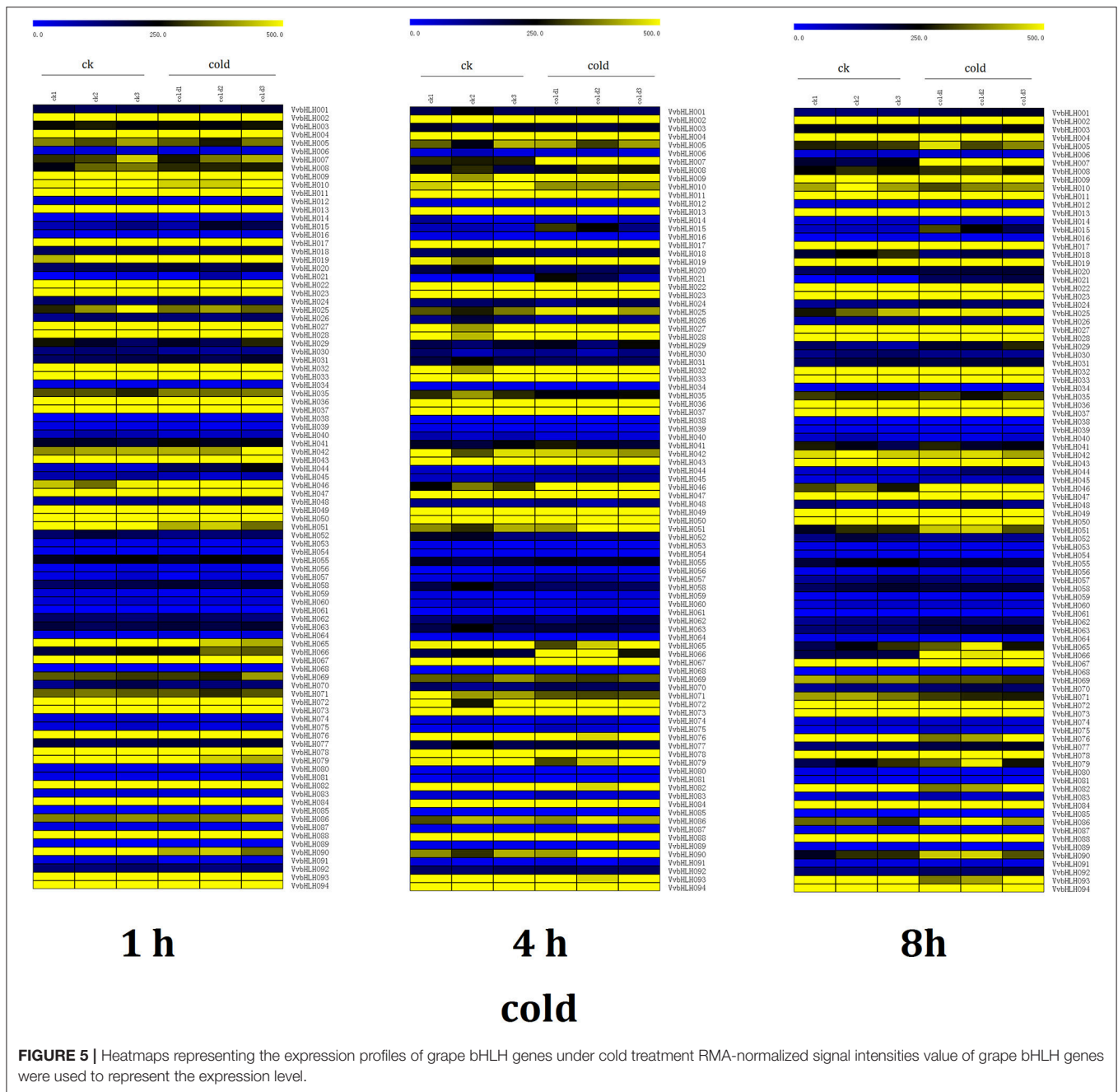


to insects, response to chitin, response to abscisic acid stimuli, and response to desiccation. Therefore, these bHLHs genes are also involved in response to abiotic or biotic stimuli.

In our study, the promoters of the three bHLH genes contained HSE elements (Table S9), suggesting that they could be regulated by other Hsfs. Their promoters also contained

MYB binding sites (MBS/MER) which are involved in drought response and G-box elements (Table S9), indicating that they could be regulated by MYB and bHLH transcription factors. The promoters of VvbHLH003 and VvbHLH007 contained ABRE elements, which are involved in ABA-dependent or independent stress tolerance (Chen et al., 2012).





## The Majority of Promoters of Genes Involved in Flavonoid and Anthocyanin Biosynthetic Pathways Present G-Box/E-Box Elements

Recently, Malacarne et al. (2015) and Costantini et al. (2015) identified genes with a role in the regulation of anthocyanin and flavonol content and composition including 64 genes involved in flavonoid biosynthesis, such as Myb domain protein 12 VIT\_07s0005g01210 (which is involved in the regulation of flavonol biosynthesis), VIT\_11s0052g01600 (UDP-glucose flavonoid 3-O-glucosyltransferase), VIT\_11s0052g01630

(UDP-glucose flavonoid 3-O-glucosyltransferase), and VIT\_16s0098g00860 (flavanone 3-hydroxylase). 60 of these gene promoters contained G-box or E-box elements (Table S8). Forty-four genes are involved in anthocyanin biosynthesis such as VIT\_02s0033g00370 (VvMYBA4), VIT\_02s0033g00380 (VvMYBA2), VIT\_02s0033g00390 (VvMYBA2), VIT\_02s0033g00410 (VvMYBA1), VIT\_02s0033g00430, VIT\_02s0033g00440, VIT\_02s0033g00450 (VvMYBA3), VIT\_08s0007g03560 (Anthocyanin membrane protein 1), and VIT\_06s0009g02830 (flavonoid 3', 5'-hydroxylase). 40 of these gene promoters contained G-box or E-box elements (Table S10).

**TABLE 3** | Difference expression gene between ck and cold treatment.

Gene name	Value of ck			ck'	Value of cold treatment			Cold'	Fold	P-value
	ck1	ck2	ck3		cold1	cold2	cold3			
<b>DEG BETWEEN CK AND COLD TREATMENT 1H</b>										
VvbHLH021	2.10	14.28	26.61	14.33	11.89	86.70	29.95	42.85	2.99	0.16
VvbHLH039	0.73	3.78	8.47	4.33	12.17	15.49	24.80	17.49	4.04	0.01
VvbHLH045	62.18	42.87	44.08	49.71	162.72	176.20	239.47	192.80	3.88	0.02
VvbHLH047	445.19	361.14	489.53	431.95	618.86	1,104.60	831.11	851.52	1.97	0.07
VvbHLH061	1.27	1.11	2.64	1.67	0.69	1.57	5.86	2.71	1.62	0.23
VvbHLH066	187.19	192.55	228.94	202.89	214.10	361.45	342.72	306.09	1.51	0.06
VvbHLH068	1.93	2.05	4.95	2.98	9.26	2.56	16.76	9.53	3.20	0.09
<b>DEG BETWEEN CK AND COLD TREATMENT 4H</b>										
VvbHLH021	9.10	17.13	8.83	11.69	256.28	174.36	66.27	165.64	7.51	0.07
VvbHLH045	55.38	18.50	49.49	41.13	75.06	89.14	92.29	85.50	1.86	0.16
VvbHLH007	298.12	276.57	280.45	285.05	1,175.30	925.11	494.27	864.89	1.76	0.05
VvbHLH047	249.06	375.78	305.60	310.15	1,369.50	950.03	649.32	989.62	3.19	0.05
VvbHLH015	67.87	71.65	47.21	62.24	305.80	252.29	119.51	225.87	2.53	0.04
<b>DEG BETWEEN CK AND COLD TREATMENT 1H</b>										
VvbHLH021	3.91	3.85	1.39	3.05	158.36	155.58	181.99	165.31	54.16	0.00
VvbHLH083	32.09	59.85	32.27	41.40	80.60	45.14	66.48	64.07	1.55	0.18
VvbHLH046	28.12	31.90	48.93	36.32	72.78	60.65	70.58	68.00	1.87	0.02
VvbHLH090	217.47	296.32	293.81	269.20	450.02	458.91	332.50	413.81	1.54	0.06
VvbHLH052	217.47	296.32	293.81	269.20	450.02	458.91	332.50	413.81	1.54	0.06
VvbHLH047	321.52	387.41	271.73	326.89	1,086.20	857.72	1,041.80	995.24	3.04	0.01
VvbHLH065	167.85	238.48	299.32	235.22	346.41	490.64	268.64	368.56	1.57	0.13
VvbHLH079	167.85	238.48	299.32	235.22	346.41	490.64	268.64	368.56	1.57	0.13
VvbHLH066	171.62	177.90	179.35	176.29	561.44	480.46	539.36	527.09	2.99	0.00
VvbHLH068	3.14	8.76	9.52	7.14	10.97	9.32	11.98	10.76	1.51	0.12
VvbHLH015	58.75	65.04	58.58	60.79	323.91	251.12	170.73	248.59	4.09	0.03
VvbHLH011	1,558.90	1,958.30	1,886.10	1,801.10	3,256.80	2,740.20	3,764.40	3,253.80	1.81	0.03
VvbHLH029	155.24	117.41	84.02	118.89	205.26	203.01	288.18	232.15	1.95	0.07

ck', average value of gene expression in ck; cold', average value of gene expression in cold treatment grape; DEG, different expression gene; fold, cold'/ck'.

## DISCUSSION

### Expansion of the Grape bHLH Gene Family May be Drive by Gene Duplication

The basic helix-loop-helix (bHLH) transcription factor family is very large in eukaryotes (Ledent and Vervoort, 2001; Toledo-Ortiz et al., 2003) and many bHLH genes (100–170) have been identified in land plants (Carretero-Paulet et al., 2010). For example, 162 bHLH genes have been identified in *Arabidopsis* (Bailey et al., 2003), 167 bHLH genes in rice (Li et al., 2006), and 152 bHLH genes in tomato (Wang et al., 2015). In our study, we identified 94 bHLH genes in *V. vinifera*. The gene number of the grape bHLH gene family is large. Previous studies in tomato have shown that 103 genes were generated by gene duplication (Sun et al., 2015), indicating that the expansion of the tomato bHLH gene family is mainly driven by duplication. In *V. vinifera*, we found 31 duplicated gene pairs, showing that the expansion of the grape bHLH

gene family may also be driven by duplication. A collinearity analysis showed that many duplicated gene-pair blocks were collinear and only one tandem duplication was found. Therefore, gene duplications of the grape bHLH gene family may mainly derive from chromosome segmental duplication or a large-scale duplication event. Both A whole genome replication event ( $\gamma$  event) and whole genome duplication event (WGD) could lead to chromosome segmental duplication or a large-scale duplication event. *Vitis vinifera* is thought to have undergone an ancient whole genome replication event ( $\gamma$  event), but not a WGD event (Jaillon et al., 2007). Therefore, the duplication in the grape bHLH gene family may mainly derive from the  $\gamma$  event. The  $\gamma$  event could have led to the triplicated arrangement in plant genomes and the generation of the triplicated gene group (Jaillon et al., 2007). In *V. vinifera*, we identified triplicated bHLH gene groups. Triplicated gene groups can be also found in many plants such as in *Arabidopsis* (Guo et al., 2017).

Both rice and *Arabidopsis* have undergone  $\gamma$  events and at least two WGD events, and grapes are believed to have only undergone a  $\gamma$  event (Jaillon et al., 2007). This may be the reason that grape contained fewer bHLH genes (94) compared with *Arabidopsis* (162) and rice (167).

## The Functional Divergence in the bHLH Family and Conserved Functions in Triplicated Gene Groups

The phylogenetic analysis showed that the grape bHLH gene family contained 15 clades (Figure 1). A previous study has shown that the *Arabidopsis* bHLH gene family contained 15 clades and some orphans (Feller et al., 2011), rice contained 22 clades (Li et al., 2006), and tomato contained 26 clades (Wang et al., 2015). The phylogenetic analysis showed that 17 grape bHLHs were tightly grouped with the *AtbHLHs* (Figure S1). A previous study showed that 43 tomato bHLHs were tightly grouped with the *AtbHLHs*. This indicates that the number of orthologous pairs between grape and *Arabidopsis* or between tomato and *Arabidopsis* is not large compared with the number of the bHLH gene family members. Different numbers of clades in different species and few orthologous pairs between different species showed that interspecific divergence existed in plant bHLH gene families.

Our Nr-annotation showed that the grape bHLH family contained some homologous bHLHs previously reported in *Arabidopsis*. However, the annotation did not contain the homologs of many bHLHs in *Arabidopsis*. For example, the grape bHLH family does not contain the homologs of FIT (involved in the regulation of iron uptake), AIB (involved in ABA signaling), or RHD6 (involved in root hair formation) in *Arabidopsis*. This supports the view that interspecific functional divergence exists in plant bHLH gene families.

In the grape bHLH gene family, there are a total of 464 exons; 215 exons are symmetric. Among the 331 introns found in the bHLH genes, 276 are phase 0. An excess of symmetric exons and phase 0 introns are likely to facilitate exon shuffling, recombinational fusion, and protein domain exchange (Gilbert, 1987; Patthy, 1987; Zhang et al., 2015). Therefore, our analysis of the bHLH gene structures in grapes strongly indicates a large diversity of bHLH transcription factors, which was similarly found in *Salvia miltiorrhiza* (Zhang et al., 2015). Our GO annotation and Nr annotation of the grape bHLH family showed that the grape bHLH family may be involved in many functions related to plant development and other biological processes (Tables S1, S2). Different grape bHLH genes had different expression modes in grape, showing that they may play roles in different tissues or biological processes too (Figures 2, 5, Figures S4, S5). Intraspecific functional divergence exists in the grape bHLH family and the functional divergence may be related to the structures of bHLH genes. In other species, such as *Arabidopsis* (Feller et al., 2011) and rice (Li et al., 2006), many different functions were identified in their respective bHLH families. This is consistent with our findings for *V. vinifera*.

All of the triplicate gene groups in the grape bHLH family underwent purifying selection (Table 1). Purifying selection may

generate genes with conserved functions or pseudogenization (Zhang, 2003). Based upon our Nr and GO function annotations and motif analysis, we found that the genes may be functionally conserved within a triplicated gene group. The motifs of one bHLH family member are also similar to the other members in the same triplicated gene group (Figure S3), indicating that function divergence was weak within triplicated bHLH gene groups in grape.

## RSCU Values of Codons Were Correlated to tRNA Copy Numbers in Grape

In grape bHLH genes, many codons with larger RSCU values had larger corresponding tRNA copy numbers in grape, such as the synonymous codons in glu, gln, cys, thr, val, leu, pro, ser, and lys (Figure 4). The translation efficiency constrains codon choice; therefore, the frequency of codon usage is positively correlated with tRNA availability (Ikemura, 1981). The above result showed that the codon usage bias of some grape bHLH genes may be related to the translation efficiency.

## Grape bHLH Genes May Play Important Roles in Abiotic Stress Tolerance

In our study, we found 22 grape bHLH genes could be induced by PEG treatment and 17 grape bHLH genes could be induced by cold stress treatment. A homolog of bHLH92 (VvbHLH021) and a homolog of the MYC protein (VvbHLH007) could be significantly induced by both PEG and cold treatment. A previous study showed bHLH92 were involved in abiotic stress tolerance (Jiang et al., 2009). MYC proteins are known to function as activators in one of the ABA-dependent regulatory systems (Abe et al., 2003; Valliyodan and Nguyen, 2006). Over-expression of *AtMYC2* resulted in an ABA-hypersensitive phenotype and improved osmotic stress tolerance of transgenic plants (Shinozaki and Yamaguchi-Shinozaki, 2007; Gao et al., 2011). In the grape bHLH family, we found three MYC genes, but only VvbHLH007 (MYC2) could be significantly induced by PEG or cold treatment.

The CBF cold-responsive pathway is related to cold acclimation and plays a key role in the CBF family (Thomashow, 2001). ICEs, a bHLH member that induces the expression of CBF, could induce the CBF family when temperatures are cold (Zarka et al., 2003). ICE1 can bind to MYC recognition elements in the CBF3 promoter and is important for the expression of CBF3 during cold acclimation (Chinnusamy et al., 2003). ICE2 participates in the response to deep freezing through the cold acclimation-dependent pathway in *Arabidopsis thaliana* (Gagne et al., 2002; Smalle and Vierstra, 2004). In grape, we found six ICE genes, but none were up-regulated under the PEG or cold treatments.

In our study, most of the PEG or cold induced grape bHLH gene promoters contained HSE elements (Table S7), suggesting these genes could be regulated by Hsfs. Most grape bHLH gene promoters contained MYB binding sites, which are involved in drought response (Table S7). This indicated that they could be regulated by MYB transcription factors under drought stress. Most grape bHLH gene promoters contained G-box elements, indicating that they may be regulated by other grape bHLHs.

Many grape bHLH gene promoters contained ABRE and DRE elements, which are involved in ABA-dependent or independent stress tolerance (Chen et al., 2012). Therefore, these genes may play important roles in gene regulation in response to different stresses in grape.

Additionally, we found seven bHLH genes that were differentially expressed between a CBF4 over-expression transgenic grape compared with the control (**Figure S5**; **Table S7**). The promoter of VvbHLH037 and VvbHLH073 contained the CRT/DRE (A/GCCGAC) element (**Table S7**), indicating they may be regulated by the CBF4 protein.

The above results indicated that some grape bHLH family members may be regulated by the transcription factors CBF, Myb, Hsf, or other bHLHs, which could be regulated by other transcript factors or induced by abiotic stress. Some transcriptome data showed that the expressions of transcript factor gene families will change in response to abiotic stressors (Rienth et al., 2014, 2016; Rocheta et al., 2014). Transcriptome data from the grape cv. Trincadeira showed the MYB4-like gene could be up-regulated under water deficit (Rocheta et al., 2016). Hsf could also respond to cold, salinity, and drought stress (Kotak et al., 2007; Swindell et al., 2007; Hu et al., 2015; Wang et al., 2017). Therefore, Myb, Hsf, bHLH, CBF, and other transcript factors may interact creating a coordinated network, and play roles in abiotic stress response. Some bHLHs were found to be related to abiotic stress response in this study and may play roles in abiotic stress response and tolerance.

## Grape bHLH Genes May Play Important Roles in Flavonoid and Anthocyanin Biosynthesis Pathways

Flavonols are an important class of bioactive compounds in almost all higher plants such as grapes. They are most abundant in the flowers and skins of grape berries (Hmamouchi et al., 1996), thus are relevant to viticulture and enology. Flavonols, including anthocyanins and condensed tannins, are produced by the flavonoid biosynthetic pathway, and they represent one of the most important classes among flavonoids in terms of concentration (Flamini et al., 2013). Anthocyanins are the largest class of flavonoids (Welch et al., 2008), and constitute one of the most important families of secondary metabolites. In response to UV light, cold, drought, and other abiotic stressors, anthocyanins have been shown to be synthesized as protective compounds, and can also function to attract pollinators (Costantini et al., 2015).

Here, we found 60 of 64 genes in flavonoid and 40 of 44 genes in anthocyanin biosynthesis pathways had promoters that contain G-box or E-box elements. bHLH family members could bind to E-box or G-box cis elements and regulate gene expression (Atchley et al., 1999; Massari and Murre, 2000; Toledo-Ortiz et al., 2003; Li et al., 2006). Therefore, most of these genes may be regulated by grape bHLH family members. Previous studies have shown that MYB and WD40 transcription factors could interact with bHLHs to be involved with the biosynthesis of anthocyanins (Mol et al., 1998; Nesi et al., 2000; Winkel-Shirley, 2001) and flavonols (Walker et al., 1999; Nesi et al., 2000; Baudry et al., 2004). Additionally, G-box is also a cis-acting regulatory element

involved in light responsiveness (Giuliano et al., 1988). Light could affect the flavonoid and anthocyanin biosynthesis and the expression of related genes (Azuma et al., 2012; Costantini et al., 2015). The G-box element may be the connector between light and flavonoid and anthocyanin biosynthesis.

We found that some bHLH genes were highly expressed in mid-ripening berry flesh, ripening berry flesh or mid-ripening berry skin (VvbHLH075, VvbHLH088, and VvbHLH048; **Figure 4**). Therefore, they may play an important role in fruit ripening and in flavonoid and anthocyanin biosynthesis pathways. Additionally, we found three genes that are potentially related to anthocyanin or flavonol biosynthesis based on GO and Nr function annotation. These three genes were expressed moderately in mid-ripening berry flesh, ripening berry flesh or mid-ripening berry skin.

The promoters of VvbHLH075, VvbHLH088, VvbHLH048, and three potentially related genes (VvbHLH003, VvbHLH007, and VvbHLH10) contained cis-acting elements involved in gibberellin-responsiveness (TATC-box/P-box), HSE, MYB binding sites (MBS/MER), and G-box elements. This showed that they may be involved in the gibberellin signaling pathway, and could be regulated by MYBs, Hsfs, and other bHLHs transcription factors. The promoters of VvbHLH003 and VvbHLH007 contained ABRE element which are involved in ABA-dependent or independent stress tolerance (Chen et al., 2012; **Table S9**). Therefore, these genes could also play important roles in response to different stresses in grape. The three genes that are potentially related to anthocyanin or flavonol biosynthesis are also predicted to have other functions involved in response to abiotic stressors based on the GO and Nr function annotation. They may not only be related to anthocyanin or flavonol biosynthesis but also to stressor/stimuli such as VvbHLH007 (MYC2). A previous study has shown that abiotic stressors (such as cold or heat stress) may regulate anthocyanin or flavonoid biosynthesis-related genes (Azuma et al., 2012; Rienth et al., 2014). It is possible that these biotic or abiotic stressors are related to the anthocyanin or flavonol biosynthesis in *V. vinifera*, and we hypothesize that bHLH members may be a connector between stressor and secondary metabolite biosynthesis in grape. Future research should seek to understand the relationships between the anthocyanin and/or flavonoid biosynthesis pathways and abiotic or other stimuli. Further studies on anthocyanin and flavonoid biosynthesis pathways and their stimuli will positively impact viticulture and breeding.

## CONCLUSIONS

Genome-wide identification and comparison of grape bHLHs with other plant species revealed that grape bHLH gene family is a large family. Chromosome segmental duplication or a large-scale duplication event played important roles in the expansion of grape bHLH gene family. Our study provided evidence for the roles of grape bHLH family members in many biological processes including abiotic stress tolerance and secondary metabolite biosynthesis. These results are useful for better understanding the complexity of the gene family and the

relation between stressor and secondary metabolite biosynthesis in grape.

## AUTHOR CONTRIBUTIONS

PW, YW, and FR designed the study, PW wrote the manuscript. YL and QZ finalized the figures and tables. LS, HG, XJ and XW carried out most of the experiment, data analysis, and wrote the method section of the manuscript.

## ACKNOWLEDGMENTS

This work was supported by the Agricultural Scientific and Technological Innovation Project of Shandong Academy of Agricultural Sciences (CXGC2016D01), Agricultural Scientific and Technological Innovation Project of Shandong Academy of Agricultural Sciences-Cultivating Project for National Natural Science Foundation of China in 2018 “Identification and Function Research of *Vitis vinifera* and *Vitis amurensis* Cold Stress Response-Related microRNAs,” Major Agricultural Application Technology Innovation Project of Shandong Province “Research and Application of Precision Control of Maturation and Product Innovation of Featured Brewing Grape,” Major Agricultural Application Technology Innovation Project of Shandong Province “Development of Landmark Wines and Integrated Application of Key Technologies in Shandong Province,” and Fruit Innovation Team of Modern Agricultural Industry Technology System in Shandong Province-Jinan Comprehensive Test Station (SDAIT-06-21).

## REFERENCES

- Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell* 15, 63–78. doi: 10.1105/tpc.006130
- Atchley, W. R., Terhalle, W., and Dress, A. (1999). Positional dependence, cliques, and predictive motifs in the bHLH protein domain. *J. Mol. Evol.* 48, 501–516. doi: 10.1007/PL00006494
- Azuma, A., Yakushiji, H., Koshita, Y., and Kobayashi, S. (2012). Flavonoid biosynthesis-related genes in grape skin are differentially regulated by temperature and light conditions. *Planta* 236, 1067–1080. doi: 10.1007/s00425-012-1650-x
- Bailey, P. C., Martin, C., Toledo-Ortiz, G., Quail, P. H., Huq, E., Heim, M. A., et al. (2003). Update on the basic helix-loop-helix transcription factor gene family in *Arabidopsis thaliana*. *Plant Cell* 15, 2497–2501. doi: 10.1105/tpc.151140
- Bailey, T. L., Williams, N., Mistleh, C., and Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34(Suppl. 2), W369–W373. doi: 10.1093/nar/gkl198
- Baudry, A., Heim, M. A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J.* 39, 366–380. doi: 10.1111/j.1365-313X.2004.02138.x
- Carretero-Paulet, L., Galstyan, A., Roig-Villanova, I., Martínez-García, J. F., Bilbao-Castro, J. R., and Robertson, D. L. (2010). Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in *Arabidopsis*, poplar, rice, moss, and algae. *Plant Physiol.* 153, 1398–1412. doi: 10.1104/pp.110.153593

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00064/full#supplementary-material>

**Figure S1** | Neighbor-joining phylogenetic analysis of grape and Arabidopsis bHLH family.

**Figure S2** | Gene structure of grape bHLH family members.

**Figure S3** | Motifs identified from grape bHLH family members.

**Figure S4** | Heatmaps representing the expression profiles of grape bHLH genes under PEG treatment RMA-normalized signal intensities value of grape bHLH genes were used to represent the expression level.

**Figure S5** | Heatmaps representing the expression profiles of grape bHLH genes in CBF4 overexpression grape RMA-normalized signal intensities value of grape bHLH genes were used to represent the expression level.

**Table S1** | Nr function annotation of grape bHLH genes.

**Table S2** | GO function annotation of grape bHLH genes.

**Table S3** | Conserved sequence of motifs from grape bHLH family members.

**Table S4** | Duplication gene-pairs of grape bHLH gene family.

**Table S5** | List of tissue abbreviations.

**Table S6** | Difference expression genes between ck and PEG treatment.

**Table S7** | Cis-regulation elements of difference expression genes between ck and PEG, cold treatment and CBF4 over-expression grape.

**Table S8** | Difference expression genes between ck and CBF4 overexpression transgenic grape.

**Table S9** | Cis-regulation elements of genes potentially related to anthocyanin or flavonol biosynthesis.

**Table S10** | Promoter sequences of genes are involved in flavonoid biosynthesis and anthocyanin biosynthesis.

- Chen, L., Chen, Y., Jiang, J., Chen, S., Chen, F., Guan, Z., et al. (2012). The constitutive expression of *Chrysanthemum dichrum* ICE1 in *Chrysanthemum grandiflorum* improves the level of low temperature, salinity and drought tolerance. *Plant Cell Rep.* 31, 1747–1758. doi: 10.1007/s00299-012-1288-y
- Chen, S. L., Lee, W., Hottes, A. K., Shapiro, L., and McAdams, H. H. (2004). Codon usage between genomes is constrained by genome-wide mutational processes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3480–3485. doi: 10.1073/pnas.0307827100
- Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B. H., Hong, X., Agarwal, M., et al. (2003). ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. *Genes Dev.* 17, 1043–1054. doi: 10.1101/gad.1077503
- Choi, H., and Oh, E. (2016). PIF4 integrates multiple environmental and hormonal signals for plant growth regulation in *Arabidopsis*. *Mol. Cells* 39, 587–593. doi: 10.14348/molcells.2016.0126
- Costantini, L., Malacarne, G., Lorenzi, S., Troggo, M., Mattivi, F., Moser, C., et al. (2015). New candidate genes for the fine regulation of the colour of grapes. *J. Exp. Bot.* 66, 4427–4440. doi: 10.1093/jxb/erv159
- Fasoli, M., Dal Santo, S., Zenoni, S., Tornielli, G. B., Farina, L., Zamboni, A., et al. (2012). The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. *Plant Cell Online* 24, 3489–3505. doi: 10.1105/tpc.112.100230
- Feller, A., Machemer, K., Braun, E. L., and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* 66, 94–116. doi: 10.1111/j.1365-313X.2010.04459.x
- Feng, H. L., Ma, N. N., Meng, X., Zhang, S., Wang, J. R., Chai, S., et al. (2013). A novel tomato MYC-type ICE1-like transcription factor, SlICE1a, confers cold, osmotic and salt tolerance in transgenic tobacco. *Plant Physiol. Biochem.* 73, 309–320. doi: 10.1016/j.plaphy.2013.09.014

- Flamini, R., Mattivi, F., De Rosso, M., Arapitsas, P., and Bavarese, L. (2013). Advanced knowledge of three important classes of grape phenolics: anthocyanins, stilbenes and flavonols. *Int. J. Mol. Sci.* 14, 19651–19669. doi: 10.3390/ijms141019651
- Gagne, J. M., Downes, B. P., Shiu, S. H., Durski, A. M., and Vierstra, R. D. (2002). The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11519–11524. doi: 10.1073/pnas.162339999
- Gao, J. J., Zhang, Z., Peng, R. H., Xiong, A. S., Xu, J., Zhu, B., et al. (2011). Forced expression of Mdm10, a myb transcription factor gene from apple, enhances tolerance to osmotic stress in transgenic Arabidopsis. *Mol. Biol. Rep.* 38, 205–211. doi: 10.1007/s11033-010-0096-0
- Gilbert, W. (1987). The exon theory of genes. *Cold Spring Harb. Symp. Quant. Biol.* 52, 901–905. doi: 10.1101/SQB.1987.052.01.098
- Giuliano, G., Pichersky, E., Malik, V. S., Timko, M. P., Scolnik, P. A., and Cashmore, A. R. (1988). An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. *Proc. Natl. Acad. Sci. U.S.A.* 85, 7089–7093. doi: 10.1073/pnas.85.19.7089
- Guo, Y., Liu, J., Zhang, J., Liu, S., and Du, J. (2017). Selective modes determine evolutionary rates, gene compactness and expression patterns in Brassica. *Plant J.* 91, 34–44. doi: 10.1111/tj.13541
- Haake, V., Cook, D., Riechmann, J. L., Pineda, O., Thomashow, M. F., and Zhang, J. Z. (2002). Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. *Plant Physiol.* 130, 639–648. doi: 10.1104/pp.006478
- Hershberg, R., and Petrov, D. A. (2008). Selection on codon bias. *Annu. Rev. Genet.* 42, 287–299. doi: 10.1146/annurev.genet.42.110807.091442
- Hichri, I., Heppel, S. C., Pillet, J., Léon, C., Czermel, S., Delrot, S., et al. (2010). The basic helix-loop-helix transcription factor MYC1 is involved in the regulation of the flavonoid biosynthesis pathway in grapevine. *Mol. Plant* 3, 509–523. doi: 10.1093/mp/ssp118
- Hmamouchi, M., Es-Safi, N., Lahrichi, M., Fruchier, A., and Essassi, E. M. (1996). Flavonol and flavonols in leaves of some Moroccan *Vitis vinifera* cultivars. *Am. J. Enol. Vitic.* 47, 186–192
- Hu, Y., Han, Y. T., Wei, W., Li, Y. J., Zhang, K., Gao, Y. R., et al. (2015). Identification, isolation, and expression analysis of heat shock transcription factors in the diploid woodland strawberry *Fragaria vesca*. *Front. Plant Sci.* 6:736. doi: 10.3389/fpls.2015.00736
- Huang, G.-T., Ma, S.-L., Bai, L.-P., Zhang, L., Ma, H., Jia, P., et al. (2012). Signal transduction during cold, salt, and drought stresses in plants. *Mol. Biol. Rep.* 39, 969–987. doi: 10.1007/s11033-011-0823-1
- Huang, H. B. (1980). A preliminary evaluation of climatic regions for grape production in North China. *J. Chin. Agric. Univ.* 2, 43–51.
- Huang, X., Li, K., Jin, C., and Zhang, S. (2015). ICE1 of *Pyrus ussuriensis* functions in cold tolerance by enhancing PuDREBa transcriptional levels through interacting with PuHHP1. *Sci. Rep.* 5:17620. doi: 10.1038/srep17620
- Ikemura, T. (1981). Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. *J. Mol. Biol.* 146, 1–21. doi: 10.1016/0022-2836(81)90363-6
- Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–467. doi: 10.1038/nature06148
- Jiang, Y., Yang, B., and Deyholos, M. K. (2009). Functional characterization of the Arabidopsis bHLH92 transcription factor in abiotic stress. *Mol. Genet. Genomics* 282, 503–516. doi: 10.1007/s00438-009-0481-3
- Kawabe, A., and Miyashita, N. T. (2003). Patterns of codon usage bias in three dicot and four monocot plant species. *Genes Genet. Syst.* 78, 343–352. doi: 10.1266/ggs.78.343
- Knight, R. D., Freeland, S. J., and Landweber, L. F. (2001). A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. *Genome Biol.* 2:research0010.1. doi: 10.1186/gb-2001-2-4-research0010
- Kotak, S., Larkindale, J., Lee, U., von Koskull-Döring, P., Vierling, E., and Scharf, K. D. (2007). Complexity of the heat stress response in plants. *Curr. Opin. Plant Biol.* 10, 310–316. doi: 10.1016/j.pbi.2007.04.011
- Larracuente, A. M., Sackton, T. B., Greenberg, A. J., Wong, A., Singh, N. D., Sturgill, D., et al. (2008). Evolution of protein-coding genes in Drosophila. *Trends Genet.* 24, 114–123. doi: 10.1016/j.tig.2007.12.001
- Ledent, V., and Vervoort, M. (2001). The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. *Genome Res.* 11, 754–770. doi: 10.1101/gr.177001
- Lee, B. H., Henderson, D. A., and Zhu, J. K. (2005). The Arabidopsis cold-responsive transcriptome and its regulation by ICE1. *Plant Cell* 17, 3155–3175. doi: 10.1105/tpc.105.035568
- Li, W. H. (1987). Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. *J. Mol. Evol.* 24, 337–345. doi: 10.1007/BF02134132
- Li, X., Duan, X., Jiang, H., Sun, Y., Tang, Y., Yuan, Z., et al. (2006). Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and Arabidopsis. *Plant Physiol.* 141, 1167–1184. doi: 10.1104/pp.106.080580
- Li, X., Song, H., Kuang, Y., Chen, S., Tian, P., Li, C., et al. (2016). Genome-wide analysis of codon usage bias in *Echinochloa festucae*. *Int. J. Mol. Sci.* 17:1138. doi: 10.3390/ijms17071138
- Liu, W., Tai, H., Li, S., Gao, W., Zhao, M., Xie, C., et al. (2014). bHLH122 is important for drought and osmotic stress resistance in Arabidopsis and in the repression of ABA catabolism. *New Phytol.* 201, 1192–1204. doi: 10.1111/nph.12607
- Liu, Y., Ji, X., Nie, X., Qu, M., Zheng, L., Tan, Z., et al. (2015). Arabidopsis AtbHLH12 regulates the expression of genes involved in abiotic stress tolerance by binding to their E-box and GCG-box motifs. *New Phytol.* 207, 692–709. doi: 10.1111/nph.13387
- Malacarne, G., Costantini, L., Collier, E., Battilana, J., Velasco, R., Vrhovsek, U., et al. (2015). Regulation of flavonol content and composition in (Syrah × Pinot Noir) mature grapes: integration of transcriptional profiling and metabolic quantitative trait locus analyses. *J. Exp. Bot.* 66, 4441–4453. doi: 10.1093/jxb/erv243
- Man, L., Xiang, D., Wang, L., Zhang, W., Wang, X., and Qi, G. (2017). Stress-responsive gene RsICE1 from *Raphanus sativus* increases cold tolerance in rice. *Protoplasma* 254, 945–956. doi: 10.1007/s00709-016-1004-9
- Massari, M. E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* 20, 429–440. doi: 10.1128/MCB.20.2.429-440.2000
- Matus, J. T., Poupin, M. J., Cañón, P., Bordeu, E., Alcalde, J. A., and Arce-Johnson, P. (2010). Isolation of WDR and bHLH genes related to flavonoid synthesis in grapevine (*Vitis vinifera* L.). *Plant Mol. Biol.* 72, 607–620.
- Mol, J., Grotewold, E., and Koes, R. (1998). How genes paint flowers and seeds. *Trends Plant Sci.* 3, 212–217. doi: 10.1016/S1360-1385(98)01242-4
- Murre, C., McCaw, P. S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56, 777–783. doi: 10.1016/0092-8674(89)90682-X
- Nesi, N., Debeaujon, I., Jond, C., Pelletier, G., Caboche, M., and Lepiniec, L. (2000). The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in Arabidopsis siliques. *Plant Cell* 12, 1863–1878. doi: 10.1105/tpc.12.10.1863
- Pathy, L. (1987). Intron-dependent evolution: preferred types of exons and introns. *FEBS Lett.* 214, 1–7. doi: 10.1016/0014-5793(87)80002-9
- Plotkin, J. B., and Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* 12, 32–42. doi: 10.1038/nrg2899
- Rienth, M., Torregrosa, L., Luchoire, N., Chatbanyong, R., Lecourieux, D., Kelly, M. T., et al. (2014). Day and night heat stress trigger different transcriptional responses in green and ripening grapevine (*Vitis vinifera*) fruit. *BMC Plant Biol.* 14:108. doi: 10.1186/1471-2229-14-108
- Rienth, M., Torregrosa, L., Sarah, G., Ardisson, M., Brillouet, J. M., and Romieu, C. (2016). Temperature desynchronizes sugar and organic acid metabolism in ripening grapevine fruits and remodels their transcriptome. *BMC Plant Biol.* 16:164. doi: 10.1186/s12870-016-0850-0
- Rizzon, C., Ponger, L., and Gaut, B. S. (2006). Striking similarities in the genomic distribution of tandemly arrayed genes in Arabidopsis and rice. *PLoS Comput. Biol.* 2:e115. doi: 10.1371/journal.pcbi.0020115
- Rocheta, M., Becker, J. D., Coito, J. L., Carvalho, L., and Amâncio, S. (2014). Heat and water stress induce unique transcriptional signatures of heat-shock proteins and transcription factors in grapevine. *Funct. Integr. Genomics* 14, 135–148. doi: 10.1007/s10142-013-0338-z
- Rocheta, M., Coito, J. L., Ramos, M. J., Carvalho, L., Becker, J. D., Carbonell-Bejerano, P., et al. (2016). Transcriptomic comparison between two *Vitis*

- vinifera* L. varieties (Trincadeira and Touriga Nacional) in abiotic stress conditions. *BMC Plant Biol.* 16:224. doi: 10.1186/s12870-016-0911-4
- Shinozaki, K., and Yamaguchi-Shinozaki, K. (2007). Gene networks involved in drought stress response and tolerance. *J. Exp. Bot.* 58, 221–227. doi: 10.1093/jxb/erl164
- Smalle, J., and Vierstra, R. D. (2004). The ubiquitin 26S proteasome proteolytic pathway. *Annu. Rev. Plant Biol.* 55, 555–590. doi: 10.1146/annurev.arplant.55.031903.141801
- Sun, H., Fan, H. J., and Ling, H. Q. (2015). Genome-wide identification and characterization of the bHLH gene family in tomato. *BMC Genomics* 16:9. doi: 10.1186/s12864-014-1209-2
- Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34, W609–W612. doi: 10.1093/nar/gkl315
- Swindell, W. R., Huebner, M., and Weber, A. P. (2007). Transcriptional profiling of Arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. *BMC Genomics* 8:125. doi: 10.1186/1471-2164-8-125
- Thomashow, M. F. (2001). So what's new in the field of plant cold acclimation? Lots!. *Plant Physiol.* 125, 89–93. doi: 10.1104/pp.125.1.89
- Toledo-Ortiz, G., Huq, E., and Quail, P. H. (2003). The Arabidopsis basic/helix-loop-helix transcription factor family. *Plant Cell* 15, 1749–1770. doi: 10.1105/tpc.013839
- Valliyanadan, B., and Nguyen, H. T. (2006). Understanding regulatory networks and engineering for enhanced drought tolerance in plants. *Curr. Opin. Plant Biol.* 9, 189–195. doi: 10.1016/j.pbi.2006.01.019
- Walker, A. R., Davison, P. A., Bolognesi-Winfield, A. C., James, C. M., Srinivasan, N., Blundell, T. L., et al. (1999). The Transparent Testa GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *Plant Cell* 11, 1337–1349. doi: 10.1105/tpc.11.7.1337
- Wang, J., Hu, Z., Zhao, T., Yang, Y., Chen, T., Yang, M., et al. (2015). Genome-wide analysis of bHLH transcription factor and involvement in the infection by yellow leaf curl virus in tomato (*Solanum lycopersicum*). *BMC Genomics* 16:39. doi: 10.1186/s12864-015-1249-2
- Wang, P., Gao, C., Bian, X., Zhao, S., Zhao, C., Xia, H., et al. (2016). Genome-wide identification and comparative analysis of cytosine-5 DNA methyltransferase and demethylase families in wild and cultivated peanut. *Front. Plant Sci.* 7:7. doi: 10.3389/fpls.2016.00007
- Wang, P., Song, H., Li, C., Li, P., Li, A., Guan, H., et al. (2017). Genome-Wide dissection of the heat shock transcription factor family genes in arachis. *Front. Plant Sci.* 8:106. doi: 10.3389/fpls.2017.00106
- Welch, C. R., Wu, Q., and Simon, J. E. (2008). Recent advances in anthocyanin analysis and characterization. *Curr. Anal. Chem.* 4, 75–101. doi: 10.2174/157341108784587795
- Winkel-Shirley, B. (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.* 126, 485–493. doi: 10.1104/pp.126.2.485
- Wright, F. (1990). The 'effective number of codons' used in a gene. *Gene* 87, 23–29. doi: 10.1016/0378-1119(90)90491-9
- Yadav, C. B., Muthamilarasan, M., Dangi, A., Shweta, S., and Prasad, M. (2016). Comprehensive analysis of SET domain gene family in foxtail millet identifies the putative role of SiSET14 in abiotic stress tolerance. *Sci. Rep.* 6:32621. doi: 10.1038/srep32621
- Yang, T., Zhang, L., Zhang, T., Zhang, H., Xu, S., and An, L. (2005). Transcriptional regulation network of cold-responsive genes in higher plants. *Plant Sci.* 169, 987–995. doi: 10.1016/j.plantsci.2005.07.005
- Zarka, D. G., Vogel, J. T., Cook, D., and Thomashow, M. F. (2003). Cold induction of Arabidopsis CBF genes involves multiple ICE (inducer of CBF expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. *Plant Physiol.* 133, 910–918. doi: 10.1104/pp.103.027169
- Zhang, J. (2003). Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18, 292–298. doi: 10.1016/S0169-5347(03)00033-8
- Zhang, W. J., Zhou, J., Li, Z. F., Wang, L., Gu, X., and Zhong, Y. (2007). Comparative analysis of codon usage patterns among mitochondrion, chloroplast and nuclear genes in *Triticum aestivum* L. *J. Integr. Plant Biol.* 49, 246–254. doi: 10.1111/j.1744-7909.2007.00404.x
- Zhang, X., Luo, H., Xu, Z., Zhu, Y., Ji, A., Song, J., et al. (2015). Genome-wide characterisation and analysis of bHLH transcription factors related to tanshinone biosynthesis in *Salvia miltiorrhiza*. *Sci. Rep.* 5:11244. doi: 10.1038/srep11244

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Wang, Su, Gao, Jiang, Wu, Li, Zhang, Wang and Ren. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.