Genome-wide identification of HSF Family in Peach and Functional analysis of *PpHSF5* involvement in Root and Aerial Organ Development

Bin Tan [Equal first author, 1, 2], **Liu Yan** [Equal first author, 1, 2], **Huannan Li** [1, 2], **Xiaodong Lian** [1, 2], **Jun Cheng** [1, 2], **Wang Wei** [1, 2], **Xianbo Zheng** [1, 2], **Xiaobei Wang** [1, 2], **Jidong Li** [1, 2], **Xia Ye** [1, 2], **Langlang Zhang** [1, 2], **Zhiqian Li** [1, 2], **Jiancan Feng** [Corresp. 1, 2]

[1] College of Horticulture, Henan Agricultural University, Zhengzhou, China

[2] Henan Key Laboratory of Fruit and Cucurbit Biology, Zhengzhou, China

Corresponding Author: Jiancan Feng
Email address: jcfeng@henau.edu.cn

Background. Heat shock factors (HSFs) play important roles during normal plant growth and development and when plants respond to diverse stressors. Although most studies have focused on the involvement of HSFs in the response to abiotic stresses, especially in model plants, there is little research on their participation in plant growth and development or on the HSF (PpHSF) gene family in peach (*Prunus persica*).

Methods. DBD (PF00447), the HSF characteristic domain, was used to search the peach genome and identify *PpHSFs*. Phylogenetic, multiple alignment and motif analyses were conducted using MEGA 6.0, ClustalW and MEME, respectively. The function of *PpHSF5* was confirmed by overexpression of *PpHSF5* into Arabidopsis.

Results. Eighteen *PpHSF* genes were identified within the peach genome. The *PpHSF* genes were nonuniformly distributed on the peach chromosomes. Seventeen of the *PpHSFs* (94.4%) contained one or two introns, except *PpHSF18*, which contained three introns. The in silico-translated PpHSFs were classified into three classes (PpHSFA, PpHSFB and PpHSFC) based on multiple alignment, motif analysis and phylogenetic comparison with HSFs from *Arabidopsis thaliana* and *Oryza sativa*. Dispersed gene duplication (DSD at 67%) mainly contributed to HSF gene family expansion in peach. Promoter analysis showed that the most common cis-elements were the MYB (abiotic stress response), ABRE (ABA-responsive) and MYC (dehydration-responsive) elements. Transcript profiling of 18 *PpHSFs* showed that the expression trend of *PpHSF5* was consistent with shoot length changes in the cultivar ‘Zhongyoutao 14’. Further analysis of the *PpHSF5* was conducted in 5-year-old peach trees, *Nicotiana benthamiana* and *Arabidopsis thaliana*, respectively. Tissue-specific expression analysis showed that *PpHSF5* was expressed predominantly in young vegetative organs (leaf and apex). Subcellular localization revealed that PpHSF5 was located in the nucleus in *N. benthamiana* cells. Two transgenic Arabidopsis lines were obtained that overexpressed *PpHSF5*. The root length and the number of lateral roots in the transgenic seedlings were significantly less than in WT seedlings and after cultivation for three weeks. The transgenic rosettes were smaller than those of the WT at 2-3 weeks. The two transgenic lines exhibited a dwarf phenotype three weeks after transplanting, although there was no significant difference in the number of internodes. Moreover, the PpHSF5-OE lines exhibited enhanced thermotolerance. These results indicated that PpHSF5 might be act as a suppresser of growth and development of root and aerial organs.

Genome-wide Identification of HSF Family in Peach and Functional Analysis of *PpHSF5* involvement in Root and Aerial Organ Development

Bin Tan[1,2]*, Liu Yan[1,2]*, Huannan Li[1,2], Xiaodong Lian[1,2], Jun Cheng[1,2], Wei Wang[1,2], Xianbo Zheng[1,2], Xiaobei Wang[1,2], Jidong Li[1,2], Xia Ye[1,2], Langlang Zhang[1,2], Zhiqian Li[1,2] and Jiancan Feng[1,2]

[1] College of Horticulture, Henan Agricultural University, Zhengzhou 450002, China;

[2] Henan Key Laboratory of Fruit and Cucurbit Biology, Zhengzhou 450002, China

* These authors contributed equally to this work.

Corresponding Author:

Jiancan Feng[1,2]

Nongye Road No. 63, Zhengzhou, Henan Province, 450002, China

Email address: jcfeng@henau.edu.cn

Abstract

Background. Heat shock factors (HSFs) play important roles during normal plant growth and development and when plants respond to diverse stressors. Although most studies have focused on the involvement of HSFs in the response to abiotic stresses, especially in model plants, there is little research on their participation in plant growth and development or on the HSF (PpHSF) gene family in peach (*Prunus persica*).

Methods. DBD (PF00447), the HSF characteristic domain, was used to search the peach genome and identify *PpHSFs*. Phylogenetic, multiple alignment and motif analyses were conducted using MEGA 6.0, ClustalW and MEME, respectively. The function of *PpHSF5* was confirmed by overexpression of *PpHSF5* into Arabidopsis.

Results. Eighteen *PpHSF* genes were identified within the peach genome. The *PpHSF* genes were nonuniformly distributed on the peach chromosomes. Seventeen of the *PpHSFs* (94.4%) contained one or two introns, except *PpHSF18*, which contained three introns. The in silico-translated PpHSFs were classified into three classes (PpHSFA, PpHSFB and PpHSFC) based on multiple alignment, motif analysis and phylogenetic comparison with HSFs from *Arabidopsis thaliana* and *Oryza sativa*. Dispersed gene duplication (DSD at 67%) mainly contributed to HSF gene family expansion in peach. Promoter analysis showed that the most common cis-elements were the MYB (abiotic stress response), ABRE (ABA-responsive) and MYC (dehydration-responsive) elements. Transcript profiling of 18 *PpHSFs* showed that the expression trend of *PpHSF5* was consistent with shoot length changes in the cultivar ‘Zhongyoutao 14’. Further

analysis of the *PpHSF5* was conducted in 5-year-old peach trees, *Nicotiana benthamiana* and *Arabidopsis thaliana*, respectively. Tissue-specific expression analysis showed that *PpHSF5* was expressed predominantly in young vegetative organs (leaf and apex). Subcellular localization revealed that PpHSF5 was located in the nucleus in *N. benthamiana* cells. Two transgenic Arabidopsis lines were obtained that overexpressed *PpHSF5*. The root length and the number of lateral roots in the transgenic seedlings were significantly less than in WT seedlings and after cultivation for three weeks. The transgenic rosettes were smaller than those of the WT at 2-3 weeks. The two transgenic lines exhibited a dwarf phenotype three weeks after transplanting, although there was no significant difference in the number of internodes. Moreover, the PpHSF5-OE lines exhibited enhanced thermotolerance. These results indicated that PpHSF5 might be act as a suppresser of growth and development of root and aerial organs.

Key Words Heat shock factors family, Root development, Peach (*Prunus persica*), Aerial organ, Functional identification, *PpHSF5*

Introduction

Plant growth and development are affected by a range of abiotic stress, including cold, heat, salinity and drought stress (*Guo et al., 2016*). Heat shock factors (HSFs) act with heat shock proteins (HSPs) as key transcriptional activators during responses to abiotic stress (*Hu et al., 2009*). Recent studies indicated that HSFs act as key components of signal transduction in response to different abiotic stresses in plants (*Guo et al., 2016*; *Scharf et al., 2012*).

HSFs in plant genomes can be identified by a conserved DNA-binding domain (DBD). The DBD domain is located in the N-terminal of all HSFs and specifically binds to heat stress (HS) motifs in the promoters of target genes (*Wang et al., 2018*). The adjacent HR-A/B region is linked to the DBD by a connector of variable length (15-80 amino acid residues) that contains a bipartite heptad pattern of hydrophobic amino acid residues, which constitutes a coiled-coil domain for protein interaction. According to the number of amino acid residues inserted into the HR-A/B region, HSFs are divided into three main groups, each with subgroups, namely HSFA (A1-A9), HSFB (B1-B5) and HSFC (C1-C2) (*Koskull-Doring, Scharf&Nover, 2007*; *Yang et al., 2014*). HSFA members contain an acidic motif (AHA activation domain) at their C-terminus and act as transcriptional activators. The members of HSFB as transcriptional repressors.

In a wide range of plants, a number of HSFs have been shown to be involved in resistance to heat (*Guo et al., 2016*) and other abiotic or biotic stresses (*Yu et al., 2019*). Of the 21 HSF family members in Arabidopsis, a number act as pioneer regulators of the response to heat shock. HSFA1a, HSFA1b, HSFA1d, HSFA1e and HSFA2 play active regulatory roles in the response

to HS in plants (*Busch et al., 2005*; *Nishizawa et al., 2006*). In Arabidopsis, the assembly of the HSFA1/A2 super-activated complex regulates heat stress (HS) genes (*Chan-Schaminet et al., 2009*). *HSFB1* and *HSFB2b* participate in disease resistance regulation of Arabidopsis and expression of *Pdf1.2* (*Kumar et al., 2009*). *OsHSFB4b* and *OsHSFA2c* participate in the regulation of the heat shock response by regulating the expression of HSP100 (*Singh et al., 2012*). *OsHSFC1b* is related to the regulation of salt stress and plant development (*Schmidt et al., 2012*).

Several HSFs are stress-inducible transcriptional factors that participate in the growth and development of root and aerial organs in plant. Overexpression of *AtHsfB4* in Arabidopsis induces specific effects on root development, resulting in shortened roots (*Begum et al., 2013*). The over-expression of *BhHsf1* conferred growth retardation of aerial organs, producing a dwarf phenotype, although the primary roots were not obviously different from those of wild type (*Zhu et al., 2009*). Transgenic Arabidopsis plants with strong expression of *AtHsfA3* and *AtHsfA2* showed a severely dwarfed phenotype and increased tolerance to heat (*Ogawa et al., 2007; Yoshida et al., 2008*). The thermotolerant phenotype was also observed in the cotyledons, rosette leaves, inflorescence stems and seeds of transgenic Arabidopsis plants expressing *OsHsfA2e* (*Yokotani et al., 2008*).

The HSF family have been analyzed genome-wide in several plants, such as rice (*Oryza sativa*), Arabidopsis (*Arabidopsis thaliana*), cotton (*Gossypium hirsutum*), soybean (*Glycine max*), wheat (*Triticum aestivum*), pepper (*Capsicum annuum*), poplar (*Populus trichocarpa*),

Brassica napus, grape (*Vitis vinifera*) and Tartary buckwheat (*Fagopyrum tataricum*) (*Nover et al., 2001; Chauhan et al., 2011*; *Wang et al., 2014*; *Li et al., 2014*; *Guo et al., 2015*; *Xue et al., 2014*; *Zhang et al., 2016*; *Zhu et al., 2017*; *Liu et al.,2018; Liu et al., 2019*). Peach (*Prunus persica L.*) is an important economical crop and a popular fruit with consumers. However, there are limited studies on peach HSFs. To remedy this, we identified the *HSF* family in the peach genome and conducted bioinformatics analysis of the 18 identified *PpHSFs*. Based on the latest transcriptome data (*Lian et al., 2020*), the expression patterns of the *PpHSF* genes were analyzed during development of the cultivar 'Zhongyoutao 14'. 'Zhongyoutao 14' (derived from 'SD9238'), is a temperature-sensitive peach that exhibits a shorter internode length and a smaller canopy when grown below 30°C (*Lu et al., 2016*). *PpHSF5* was further analyzed and found to function in the development of the root and aerial organs. Furthermore, the thermotolerant phenotype was analyzed in newly obtained transgenic Arabidopsis plants expressing *PpHSF5*. The genome-wide analysis of *PpHSF* gene family offers a basis for further investigation into the function and evolutionary history of peach *HSFs* and provides candidate genes for peach molecular breeding.

Materials & Methods

Plant Materials

Established peach trees (*Prunus persica*) cultivar 'Zhongyoutao 14' ('Maotao' as rootstock) have

been grown for 5 years at the Experimental Station of the Horticulture College, Henan Agricultural University (Zhengzhou, China). Samples from the apex, young and mature leaves, self-pollinated embryos, and fruit were collected, frozen in liquid nitrogen and stored at -80°C. Leaves from *Nicotiana benthamiana* were used for subcellular location of PpHSF5. *Arabidopsis thaliana* (L.) Heynh Columbia 0 (Col-0) was used for transformation with *PpHSF5*.

Identification and Chromosomal Location of HSF Genes in Peach

The hidden Markov model (HMM) of the DBD domain (PF00447), characteristic of HSFs, was downloaded from the Pfam website (*Finn et al., 2016*) and used to identify HSF genes in peach. The peach genome files (v2.1) were downloaded from JGI database (https://phytozome.jgi.doe.gov/pz/portal.html) (*Verde et al., 2013*), HSF protein sequences were obtained in peach genome by BLASTP and hmmsearch function, and then the DBD domain were further identified by Pfam analysis. The peach HSF gene and protein sequences were extracted from Phytozome v12.1. *PpHSF* genes were named according to physical location on the chromosomes. Positional information was retrieved from peach genome annotations obtained from Phytozome v12.1, and chromosome locations of the *PpHSFs* were drawn using the Circos software (*Krzywinski et al., 2009*). The isoelectric points and other physical properties were approximated from ExPASy (http://web.expasy.org/compute_pi). Gene structures were predicted using the Gene Structure Display Server 2.0 (http://gsds.cbi.pku.edu.cn/).

Phylogenetic and Motif Analysis of PpHSFs

The amino acid sequences of 21 AtHSFs (*Arabidopsis thaliana*), 25 OsHSFs (*Oryza sativa*) and 18 PpHSFs (*Prunus persica*) were gathered from Phytozome v12.1 using ClustalW with system default settings. The phylogenetic trees were formulated by the maximum likelihood method (ML) with Jones-Taylor-Thornton (JTT) model in MEGA 6.0 (http://www.megasoftware.net/download_form).. Conserved motifs of HSF proteins in peach were identified using the MEME tool (http://meme.nbcr.net/meme/cgi-bin/meme.cgi) with default parameters in normal operation mode. The subcellular localization was predicted with Plant-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/#).

Gene Duplication and *Cis*-Element Analysis of *PpHSFs*

Gene duplication was analyzed using MCScanX (*Wang et al., 2012)*. Genomic DNA sequences (2000 bps upstream of the start codons) for each *PpHSF* were obtained from the peach genome and skimmed in the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) for cis-acting elements analysis of the promoter in *PpHSFs*.

Gene Expression Analysis of *PpHSFs*

The FPKM (fragments per kilobase of exon per million fragments mapped) values of the 18 *PpHSFs* (Supplemental Table S1-1) were obtained from our previous study of gene expression in shoots at four critical growth stages, namely initial period (IP), initial elongation period (IEP), rapid growth period (RGP) and stable growth period (SGP) of temperature-sensitive peach

147 cultivar ‘Zhongyoutao 14’ (*Lian et al., 2020*). The average maximum temperature of previous
148 week (AMTPW) began to be higher than 30° C in the first day of RGP *(Lian et al., 2020)*. The
149 heat map was generated by TBtools (*Chen et al., 2020*).

150 **Quantitative Real-time PCR Analysis of *PpHSF5***

151 Total RNA of different tissues from ‘Zhongyoutao 14’ peach and leaves from T_2 transgenic
152 Arabidopsis lines was isolated using the Spin Column Plant Total RNA Purification Kit
153 (ShengGong, Shanghai, China). The cDNA was synthesized using FastQuant RT Kit (Tiangen
154 Biotech, Beijing, China). qRT-PCR was implemented using an ABI PRISM 7500 FAST
155 Sequence Detection System (Applied Biosystems, Madrid, CA, USA) with SYBR Select Master
156 Mix (Applied Biosystems, USA). Primers of *PpHSF5* were designed using Primer Premier 5.0.
157 *PpGAPDH* (Prupe.1G234000) and *AtUBC* (AT5G25760) were used as constitutive controls for
158 either tissue-specific expression in peach or expression analysis in transgenic Arabidopsis,
159 respectively. Primers are shown in Supplemental Table S1-2. The reaction mixture was as
160 follows: 1 μL cDNA template (200 ng/μL), 1 μL of each primer (10 μM), 10 μL SYBR Premix
161 and 7 μL ddH_2O. Melting curve analysis was performed after the end of 40 cycles to insure
162 proper amplification of the target. During the melting process, fluorescence readings were
163 continuously collected from 60-90°C at a heating rate of 0.5 °C s^{-1}. All analysis was repeated
164 three times using biologically replicates. The relative expression levels of *PpHSF5* were
165 calculated as $2^{-\Delta\Delta CT}$ method (*Schmittgen et al., 2008*). The relative expression levels of *PpHSF5*
166 was calculated in SPSS using ANOVA at significance levels of $P < 0.05$.

167 **Subcellular Localization of PpHSF5**

168 *PpHSF5* without the termination codon was amplified by PCR using cDNA from 'Zhongyoutao
169 14' as the template (Primer details in Supplemental Table S1-2). This coding region was cloned
170 into the pSAK277-GFP vector to construct PpHSF5::GFP fusion proteins that were driven by the
171 CaMV 35S promoter. The p35SPpHSF5::GFP and p35SGFP (control) vectors were transformed
172 into *Agrobacterium tumefaciens* strain GV3101, which were then injected into leaves of *N.*
173 *benthamiana* according to previously published protocols (*Sparkes et al., 2006*). The leaves were
174 observed 48-72 h after injection using laser scanning confocal microscopy (Zeiss LSM700).

175 **Construction of Expression Vectors for Plant Transformation**

176 The CDS of *PpHSF5* was PCR-amplified and cloned into the pSAK277 vector using the
177 restriction enzymes *Xho* I and *Xba* I (Primer details in Supplemental Table S1-2). The
178 *p35S::PpHSF5* vector was transformed into *Agrobacterium tumefaciens* strain GV3101. The
179 floral dip method was used to transform *Arabidopsis thaliana* (Col-0) (*Chung et al., 2000*).

180 **Phenotype of Overexpression *PpHSF5* in Arabidopsis**

181 The seeds from T_2 transgenic Arabidopsis lines were sterilized by 6.25% NaClO for 5 min, and
182 then washed in sterilized ddH_2O. The seeds were cultured at 4°C for 2 d and then transferred
183 onto MS solid medium under 16/8 h light/dark cycle for one week on square plates. Three
184 biological replicates (with three seedlings of each lines per square plate) were used for
185 observation of root phenotype. The roots of different transgenic lines with three plants per line

were measured by a LA2400 Scanner at three weeks to determine the growth status. The leaves were cut from the transgenic seedlings for gene expression analysis. Other seedlings, germinated on agar and grown for one week, were transferred into the soil and grown under normal conditions. The length and width of rosettes (four leaves per plants, five repetition) and number of rosettes (five plants per line) in different transgenic lines and WT were measured and photographed at two weeks and at three weeks after transplanting, respectively. Moreover, the morphology of transgenic lines and WT, including the height of plants (eight plants per line) and the number of branches and blooms (five plants per line) were recorded, three weeks after transplanting.

Heat Stress Treatment

For performing heat stress treatment on the seeds germination and plants grown on the agar medium, seeds of WT and transgenic Arabidopsis lines sown on MS medium at 4°C for 2 d and in darkness for 2 d (22°C) were exposed to HS stress at 46°C for 30 min, and then were transferred into a climate chamber (22°C, 16 h light/8 h dark cycles). After HS treatment, the germination of seeds were counted daily and photographed. More than 50 seeds of each line were used in each plate with three replications. Difference in HS stress was confirmed using *t*-test.

Statistical Analysis

Data were analyzed by ANOVA, Tukey HSD[a] and Duncan[a]'s multiple range tests (at $P < 0.05$)

using IBM SPSS Statistics 20 (SPSS, USA).

Results

Genome-wide Identification, Chromosomal Distribution and Gene Structures of HSF Genes in Peach

Eighteen *HSF* gene family members were identified from the peach genome and then named *PpHSF1* to *PpHSF18* according to their physical locations (Table 1 and Supplemental Table S2-1). The *PpHSF* coding sequences ranged from 591 bp (*PpHSF8*) to 1608 bp (*PpHSF14*). In silico-translated PpHSF proteins showed divergent lengths [196 to 535 amino acids (aa)] with different molecular weights (22.36~59.56 kDa) and isoelectric points (4.67 to 8.75) (Table1). All PpHSFs were predicted to be nuclear-localized proteins.

Seven of the 8 peach chromosomes contained at least one *PpHSF*, with the exception being chromosome 6 (Fig. 1 and Table 1). Five *PpHSFs* were located on chromosome 1 (*PpHSF1-5*), and another five (*PpHSF13-17*) on chromosome 7. Chromosomes 2, 3 and 8 carried only one *PpHSF* gene each, while chromosome 5 had two, and chromosome 4 had three. The above results indicated that *PpHSFs* were unevenly distributed across the peach chromosomes.

The structural differences of the *PpHSF* genes were also analyzed. The number of introns ranged from one to three among the *PpHSFs*. The majority of the *PpHSFs* (66.67%) contained one intron, 27.78% contained two introns, and only *PpHSF18* contained three introns

(Supplemental Figure S1 and Supplemental Table S2-2). Interestingly, both *PpHSF18* and *PpHSF12* has predicted introns in the 5'-UTR and 3'-UTR, respectively.

Gene Duplication Pattern Analysis of *PpHSFs*

To explain the expansion of the *PpHSFs* gene family, the gene duplication patterns of the *PpHSFs* were analyzed and compared across the peach genome (Supplemental Table S3). There were only two patterns of gene duplications, with 67% of the gene pairs derived from dispersed gene duplication (DSD) and the remaining gene pairs derived from whole-genome duplication (WGD). Three syntenic pairs were identified, and all originated from WGD. The syntenic genes were located on different chromosomes from their partner (Fig. 1).

Classification, Phylogenetic and Motif Analyses of PpHSFs

Among plant species, there are two characteristic amino acid domains in the HSF family, the DBD and adjacent HR-A/B region (*Nover et al., 2001*). The PpHSFs were divided into three classes (PpHSFA, PpHSFB and PpHSFC), according to the number of amino acids between part A and part B of the HR-A/B domain (Fig. 2A). Multiple sequence alignment analysis of the PpHSF proteins indicated that an insertion of 21 amino acids was found in Class A (11 of the PpHSFs) and a shorter insertion of 14 amino acid in Class C (1 of the PpHSFs) between the HR-A and HR-B regions. Six of the PpHSFs had no aa insertion between the two domains (Class B).

Phylogenetic analysis among the HSF proteins from three plant species, namely 21 AtHSFs (*Arabidopsis thaliana*), 25 OsHSFs (*Oryza sativa*) and 18 PpHSFs (*Prunus persica*), was

conducted by constructing a phylogenetic tree. According to the phylogenetic tree, the 64 HSFs derived from the three plant species were divided into three classes and 15 subclasses (Fig. 2B). The peach proteins sorted into the classes of HSFs, within (11 members) in class HSFA, six in HSFB, and one in HSFC. Class A included nine subclasses (A1-A9), the largest number of subclasses. The PpHSFs were grouped into eight of the Class A subclasses, with no PpHSF in Class A7. Class B consisted of 18 total members and was divided into four subclasses (B1-B4). It is noteworthy that PpHSF8 clustered with Class B but as a single branch. Only six members were clustered into Class C, with two subclasses (C1-C2). No PpHSFs clustered with subclass C2.

The conserved motifs in the PpHSF proteins were analyzed using MEME. The results revealed that PpHSFs contained ten conserved motifs (Fig. 2C and Supplemental Table S4). Motifs 1-3 were found in the N-terminals (the most conserved region) of each PpHSF. Motif 4 was found in Class A and Class B. Motif 5, which was found between the HR-A and HR-B regions, was observed in Class A and Class C. The motif analysis was consistent with the multiple sequence alignment and phylogenetic analyses.

Analysis of the *Cis*-Acting Regulatory Elements in the *PpHSF* Gene Promoters

The cis-acting elements within the promoters of the 18 *PpHSFs* were analyzed using PlantCARE. Every promoter contained at least two MYB elements (abiotic stress response) (Table 2). All but one promoter contained an ABRE (ABA-responsive element). The next most common elements were MYC elements (dehydration-responsive) (in 88.8% of the promoters), CGTCA- and

TGACG- motifs (83.3%), and ARE elements (anaerobic induction) (77.8%). ERE (ethylene-responsive element), MBS (drought inducible), MRE and P-box elements were also present in the promoters of some *PpHSFs*. The TCA-motif was observed in only five *PpHSFs*, namely *PpHSF1*, *PpHSF2*, *PpHSF5*, *PpHSF6* and *PpHSF13*. Previous studies reported several elements, including MYB, ABRE, MYC, play vital roles in stress responses in plants (*He et al., 2012*; *Li et al., 2012*). The different *cis-elements* in the promoter regions of these *PpHSFs* implied that the *PpHSFs* may function in plant development and stress responses.

The Expression Patterns of *PpHSFs* during shoot elongation in 'Zhongyoutao 14'

Based on our previous RNA-seq analysis (*Lian et al., 2020*), the expression patterns of *PpHSFs* were compared in four critical stages of shoot elongation of 'Zhongyoutao 14' grown under elevated temperature in the field (Fig. 3). Most of *PpHSFs* belonging to the A and C classes (except *PpHSF4* and *PpHSF11*) were maintained at lower expression level. The *PpHSFs* of B class exhibit diverse expression patterns. The FPKM values of *PpHSF8* and *PpHSF13* remained almost unchanged at the four stages. The transcripts of another three *PpHSFs* (*PpHSF15*, *PpHSF6* and *PpHSF12*) were present at lower levels during the IEP stage and then slightly increased during the RGP and SGP stages. The expression of level of *PpHSF5* showed higher in IEP stage and increased from the RGP to SGP stages. *PpHSF5* might participate in temperature-induced shoot growth of temperature-sensitive peach.

Expression Analysis of *PpHSF5* and Subcellular Localization of PpHSF5

The relative expression of *PpHSF5* was investigated by qRT-PCR in different organs of 'Zhongyoutao 14' (Fig. 4; supplement Table S5-1). The results showed that *PpHSF5* were expressed predominantly in young vegetative organs (leaves and apex), but barely detectable in embryos and mature leaves. This suggested that *PpHSF5* might participated in the growth and development of plants. The 35S::PpHSF5-GFP signal was evident in the cellular nucleus in *N. benthamiana* cells, indicating a nuclear localization (Fig. 5). The result was in concurrence with the prediction from Plant-mPLoc of subcellular localization (Table 1).

Overexpression of *PpHSF5* in Arabidopsis Results in Dwarf Phenotypes

To investigate the function of *PpHSF5*, an overexpression vector with *PpHSF5* was transformed into Arabidopsis. The phenotype of two transgenic lines and WT were recorded (Fig. 6). One week after germination on agar medium, the transgenic lines had shorter roots and a smaller number of lateral roots than WT seedlings (Fig. 6A). The average root length in WT was 7.13 cm, in transgenic line L1 was 3.08 cm, and in L2 was 3.50 cm (Fig. 6A and B; Supplemental Table S6-1). Two weeks after transplantation, there was no difference in the number of rosette leaves between the transgenic lines and WT (Fig. 6C and D-a and Supplemental Table S6-2), although the rosette leaves were significantly shorter and narrower in the transgenic lines (the average length and width; Fig. 6D-b, c and Supplemental Tables S6-3 and S6-4). The mRNA levels in the PpHSF5-OE lines were obviously higher than WT plants (Fig. 6D-d and Supplemental Table S5-2).

Three weeks after transplanting, the soil-grown transgenic lines had fewer rosette leaves and

302 the leaves were shorter and narrower than those in WT plants (Fig. 6F). Moreover, the two
303 transgenic lines (L1 and L2) exhibited a dwarf phenotype (Fig. 6E and F). The average height of
304 L1 (16.83 cm) was 40% shorter than that of the WT (26.77 cm). The number of rosette branches
305 was much greater in WT than in transgenic lines, which had just one flowering stalk (Fig. 6F).
306 There was no significant difference in the number of internodes (Fig. 6E and F), indicating that
307 the dwarf phenotype of the transgenic lines might be caused by shorter internode length.

308 Shorter roots were also observed in the transgenic lines for cultivation three weeks after
309 transplanting (Fig. 6G and H). Root length and root volume were significantly lower in
310 transgenic lines compared to WT (Fig. 6G and H, Supplemental Table S6-12, 13). The average
311 length of roots in Line 1 was 219.34 cm, which was 54% of the length in WT plants. The root
312 volumes in the transgenic lines (Line 1 was 0.19 cm^3, Line 2 was 0.36 cm^3) was no more than
313 20% of that in WT (1.95 cm^3). Other root indexes output by the root scanner were also less in the
314 PpHSF5-OE lines, including the forks, tips and crossings of roots (Fig. 6G and H, Supplemental
315 Table S6-14, 15, 16). Between the two transgenic lines, the higher expression level of *PpHSF5*
316 in L1 resulted in more obvious phenotypes compared to PpHSF5-OE L2 and WT (Fig. 6A, C,
317 and E and Supplemental Table S5-2). The above results indicated that *PpHSF5* might participate
318 in plant growth and development and that overexpression of *PpHSF5* results in a dwarf
319 phenotype in transgenic Arabidopsis.

320 ## PpHSF5-OE Lines Exhibit Enhanced Thermotolerance

321 The thermotolerance of PpHSF5-OE lines was assayed with that of WT (Fig 7 and

Supplemental Table S7). As shown in Fig 7B and F, only 8.3% WT seeds germinated, whereas more than 93.3% of the transgenic seeds germinated after HS treatment 1 d. Nearly half of the WT seeds germinated after HS treatment 3 d, whereas 100% of the transgenic seeds were germinated (Fig 7C and F). After HS treatment 5 d and 7 d, 68.4% and 82.6% of WT seeds germinated, respectively (Fig 7D, E and F). Compared to WT seedlings, the PpHSF5-OE seedlings exhibited green cotyledons and vigor growth (Fig 7C, D and E). These results suggested that the overexpression of *PpHSF5* improves thermotolerance of PpHSF5-OE lines.

Discussion

Peach Contains Fewer HSF Gene Family Members among Several Plant Species

HSFs play vital roles in plant growth and defense. Through plant genome sequencing, *HSF* gene family members have been identified in several model organisms and more than 20 plant species (Supplemental Table S8). Only a single *HSF* was detected in yeast, nematodes and flies (*Nakai,1999*; *Nover,1996*). In this study, 18 *HSF* genes were identified in peach, which is less than in most other plant species, but more than in tea (*Camellia sinensis*), strawberry (*Fragaria vesca*), Chinese plum (*Prunus salicina*) and carnation (*Dianthus caryophyllus*) (*Hu et al., 2015*; *Liu et al., 2016*; *Qiao et al., 2015; Li et al., 2019*).

HSFs in each subgroup are highly similar to each other across a variety of plants. Among these species, Class A contains the largest number of HSFs, followed by Class B, and then Class

C. The same phenomenon was also observed in peach, which contained 11 *HSFAs*, six *HSFBs* and one *HSFCs*.

The *HSF* Gene Family Expanded along with DSD in Peach

The number of *HSF*s expanded markedly during plant evolution. The analysis of 51 representative species indicated that the HSF gene family largely expanded along with WGD during plant evolution (*Wang et al., 2018*). In Chinese white pear (*Pyrus bretschneideri*), most *PbHSF* expansions dated back to a recent WGD (Qiao et al., 2015). On the other hand, *GmHSFs* in cotton expanded along single gene duplication events (*Wang et al., 2014*). Here, DSD (67%) was the primary type of duplication for the HSF gene family in peach. The same phenomenon was also seen for the E3 ligase gene family in peach *(Tan et al., 2019*). It is probably that peach has not undergone a recent WGD (*Verde et al., 2013*).

HSF Gene Family was Classified into Three Classes

Plant HSF proteins contain a few conserved characteristic domain (*Guo et al., 2016*). Generally, HSF families in plant species can be divided into three subfamilies, termed HSFA, HSFB, and HSFC (*Liu et al., 2018; Wang et al., 2018*). The classification of the PpHSF family was consistent with that in other plant species (Supplemented Table 7). Multiple sequence alignments revealed that an insertion occurred in the DBD domain near the N-terminus in the PpHSFA and PpHSFC groups. Like in other plants, the PpHSFA and PpHSFC genes contained inserted coding sequence for 21 and 7 aa in the HR-A/B region, respectively, while the HR-A/B region of

PpHSFB was compact (*Nover, 1996*; *Scharf et al., 2012*). The organization, composition, number of conserved motifs in the HSFs differed among plant species (*Wang et al., 2018*). In Chinese whit pear, Class A in PbHSFs contained the most conversed motifs, followed by class B and then class C (*Qiao et al., 2015*). In this study, the number of motifs in the different classes was consistent with those in Chinese white pear. This also showed that members of the same class often have similar sequence structures in peach. For example, motif 5 was present only in PpHSFA and PpHSFC, while all Class B and Class A HSFs contain motif 4. The presence of these motifs may lead to functional group specificity. The similar classifications of HSF families in diverse plants showed that the HSF family was highly conserved during long-term evolution.

PpHSF5 **Acts as Repressor of Organ Size in Plants**

In plants, organ size is primarily controlled by internal developmental signals (*Mizukami, 2001; Dubrovsky et al., 2006; Spradling , 2001; West et al., 2004)*. Previous research in the model organism *Arabidopsis thaliana indicates* that plant hormones and transcription factors, including HSFs, play crucial roles in growth and development *(Petricka et al., 2012; Begum et al., 2013*). HSFs as key transcription factors protect plants from various abiotic stresses and then participate in the growth and development (*Guo et al., 2016*). For example, *OsHsfA1a*, *OsHsfA1b* and *OsHsfA1d* are the main positive regulators of gene expression on heat stress-responsive, and four HSFA proteins play significant roles in gene expression of plant growth and development (*Yoshida et al., 2011*). In poplar (*Populus trichocarpa*), the transcripts of three *PtHsfs* in the B4 subfamily (-B4b, -B4c and -B4d) were maintained at higher levels during the leaf expansion

stages (*Liu et al., 2019*). In carnation (*Dianthus caryophyllus*), five *DcaHsfs*, namely *DcaHsf-A1, A2a, A9a, B2a, B3a*, were involved in early flowering stages (*Li et al., 2019*). Transgenic Arabidopsis plants overexpressing *AtHSFB4* contained massively enhanced levels of *AtHSFB4* mRNAs and exhibited shorter roots (*Begum et al., 2013*). In this study, overexpression of *35S:PpHSF5* in Arabidopsis resulted in not only shorter roots but also in lesser root volume and fewer lateral roots and root forks compared to WT.

The root system of a plant is instrumental to its growth and productivity because it is responsible for the extraction of water and mineral nutrients from the soil and their transport to aboveground parts of the plant (*Hochholdinger, 1998*). In this study, the *35S:PpHSF5* transgenic lines produced smaller aerial organs compared with WT. For example, the size (length and width) of rosette leaves were smaller than WT two and three weeks after transplanting, while the number of rosette leaves was not affected. The height of the overexpression lines was significantly lower than that in WT, while the number of internodes was not. Overexpression of *OsHsfA2e* in rice caused a dwarf phenotype (*Yokotani et al., 2008*). In plants overexpressing *BhHsf1*, the reduced organ size was mainly attributed to decreased cell proliferation (*Zhu et al., 2009*). The overexpression of *PpHSF5* in peach suggested that the dwarf phenotype of transgenic plants was caused by shorter internodes.

It is still unknown how *PpHSF5* regulates root and aerial organs development. *PpHSF5* is homologous to *AtHSFB4* and thus may play similar roles in root development. Confocal laser scanning of roots in *AtHSFB4*-overexpression transgenic lines showed that ectopic division of

400 the lateral root cap cells (LRC) occurred (*Begum et al., 2013*). Previous studies indicated that
401 auxin acts in the production of lateral root primordium (LR) (*Casimiro et al., 2003; West et al.,*
402 *2004*). In the promoter of *PpHSF5*, there are three *cis-acting* regulatory elements that contain the
403 auxin-inducible TGACG-motif. Two auxin-inducible TGA-box elements in the *GmGH3*
404 promoter were strong binding sites of plant nuclear proteins and improved the auxin inducibility
405 of the *GmGH3* promoter (*Zhan-Bin et al., 1994*). Moreover, the HS assays indicated PpHSF5-
406 OE lines exhibited enhanced thermotolerance compared to WT. Similarity results were observed
407 in transgenic Arabidopsis plants with *AtHsfA3* and rice plants with *OsHsfA2e* (*Ogawa et al.,*
408 *2007;Yokotani et al., 2008*). Therefore, *PpHSF5* might be as a responsive factor for temperature
409 change and involved in auxin signal transduction due to the TGA motifs in its promoter and
410 might serve to negatively regulate root elongation and lateral root development, ultimately
411 affecting the growth of aboveground parts of the plant.

412

413 Conclusions

414 In this report, 18 *PpHSF* genes were discovered in peach and found to be nonuniformly
415 distributed on the peach chromosomes. The PpHSF family could be classified into three classes
416 (PpHSFA, PpHSFB and PpHSFC) through multiple alignment, motif analysis and phylogenetic
417 comparison. The expansion of the HSF gene family in peach occurred through DSD (67%) and
418 WGD (33%). *PpHSF5* was expressed in diverse tissues and organs of the peach cultivar
419 ‘Zhongyoutao 14’, with higher levels in young vegetative organs (leaf and apex). Transgenic

Arabidopsis lines overexpressing *PpHSF5* showed massively enhanced levels of *PpHSF5*. Ectopic expression *PpHSF5* repressed the length and number of roots, length and width of rosette leaves, and the height of plants, and enhanced thermtolerance in Arabidopsis after heat stress treatment. Our results further supplied functional and annotation information of the HSF gene family in general and revealed potential roles, outside of the response to heat stress, for *PpHSF5* during plant development.

Acknowledgements

The authors would like to thank Anita K. Snyder for critical reading the manuscript.

References

Begum T, Reuter R, Schoffl F. 2013. Overexpression of AtHsfB4 induces specific effects on root development of Arabidopsis. *Mech Dev,* **130(1)**: 54-60 DOI: 10.1016/j.mod.2012.05.008.

Busch W, Wunderlich M, Schoffl F. 2005. Identification of novel heat shock factor-dependent genes and biochemical pathways in Arabidopsis thaliana. *Plant Journal,* 41(1):1-14

DOI: 10.1111/j.1365-313X.2004.02272.x.

Casimiro I, Beeckman T, Graham N, Bhalerao R, Zhang H, Casero P, Bennett MJ. 2003. Dissecting Arabidopsis lateral root development. *Trends in Plant Science,* **8(4)**: 165-171.

DOI: 10.1016/s1360-1385(03)00051-7.

Chan-Schaminet KY, Baniwal SK, Bublak D, Nover L, Scharf KD. 2009. Specific interaction between tomato HsfA1 and HsfA2 creates hetero-oligomeric superactivator complexes for synergistic activation of heat stress gene expression. *Journal of Biological Chemistry,* **284(31)**: 20848-20857

DOI: 10.1074/jbc.M109.007336.

Chauhan H, Khurana N, Agarwal P, Khurana P. 2011. Heat shock factors in rice (*Oryza sativa L.*): Genome-wide expression analysis during reproductive development and abiotic stress. *Molecular Genetics and Genomics,* **286(2)**: 171-187 DOI: 10.1007/s00438-011-0638-8.

Chen C, Chen H, Zhang Y, Thomas H R, Frank M H, He Y, Xia R. 2020. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data[J]. Molecular Plant, **13(8)**:1194-1202.

DOI: 10.1016/j.molp.2020.06.009

Chung MH, Chen MK, Pan SM. 2000. Floral spray transformation can efficiently generate *Arabidopsis* transgenic plants. *Transgenic Research*, **9(6)**:471-486 DOI: 10.1023/A:1026522104478.

Dubrovsky JG, Gambetta GA, Hernandez-Barrera A, Shishkova S, Gonzalez I. 2006. Lateral root initiation in *Arabidopsis*: developmental window, spatial patterning, density and predictability. *Annals of Botany,* **97(5)**: 903-915 DOI: 10.1093/aob/mcj604.

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. The Pfam protein families database: towards a more sustainable future. *Nucleic acids research,* **44(D1)**: 279-285 DOI: 10.1093/nar/gkv1344.

Guo M, Liu JH, Ma X, Luo DX, Gong ZH, Lu MH. 2016. The plant heat stress transcription factors (HSFs): structure, regulation, and function in response to abiotic stresses. *Frontiers in Plant Science,* **7**: 114-127 DOI: 10.3389/fpls.2016.00114.

Guo M, Lu JP, Zhai YF, Chai WG, Gong ZH, Lu MH. 2015. Genome-wide analysis, expression profile of heat shock factor gene family (CaHsfs) and characterisation of CaHsfA2 in pepper (*Capsicum annuum L.*). *BMC Plant Biology,* **15(1)**: 151-171 DOI: 10.1186/s12870-015-0512-7.

Hahn A, Bublak D, Schleiff E, Scharf KD. 2011. Crosstalk between Hsp90 and Hsp70 chaperones and heat stress transcription factors in tomato. *Plant Cell,* **23(2)**: 741-755 DOI: 10.1105/tpc.110.076018.

He Y, Li W, Lv J, Jia Y, Wang M, Xia G. 2012. Ectopic expression of a wheat MYB transcription factor gene, *TaMYB73*, improves salinity stress tolerance in *Arabidopsis thaliana*. *Journal of Experimental Botany,* **63(3)**: 1511-1522 DOI: 10.1093/jxb/err389.

Hochholdinger F, Gunter Feix. 1998. Early post-embryonic development is specifically affected in the maize mutant *lrt1*. *The Plant Journal,* **16(2)**: 247-255 DOI: 10.1046/j.1365-313x.1998.00280.x.

Hu W, Hu G, Han B. 2009. Genome-wide survey and expression profiling of heat shock proteins and heat shock

factors revealed overlapped and stress specific response under abiotic stresses in rice. *Plant Science,* **176(4)**: 583-590 DOI: 10.1016/j.plantsci.2009.01.016.

Hu Y, Han YT, Wei W, Li YJ, Zhang K, Gao YR, Feng JY. 2015. Identification, isolation, and expression analysis of heat shock transcription factors in the diploid woodland strawberry *Fragaria vesca*. *Frontiers in Plant Science,* **6**: 736-751 DOI: 10.3389/fpls.2015.00736.

Kotak S, Vierling E, Baumlein H, von Koskull-Doring P. 2007. A novel transcriptional cascade regulating expression of heat stress proteins during seed development of *Arabidopsis*. *Plant Cell,* **19(1)**: 182-195 DOI: 10.1105/tpc.106.048165.

Koskull-Doring vP, Scharf KD, Nover L. 2007. The diversity of plant heat stress transcription factors. *Trends Plant Science,* **12(10)**: 452-457 DOI: 10.1016/j.tplants.2007.08.014.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Marra, MA. 2009. Circos: An information aesthetic for comparative genomics. *Genome Research,* **19(9)**: 1639-1645 DOI: 10.1101/gr.092759.109.

Kumar M, Busch W, Birke H, Kemmerling B, Nurnberger T, Schoffl F. 2009. Heat shock factors HsfB1 and HsfB2b are involved in the regulation of *Pdf1.2* expression and pathogen resistance in *Arabidopsis*. *Molecular Plant,* **2(1)**: 152-165 DOI: 10.1093/mp/ssn095.

Li PS, Yu TF, He GH, Chen M, Zhou YB, Chai SC, Ma YZ. 2014. Genome-wide analysis of the Hsf family in soybean and functional identification of *GmHsf-34* involvement in drought and heat stresses. *BMC Genomics,* **15(1)**: 1009-1025 DOI: 10.1186/1471-2164-15-1009.

Li W, Cui X, Meng Z, Huang X, Xie Q, Wu H, Liang W. 2012. Transcriptional regulation of *Arabidopsis MIR168a* and *ARGONAUTE1* homeostasis in abscisic acid and abiotic stress responses. *Plant Physiol,* **158(3)**: 1279-1292 DOI: 10.1104/pp.111.188789.

Li W, Wan X L, Yu J Y, Wang K L, Zhang J. 2019. Genome-wide identification, classification, and expression analysis of the Hsf gene family in Carnation (*Dianthus caryophyllus*). *International Journal of Molecular Sciences,* **20(20)**: 5233-5251 DOI: 10.3390/ijms20205233.

Lian X D, Tan B, Yan L, Jiang C, Cheng J, Zheng X B, Feng J C. 2020. Transcript profiling provides insights into molecular processes during shoot elongation in temperature-sensitive peach (*Prunus persica*). *Scientific Reports,* **10(1)**: 7801-7813 DOI: 10.1038/s41598-020-63952-2.

Liu G, Chai F, Wang Y, Jiang J, Duan W, Wang Y, Wang L. 2018. Genome-wide identification and classification of HSF family in Grape, and their transcriptional analysis under heat acclimation and heat

stress. *Horticultural Plant Journal,* **4(4)**: 133-143 DOI: 10.1016/j.hpj.2018.06.001.

Liu M, Huang Q, Sun W, Ma Z, Huang L, Wu Q, Tang, Z, Bu T, Li C, Chen H. 2019. Genome-wide investigation of the heat shock transcription factor (Hsf) gene family in Tartary buckwheat (*Fagopyrum tataricum*). BMC Genomics, **20(1)**: 871-888 DOI: 10.1186/s12864-019-6205-0.

Liu ZW, Wu ZJ, Li XH, Huang Y, Li H, Wang YX, Zhuang J. 2016. Identification, classification, and expression profiles of heat shock transcription factors in tea plant (*Camellia sinensis*) under temperature stress. *Gene,* **576(1 Pt 1)**: 52-59 DOI: 10.1016/j.gene.2015.09.076.

Liu B, Hu J, Zhang J. 2019. Evolutionary divergence of duplicated *Hsf* genes in *Populus*. *Cells*, **8(5)**: 438-463 DOI: 10.3390/cells8050438.

Lu ZH, Niu L, Chagné D, Cui GC, Pan L, Foster T, Zhang RP, Zeng WF, Wang ZQ. 2016. Fine mapping of the temperature-sensitive *semi-dwarf* (*Tssd*) locus regulating the internode length in peach *(Prunus persica*). Molecular Breeding, **36(2)**. DOI: 10.1007/s11032-016-0442-6

Nakai A. 1999. New aspects in the vertebrate heat shock factor system: Hsf3 and Hsf4. *Cell Stress and Chaperones*, **4(2)**: 86-93 DOI: 10.1054/csac.1998.0002.

Nishizawa-Yokoi A, Nosaka R, Hayashi H, Tainaka H, Maruta T, Tamoi M, Shigeoka S. 2011. HsfA1d and HsfA1e involved in the transcriptional regulation of *HsfA2* function as key regulators for the Hsf signaling network in response to environmental stress. *Plant Cell Physiol,* **52(5)**: 933-945 DOI: 10.1093/pcp/pcr045.

Nishizawa A, Yabuta Y, Yoshida E, Maruta T, Yoshimura K, Shigeoka S. 2006. Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *The Plant Journal,* **48(4)**: 535-547 DOI: 10.1111/j.1365-313X.2006.02889.x.

Nover L. 1996. The Hsf world: classification and properties of plant heat stress transcription factors. *Cell Stress and Chaperones,* **1(4)**: 215-224.DOI: 10.1379/1466-1268(1996)001<0215:THWCAP>2.3.CO;2.

Nover L, Bharti K, Doring P, Mishra SK, Ganguli A, Scharf KD. 2001. *Arabidopsis* and the heat stress transcription factor world: how many heat stress transcription factors do we need? *Cell Stress Chaperones,* **6(3)**: 177-189 DOI: 10.2307/1601759.

Ogawa D, Yamaguchi K, Nishiuchi T. 2007. High-level overexpression of the Arabidopsis *HsfA2* gene confers not only increased themotolerance but also salt/osmotic stress tolerance and enhanced callus growth. J Exp Bot, 58(12), 3373-3383. DOI: 10.1093/jxb/erm184

Petricka JJ, Winter CM, Benfey PN. 2012. Control of *Arabidopsis* root development. *Annual Review of Plant Biology,* **63(1)**: 563-590 DOI: 10.1146/annurev-arplant-042811-105501.

Qiao X, Li M, Li L, Yin H, Wu J, Zhang S. 2015. Genome-wide identification and comparative analysis of the heat shock transcription factor family in Chinese white pear (*Pyrus bretschneideri*) and five other Rosaceae species. *BMC Plant Biology,* **15(1)**: 12-28 DOI: 10.1186/s12870-014-0401-5.

Scharf KD, Berberich T, Ebersberger I, Nover L. 2012. The plant heat stress transcription factor (Hsf) family: structure, function and evolution. *Biochim et Biophys Acta,* **1819(2)**: 104-119

DOI: 10.1016/j.bbagrm.2011.10.002.

Schmidt R, Schippers JHM, Welker A, Mieulet D, Guiderdoni E, Mueller-Roeber B. 2012. Transcription factor OsHsfC1b regulates salt tolerance and development in *Oryza sativa* ssp. *japonica. AoB PLANTS,* **2012**: 1-17 DOI: 10.1093/aobpla/pls011.

Singh A, Mittal D, Lavania D, Agarwal M, Mishra RC, Grover A. 2012. OsHsfA2c and OsHsfB4b are involved in the transcriptional regulation of cytoplasmic *OsClpB (Hsp100)* gene in rice (*Oryza sativa L.*). *Cell Stress Chaperones,* **17(2)**: 243-254 DOI: 10.1007/s12192-011-0303-5.

Schmittgen T D, Livak K J. 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nature Protocols,* **3(6)**: 1101-1108 DOI: 10.1038/nprot.2008.73.

Sparkes IA, Runions J, Kearns A, Hawes C. 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols,* **1(4)**: 2019-2025 DOI: 10.1038/nprot.2006.286.

Spradling A, Drummond-Barbosa D, Kai T. 2001. Stem cells find their niche. *Nature,* **414(6859)**: 98-104 DOI: 10.1038/35102160.

Tan B, Lian X, Cheng J, Zeng W, Zheng X, Wang W, Feng J. 2019. Genome-wide identification and transcriptome profiling reveal that E3 ubiquitin ligase genes relevant to ethylene, auxin and abscisic acid are differentially expressed in the fruits of melting flesh and stony hard peach varieties. *BMC Genomics,* **20(1)**: 892-906 DOI: 10.1186/s12864-019-6258-0.

Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Rokhsar DS. 2013. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nature Genetics,* **45(5)**: 487-494 DOI: 10.1038/ng.2586.

Wang J, Sun N, Deng T, Zhang L, Zuo K. 2014. Genome-wide cloning, identification, classification and functional analysis of cotton heat shock transcription factors in cotton (*Gossypium hirsutum*). *BMC Genomics,* **15(1):** 961-979 DOI: 10.1186/1471-2164-15-961.

Wang X, Shi X, Chen S, Ma C, Xu S. 2018. Evolutionary origin, gradual accumulation and functional divergence

of heat shock factor gene family with plant evolution. *Frontiers in Plant Science,* **9**: 71-87. DOI: 10.3389/fpls.2018.00071.

Wang Y, Tang H, Debarry JD, Tan X, Li JP, Wang XY, Lee T, Jin HZ, Marler B, Guo H, Kissinger JC, Paterson AH. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*, **40(7)**, e49. DOI: 10.1093/nar/gkr1293

West G, Inzé D, Beemster GTS. 2004. Cell cycle modulation in the response of the primary root of Arabidopsis to salt stress. *Plant Physiol,* **135(2)**: 1050-1058 DOI: 10.1104/pp.104.040022.

Xue GP, Sadat S, Drenth J, McIntyre CL. 2014. The heat shock factor family from *Triticum aestivum* in response to heat and other major abiotic stresses and their role in regulation of heat shock protein genes. *Journal of Experimental Botany,* **65(2)**: 539-557 DOI: 10.1093/jxb/ert399.

Yang Z, Wang Y, Gao Y, Zhou Y, Zhang E, Hu Y, Xu C. 2014. Adaptive evolution and divergent expression of heat stress transcription factors in grasses. *BMC Evolutionary Biology,* **14(1)**, 147-158

DOI: 10.1186/1471-2148-14-147.

Yokotani N, Ichikawa T, Kondou Y, Matsui M, Hirochika H, Iwabuchi M, Oda K. 2007. Expression of rice heat stress transcription factor *OsHsfA2e* enhances tolerance to environmental stresses in transgenic Arabidopsis. Planta, **227(5)**, 957-967. DOI: 10.1007/s00425-007-0670-4

Yoshida T, Ohama N, Nakajima J, Kidokoro S, Mizoi J, Nakashima K, Yamaguchi-Shinozaki K. 2011. Arabidopsis *HsfA1* transcription factors function as the main positive regulators in heat shock-responsive gene expression. *Molecular Genetics and Genomics, 286*(5-6): 321-332

DOI: 10.1007/s00438-011-0647-7.

Yu X Y, Yao Y, Hong Y H, Hou P Y, Li C X, Xia Z Q, Chen Y H. 2019. Differential expression of the Hsf family in cassava under biotic and abiotic stresses. *Genome,* **62(8)**: 563-569 DOI: 10.1139/gen-2018-0163.

Zhan-Bin Liu TU, Xiangyang Shi, Gretchen Hagen, and Tom J. Guilfoyle. 1994. Soybean *GH3* promoter contains multiple auxin-inducible elements. *The Plant Cell,* **6(5)**: 645-657 DOI: 10.1105/tpc.6.5.645.

Zhang J, Jia H, Li J, Li Y, Lu M, Hu J. 2016. Molecular evolution and expression divergence of the Populus *euphratica Hsf* genes provide insight into the stress acclimation of desert poplar. *Scientific Reports,* **6(1)**: 30050-30063 DOI: 10.1038/srep30050.

Zhu X, Huang C, Zhang L, Liu H, Yu J, Hu Z, Hua W. 2017. Systematic analysis of Hsf family genes in the *Brassica napus* genome reveals novel responses to heat, drought and high CO_2 stresses. *Frontiers in Plant*

590 *Science,* **8**: 1174 DOI: 10.3389/fpls.2017.01174.

591

Figure 1

Chromosomal location of *HSF* genes in peach (*PpHSFs*).

Three syntenic pairs are linked by red lines.

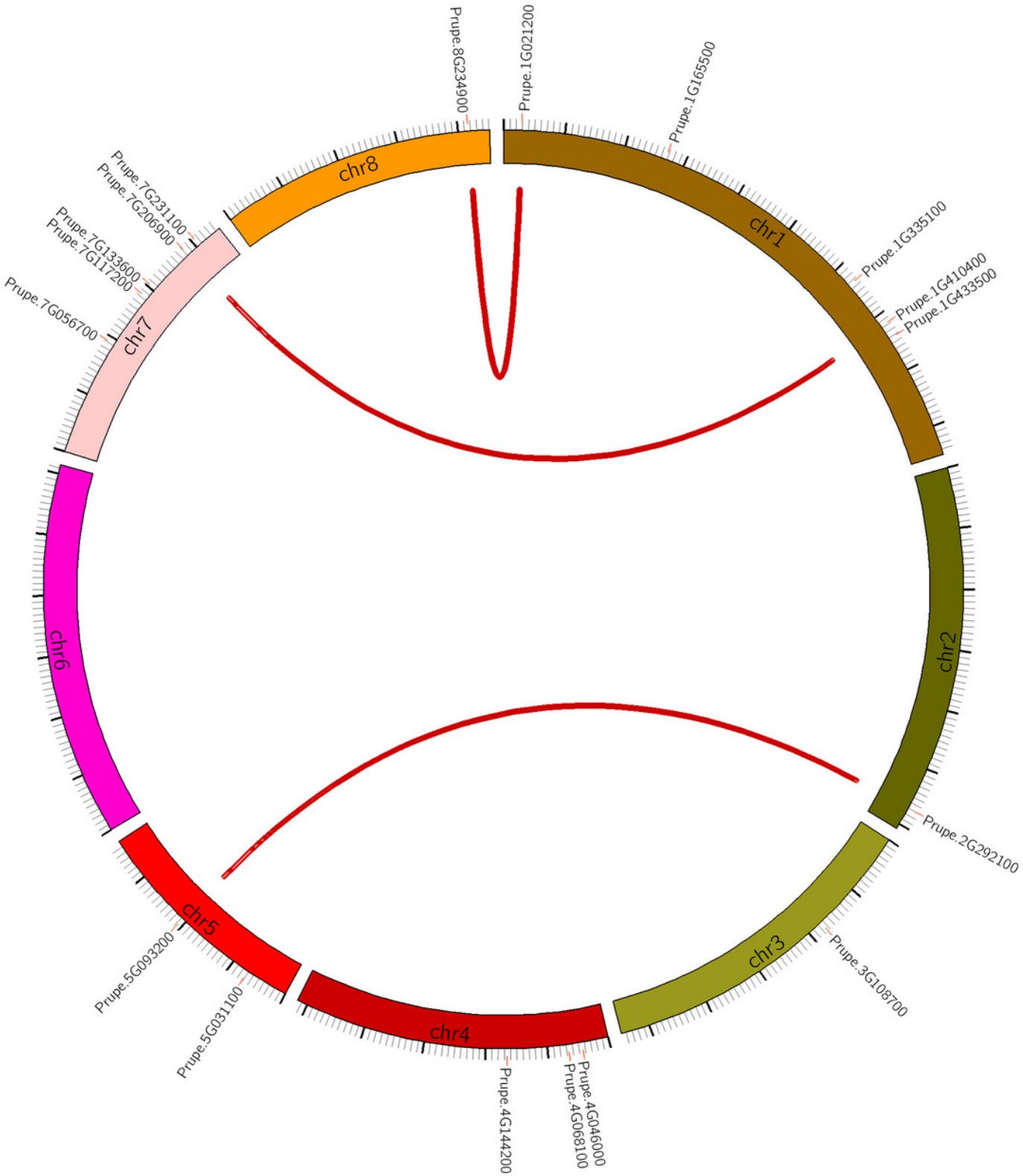
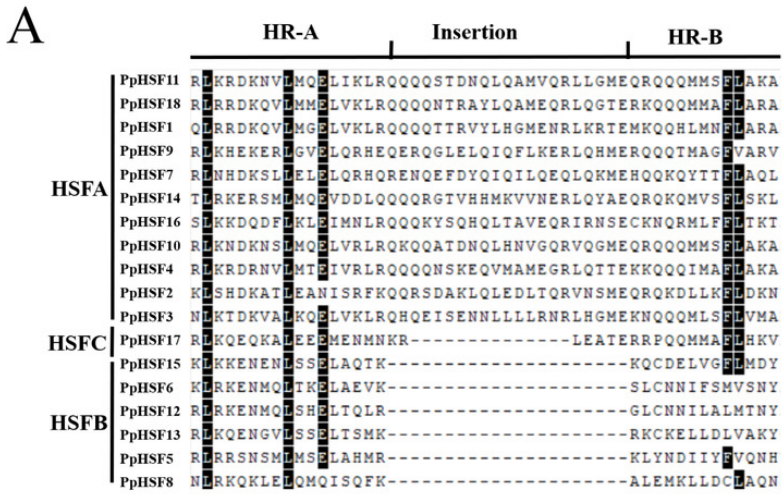
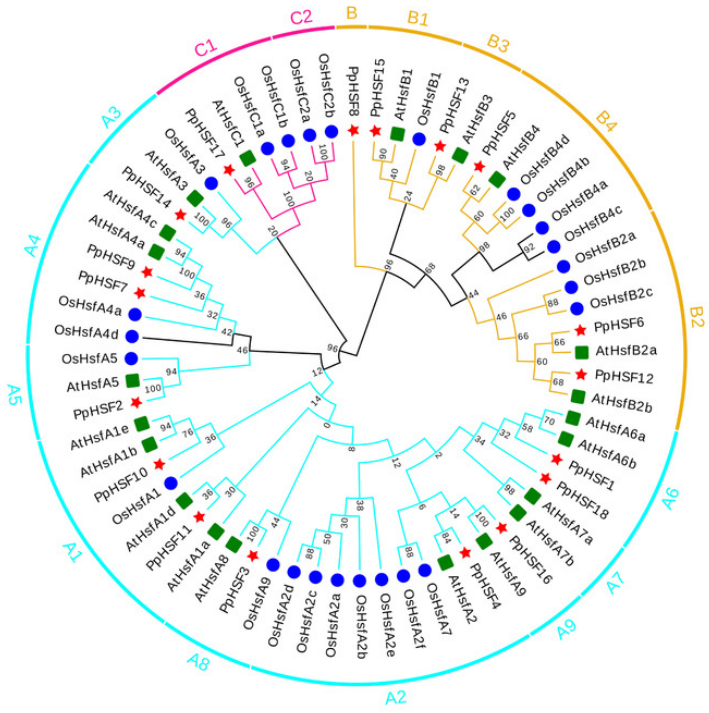

Prupe.8G234900
Prupe.1G021200
Prupe.1G165500
Prupe.1G335100
Prupe.1G410400
Prupe.1G433500
Prupe.2G292100
Prupe.3G108700
Prupe.4G046000
Prupe.4G068100
Prupe.4G144200
Prupe.5G031100
Prupe.5G093200
Prupe.7G056700
Prupe.7G117200
Prupe.7G133600
Prupe.7G206900
Prupe.7G231100
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8

Figure 2

Multiple sequence alignment of the HR-A/B regions (OD), conserved motif and phylogenetic analysis of PpHSFs

A. Multiple sequence alignment of the HR-A/B regions, from the start of the DNA-binding domain to the end of the HR-A/B region, of the HSF proteins were aligned with MEGA 6. **B.** hylogenetic tree of HSFs from *Prunus persica* (Pp, red star), *Oryza sativa* (Os, blue circle) and *Arabidopsis thaliana* (At, green square) constructed by maximum likelihood method (ML) with Jones-Taylor-Thornton (JTT) model in MEGA 6.0. Both locus ID and subclass numbers are listed. **C.** Analysis of conserved motifs in the HSF gene family in peach. Proteins are organized according to the groups in Figure 2A. Ten motifs were found in the protein sequences as shown in Supplemental Table S4.


A
HR-A
Insertion
HR-B
HSFA
HSFC
HSFB
PpHSF11
PpHSF18
PpHSF1
PpHSF9
PpHSF7
PpHSF14
PpHSF16
PpHSF10
PpHSF4
PpHSF2
PpHSF3
PpHSF17
PpHSF15
PpHSF6
PpHSF12
PpHSF13
PpHSF5
PpHSF8


B
A1
A2
A3
A4
A5
A6
A7
A8
A9
B
B1
B2
B3
B4
C1
C2

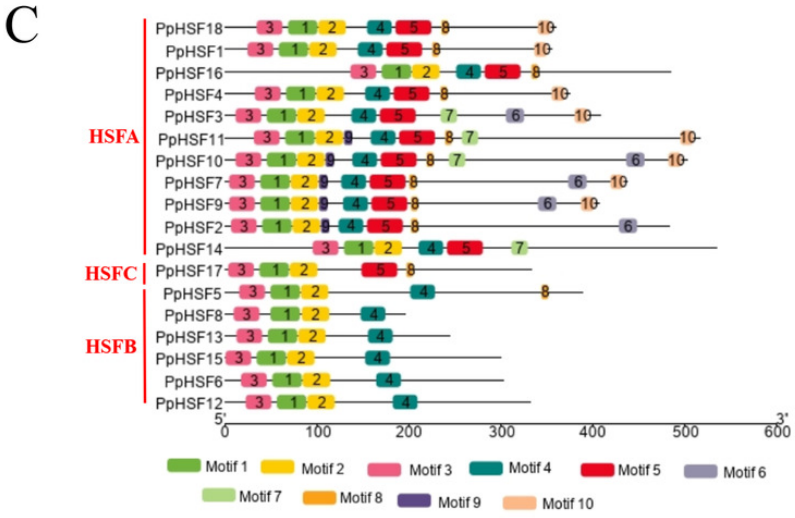

C
HSFA
HSFC
HSFB
5'
3'
0
100
200
300
400
500
600
Motif 1
Motif 2
Motif 3
Motif 4
Motif 5
Motif 6
Motif 7
Motif 8
Motif 9
Motif 10

Figure 3

Heatmap of transcript levels of HSF genes in peach.

Transcriptome data were used to measure the expression level of *PpHSFs*. The gene names on the right are organized according to the different subclasses. Samples were harvest from shoots at the IP (initial period), IEP (initial elongation period), RGP (rapid growth period), and SGP (stable growth period), which are four key growth stages during temperature-sensitive peach shoot development. Color scale at the top represents FPKM values. Blue indicates low expression and red indicates high expression. Heatmap was generated using TBtools.

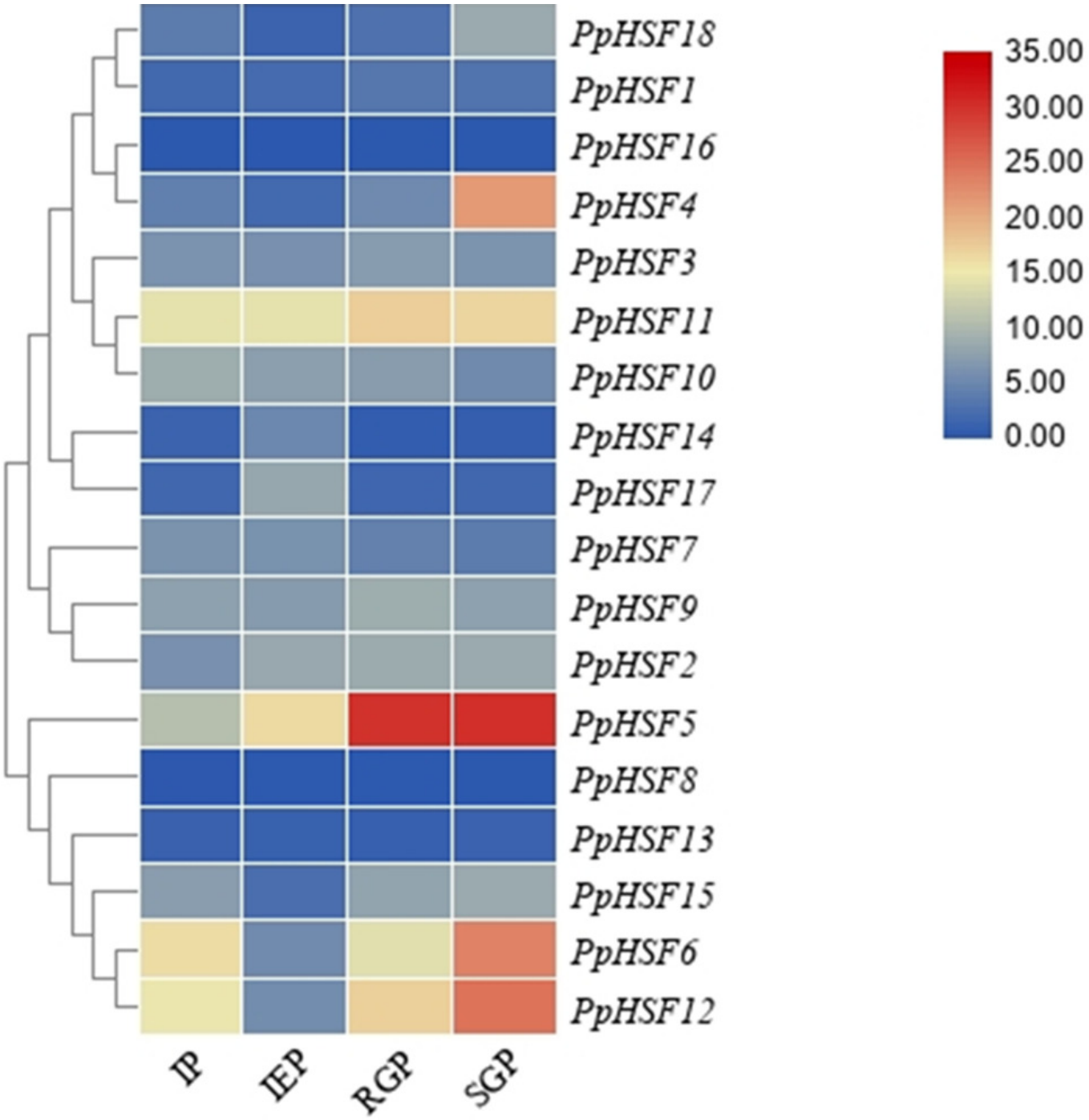

PpHSF18
PpHSF1
PpHSF16
PpHSF4
PpHSF3
PpHSF11
PpHSF10
PpHSF14
PpHSF17
PpHSF7
PpHSF9
PpHSF2
PpHSF5
PpHSF8
PpHSF13
PpHSF15
PpHSF6
PpHSF12
IP
IEP
RGP
SGP
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00

Figure 4

Relative expression of *PpHSF5* in different tissues of ‘Zhongyoutao 14’ peach.

Established plants were grown under normal conditions. The analyzed tissues include the apex, flower, embryo, young leaf, and mature leaf,which harvested at the same time. The relative expression levels were calculated using the $2^{-\Delta\Delta CT}$ method.

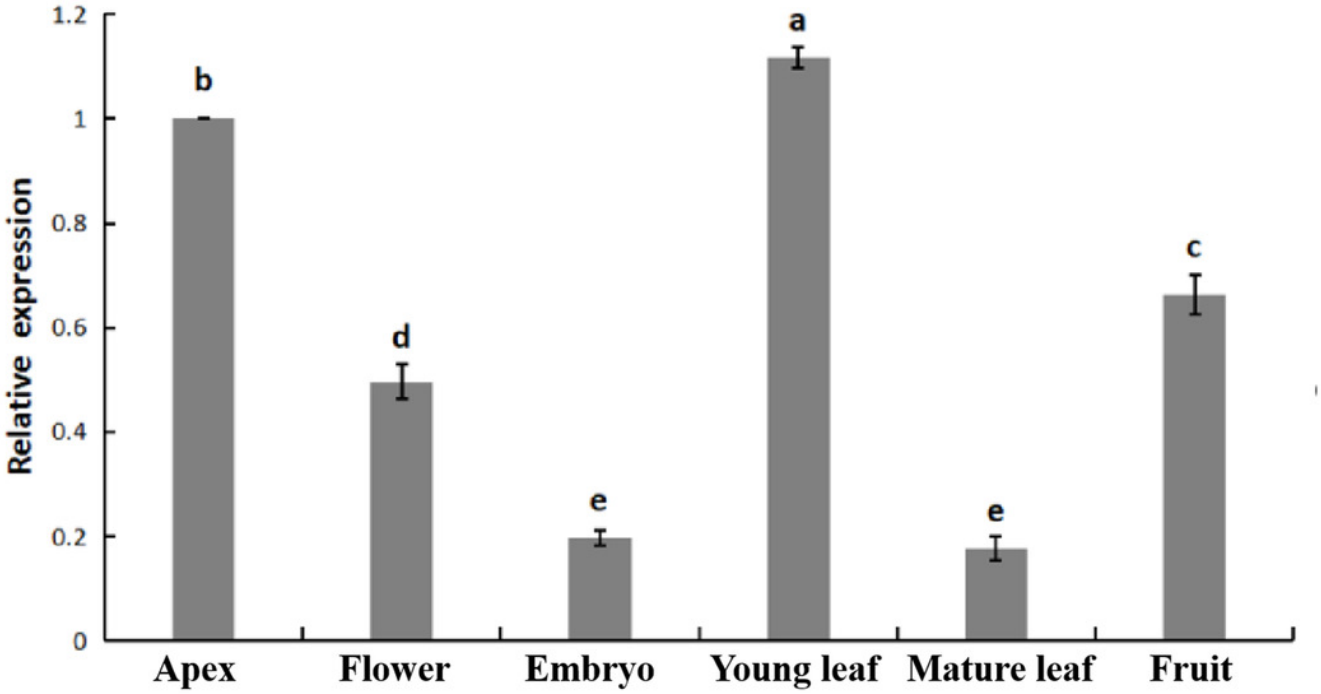


Figure 5

Subcellular localization of PpHSF5 in *N. benthamiana* epidermal cells.

A and **D**: Images of green fluorescence from the GFP protein and the PpHSF5-GFP fusion protein in tobacco cells under the confocal microscope; **B** and **E**: Bright field image of tobacco epidermal cells; **C**: Overlay of A and B; **F**: Overlay of D and E.

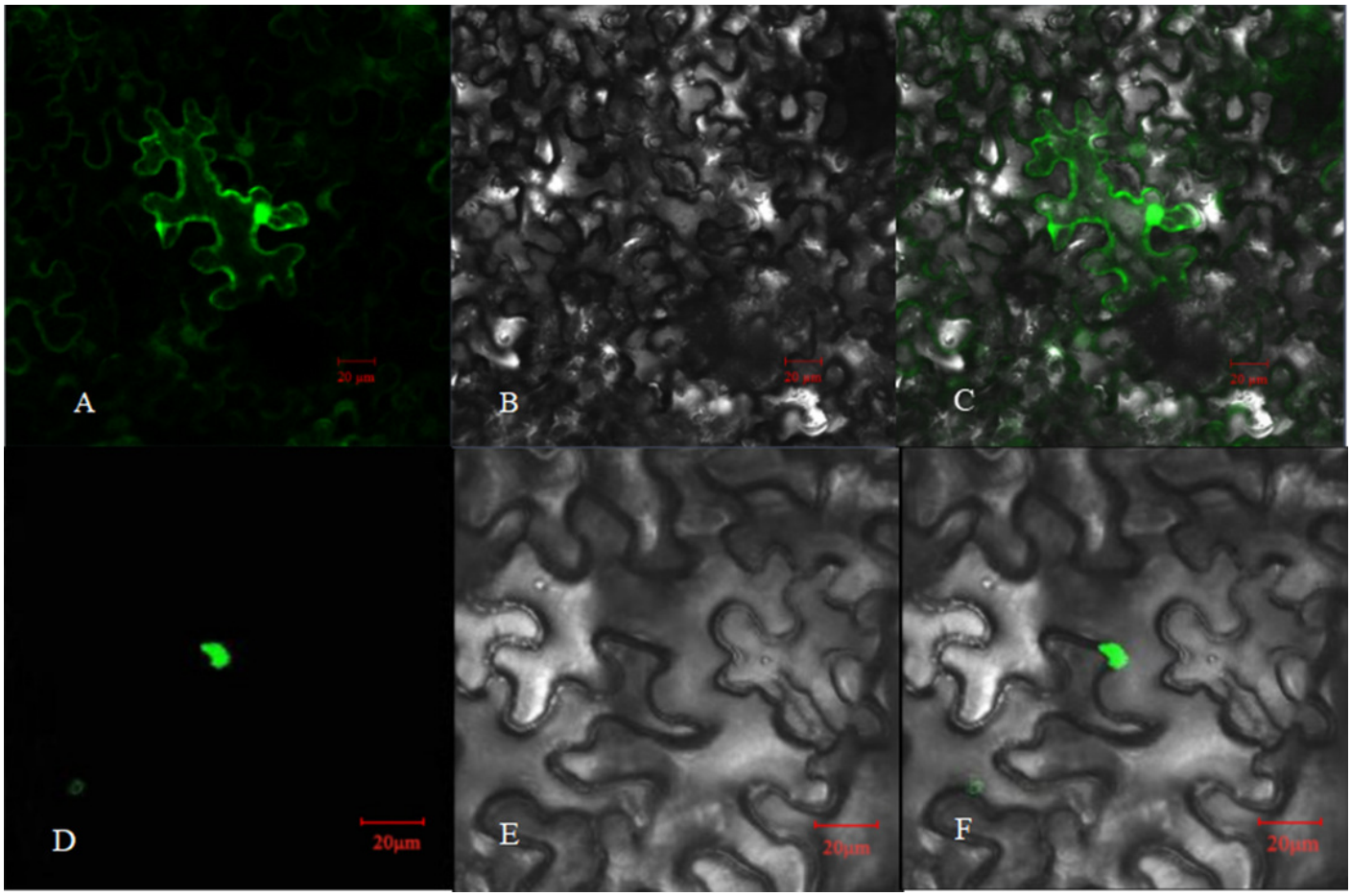


Figure 6

Phenotypic and expression analysis of transgenic Arabidopsis over-expressing *PpHSF5*

A. Phenotype of T_2 transgenic plants from two lines over-expressing *PpHSF5* after cultivation for one week. **B**. Root length of T_2 transgenic plants over-expressing *PpHSF5*. Three plants were measured in each biological replicate. **C**. Phenotype of T_2 transgenic plants from two lines over-expressing *PpHSF5* after cultivation for two weeks. Seeds were transferred to soil after germination and growth on agar for five days. **D**. The morphology and relative expression of T_2 transgenic plants with *PpHSF5* and WT after cultivation for two weeks. The number of rosettes (a), length (b) and width (c) after cultivation for two weeks. Relative expression of *PpHSF5* in transgenic Arabidopsis plants carrying *p35S*:*PpHSF5* (d); **E**. Phenotype of T_2 transgenic plants over-expressing *PpHSF5* after cultivation in soil for three weeks. **F**. The length, width and number of rosette leaves, number of internodes and flower stalks, and the height of plants after cultivation for three weeks. **G**. Phenotype of T_2 transgenic plant roots over-expressing *PpHSF5* after cultivation for three weeks. **H**. The root length, volume and other indexes were scanned after cultivation for three weeks.

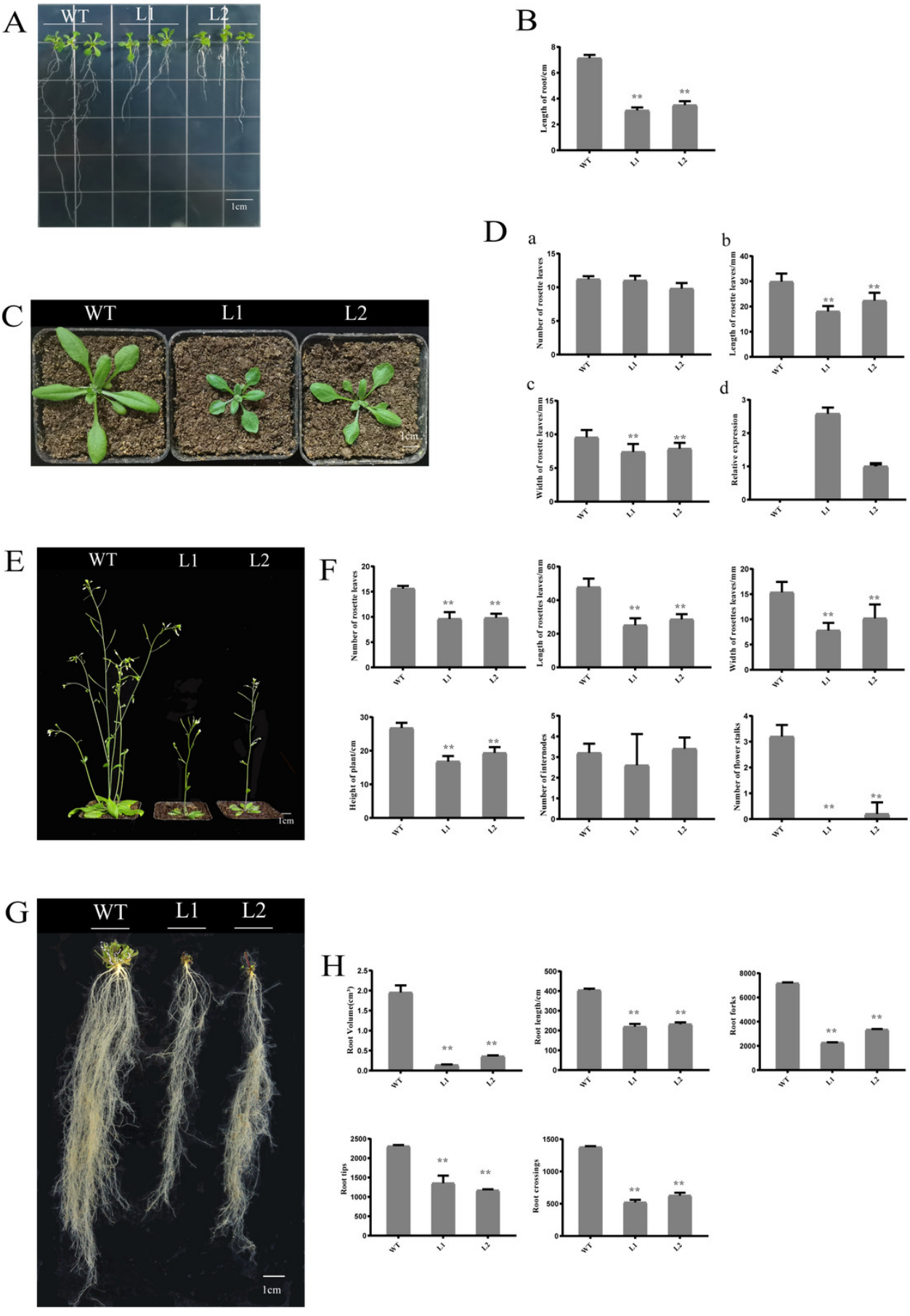

A
WT
L1
L2
1cm
B
Length of root/cm
C
WT
L1
L2
1cm
D
a
Number of rosette leaves
b
Length of rosette leaves/mm
c
Width of rosette leaves/mm
d
Relative expression
E
WT
L1
L2
1cm
F
Number of rosette leaves
Length of rosettes leaves/mm
Width of rosettes leaves/mm
Height of plant/cm
Number of internodes
Number of flower stalks
G
WT
L1
L2
1cm
H
Root Volume(cm³)
Root length/cm
Root forks
Root tips
Root crossings

Figure 7

Thermotolerance of the *p35S::PpHSF5* plants.

A: Five-day-old seedlings of wild type and the *p35S::PpHSF5* plants were treated at 46 ℃ for 30 min. Photographs were taken before HS treatment. **B.** Photographs were taken after 1d in 22 ℃. **C:** Photographs were taken after 3d in 22 ℃. **D:** Photographs were taken after 5d in 22 ℃.**E:** Photographs were taken after 7d in 22 ℃. **F:** Comparison of germination rate among wild-type, *p35S::PpHSF5* transgenetic plants after HS treatment. The number of germinated plants was counted daily after HS treatment. For three replication, more than 50 seedlings were used each lines (*t*-test signifificant at $P < 0.05$ and $P < 0.01$, respectively)**.**

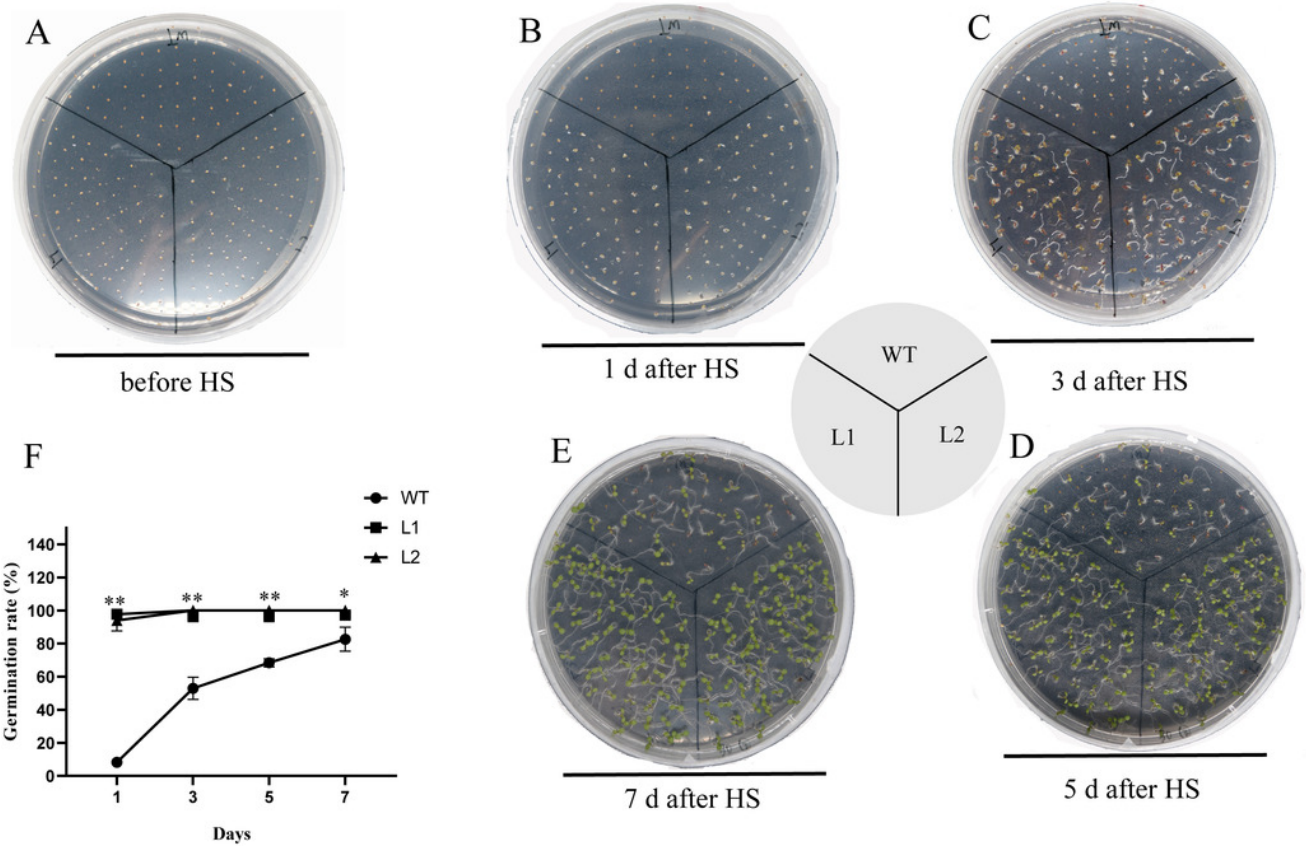


Table 1(on next page)

Basic information of *PpHSF* gene family members

Gene name	Gene ID	Length of CDS (bp)	No. of amino acids (aa)	Molecular weight (Da)	Predicted isoelectric point (PI)	Chromosome location	Subcellular localization
PpHSF1	Prupe.1G021200	1068	355	41166.82	4.9	1	Nucleus
PpHSF2	Prupe.1G165500	1452	483	54128.99	5.61	1	Nucleus
PpHSF3	Prupe.1G335100	1227	408	46017.53	4.81	1	Nucleus
PpHSF4	Prupe.1G410400	1125	374	41952.84	4.95	1	Nucleus
PpHSF5	Prupe.1G433500	1170	389	43557.76	7.17	1	Nucleus
PpHSF6	Prupe.2G292100	912	303	33900.98	5.19	2	Nucleus
PpHSF7	Prupe.3G108700	1314	437	49855.78	5.13	3	Nucleus
PpHSF8	Prupe.4G046000	591	196	22364.46	8.75	4	Nucleus
PpHSF9	Prupe.4G068100	1224	407	46118.31	5.23	4	Nucleus
PpHSF10	Prupe.4G144200	1512	503	56052.61	4.78	4	Nucleus
PpHSF11	Prupe.5G031100	1551	516	56261.49	4.67	5	Nucleus
PpHSF12	Prupe.5G093200	996	331	36068.73	4.75	5	Nucleus
PpHSF13	Prupe.7G056700	735	244	28001.34	5.75	7	Nucleus
PpHSF14	Prupe.7G117200	1608	535	59567.89	4.98	7	Nucleus
PpHSF15	Prupe.7G133600	900	299	33339.83	5.07	7	Nucleus
PpHSF16	Prupe.7G206900	1458	485	53632.01	5.07	7	Nucleus
PpHSF17	Prupe.7G231100	1002	333	37851.52	5.68	7	Nucleus
PpHSF18	Prupe.8G234900	1080	359	40936.56	5.58	8	Nucleus

1

Table 2(on next page)

Cis-elements in the promoters of eighteen *PpHSF* genes

	ABRE	ARE	CGTCA-motif	ERE	MBS	MRE	MYB	MYC	P-box	TGACG-motif	TCA-element	LTR	TGA-element
PpHSF1	1	3	2	2	1	1	4	3	2	2	2	-	-
PpHSF2	1	2	2	1	-	-	3	1	-	2	1	3	1
PpHSF3	1	1	-	2	1	1	4	2	-	1	-	1	1
PpHSF4	3	6	3	-	-	-	3	-	1	3	-	2	-
PpHSF5	5	-	3	-	2	-	7	-	1	3	3	-	-
PpHSF6	4	-	3	-	1	-	7	5	2	3	1	-	-
PpHSF7	1	5	-	1	-	-	6	3	-	-	-	1`	1
PpHSF8	3	3	3	1	1	1	3	8	1	3	-	-	3
PpHSF9	5	5	4	-	-	1	13	4	-	4	-	2	2
PpHSF10	3	1	5	2	-	-	2	6	-	5	-	-	2
PpHSF11	-	1	1	-	2	-	7	8	-	1	-	-	-
PpHSF12	6	-	1	1	1	1	13	3	-	1	-	-	1
PpHSF13	3	4	1	-	-	-	4	5	-	1	1	1	-
PpHSF14	3	4	3	-	3	-	11	3	-	3	-	1	-
PpHSF15	8	2	4	1	1	1	6	5	-	4	-	-	-

PpHSF16	4	4	1	1	-	2	4	11	-	-	-	-	-
PpHSF17	11	2	2	1	1	-	4	6	-	2	-	-	-
PpHSF18	13	3	-	1	1	2	2	4	-	-	-	1	-

1