Genome-wide analysis and expression profiling of the heat shock transcription factor gene family in Physic Nut (*Jatropha curcas* L.)

Lin Zhang, Wei Chen and Ben Shi

School of Environmental Engineering and Chemistry, Luoyang Institute of Science and Technology, Luoyang, Henan, China

ABSTRACT

The heat shock transcription factor (Hsf) family, identified as one of the important gene families, participates in plant development process and some stress response. So far, there have been no reports on the research of the *Hsf* transcription factors in physic nut. In this study, seventeen putative *Hsf* genes identified from physic nut genome. Phylogenetic analysis manifested these genes classified into three groups: A, B and C. Chromosomal location showed that they distributed eight out of eleven linkage groups. Expression profiling indicated that fourteen *JcHsf* genes highly expressed in different tissues except *JcHsf1*, *JcHsf6* and *JcHsf13*. In addition, induction of six and twelve *JcHsf* genes noted against salt stress and drought stress, respectively, which demonstrated that the *JcHsf* genes are involved in abiotic stress responses. Our results contribute to a better understanding of the *JcHsf* gene family and further study of its function.

INTRODUCTION

The heat shock transcription factor (Hsf) as the direct transcriptional activator of genes plays an important role in regulating plant growth, development and in response to abiotic stresses, such as heat, drought and salt stress (*Von Koskull-Döring, Scharf & Nover, 2007*). Under heat stress condition, Hsfs functioned as the molecular chaperone protects cells structure in protein folding and assembly process (*Hu, Hu & Han, 2009*).

Hsfs well-known as a group of DNA-binding proteins, specially recognize the binding motifs AGAAnnTTCT which conserved in the promoter regions of Hsf-inducible genes (*Guo et al., 2016*). Despite a considerable variety in size and sequence, the main domains of Hsf in eukaryotes are conservative. A typical Hsf was composed of the DNA binding domain (DBD), the oligomerization domain (OD), nuclear localization signal (NLS), nuclear export signal (NES), repressor domain (RD) and C-terminal activator peptide motif (AHA) (*Nover et al., 2001*; *Guo et al., 2016*). Since the first *Hsf* gene was cloned and characterized in the yeast (*Wiederrecht, Seto & Parker, 1988*), more and more Hsf proteins had been cloned and characterized in plants. The plants Hsfs were separated into three groups (A, B and C), and further divided into 16 subgroups (A group: A1-A9; B group:

Corresponding author
Wei Chen, parddy@hotmail.com

B1-B5; C group: C1-C2) (*Nover et al., 2001*; *Hu, Hu & Han, 2009*; *Scharf et al., 2012*). Both AHA and adjacent NES motifs in the C-terminal are specific of Hsf A, but absent in Hsf B and Hsf C (*Nover et al., 2001*). Increasing research indicated that HSF is involved in plant development, biotic and abiotic stress. For example, overexpression of *HsfA1a, HsfA1b, HsfA2* or *Hsf3* in *Arabidopsis* promoted its stress tolerance (*Qian et al., 2014*; *Albihlal et al., 2018*; *Ogawa, Yamaguchi & Nishiuchi, 2007*; *Prandl et al., 1998*). Similar stress tolerance phenotypes were also found in tobacco (*Personat et al., 2014*). In contrast, the *hsfa1a/b/d/e* quadruple mutant exhibited a severe growth retardation phenotype (*Yoshida et al., 2011*). Additionally, AtHsfA6a-overexpressing *Arabidopsis* demonstrate a significantly decreased germination under ABA treatment (*Hwang et al., 2014*). It is worth noting that *HsfA2*, which shared similar structure with *HsfA1* but with a different expression profile, was strongly induced under long-term or cycled heat stress conditions in *Arabidopsis* (*Charng et al., 2007*; *Nishizawa-Yokoi et al., 2009*). This indicated that even the orthologous genes play a divergent role in different plants.

Based on released genome sequences, genome-wide analyses of Hsf family performed in various plant species, such as *Arabidopsis* (*Guo et al., 2008*), rice (*Guo et al., 2008*; *Mittal et al., 2009*; *Hu, Hu & Han, 2009*), pepper (*Guo et al., 2015*), Chinese cabbage (*Huang et al., 2015*), maize (*Lin et al., 2011*), *Triticum aestivum* (*Xue et al., 2014*), tea plant (*Camellia sinensis*) (*Liu et al., 2016*), soybean (*Glycine max*) (*Chung, Kim & Lee, 2013*), *Brassica oleracea* (*Lohani et al., 2019*), *Salix suchowensis* (*Zhang et al., 2015a*) and sesame (*Dossa, Diouf & Cisse, 2016*). These investigations are of great significance for the function analysis of *Hsfs* refers to plant development and in response to abiotic stresses.

Physic nut (*Jatropha curcas* L.), which is well-known as a renewable resource for biodiesel production, has a high tolerance to drought and salt stress (*King et al., 2009*; *Divakara et al., 2010*). To date, the genome sequence and expression profile of physic nut under several abiotic stresses were available (*Wu et al., 2015*; *Zhang et al., 2014*; *Zhang et al., 2015b*). Therefore, the Hsf family in physic nut could be characterized at the molecular level. However, no detailed study of *JcHsf* family genes performed. In this study, a total of 17 putative *JcHsf* genes were identified. The classification, phylogenetic reconstruction, chromosome distribution, gene structure and conserved motifs of the *JcHsfs* were predicted and analyzed. In addition, the expression profile of *JcHsf* genes was analyzed under normal condition and in response to salt and drought stress. Taken together, these results these results provide important information for further study of functional genes in the physic nut Hsf family and will provide insights into the functional analysis of other plants *Hsf* genes.

MATERIALS & METHODS

Plant materials and treatment

The inbred physic nut cultivar GZQX0401 were used in this study. The seed germination, planting conditions, stress treatment and material collection were performed as described previously (*Zhang et al., 2014*; *Zhang et al., 2015b*).

RNA isolation and gene expression analysis

Total RNA was extracted from ∼100 mg leave samples using the modified CTAB method (*Zhang et al., 2014*). The RNA was then reverse transcript into cDNA by M-MLV reverse transcriptase (Promega). For qRT-PCR analysis, primers used in this study were designed by primer 6.0 (Table S1). Each reaction was performed using TaKaRa Ex Taq HS Kit according to the instructions of the manufactor. The expression levels were calculated using the $2^{-\Delta\Delta CT}$ method. Each PCR assay was run in duplicate for three independent biological replicates. For gene expression profiles analysis, the number of expressed tags was calculated and then normalized to TPM (number of transcripts per million tags). Heatmap of Hsf family expression in different tissues was constructed based on the number of TPMs. While the heatmap of Hsf family expression under salt and drought stress was constructed by log2 conversion of TPM values.

Identification and characteristics of *Hsf* genes in physic nut

To identify the *Hsf* family genes of physic nut, the HSF protein sequences of in *Arabidopsis* and rice was used as query sequences to execute BLASTP search against the physic nut genome. Target gene sequences were selected with *e*-value cut-off less than $1e^{-10}$. The Hsf protein sequences of *Arabidopsis* and rice were downloaded from TAIR (http://www.arabidopsis.org/) and the website (http://plntfdb.bio.uni-potsdam.de/v3.0/), respectively. In addition, the HMM file built based on Hsf domain (PF00447) was used to perform HMM searches against the local protein databases of the physic nut by HMMER3 (*Eddy & Pearson, 2011*). In order to confirm the accuracy of identified genes, the SMART program was used to detect DBD domains and coiled-coil structures (SMART: http://smart.embl-heidelberg.de/). Those protein sequences lacking the DBD domain or a coiled-coil structure were removed. Next, Clustal X was aligned to remove redundant sequences of the confirmed JcHsf sequences (*Larkin et al., 2007*). Finally, the online ExPasy program (http://www.expasy.org/tools/) was applied to analysis the length, molecular weight and isoelectric point parameters of each JcHsf protein.

Structure and motif analysis of *JcHsf* genes

The exon/intron structures of *JcHsf* genes were elucidated by Gene Structure Display Server (GSDS, http://gsds.cbi.pku.edu.cn/) (*Guo et al., 2007*). Conserved motifs were analyzed using the MEME Suite version 5.0.5 (http://meme-suite.org/tools/meme; *Bailey et al., 2009*). The parameters were set as the default value except the numbers of different motifs, which was set to 30.

Chromosomal locations, multiple sequence alignment and phylogenetic analysis of *JcHsf* genes

All identified *JcHsf* genes were mapped on the eight out of eleven linkage group base on physic nut genome database. The maximum likelihood mapping algorithm and Kosambi mapping function are used to calculate the map distance within cM (*Wu et al., 2015*). The linkage map of *JcHsf* genes was mapped using the Map-Chart software package. Amino acid sequences of Hsf protein were aligned by Clustal X (version 1.83). GeneDoc was then used to manually edit the results. The phylogenetic tree was constructed among

Table 1 Accession members and characteristics of 17 *JcHsf* genes in physic nut.

Genes	Gene ID	Protein length (aa)	pI	MW (kDa)	Location
JcHSF1	JCGZ_02229	359	5.42	41.4	LG1
JcHSF2	JCGZ_01081	495	5.93	55.2	LG1
JcHSF3	JCGZ_17108	517	4.81	56.3	LG3
JcHSF4	JCGZ_04789	358	4.74	41.1	LG3
JcHSF5	JCGZ_21617	406	5.04	46.2	LG3
JcHSF6	JCGZ_21433	214	9.05	24.8	LG3
JcHSF7	JCGZ_26137	311	5.44	34.4	LG5
JcHSF8	JCGZ_21103	242	7.59	28.3	LG6
JcHSF9	JCGZ_06744	560	4.77	62.3	LG6
JcHSF10	JCGZ_02539	497	5.19	55.2	LG6
JcHSF11	JCGZ_12391	288	5.54	31.2	LG7
JcHSF12	JCGZ_21870	357	5.15	40.8	LG8
JcHSF13	JCGZ_03936	425	4.91	48.5	LG9
JcHSF14	JCGZ_20639	333	5.45	36.6	LG9
JcHSF15	JCGZ_07238	320	5.94	35.7	LG11
JcHSF16	JCGZ_07430	473	4.85	52.5	LG11
JcHSF17	JCGZ_07655	380	4.67	42.7	LG11

Arabidopsis thaliana, *Oryza sativa*, *vitis vinifera* and *Jatropha curcas* L. using the maximum likelihood method in MEGA5 with 1,000 replicates (*Tamura et al., 2011*). The HSF proteins from four plant species were listed in Table S2.

RESULTS

Identification and chromosomal localization of *JcHsf* gene family

Seventeen (17) putative genes of *JcHsf* family were identified in the present study, and they were named *JcHsf1-17* from top to bottom according to their position on the physic nut linkage groups (LGs) 1 to 11. The accession numbers in GenBank and detailed information of *JcHsf* gene family were listed in Table 1. The JcHsf protein lengths ranged from 214 aa (JcHsf 7) to 560 aa (JcHsf 11), The molecular masses of the JcHsf proteins were predicted between 28.3 to 62.3 KDa, and the theoretical pIs were ranged from 4.67 to 9.05 (Table 1). Chromosomal location indicated that no *JcHsf* was mapped on LGs 2, 4 and 10, and 17 *JcHsf* genes were not randomly distributed in the other 8 LGs (*Wu et al., 2015*) (Fig. 1). Large difference in number of *JcHsf* genes located in each LG. The result showed four *JcHsf* genes on LG3, three on LG6 and 11, two on LG1 and 9, and one on LG5, 7 and 8 (Fig. 1).

Phylogenetic and structures analysis of *JcHsf* gene family

To survey the evolutionary relationships of *JcHsf* gene family, protein sequences from other three well-studied and representative species, including a dicot *Arabidopsis*, a monocot rice and grape, were selected to construct a phylogenetic tree (Fig. 2). Similar to the Hsfs from other three plant species, JcHsfs were divided into three groups (Group
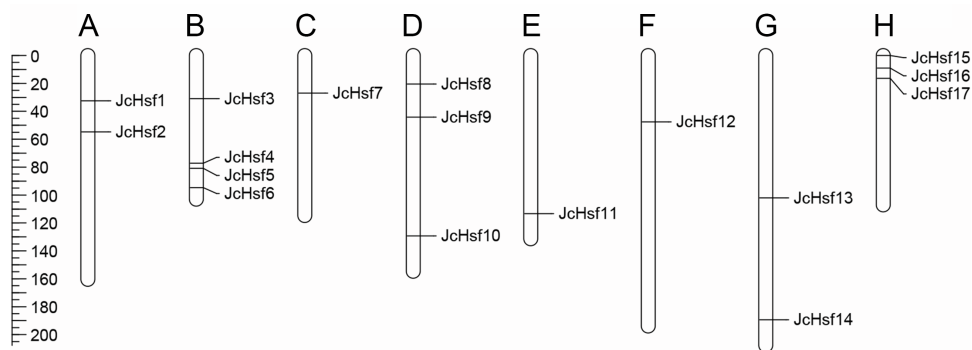
Figure 1 Distribution of *JcHsf* genes on physic nut chromosomes according to the linkage map. (A) Linkage group 1. (B) Linkage group 3. (C) Linkage group 5. (D) Linkage group 6. (E) Linkage group 7. (F) Linkage group 8. (G) Linkage group 9. (H) Linkage group 11. In total, 17 *JcHsf* genes were mapped to eight linkage groups (LGs). The scale is in centiMorgans (cM).

A, B and C) according to previous study on AtHsfs (*Guo et al., 2008*). The largest group A was then classified into nine subgroups (A1-A9), including 11 members (JcHsf1, JcHsf2, JcHsf3, JcHsf4, JcHsf5, JcHsf9, JcHsf10, JcHsf12, JcHsf13, JcHsf16 and JcHsf17), accounted for 64.7% of total JcHsfs. The next group B was classified into four subgroups (B1-B4), with five members (JcHsf6, JcHsf7, JcHsf8, JcHsf11 and JcHsf14), accounted for 29.4%. Only JcHsf15 was classified into the smallest group C, which represented 5.9%. Results showed a close evolutionary relationship with dicotyledons plants.

During the evolution process, the exon/intron structural within a gene family appears divergence, which of great significance to study the evolution of gene families. So the exon/intron structures of *JcHsf* genes were analyzed in the present study. It was shown that all *JcHsfs* had only one intron, illustrating a very highly conserved exon/intron splicing arrangement exists in physic nut (Fig. 3). Although physic nut *Hsf* genes shared same intron number, the length of intron was different in the groups. For instance, in the subgroup A1, the intron length of *JcHsf3* was much longer than that of *JcHsf10*.

Conserved domains and motifs of JcHSF proteins

To well investigate the gene structure in the aspects of functional divergence, the conserved motifs/domains of the JcHsfs were analyzed based on the reported information about AtHsfs and OsHsfs (*Guo et al., 2008*), which contribute to discern a serious about putative functional domains for all JcHsfs. DBD, HR-A/B, NLS, NES, AHA, and RD conserved domains were confirmed in JcHsf proteins (Table 2). Multiple alignment analyses clearly showed that the DBD, most conserved section of JcHsfs, was located in the N-terminal region (Fig. 4). Based on the differences between the HR-A and HR-B regions, three types of Hsfs were identified in the physic nut (Fig. 5).

The motif of JcHsf protein was also analysis using the MEME motif search tool (Fig. 6). Among the ten detected motifs, motifs 1, 2, and 4 including the highly conserved DBD existed in all the JcHsfs (Fig. 7). The motif 6 representing the HR-A/B region was found in all class B JcHsfs, while which was replaced by motif 3 in classes A and C. Most JcHsf
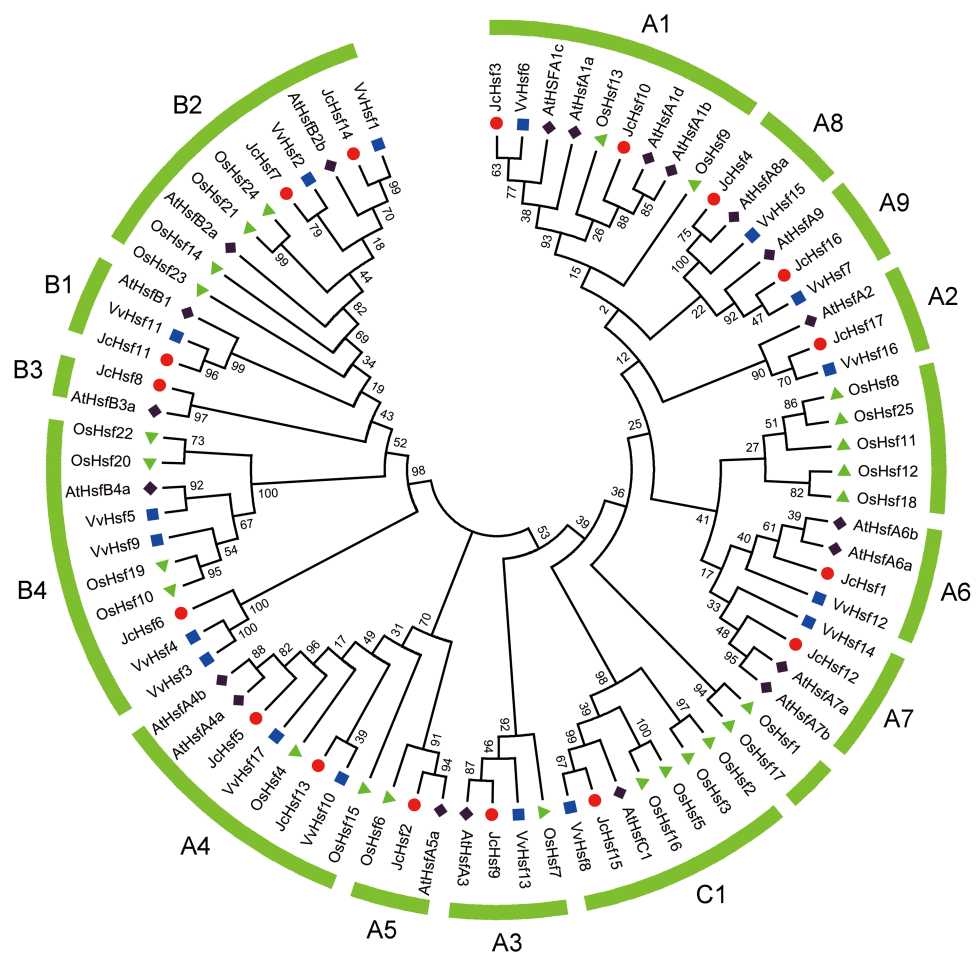
Figure 2 **Phylogenetic relationships of *Arabidopsis*, rice, grape and physic nut Hsf proteins.** The unrooted tree was constructed, using the MEGA5.0 program, by the neighbor-joining method. Bootstrap values were calculated for 1,000 replicates.

proteins contain Motifs 8 representing NLS except JcHsf4 and JcHsf9. Six proteins in the subgroup A1, A2, A4 and A6 had motif nine which consisted of NES. In additionally, motifs 10 containing AHA were discovered in the C-terminus of most groups A JcHsfs. Moreover, some unknown motifs were also detected in this study.

Expression profiles of *JcHsf* genes in different tissues

To gain insight into the function of *JcHsf* genes, their expression patterns in different tissues were analyzed. Results showed that *JcHsf1* was not expressed in all the tissues, *JcHsf6* and *JcHsf13* with low expression levels were shown in different tissues and developing seeds. Of the other fourteen *JcHsf* genes, four (*JcHsf3*, *5*, *11* and *8*), two (*JcHsf7*, *17*) and three (*JcHsf2*, *9* and *12*) showed the highest expression level in the root, stems and leaves, respectively. Four (*JcHsf4*, *9*, *10* and *15*) showed a much higher expression level in the seed early developing stage (S1–S4). Six (*JcHsf2*, *11*, *14* and *16*) showed much higher
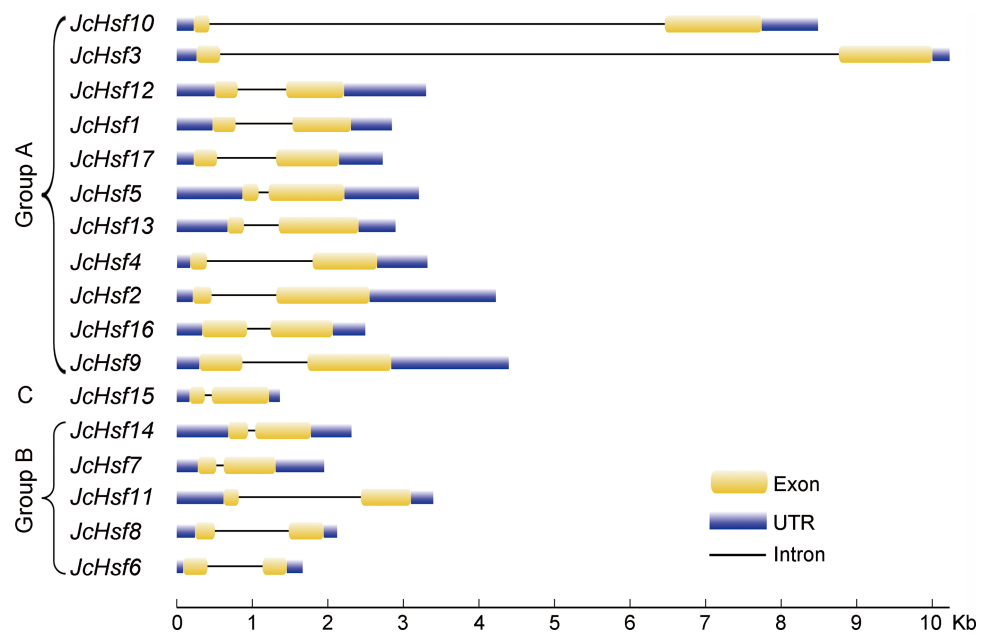
Figure 3 **Intron-exon structure of *JcHsf* genes.** Yellow boxes indicate the exon regions, black lines indicate introns and blue boxes indicate upstream and downstream UTR regions. The lengths of the boxes and lines were scaled according to the length of the genes.

Table 2 **Functional domains of physic nut *Hsf* genes.**

Gene name	Group	DBD	NLS	NES	AHA	RD
JcHsf1	A6	41-134	(226)KDKMK7KKRRR	(347)LVEQLGYL	(319)EGFWDDLLNE	
JcHsf2	A5	23-116	(216)KKRR	(361)LNLTL	(443)DVFWEQFLTE	
JcHsf3	A1	44-137	(242)NRR5KKRRLK	(503)LTEKMGLL		
JcHsf4	A8	12-106	(205)KENNWR		(298)DSAWEQFLL	
JcHsf5	A4	11-104	(206)RKRR	(162)MQALKE	(342)DVFWEQFL	
JcHsf9	A3	130-223	(324) RMNRKFVK			
JcHsf10	A1	9-102	(204)KR6KKRR	(483)LTEQMELL	(432)DVFWEQFL	
JcHsf12	A7	41-134	(227)KDKRK7KKRRR	(345)LAERLGYL	(322)DGFWDELLSE	
JcHsf13	A4	12-105	(203)NKKRR	(161)IMSLCE	(365)DHFWEYFLTE	
JcHsf16	A9	138-231	(320)KRKRQR8KKRR	(447)LELEDL		
JcHsf17	A2	42-135	(230)KK9KRR	(369)LVDQMGYL	(282)ETLFSAALDD	
JcHsf7	B2	21-114	(257)KRGR			(249)LFGV
JcHsf8	B3	27-120	(179)KRKCK	(207)PKLFGVRL		(209) LFGV
JcHsf11	B1	7-100	(264)KTKKRGR	(254)FKLFGVLL		(256) LFGV
JcHsf14	B2	27-120	(282)KRMRK			(274) LFGV
JcHsf6	B4	34-140	(208)NKIRRL			
JcHsf15	C1	7-100	(184)KKRR			

Figure 4 Multiple sequence alignment of the DBD domains of 17 members of the JcHsf protein family.
Multiple alignment results clearly revealed highly conserved DBD domains among physic nut *Hsf* genes.

Figure 5 Multiple sequence alignment of the HR-A/B regions of the Hsf protein family in physic nut.
The scheme at the top depicts the locations and boundaries of the HR-A core, insert, and HR-B regions
within the HRA/B regions.

expression level in the seed late developing stage (S5–S7). Two (*JcHsf3* and *17*) showed
sustained high expression throughout the seed development stages (Fig. 8).

Expression profiles of *JcHsf* genes under drought and salt stress

In order to identify the expression pattern of *JcHsf* genes in response to salt and drought
stress, the reported expression profile data of these genes in leaves 2 h, 2d and 7d under
salt stress (*Zhang et al., 2014*) and 1d, 4d and 7d under drought stress (*Zhang et al., 2015b*)
were analyzed (Fig. 9). The results showed that six genes (*JcHsf4*, *8*, *10*, *11*, *13* and *15*) were
significantly up-regulated and two genes (*JcHsf 6* and *12*) were down regulated in both
salt and drought stress condition. In addition, the other six genes (*JcHsf2*, *3*, *5*, *14*, *16*, *17*)
highly up-regulated in response to drought were also identified.
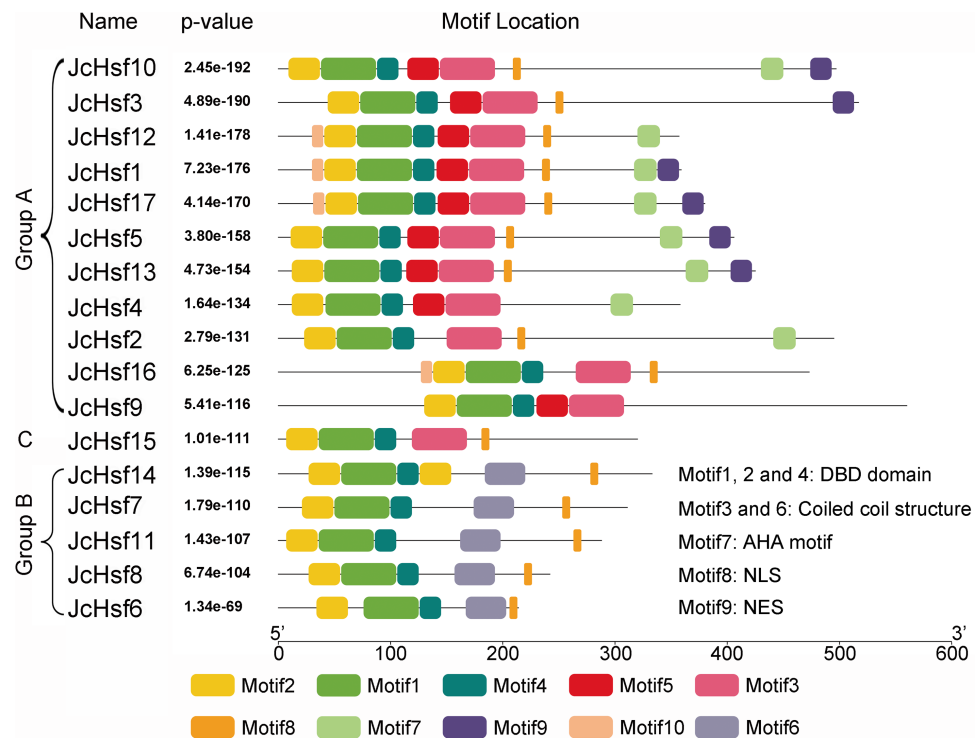
Figure 6 **Distribution of conserved motifs in the JcHsf family members.** The names of all members of the gene clusters defined and the combined *P* values are shown on the left side of the figure, while motif sizes are indicated at the bottom of the figure. Different motifs are indicated by different color numbered 1–10. For details of motifs, refer to Table S3.

To verify the *JcHsf* genes expression pattern from the RNA-seq data, the qRT-PCR were performed to detect the transcriptional level of group B *JcHsf* (*6*, *7*, *8*, *11* and *14*) in leaves under salt and drought stress (Fig. 10). The results were basically consistent with the changes in the expression of RNA-seq, indicating that the data of RNA-seq were generally accurate.

DISCUSSION

Physic nut has been emerged as a renewable resource for biodiesel production, which can be grown on the barren soil (*King et al., 2009*; *Divakara et al., 2010*). Until now, little is known about the response mechanism to abiotic stress in physic nut. More and more evidence suggests that Hsfs play central roles in plant developmental and response to abiotic stresses (*Lin et al., 2018*; *Lohani et al., 2019*). However, there have been no reports of any studies on the *Hsf* genes in physic nut, so it is necessary to investigate the new *Hsf* genes in physic nut. In this study, 17 *Hsf* genes were found in physic nut genome (Table 1) (*Wu et al., 2015*).

Phylogenetic tree suggested that, JcHsf proteins could be divided into three groups (Fig. 2), similar to *Arabidopsis* and other plants (*Guo et al., 2008*; *Lin et al., 2011*; *Liu et al., 2016*). The *Hsf* gene numbers of some specific subgroups in physic nut were different

	E-value	Sites	Width
Motif1	1.1e-599	17	50
Motif2	1.1e-599	16	29
Motif3	1.1e-599	12	50
Motif4	1.1e-599	17	20
Motif5	1.1e-599	8	29
Motif6	1.1e-599	5	37
Motif7	1.1e-599	8	21
Motif8	1.1e-599	15	8
Motif9	1.1e-599	4	20
Motif10	1.1e-599	4	11

Figure 7 Motif logo, the amino acid composition of each conserved motif. Amino acid sequences of each motif identified by MEME tools. The font size indicates the frequency of the corresponding amino acid. Motif location and combined *E*-value are shown.

from *Arabidopsis*. For example, the number of subgroup A1, A6 and A7 in physic nut were less than that in *Arabidopsis*. One possible reason is that during the early stages of evolution, *JcHsf* genes have not yet experienced chromosome fragment replicate events (*Wu et al., 2015*). Another possible reason is that the *Arabidopsis* may acquire the *Hsf* genes while the physic nut lost it from their common ancestor (*Wu et al., 2015*).

Gene's structure analyses indicated that all *Hsf* genes were found to contain only one intron (Fig. 3). Although physic nut *Hsf* genes shared same intron number, the intron length differed across the groups, which revealed the very highly conserved exon/intron splicing arrangement. Notably, *JcHsf4* and *JcHsf7* in the subgroup A1 has a long intron which made them different from other *JcHsf* gene members because of their larger size (9,741 bp and 7,519 bp respectively). In general, the gene structure of members in the same class has the similar domain or motif. In this study, although motif 3 and 6 were included in coiled coil structure, but motif 3 was only presented in group A and C, while motif 6 was just detected in group B (Fig. 6). These motifs were specific to some group, which are awaited for participating in group-specific function. Additionally, there are no AHA motifs required for transcriptional activity in the three members of group A (JcHsf3, JcHsf9 and JcHsf16). Previous studies showed that these proteins may gain function through binding to other menbers of group A to form hetero-oligomers (*Guo et al., 2008*).

Gene's expression profiles are often related to their functions (*Guo et al., 2008*). In the present study, the expression profiles of each *JcHsf* gene in roots, stems, leaves and
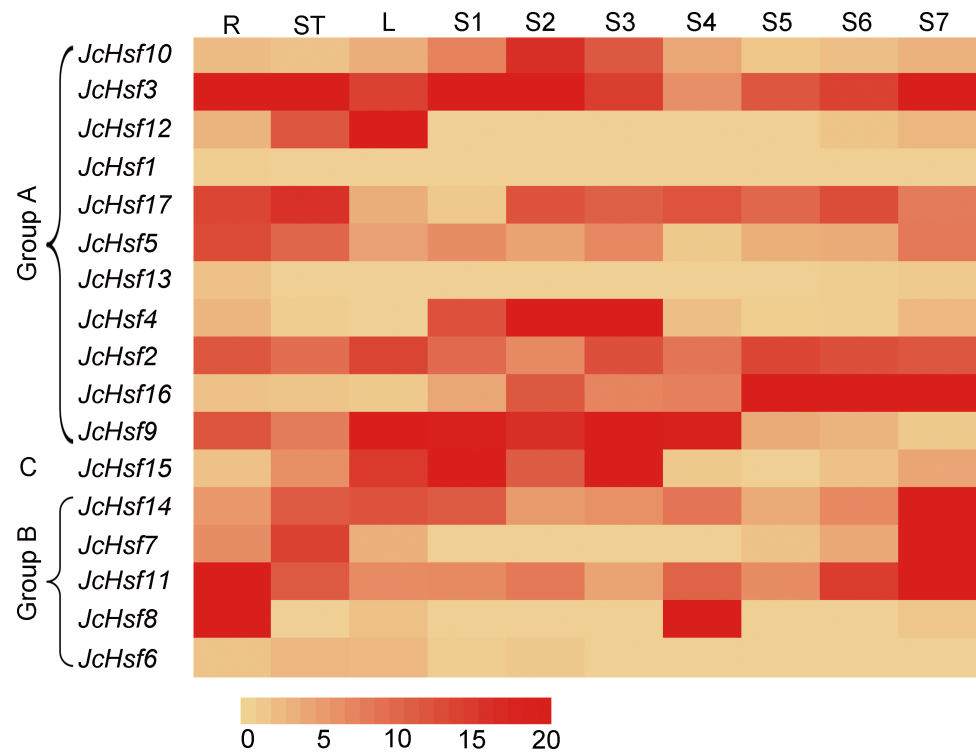
Figure 8 **Expression analysis of the *JcHsf* genes in different tissues.** R: roots, St: stem cortex, L: leaves S1–S7 correspond to seven different developing physic nut seeds: on 14, 19, 25, 29, 35, 41, and 45 days after pollination. The color scale represents signal values.

developing seeds were investigated (*Jiang et al., 2012*). Most *JcHsfs* were detected highly expressed in different tissues and developing seeds. It illustrated that *Hsf* genes are critical regulators involved in plant growth and development (Fig. 8). In addition, *JcHsf10* expression was highest in S2 stage of developing seeds (Fig. 8); its homolog *AtHsfA1b* was reported to regulate multiple developmental genes under heat stress in *Arabidopsis* (*Albihlal et al., 2018*). Therefore, it can be inferred that, at the early stage of developing seed; high level of *JcHsf10* may take part in regulating seed development of physic nuts under heat stress.

Studies on genome-wide expression analysis of different plants showed that the expression level of some *Hsf* genes changed in response to different abiotic stresses (*Mittal et al., 2009*; *Guo et al., 2015*; *Dossa, Diouf & Cisse, 2016*; *Liu et al., 2016*; *Lohani et al., 2019*). Consistent with this, 14 *JcHsf* genes out of 17 were up- or down-regulated under at least one stress condition (Fig. 9). Among these 14 genes, 11 genes showed at least 2 fold up- or down-regulated in at least one-time point in response to both salt and drought stimuli (Fig. 9), which is consisted with previous studies on *Hsfs* in various plant species (*Chauhan et al., 2011*; *Li et al., 2014*; *Dossa, Diouf & Cisse, 2016*). Overexpression *Arabidopsis HsfA1a* had a positive effect on tolerance to various stressors by acting the inducible heat shock protein expression (*Qian et al., 2014*). Its physic nut homolog *JcHsf3* was highly expressed in all tested tissues (Fig. 8) and up-regulated significantly at 7d in
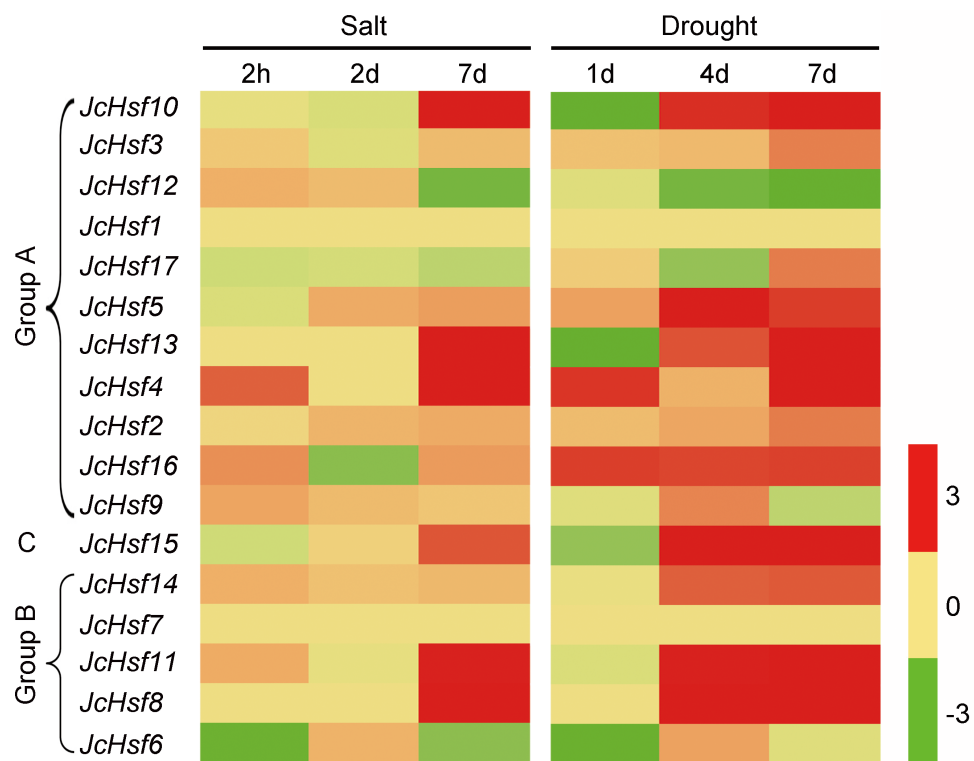
Figure 9 Expression analysis of *JcHsf* genes in leaves under salt and drought stress. Log$_2$ based values (signal values for treatment/signal value for control) were used to create the heat map based on transcriptomic data for *Hsfs*. The color scale represents signal values.

leaves after drought stress (Fig. 9), suggesting a role response to drought stress. *AtHsfA2* is essential for acquiring thermotolerance in *Arabidopsis* (*Charng et al., 2007*), and its homolog *JcHsf17* presented high level in roots, stems and after S2 stage of developing seeds (Fig. 8), and also up-regulated significantly at 7d in leaves after drought stress (Fig. 9). The result suggests that *JcHsf17* may play roles in both thermo and drought tolerance in physic nut. In addition, *AtHsfB1* and *AtHsfB2b*, took part in regulating the expression of defensin gene *Pdf1.2*, plays an important role in pathogen resistance (*Kumar et al., 2009*), and their homolog in physic nut were *JcHsf11* and *JcHsf14*, respectively. *JcHsf11* and *JcHsf14*, both classified into group B, presented similar expression pattern under normal and abiotic condition (Figs. 8 and 9). It indicated that *JcHsf11* and *JcHsf14* may involve in pathogen resistance in physic nut. In brief, we speculate that *JcHsf* genes may participate in many aspects of the developmental process in physic nut, and their roles deserve further study.

CONCLUSIONS

In summary, 17 *JcHsf* genes were identified from the physic nut genome, and gene information provided, including chromosomal localization, phylogenetic and structure analysis. Most *JcHsf* genes showed a differential expression pattern, indicating the great
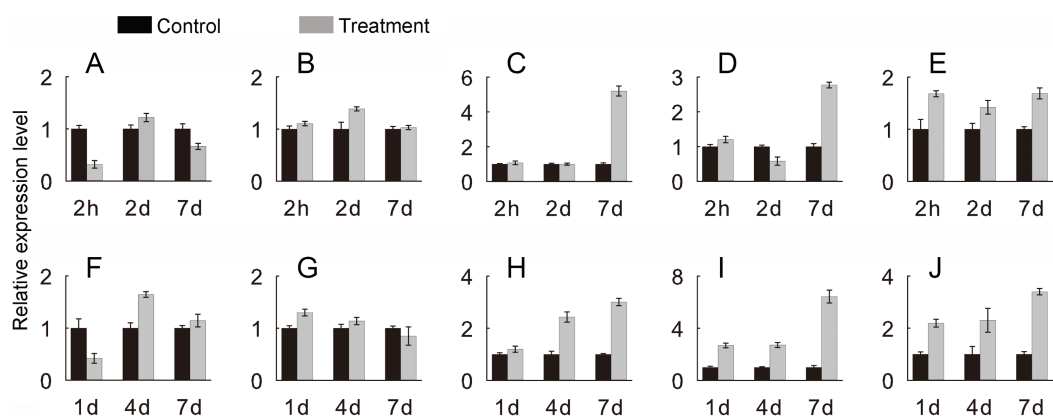
Figure 10 **Expression level of selected *JcHsf* genes.** qRT-PCR analysis of selected genes in leaves after salt and drought treatments at different time points. (A) *JcHsf6* under salt stress. (B) *JcHsf7* under salt stress. (C) *JcHsf8* under salt stress. (D) *JcHsf11* under salt stress. (E) *JcHsf14* under salt stress. (F) *JcHsf6* under drought stress. (G) *JcHsf7* under drought stress. (H) *JcHsf8* under drought stress. (I) *JcHsf11* under drought stress. (J) *JcHsf14* under drought stress. Relative expression was normalized with *JcActin* as the reference gene. Bars show standard deviations of the repeats. The experiment involved three biological replicates, each with two technological replicates.

significance of *JcHsf* in plant development and stress response. Although the exact function of the *JcHsf* genes cannot still be indicated, this study lays a foundation for further study on its potential role in regulating developmental process and stress response in physic nut.

ACKNOWLEDGEMENTS

We would like to thank Professor Guo-jiang Wu in South China Botanical Garden, Chinese Academy of Sciences for his help in providing expression profile data and Mr. Yue-Hui Tang in Zhoukou Normal University for his help on bioinformatics analysis.

Funding

This work was funded by the Education Department of Henan province (No. 16A180030), the Science and Technology Department of Henan province (No. 172102110103), the National fund pre-research projects (No. 2018YZ10, 2018YZ14) and the Scientific Research Projects for High Level Talents (No. 2014BZ02, 2014BZ03 and 2014BZ07) of Luoyang Institute of Science and Technology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Education Department of Henan province: 16A180030.
Science and Technology Department of Henan province: 172102110103.

Luoyang Institute of Science and Technology: 2018YZ10, 2018YZ14, 2014BZ02, 2014BZ03, 2014BZ07.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Lin Zhang conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Wei Chen conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Ben Shi performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

Data is available at NCBI SRA: PRJNA257901, PRJNA244896.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.8467#supplemental-information.

REFERENCES

Albihlal WS, Irabonosi O, Blein T, Persad R, Chernukhin I, Crespi M, Bechtold U, Mullineaux PM. 2018. *Arabidopsis* heat shock transcription factor A1b regulates multiple developmental genes under benign and stress conditions. *Journal of Experimental Botany* **69**:2847–2862 DOI 10.1093/jxb/ery142.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* **37**:W202–W208 DOI 10.1093/nar/gkp335.

Charng YY, Liu HC, Liu NY, Chi WT, Wang CN, Chang SH, Wang TT. 2007. A heat-inducible transcription factor, HsfA2, is required for extension of acquired thermotolerance in *Arabidopsis*. *Plant Physiology* **143**:251–262 DOI 10.1104/pp.106.091322.

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**:171–187 DOI 10.1007/s00438-011-0638-8.

Chung E, Kim KM, Lee JH. 2013. Genome-wide analysis molecular characterization of heat shock transcription factor family in *Glycine max*. *Journal of Genetics & Genomics* **40**:127–135 DOI 10.1016/j.jgg.2012.12.002.

Divakara BN, Upadhyaya HD, Wani SP, Gowda CLL. 2010. Biology and genetic improvement of *Jatropha curcas* L.: a review. *Applied Energy* **87**:732–742 DOI 10.1016/j.apenergy.2009.07.013.

Dossa K, Diouf D, Cisse N. 2016. Genome-wide investigation of hsf genes in sesame reveals their segmental duplication expansion and their active role in drought stress response. *Frontiers in Plant Science* **7**:Article 1522.

Eddy SR, Pearson WR. 2011. Accelerated profile HMM searches. *PLOS Computational Biology* **7**:e1002195 DOI 10.1371/journal.pcbi.1002195.

Guo AY, Zhu QH, Chen X, Luo JC. 2007. GSDS: a gene structure display server. *Hereditas* **29**:1023–1026.

Guo J, Wu J, Ji Q, Wang C, Luo L, Yuan Y, Wang Y, Wang J. 2008. Genome-wide analysis of heat shock transcription factor families in rice and *Arabidopsis*. *Journal of Genetics & Genomics* **35**:105–118 DOI 10.1016/S1673-8527(08)60016-8.

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 stress. *Frontiers in Plant Science* **7**:Article 114.

Guo M, Lu JP, Zhai YF, Chai WG, Gong ZH, Lu MH, Guo M, Lu JP, Zhai YF, Chai WG. 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(2015-06-19))**:151 DOI 10.1186/s12870-015-0512-7.

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**:583–590 DOI 10.1016/j.plantsci.2009.01.016.

Huang XY, Tao P, Li BY, Wang WH, Yue ZC, Lei JL, Zhong XM. 2015. Genome-wide identification, classification, and analysis of heat shock transcription factor family in Chinese cabbage (*Brassica rapa pekinensis*). *Genetics & Molecular Research* **14**:2189–2204 DOI 10.4238/2015.March.27.5.

Hwang SM, Kim DW, Woo MS, Jeong HS, Son YS, Akhter S, Choi GJ, Bahk JD. 2014. Functional characterization of *Arabidopsis* HsfA6a as a heat-shock transcription factor under high salinity and dehydration conditions. *Plant Cell & Environment* **37**:1202–1222 DOI 10.1111/pce.12228.

Jiang H, Wu P, Zhang S, Song C, Chen Y, Li M, Jia Y, Fang X, Chen F, Wu G. 2012. Global analysis of gene expression profiles in developing physic nut (*Jatropha curcas* L.) seeds. *PLOS ONE* **7**:e36522 DOI 10.1371/journal.pone.0036522.

King AJ, He W, Cuevas JA, Freudenberger M, Ramiaramanana D, Graham IA. 2009. Potential of *Jatropha curcas* as a source of renewable oil and animal feed. *Journal of Experimental Botany* **60**:2897–2905 DOI 10.1093/jxb/erp025.

Kumar M, Busch W, Birke H, Kemmerling B, Nürnberger T, Schöffl 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**:152–165 DOI 10.1093/mp/ssn095.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**:2947–2948 DOI 10.1093/bioinformatics/btm404.

Li PS, Yu TF, He GH, Chen M, Zhou YB, Chai SC, Xu ZS, 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**:1009 DOI 10.1186/1471-2164-15-1009.

Lin KF, Tsai MY, Lu CA, Wu SJ, Yeh CH. 2018. The roles of *Arabidopsis* HSFA2, HSFA4a, and HSFA7a in the heat shock response and cytosolic protein response. *Botanical Studies* **59**:15.

Lin YX, Jiang HY, Chu ZX, Tang XL, Zhu SW, Cheng BJ. 2011. Genome-wide identification, classification and analysis of heat shock transcription factor family in maize. *BMC Genomics* **12**:76 DOI 10.1186/1471-2164-12-76.

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**:52–59 DOI 10.1016/j.gene.2015.09.076.

Lohani N, Golicz AA, Singh MB, Bhalla PL. 2019. Genome-wide analysis of the Hsf gene family in *Brassica oleracea* and a comparative analysis of the Hsf gene family in B. oleracea, *B. rapa* and *B. napus*. *Functional and Integrative Genomics* **19**:515–531 DOI 10.1007/s10142-018-0649-1.

Mittal D, Chakrabarti S, Sarkar A, Singh A, Grover A. 2009. Heat shock factor gene family in rice: genomic organization and transcript expression profiling in response to high temperature, low temperature and oxidative stresses. *Plant Physiology & Biochemistry* **47**:785–795 DOI 10.1016/j.plaphy.2009.05.003.

Nishizawa-Yokoi A, Yoshida E, Yabuta Y, Shigeoka S. 2009. Analysis of the regulation of target genes by an *Arabidopsis* heat shock transcription factor, HsfA2. *Journal of the Agricultural Chemical Society of Japan* **73**:890–895.

Nover N, Bharti K, Döring 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**:177–189 DOI 10.1379/1466-1268(2001)006<0177:AATHST>2.0.CO;2.

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. *Journal of Experimental Botany* **58(12)**:3373–3383.

Personat JM, Tejedor-Cano J, Prieto-Dapena P, Almoguera C, Jordano J. 2014. Co-overexpression of two heat shock factors results in enhanced seed longevity and in synergistic effects on seedling tolerance to severe dehydration and oxidative stress. *BMC Plant Biology* **14**:56 DOI 10.1186/1471-2229-14-56.

Prandl R, Hinderhofer K, Eggers-Schumacher G, Schoffl F. 1998. HSF3, a new heat shock factor from *Arabidopsis thaliana*, derepresses the heat shock response and confers thermotolerance when overexpressed in transgenic plants. *Molecular & General Genetics* **258**:269–278.
Qian J, Chen J, Liu YF, Yang LL, Zhang LM. 2014. Overexpression of *Arabidopsis* HsfA1a enhances diverse stress tolerance by promoting stress-induced Hsp expression. *Genetics & Molecular Research* **13**:1233–1243 DOI 10.4238/2014.February.27.8.

Scharf KD, Berberich T, Ebersberger I, Nover L. 2012. The plant heat stress transcription factor (Hsf) family: structure, function and evolution. *Biochimica Et Biophysica Acta* **1819**:104–119 DOI 10.1016/j.bbagrm.2011.10.002.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology & Evolution* **28**:2731–2739 DOI 10.1093/molbev/msr121.

Von Koskull-Döring P, Scharf K-D, Nover L. 2007. The diversity of plant heat stress transcription factors. *Trends in Plant Science* **12**:452–457 DOI 10.1016/j.tplants.2007.08.014.

Wiederrecht G, Seto D, Parker CS. 1988. Isolation of the gene encoding the *S. cerevisiae* heat shock transcription factor. *Cell* **54**:841–853 DOI 10.1016/S0092-8674(88)91197-X.

Wu P, Zhou C, Cheng S, Wu Z, Lu W, Han J, Chen Y, Chen Y, Ni P, Wang Y, Xu X, Huang Y, Song C, Wang Z, Shi N, Zhang X, Fang X, Yang Q, Jiang H, Chen Y, Li M, Wang Y, Chen F, Wang J, Wu G. 2015. Integrated genome sequence and linkage map of physic nut (*Jatropha curcas* L.), a biodiesel plant. *Plant Journal* **81**:810–821 DOI 10.1111/tpj.12761.

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**:539–557 DOI 10.1093/jxb/ert399.

Yoshida T, Ohama N, Nakajima J, Kidokoro S, Mizoi J, Nakashima K, Maruyama K, Kim JM, Seki M, Todaka D, Osakabe Y, Sakuma Y, SchöZ F, Shinozaki 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.

Zhang J, Li Y, Jia HX, Li JB, Huang J, Lu MZ, Hu JJ. 2015a. The heat shock factor gene family in *Salix suchowensis*: a genome-wide survey and expression profiling during development and abiotic stresses. *Frontiers in Plant Science* **6**:Article 748.

Zhang C, Zhang L, Zhang S, Zhu S, Wu P, Chen Y, Li M, Jiang H, Wu G. 2015b. Global analysis of gene expression profiles in physic nut (*Jatropha curcas* L.) seedlings exposed to drought stress. *BMC Plant Biology* **15**:17 DOI 10.1186/s12870-014-0397-x.

Zhang L, Zhang C, Wu P, Chen Y, Li M, Jiang H, Wu G. 2014. Global analysis of gene expression profiles in physic nut (*Jatropha curcas* L.) seedlings exposed to salt stress. *PLOS ONE* **9**:e97878 DOI 10.1371/journal.pone.0097878.