Genome-wide identification and expression analysis of Hsf and Hsp gene families in cucumber (Cucumis sativus L.)

Xueqian Chen1 · Zhiyuan Wang1 · Rui Tang1 · Lina Wang1 · Chunhua Chen1 · Zhonghai Ren1

Received: 30 January 2021 / Accepted: 6 August 2021 / Published online: 12 August 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract
Heat shock proteins (Hsps) are molecular chaperones that participate in plant growth and development, and in plant responses to environmental stresses, including heat stress, Hsps’ expression is monitored and regulated by specific types of transcription factors known as heat shock factors (Hsfs). Although the roles of Hsfs and Hsps in stress response have been investigated in some plants, their roles are still poorly understood in cucumber (Cucumis sativus L.). To reveal the number of Hsfs and Hsps and their members coping with various stresses in cucumber, the comprehensive analyses of Hsf and Hsp gene families were conducted. A total of 23 Hsfs and 72 Hsps were identified in the cucumber genome (v3.0), and their gene structure and motif composition were found to be relatively conserved in each subfamily. At least 23 pairs of heat shock genes may have undergone duplication in cucumber. The cis-element analysis indicated that the promotors of CsHsfs and CsHsps possess at least one hormone or stress response cis-element, suggesting that CsHsf and CsHsp genes might respond to different stress conditions. RNA-seq showed that most CsHsf and CsHsp genes were sensitive to heat stress, but not all. Interestingly, CsHsf and CsHsp genes up-regulated by heat stress were also up-regulated by NaCl, indicating that cucumber shares common pathways in the responses to heat and salt stress. In addition, our results demonstrated that the fewer introns in the stress genes, the more sensitive they are, which leads to the stronger adaptability of plants to various developmental conditions and environmental stimuli. For instance, comparing with other Hsp families, the genes of CsHsp20 family are shorter in length and have fewer introns, and 39% of them are intronless, however, most of CsHsp20s were highly induced by heat stress. These results provide valuable information to further clarify the functional characterization of the CsHsf and CsHsp genes for plant breeding purposes.

Keywords Cucumber · Heat shock factor · Heat shock protein · Genome-wide analysis · Gene expression

Introduction
During growth and development, plants are usually exposed to a variety of strenuous stresses from the surrounding environment, including not only abiotic stresses such as high temperatures, drought, salinity, but also biotic stresses such as pathogen invasion (Nejat and Mantri 2017). Plants, as sessile organisms, could not avoid the damage of adverse environments by changing their position. Therefore, a complex response and regulatory network has evolved at the biochemical, physiological, and molecular levels to adapt to adversity (Lamers et al. 2020).

Under the trend of global warming, high temperature has become a major issue for agriculture worldwide (Hayes et al. 2020). To prevent the negative effects of high temperature on plants, such as water loss, growth reduction, photosynthesis...
efficiency reduction, some genes are induced quickly by heat stress for adaptive responses (Aleem et al. 2020).

Heat shock factors (Hsfs), being a specific type of transcription factors, normally exist as inactive proteins and can be induced rapidly to modulate growth responses under high temperature (Ohama et al. 2017). Previous studies have shown that the protein structure of Hsfs with a DNA-binding domain (DBD) and an adjacent bipartite oligomerization domain (HR-A/B or OD) in the N-terminal, and a nuclear localization signal (NLS), a nuclear output signal (NES) and an activation peptide motif in the C-terminal (AHA), is very conserved (Scharf et al. 1990). Besides, according to the number of amino acid residues inserted between HR-A and HR-B, plant Hsfs can be divided into three main classes, namely: A, B, and C (Guo et al. 2016).

Heat shock proteins (Hsps), working as molecular chaperones in both plant and animal species and being regulated by Hsfs, are synthesized in large quantities to prevent incorrect aggregation of proteins when plants subject to heat stress (Ohama et al. 2017). Hsps not only are transcriptionally regulated by Hsfs, but also are able to regulate the activity of Hsfs via feedback loop caused by their physical interaction. For instance, Hsp70 represses the activity of HsfA1, while Hsp90 stimulates the DNA binding activity of HsfB1 in tomato (Hahn et al. 2011). According to their apparent molecular weight, plant Hsps are divided into five groups: small Hsps (sHsps), Hsp60, Hsp70, Hsp90, and Hsp100 (Wang et al. 2004), which are distributed in many cell components, such as cytoplasm, nucleus, plastids, mitochondria, chloroplasts, endoplasmic reticulum, and peroxisomes, protecting different organelle structures (Jacob et al. 2017).

So far, the functions of Hsf and Hsp genes in various stress responses have been reported in many plants. In Arabidopsis thaliana, it has been proved that AtHsfA1 is the central regulator of heat stress regulation and response network, which could induce the expression of many downstream heat shock genes, for example, other Hsf subclasses (A2, A3, A7, B1, and B2) (Liu and Charng 2012; Yoshida et al. 2011). AtHsfA2 not only performs an essential function in heat and osmotic stress response, but also exerts an important role in the growth and development of plants (Chang et al. 2007; Liu and Charng 2013; Ogawa et al. 2007). OsMSR3, coding a small heat shock protein, confers enhanced tolerance to cadmium and copper stresses in Arabidopsis (Cui et al. 2019). GhHsp24.7, a mitochondrial small heat shock protein in cotton regulates seed germination by inducing reactive oxygen species (ROS) via thermal sensing (Ma et al. 2019). CaHsp26 and CaHsp22.5 of sweet pepper (Capsicum annuum L.) play significant roles in the protection of PSII by enhancing photochemical activity and oxidation resistance during chilling stress (Li et al. 2018, 2012). In tobacco (Nicotiana tabacum L.), potentiation of Hsp gene expression increased the resistance exposed to NaCl stress (Li et al. 2014b). Besides, ZmHsp16.9, a cytosolic class I small heat shock protein of maize, confers heat tolerance in transgenic tobacco (Sun et al. 2012). Hsp100/ClpB located in the chloroplast plays an active role in regulating the heat resistance of tomato (Solanum lycopersicum) (Yang et al. 2006). The overexpression of SIHsp70-1 resulted in internode elongation, suggesting that it functions in the shoot growth of tomato (Vu et al. 2019). MeHsp90.9 is critical for drought-induced stress resistance in cassava by regulating hydrogen peroxide (H2O2) and abscisic acid (ABA) (Wei et al. 2020). Moreover, Hsp90 genes in barley play roles in regulating hypersensitive responses to stripe rust (Pei et al. 2015).

Cucumber (Cucumis sativus L.) is a warm-season crop, but is vulnerable to heat stress (Li et al. 2014a). High temperature is one of the main abiotic stress factors affecting cucumber yield and quality in summer cultivation or protected land production (Ding et al. 2016). However, in cucumber, it is still unknown that which Hsf and Hsp genes play important roles in heat stress tolerance. Moreover, the updated cucumber genome (‘Chinese Long’ v3) provided a good chance to comprehensively analyze Hsf and Hsp families (Li et al. 2019). In this study, we identified 23 Hsf and 72 Hsp genes including Hsp20, Hsp60, Hsp70, Hsp90, and Hsp100 in the cucumber v3 genome. Then we carried a comprehensive analysis of Hsf and Hsp genes in cucumber, including their sequence features, chromosomal locations, phylogenetic relationships, and dynamic expression patterns in response to heat, NaCl, and downy mildew stresses, which not only help to identify the function of cucumber Hsf and Hsp genes, but also provide some candidate genes for breeding new stress-resistant cucumber varieties.

Materials and methods

Identification of Hsf and Hsp genes in the cucumber genome

The cucumber genome, gene, and protein sequences were downloaded from the Cucurbit Genomics Database (http://cucurbitgenomics.org/). Cucumber ‘Chinese Long’ genome v3. To identify the putative Hsf and Hsp members in cucumber, Protein Basic Local Alignment Search Tool (BLASTP) searches were performed with an e-value ≤ 1e-2 in the Cucurbit Genomics Database using the published Hsf and Hsp protein sequences from Arabidopsis, rice, and tomato as query sequences (Chauhan et al. 2011; Guo et al. 2008; Lin et al. 2001; Sarkar et al. 2009, 2013; Scharf et al. 2001; Yang et al. 2016; Yu et al. 2016; Agarwal et al. 2001; Hu et al. 2009; Krishna and Gloor 2001; Lee et al. 2007; Liu et al. 2014; Singh et al. 2010; Swindell et al. 2007; Wang et al. 2014). Subsequently, we downloaded the Hsf (PF00447), Hsp20 (PF00011), Hsp60 (PF00118), Hsp70
(PF00012), Hsp90 (PF00183), and Hsp100 (PF02861) Hidden Markov Model (HMM) configuration files from the Pfam database (http://pfam.xfam.org/). Hsfs and Hsps in the cucumber genome database were searched using HMMER 3.0. With the default parameters, the cutoff value was set to 0.01. Additionally, name search using the words “Hsf”, “Hsp20”, “Cpn60”, “Hsp70”, “Hsp90” and “ClpB” as keywords also applied to retrieve in Cucurbit Genomics Database. Following the removal of the repetitive sequences, all sequences were reserved and submitted to CDD (https://www.ncbi.nlm.nih.gov/cdd/), Pfam, and SMART (http://smart.embl-heidelberg.de/) to confirm the Hsf/Hsp20/Cpn60/Hsp70/Hsp90/ClpB domains. All of the high-confidence and non-redundant genes were assigned as cucumber Hsfs and Hsps (CsHsfs and CsHsps). These CsHsp and CsHsf genes were named based on their positions on pseudomolecules (Su et al. 2013).

**Sequence analysis and structural characterization**

Using cNLS-Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) and NetNES 1.1 (http://www.cbs.dtu.dk/services/NetNES/) to predict the NLS and NES domains of cucumber Hsfs (Kotak et al. 2004; la Cour et al. 2004). Predict the domain of AHA based on the conserved AHA motif sequence (FWxxF/L, F/I/L) (Döring et al. 2000). Protein isoelectric point (pI), number of amino acids (AA), and molecular weight (MW) were calculated using ExPasy (https://web.expasy.org/compute_pi/) (Cai et al. 2016). WoLF PSORT (https://www.genscript.com/wolf-psort.html) was used to predict their protein subcellular localizations (Emanuelsson et al. 2000). The MEME program (Version 5.0.4, http://alternate.meme-suite.org/tools/meme) was used to identify the conserved motifs in the CsHsf and CsHsp protein sequences, with the following parameters: any number of repetitions, the maximum number of motifs: 10, and an optimum motif width of 6 to 200 amino acid residues (Bailey et al. 2009). The structures of all CsHsf and CsHsp genes were analyzed via the Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/) (Hu et al. 2015).

**Multiple sequence alignment and phylogenetic analysis**

Multiple sequence alignments of CsHsf and CsHsp proteins were performed by using the ClustalX (version 1.83) program with default parameters (Thompson et al. 1997). The full-length amino acid sequences of Hsfs and Hsps in cucumber, Arabidopsis, rice, and tomato were uploaded to the MEGA v. 7.0 software (Kumar et al. 2016). The ClustalW alignment method was applied to the sequence alignment. Phylogenetic trees were then constructed based on the alignments using the neighbour-joining (NJ) method. The trees were visualized and optimized via Evolview (https://www.evolgenius.info/evolview/#login).

**Chromosomal localization, gene duplication, and evolutionary analysis**

All the chromosomal locations of CsHsf and CsHsp genes were performed with the MapChart software to visualize the positions and relative distances of genes on cucumber chromosomes (Voorrips 2002) based on the cucumber v3 genome. Multiple Collinearity Scan toolkit (MCScanX) was used to analyze the gene duplication events, with the default settings (Wang et al. 2012). The duplication of the CsHsf and CsHsp gene was determined according to two criteria: (a) the length of the shorter aligned sequence covered > 70% of the longer sequence, and (b) the two aligned sequences shared > 70% amino acid sequence similarity (Yang et al. 2008). In the 100 kb chromosomal fragment, two genes separated by less than five intermediate genes are considered to have undergone tandem duplication (Zhang et al. 2018). Codon alignment of duplicated genes was done by MEGA v.7.0 using the ClustalW codon alignment tool, and the homologous (Ks) and non-homologous (Ka) rates of the tandem and segmental duplications of cucumber Hsf and Hsp genes were calculated using KaKs_Calculator 2.0 (Wang et al. 2010). The divergence time (T) of each duplicated gene pair was calculated as $T = Ks / 2r = 6.56 \times 10^{-9} \times 10^6$ million years ago (Mya), with Ks being the synonymous substitutions per site and r being the rate of divergence for nuclear genes from plants (Koch et al. 2000). To reveal the synteny relationship between the orthologous Hsf and Hsp genes of cucumber, Arabidopsis, and rice, syntenic analysis maps were constructed using the Dual Systeny Plotter software (https://github.com/CJ-Chen/TBtools) (Liu et al. 2017).

**Cis-element prediction and protein–protein interaction analysis**

The upstream sequences (1.5 kb) of the CsHsf and CsHsp genes-coding sequences were retrieved from the Cucurbit Genomics Database and then submitted to PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to identify hormone response elements and stress-induced components (Lescot et al. 2002). The hormone response elements include abscisic acid (ABA)-responsive element (ABRE), ethylene (ER)-responsive elements (ERE), gibberellin (GA)-responsive elements (P-box, GARE-motif, and TATC-box), salicylic acid (SA)-responsive elements (TCA-element and SARE), auxin (IAA)-responsive elements (TGA-element, AuxRR-core, and TGA-box) and jasmonate acid (MeJA)-responsive elements (CGTCA-motif and TGACG-motif). The stress-induced components include heat stress element (HSE), low-temperature responsive
element (LTR), dehydration-responsive elements (DRE), drought-responsive element (MBS), defense-responsive element (W-box), defense and stress-responsive element (TC-rich repeats), wound-responsive element (WUN-motif), hypoxia-inducible response element (ARE) and anoxic specific inducible response element (GC-motif). Protein–protein interaction (PPI) among CsHsf and CsHsp was detected by STRING (Version 11.0, https://string-db.org/cgi/input.pl) with the confidence score > 0.7 and an interaction network of CsHsf and CsHsp proteins was drawn by Cytoscape v3.7.2.

Plant materials, stress treatments, and quantitative real-time PCR analysis

The ‘Chinese long’ 9930, a typical north-China cucumber, was used for expression analyses. The selected seeds after immersion and imbibition for 12 h were sown in pots containing vermiculite and matrix and then they were cultivated in a growth chamber under day/night temperatures of 25/18 °C, with a 16 h/8 h photoperiod/dark period. Three-week-old seedlings were subjected to heat treatment at 42 °C and the leaves were collected at 0 (control), 3, and 6 h. For NaCl treatment, the seedlings were subjected to 75 mM NaCl stress for 24 h. For downy mildew treatment, the seedlings were inoculated with Pseudoperonospora cubensis for NaCl stress for 24 h. For downy mildew treatment, the seedlings were inoculated with Pseudoperonospora cubensis for 6 and 24 h. All the collected samples were quickly frozen in liquid nitrogen and stored in a −80 °C refrigerator until

Results

Identification and analysis of Hsf and Hsp genes in the cucumber genome

After HMM analysis, BLASTP and keyword search against the Cucumber Genomic Database (Version 3.0), a total of 23 Hsf and 72 Hsp genes in cucumber were identified, and among the 72 Hsps, 33, 15, 12, 6 and 6 genes belonged to Hsp20, Hsp60, Hsp70, Hsp90 and Hsp100 families, respectively. These Hsf and Hsp genes from Cucumis sativus were abbreviated as CsHsfs and CsHsps, and they were named according to their chromosomal orders. Detailed information of the CsHsf and CsHsp genes was listed in Table S1, including the gene name, gene ID, chromosome location, length of the open reading frame, number of exons, number of amino acids, molecular weight, protein isoelectric point, and subcellular localization prediction.

Structure and conserved motifs of CsHsf and CsHsp

According to the exon–intron arrangement of the coding sequence, the structural diversity in CsHsf and CsHsp genes was compared. In terms of intron number, intron, and exon length, the most closely related members of the same Hsf or Hsp subfamily share a similar gene structure (Fig. 1). Most CsHsfs have only one intron, and CsHsf-11, CsHsf-4 and CsHsf-18 have two introns, while CsHsf-17 has 11 introns, which is quite different from other CsHsfs (Fig. 1a). In the CsHsp20 family, 13 (39%) CsHsp20s are intronless, 17 (52%) have only one intron, 2 (6%) have two introns, and only CsHsp20-32 has five introns (Fig. 1b). The
Fig. 1  Phylogenetic relationship, gene structure and conserved motif analysis of CsHsfs and CsHsps. Phylogenetic tree of CsHsf (a), CsHsp20 (b), CsHsp60 (c), CsHsp70 (d), CsHsp90 (e) and CsHsp100 (f) proteins are shown in the left panel. The unrooted neighbor-joining phylogenetic tree was constructed with MEGA7.0 using the full-length amino acid sequences of CsHsf and CsHsp proteins, and the bootstrap test replicate was set as 1000 times. Exon–intron organization of CsHsf (a), CsHsp20 (b), CsHsp60 (c), CsHsp70 (d), CsHsp90 (e) and CsHsp100 (f) genes are shown in the middle panel. Yellow boxes represent exons and black lines represent introns. The upstream/downstream region of CsHsf and CsHsp genes are indicated in green boxes. The numbers of 0, 1, and 2 represent the splicing phase of intron. The length of exons can be inferred by the scale at the bottom.

Distributions of conserved motifs in CsHsf (a), CsHsp20 (b), CsHsp60 (c), CsHsp70 (d), CsHsp90 (e) and CsHsp100 (f) proteins are displayed in the right panel. Ten putative motifs are indicated in different colored boxes. The details of motifs refer to Table S2.
number of introns in different members of the CsHsp60 family varied widely (0 to 17 introns). CsHsp60-8, CsHsp60-14, and CsHsp60-15 most likely located in the mitochondria, have the largest number of introns (16 or 17). It was worth noting that CsHsp60-1 has no intron, being different from other CsHsp60 members (Fig. 1c). Except for CsHsp70-3, all CsHsp70s contain introns (Fig. 1d). The CsHsp70s (CsHsp70-9, CsHsp70-6, CsHsp70-10, and CsHsp70-1) belonging to group I are mainly located in the cytoplasm and have only one intron. The CsHsp70-7 has a relatively complex gene structure, containing 13 introns. The CsHsp90s located in different subcellular positions have different intron numbers (Fig. 1e). Among them, CsHsp90-1, CsHsp90-2 and CsHsp90-4, which are mainly located in the cytoplasm have fewer introns (2 to 3), while CsHsp90-6 and CsHsp90-5 located in the mitochondria or chloroplast have more introns (18 to 19). Except for CsHsp100-3 mainly located in the nucleus has 16 introns, other CsHsp100s have fewer introns (<10) (Fig. 1f). Therefore, it was found that the number of introns in CsHsp is closely related to their subcellular localization and evolutionary relationship.

Then, we conducted a prediction of conserved motifs shared among the related proteins in each subfamily using the Multiple Expectation Maximization for Motif Elicitation (MEME) and identified 10 putative motifs in each family. In general, most closely related members of the phylogenetic trees in different subfamilies had the similar motifs (Fig. 1). Details of these motifs were shown in Table S2.

Like Hsfs in other plants, the protein structure of CsHsfs is also highly conserved. Based on the analyses of Pfam, CDD, and SMART, we found all the CsHsf proteins contain the DBD domain composed of motif 1 and motif 2. Multiple alignments of the CsHsf protein sequences revealed a highly conserved DBD domain existed in all CsHsfs (Fig. S1A). The HR-A/B region was an essential domain in Hsfs, which was characterized by the predicted coiled-coil structure (Guo et al. 2016). The HR-A/B domain in CsHsfs was composed of two typical motifs (motif 3 and motif 4) (Fig. 1). In addition, 21 amino acids were inserted between the HR-A and HR-B regions in class A CsHsf proteins and 7 amino acids in class C (Fig. S1B). However, there was no insertion between the HR-A and HR-B regions in Class B. AHA motif (motif 6) was distinctive in the great majority of class A CsHsf proteins, and 8 (66.7%) CsHsfA proteins have AHA motif. NLS and NES are of vital importance for the intercellular distribution and interactions of Hsf proteins in the nucleus and the cytoplasm (Heerklotz et al. 2001). 17 (73%) CsHsfs contain NLS domain, and 12 (52%) NLS domain (Table S1-1).

All the CsHsp20s have a highly conserved α-crystallin domain (ACD) at the C-terminus. Multiple sequence alignment analysis and sequence logo showed that the ACD domain consisted of two conserved regions, region I and region II (Fig. S2). In addition, the consensus region I contains motif 3, and the region II motif 1, and the motif 6 and motif 4 were inserted between the consensus region I and II (Fig. 1b).

Phylogenetic analysis of CsHsf and CsHsp

We further investigated the evolutionary relationships of Hsfs and Hsps in cucumber, Arabidopsis, tomato, and rice, and based on the full-length amino acid sequences, we generated an unrooted phylogenetic tree by MEGA 7.0 using the neighbor-joining method. The CsHsf gene family could be divided into three classes: Class A (12 genes), class B (9 genes), and class C (2 genes), and each class can be further subdivided into subclasses according to the branches (Fig. S3). In addition, each CsHsp gene family could be allocated into the subfamilies according to the predicted subcellular location of the protein. For example, 130 Hsp20s from cucumber, Arabidopsis, rice and tomato, were divided into 17 different subfamilies, including CI (Cytosol I) (42 genes), CII (7 genes), CIII (3 genes), CIV (3 genes), CV (4 genes), CVI (4 genes), CVII (3 genes), CVIII (10 genes), CIX (5 genes), CX (9 genes), CXI (5 genes), MI (mitochondria I) (6 genes), MII (7 genes), MIII (4 genes), P (plastids) (7 genes), Po (peroxisome) (3 genes) and ER (endoplasmic reticulum) (8 genes) (Fig. S4). The Hsp60 family consisted of four subfamilies. According to the prediction, 24 and 6 Hsp60s of the I and IV subfamilies were mainly located in cytoplasm, 9 Hsp60s belonging to the II subfamily in mitochondria, and 15 Hsp60s of the III subfamily in chloroplast (Fig. S5). The 81 Hsp70s were divided into seven groups (I to VII), among which the I to V groups were clustered in the DnaK subfamily, and the VI and VII groups were clustered in the HSP110 / SSE subfamily (Fig. S6). The group I contained four CsHsp70s, which was the largest group mainly located in cytoplasm. Only CsHsp70-8 belonged to the group II, being involved in the metabolism of endoplasmic reticulum. The two members (CsHsp70-2 and CsHsp70-11) of group III function mainly in chloroplast and only CsHsp70-5 belonging to the group IV participates in the metabolism in mitochondrion. CsHsp70-3 belonged to the group V, being widely distributed in various subcellular regions (cytoplasm, chloroplast, mitochondria, plastid, and unknown regions). The Hsp90 family could be divided into five classes, among which the members of class I (CsHsp90-1 and CsHsp90-2) and II (CsHsp90-4) are mainly effective in the cytoplasm (Fig. S7). According to the prediction of subcellular localization, the members of class III (CsHsp90-3), IV (CsHsp90-6), and V (CsHsp90-5) play roles in the endoplasmic reticulum, mitochondrion, and chloroplast, respectively. The Hsp100 family consisted of four groups, in which the members of groups I (CsHsp100-5) and IV (CsHsp100-1 and CsHsp100-4) produce marked effects in chloroplast, and the members of groups II (CsHsp100-6) and
III (CsHsp100-2 and CsHsp100-3) make contributions to the homeostasis of mitochondria and cytoplasm, respectively (Fig. S8). Although the roles of most heat shock genes in cucumber are yet to be elucidated, Hsf and Hsp genes with conserved functions in different plants might show a tendency to aggregate to the same subgroup and have a recent common evolutionary origin.

**Chromosomal location and synteny analysis of CsHsf and CsHsp genes**

Although CsHsf and CsHsp genes are distributed on the 7 chromosomes of cucumber, the distribution appears uneven (Fig. 2). A relatively less density of CsHsf and CsHsp genes was found on chromosome 4 (10 genes), 6 (9 genes) and 7 (6 genes), while more CsHsf and CsHsp genes are located on chromosome 1 (22 genes), 2 (12 genes), 3 (18 genes) and 5 (18 genes), and most of CsHsp20s are distributed at both ends of chromosomes.

During the process of plant evolution, gene duplication, especially tandem and segmental duplication events, being the main mechanism for the expansion of gene families, made great contributions to the diversity of gene families (Kotak et al. 2004; Liu et al. 2012). Through the analysis of duplication events of CsHsf and CsHsp genes, we found that only CsHsp20 genes had tandem duplication. Among the 33 CsHsp20 genes, 16 (49%) CsHsp20 genes had tandem duplication events, resulting in the formation of 6 tandem duplication clusters (Fig. 2). On chromosome 1 and 3, three tandem duplication clusters were composed of three pairs of genes (CsHsp20-7/CsHsp20-8, CsHsp20-9/CsHsp20-10 and CsHsp20-14/CsHsp20-15). In addition, two groups of tandem duplicated genes were located on chromosome 1 and 5, each of which includes three genes (CsHsp20-3/CsHsp20-4/CsHsp20-5 and CsHsp20-21/CsHsp20-22/CsHsp20-23).

![Fig. 2](image-url) Chromosome distribution of CsHsf and CsHsp genes. The 95 heat stress genes were marked on 7 cucumber chromosomes (vertical bar). Six different heat shock gene families were indicated by different colored lines on chromosomes. Chromosome numbers are indicated at the top of each bar. Six tandem clustered genes of CsHsp20s were indicated by yellow boxes at chromosomes 1, 3 and 5. Gray lines represent gene Segmental duplications.
Only one tandem duplication cluster was composed of the four similar genes (CsHsp20-24/CsHsp20-25/CsHsp20-26/CsHsp20-27), which are present on chromosome 5. The above results showed that tandem duplications greatly promoted the expansion of the CsHsp20 gene family. Besides tandem duplication events, using BlastP and MCScanX, we also identified 13 segmental duplication events including 10 CsHsf and 14 CsHsp genes (Fig. 2).

Furthermore, to analyze the selection of the above-duplicated gene pairs, the non-synonymous to synonymous substitution ratios (Ka/Ks) were calculated (Table S3). Ka and Ks values are used to analyze the selective pressure on a protein-encoding gene as well as the approximate date of duplication events. Ka/Ks ratio = 1 is commonly used to identify the genes under neutral mutation or no selection, and Ka/Ks > 1 indicates the genes evolved under positive selection, while Ka/Ks < 1 indicates the negative purifying selection. In this study, the Ka/Ks values of 22 pairs (96%) of duplicated genes were < 1, indicating that they had experienced strong purifying selection, and only one pair of tandem duplicated genes showed a Ka/Ks ratio > 1, suggesting that these genes may have experienced positive selection. The Ks values of these duplicated gene pairs ranged from 0.2153 to 6.1528, corresponding to divergence times of 16.41 to 468.96 Mya (Table S3).

To further infer the phylogenetic mechanism of the CsHsf and CsHsp families, we performed a synteny analysis of the heat shock genes in cucumber, Arabidopsis, and rice (Fig. S9). A total of 52 pairs of heat shock genes showed the synonymous relationship between cucumber and Arabidopsis, followed by rice was 33. Many cucumber heat shock genes were homologous to both Arabidopsis and rice, and most of the homologous genes had Ka/Ks < 1 (Table S4), suggesting that these genes are essential in plant evolution and contribute greatly to maintaining the function of heat shock genes.

Expression patterns of CsHsf and CsHsp genes in different tissues

To address the expression patterns of CsHsf and CsHsp genes in different tissues, we analyzed the cucumber Illumina RNA-seq data obtained from the Cucurbit Genomics Database (http://cucurbitgenomics.org/rnaseq/home), and obtained the expression patterns of CsHsf and CsHsp genes in different tissues including root, stem, leaf, male flower, female flower, ovary (unexpanded), expanded ovary (fertilized), expanded ovary (unfertilized), and tendril (Fig. 4). The expression patterns of CsHsf and CsHsp genes varied largely in different tissues. Among the CsHsfA family, CsHsf-9 (CsHsfA1) and CsHsf-7 (CsHsfA2) were expressed at relatively high levels in all tissues, while CsHsf-4 (CsHsfA9) was not expressed in all tissues. Especially, CsHsf-23 (CsHsfA7) was more highly expressed in the tendril compared with other tissues, indicating that it might function in tendril. In CsHsfB family, CsHsf-12 and CsHsf-17 were expressed at high abundances in all tissues. In contrast, CsHsf-15 (CsHsfB3) was not expressed in all tissues. Notably, CsHsf-20 (CsHsfB1) more highly expressed in the root than in other tissues, indicating that it might play an important role in root. As for CsHsp genes, except for that CsHsp20-6, CsHsp20-24, CsHsp20-26, and CsHsp20-30 were almost not expressed in any tissue or organ, most of the CsHsp genes were expressed in at least one tissue. Almost all of

Analysis of putative cis-acting elements in the promoters of CsHsf and CsHsp genes

To identify potential cis-acting elements in the promoter regions of CsHsf and CsHsp genes, 1500 bp upstream sequences from translational start sites extracted from the cucumber genome database, were submitted to the PlantCARE database, and 12 types of hormone response elements and 9 types of stress-induced components were found (Table S5). Among the 95 genes, 60, 34, 55, 36, 50 and 38 (63%, 36%, 58%, 38%, 53% and 40%) genes have at least one type of abscisic acid (ABA)-responsive element, auxin (IAA)-responsive elements, ethylene (ER)-responsive elements, gibberellin (GA)-responsive elements, Jasmonic acid (MeJA)-responsive elements and salicylic acid (SA)-responsive elements, respectively (Fig. 3). In order to verify whether these genes can respond to hormone treatment, we performed different hormone treatments on cucumber plants, including ABA, MeJA, SA, and then detected the gene expression level by qRT-PCR. The results showed that if a gene’s promoter has a certain hormone-responsive element, indeed the gene can respond to the corresponding hormone treatment (Fig. S10). For stress-induced components, one or more ARE(s) existed in 14 (61%) CsHsf and 61 (87%) CsHsp genes, being involved in Hypoxia-inducible response. WUN-motifs were presented in 9 (39%) CsHsf and 31 (43%) CsHsp genes, participating in wound response. What’s more, heat shock elements (HSEs) were found in 12 (52%) CsHsf and 40 (56%) CsHsp genes. In addition, other elements, such as MBS, LTR, TC-rich repeats, and W-box, known to function as stress-induced components, were also found in CsHsf and CsHsp genes, being effective at variable positions and effectively responding to drought, low temperature, and biotic stresses, while GC-motif, which is useful in anoxic specific inducible response, only existed in 4 CsHsp genes (Fig. 3). The above analysis of cis-elements showed that CsHsf and CsHsp genes might respond rapidly to different stress conditions to reduce the damage caused by unfavorable environments, through maintaining physiological and metabolic balance for the normal growth of cucumber.
Expression profiles of CsHsf and CsHsp genes in response to abiotic and biotic stresses

To explore the responses of CsHsf and CsHsp genes to biotic and abiotic stresses, the expression patterns of CsHsf and CsHsp genes in response to stresses were investigated using three sets of RNA-seq data. Our RNA-seq data were used to analyze the responses of CsHsf and CsHsp genes to heat stress, and the RNA-seq data response to NaCl and downy mildew stresses were obtained from the Cucurbit Genomics Database (http://cucurbitgenomics.org/rnaseq/home) for the expression pattern of CsHsf and CsHsp genes. According to the log2 (FPKM + 1) value, the expression of many CsHsf and CsHsp genes was increased after heat treatment (Fig. 5). Interestingly, CsHsf and CsHsp genes up-regulated under heat stress were usually also up-regulated under NaCl treatment, indicating that cucumber share similar responses to heat and salt stresses. The expression of CsHsf-7 (HsfA2) increased sharply after 3 h of heat treatment compared with the control, indicating that CsHsf-7 is a very sensitive factor during the heat stress response in cucumber. Previous studies have shown that HsfA1 is a master regulator that triggers the thermal response, leading to the acquired thermotolerance in tomatoes and Arabidopsis, but it is not induced by heat stress (Mishra et al. 2002; Yoshida et al. 2011). Here, CsHsf-9 (CsHsfA1) has the same expression pattern with AtHsfA1 and SlHsfA1 (Fig. 5a), suggesting that CsHsf-9 might also be a master regulator in cucumber heat stress tolerance. The expression level of most CsHsp20s increased significantly in the first 3 h after heat treatment compared with the control and then decreased gradually in the next 3 h (Fig. 5b). After heat or NaCl treatment, the expression level of many CsHsp60s slightly decreased (Fig. 5c). The infection of downy mildew reduced the expression level of most CsHsf and CsHsp in cucumber (Fig. 5), suggesting that CsHsf and CsHsp might play contrary roles between abiotic and biotic stresses.
Potential protein–protein interactions between CsHsfs and CsHsps

To further explore the possible protein–protein interactions between CsHsfs and CsHsps, an interaction network was constructed using the STRING program. As one of the most important heat shock transcription factors, CsHsf-7 have potential interactions with CsHsp20s, CsHsp70s, CsHsp90s, and CsHsp100s, implying that CsHsf-7 is an activator of the downstream CsHsps and acts as a key regulator to improving adaptability of cucumber to various stresses (Fig. 6). According to the prediction, there are strong interactions between CsHsf-9 and many CsHp70s (CsHsp70-6, CsHsp70-9, and CsHsp70-10), reflecting previous studies that Hsp70 and Hsp90 interact with HsfA1 under normal conditions and inhibit HsfA1 activity (Hahn et al. 2011; Ohama et al. 2017). In addition to the potential regulatory relationship between CsHsfs and CsHsps, we also detected many other interactions among CsHsps of different subfamilies, such as interactions among CsHsp60, CsHsp70, and CsHsp90, suggesting that the CsHsps of different subfamilies might also be activated or inhibited through interactions with each other when cucumber was subjected to various stresses.

Discussion

The high temperature usually causes the expression changes of related genes in organisms (Ohama et al. 2017). Among the changed genes, Hsps encode a type of conserved proteins, which is synthesized rapidly by organisms under heat stress, and was first found in the salivary glands of Drosophila melanogaster (Tissiéres et al. 1974). Studies have shown that when plants are subjected to heat stress, they can develop tolerance to lethal high temperatures, in which Hsps play important roles. In addition, low temperature, drought, heavy metals, salinization, and ABA can induce plants to produce Hsps (Wang et al. 2004). The expression of Hsps is regulated by heat shock factors (Hsfs), which are the central regulators of Hsps expression and participates in heat stress response (Guo et al. 2016). Therefore, it is important to clarify the number, tissue specific expression and expression pattern induced by stresses of CsHsfs and CsHsps for understanding the mechanism of cucumber stress tolerance, especially heat tolerance. In this study, we identified 95 heat shock genes composed of 23 CsHsfs and 72 CsHsps (33 CsHsp20s, 15 CsHsp60s, 12 CsHsp70s, 6 CsHsp90s, 6 CsHsp100s) in cucumber, and then analyzed their gene structure, conserved motifs, phylogenetic relationships, chromosome distribution, duplication events, cis-acting elements, expression patterns under stress and potential protein–protein interaction among CsHsfs and CsHsps.

Phylogenetic analysis is often used to provide insights into the gene evolutionary relationships of species and to help to identify orthologs between species and paralogs within species. Based on the full-length protein sequences of Arabidopsis, rice, tomato, and cucumber, we constructed unrooted phylogenetic trees of Hsps and Hsfs, respectively. Though the total number of Hsf genes was similar to that of Arabidopsis and tomato (Guo et al. 2008), the members of some specific Hsf subclasses in cucumber were different from the other two species. For instance, the number of subclass HsfA1 members in cucumber was less, only CsHsf-9 belongs to the HsfA1 subclass, but the number in subclass HsfA4 was more than that in tomato and Arabidopsis (Fig. S3). Most of CsHsp20 genes were classified into nucleocytoplasmic subfamilies (Fig. S4), which are similar to the distribution in Arabidopsis, rice, and tomato (Sarkar et al. 2009; Scharf et al. 2001; Yu et al. 2016). Among these subfamilies, CI was the largest subfamily, containing 14 CsHsp20 genes. Since the proteins are mainly transported to cytoplasm (Nelson et al. 2014), we speculated that cytoplasm may be the main place for Hsp20 proteins as molecular chaperones to interact with the target proteins to regulate their folding, positioning, accumulation and degradation. Hsp70 gene family was divided into 7 groups, and each group contains Hsp70s from Arabidopsis, rice, and cucumber, indicating that the Hsp70 family genes occurred in the early stage of plant evolution, earlier than the divergence of dicots and monocots (Fig. S6).

In terms of intron number, intron, and exon length, the most closely related members of the same Hsf or Hsp subfamily have similar gene structures (Fig. 1). The number of introns usually affects the activity of gene transcription regulation (Le Hir et al. 2003). Some studies suggest that under various stress conditions, the splicing mechanism of RNA is often disturbed, and transcripts without introns can be transferred from the nucleus to the cytoplasm rapidly without splicing, which improves the response efficiency of plants, thus leading to the tendency to select intronless genes during evolution (Le Hir et al. 2003). The fewer introns the
were intronless (Fig. 1), and most of them were CsHsp20s. Genes are shorter in length and have fewer introns. 39% of families, the Hsp CsHsp20 family comparing with other conditions and environmental stimuli (Chung et al. 2006).

Comparing with other Hsp families, the CsHsp20 family genes are shorter in length and have fewer introns. 39% of families, the Hsp CsHsp20 family comparing with other conditions and environmental stimuli (Chung et al. 2006).

Comparing with other Hsp families, the CsHsp20 family genes are shorter in length and have fewer introns. 39% of families, the Hsp CsHsp20 family comparing with other conditions and environmental stimuli (Chung et al. 2006). NaCl and downy mildew stresses were obtained from the Cucur-

and the RNA-seq data for CsHsp genes in response to CsHsf and genes to heat stress, to analyze the response of CsHsp genes under heat, NaCl and downy mildew stresses, which may be related to mitochondrial activity and programmed cell death (PCD) (Rikhvanov et al. 2007). Different Hsps were induced to different levels under heat stress, indicating that CsHsps in different subfamilies may have distinct regulatory properties. In addition, the CsHsf and CsHsp gene promoters have multiple stress response components (Fig. 3), and there are many potential protein–protein interactions among CsHsfs and CsHsps (Fig. 6). These provide many chances to form a complex regulatory network to respond to different stress conditions in cucumber.

In the end, we believe that CsHsf-7 is a very ideal candidate gene for plant breeding of heat and salt tolerance. First, from the results of RNA-seq (Fig. 5), among all 23 hsfs, CsHsf-7 is the most strongly induced one by heat and salt stress. At the same time, from the perspective of phylogenetic relationship (Fig. S3), CsHsf-7 belongs to the A2 subfamily, and its homologous gene, AtHSFA2 in Arabidopsis plays an important role in responding to multiple stress responses. For example, the overexpression of AtHSFA2 in Arabidopsis increased heat, salt and osmotic stress tolerance (Guo et al. 2016). Therefore, we could screen the lines with higher induced CsHsf-7 among various cucumber varieties, and compare the difference of CsHsf-7 promoter based on the expression of CsHsf-7. Then, the marker linked with the highly induced CsHsf-7 could be developed for plant breeding of abiotic stress tolerance. In addition, in terms of the possible protein–protein interaction between CsHsf and CsHsp (Fig. 6), CsHsf-7 had potential interactions with multiple CsHsps, implying that CsHsf-7 might co-operate with CsHsps to play roles in abiotic stress tolerance. Then, the CsHsps, such as CsHsp70-1, 70-7 and 90-4, could be driven under CsHsf-7 promoter for their coordination in abiotic stress tolerance. In addition, CRISPR-Cas9 or RNA interference techniques could be used to knock out or knock down the genes inhibited by P. cubensis, such as CsHsf-5 and -12 (Fig. 5), to increase the resistance of cucumber to downy mildew. Of course, their specific functions need to be verified by subsequent transgenic experiments.

A large of studies have shown that Hsf and Hsp genes produced a marked effect in response to a variety of abiotic stresses (Guo et al. 2016; Scharf et al. 2012). In this study, we analyzed the dynamic expression levels of CsHsf and CsHsp genes under heat, NaCl, and downy mildew stresses by RNA-seq (Fig. 5). We have observed that most CsHsf and CsHsp genes were induced by heat stress, but not all. Interestingly, CsHsf and CsHsp genes up-regulated under heat stress were usually also up-regulated under NaCl treatment, indicating that cucumber has similar responsive mechanism between heat and salt stresses. The results showed that the expression of most heat shock genes was up-regulated, while some CsHsp60 genes were down-regulated after heat treatments, which may be related to mitochondrial activity and programmed cell death (PCD) (Rikhvanov et al. 2007). Different Hsps were induced to different levels under heat stress, indicating that CsHsps in different subfamilies may have distinct regulatory properties. In addition, the CsHsf and CsHsp gene promoters have multiple stress response components (Fig. 3), and there are many potential protein–protein interactions among CsHsfs and CsHsps (Fig. 6). These provide many chances to form a complex regulatory network to respond to different stress conditions in cucumber.

In the end, we believe that CsHsf-7 is a very ideal candidate gene for plant breeding of heat and salt tolerance. First, from the results of RNA-seq (Fig. 5), among all 23 hsfs, CsHsf-7 is the most strongly induced one by heat and salt stress. At the same time, from the perspective of phylogenetic relationship (Fig. S3), CsHsf-7 belongs to the A2 subfamily, and its homologous gene, AtHSFA2 in Arabidopsis plays an important role in responding to multiple stress responses. For example, the overexpression of AtHSFA2 in Arabidopsis increased heat, salt and osmotic stress tolerance (Guo et al. 2016). Therefore, we could screen the lines with higher induced CsHsf-7 among various cucumber varieties, and compare the difference of CsHsf-7 promoter based on the expression of CsHsf-7. Then, the marker linked with the highly induced CsHsf-7 could be developed for plant breeding of abiotic stress tolerance. In addition, in terms of the possible protein–protein interaction between CsHsf and CsHsp (Fig. 6), CsHsf-7 had potential interactions with multiple CsHsps, implying that CsHsf-7 might co-operate with CsHsps to play roles in abiotic stress tolerance. Then, the CsHsps, such as CsHsp70-1, 70-7 and 90-4, could be driven under CsHsf-7 promoter for their coordination in abiotic stress tolerance. In addition, CRISPR-Cas9 or RNA interference techniques could be used to knock out or knock down the genes inhibited by P. cubensis, such as CsHsf-5 and -12 (Fig. 5), to increase the resistance of cucumber to downy mildew. Of course, their specific functions need to be verified by subsequent transgenic experiments.
Conclusions

In summary, we identified 95 members of cucumber \textit{Hsf} and \textit{Hsp} (including \textit{Hsp20}, \textit{Hsp60}, \textit{Hsp70}, \textit{Hsp90}, and \textit{Hsp100}) gene families in this study, and then analyzed their gene structures, conserved motifs, phylogenetic relationships, chromosome distribution, duplication events, cis-acting elements, and potential protein–protein interaction. Then, their tissue specific expression patterns and induced expression patterns under heat, NaCl and downy mildew stresses were also determined using RNA-seq data.

This study provided comprehensive information on the \textit{Hsf} and \textit{Hsp} gene families in cucumber and will aid further research on the functions of \textit{CsHsf} and \textit{CsHsp} genes for improving cucumber tolerance to biotic and abiotic stresses.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10725-021-00739-z.

Author contribution ZR conceived the project. ZR, CC, XC, and ZW designed the experiments. XC, ZW, RT, and LW performed the experiments and analyzed the data. XC and ZW wrote the manuscript. ZR revised the manuscript.
Funding This work was supported by funding from the National Natural Science Foundation of China (31872950 and 31672170), the Natural Science Foundation of Shandong Province (JQ201309), the Shandong “Double Tops” Program (SYL2017YSTD06), and the ‘Taishan Scholar’ Foundation of the People’s Government of Shandong Province (ts20130932).

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

References

Agarwal M, Katiyar-Agarwal S, Sahi C, Gallie DR, Grover A (2001) Arabidopsis thaliana Hsp100 proteins: kith and kin. Cell Stress Chaperones 6(3):219–224. https://doi.org/10.1379/1466-1268(2001)006<0219:ATHPKA>2.0.CO;2

Aleem S, Sharif I, Amin M, Tahir M, Parveen N, Aslam R, Najeebullah M, Shahid MTH (2020) Heat tolerance in vegetables in the current genomic era: an overview. Plant Growth Regul 92(3):497–516. https://doi.org/10.1007/s10725-020-00658-5

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 Res 37:W202–W208. https://doi.org/10.1093/nar/gkp335

Cai B, Li Q, Xu Y, Yang L, Bi H, Ai X (2016) Genome-wide analysis of the fructose 1,6-bisphosphate aldolase (FBA) gene family and functional characterization of FBA7 in tomato. Plant Physiol Biochem 108:251–265. https://doi.org/10.1016/j.plaphy.2016.07.019

Chang 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 Physiol 143(1):251–262. https://doi.org/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. Mol Genet Genomics 286(2):171–187. https://doi.org/10.1007/s00438-011-0638-8

Chung BY, Simons C, Firth AE, Brown CM, Hellens RP (2006) The balance of nuclear import and export determines the intracellular distribution and function of tomato heat stress transcription factor HsfA2. Mol Cell Biol 21(5):1759–1768. https://doi.org/10.1128/MCB.21.5.1759-1768.2001

Hu W, Hu G, Han B (2009) Genome-wide survey and expression profiling of heat shock proteins and heat shock factors revealed upregulated and stress specific response under abiotic stresses in rice. Plant Sci 176(4):583–590. https://doi.org/10.1016/j.plantsci.2009.01.016

Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31(8):1296–1297. https://doi.org/10.1093/bioinformatics/btu817

Jacob P, Hirt H, Bendahmane A (2017) The heat-shock protein/chaperone network and multiple stress resistance. Plant Biotechnol J 15(4):405–414. https://doi.org/10.1111/pbi.12659

Koch MA, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in Arabidopsis, Arabis, and related genera (Brassicaceae). Mol Biol Evol 17(10):1483–1498. https://doi.org/10.1093/oxfordjournals.molbev.a026248

Kotak S, Port M, Ganguli A, Bicker F, Von Koskull-Döring PJT (2004) Characterization of C-terminal domains of Arabidopsis heat stress transcription factors (Hsfs) and identification of a new signature combination of plant class A Hsfs with AHA and NES motifs essential for activator function and intracellular localization. Plant J 39(1):98–112. https://doi.org/10.1111/j.1365-313X.2004.02111.x

Krishna P, Gloor G (2001) The Hsp90 family of proteins in Arabidopsis thaliana. Cell Stress Chaperones 6(3):238–246. https://doi.org/10.1379/1466-1268(2001)006<0238:thfopi>2.0.CO;2

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874. https://doi.org/10.1093/molbev/msw054

la Cour T, Kiemer L, Molgaard A, Gupta R, Skriver K, Brunak S (2004) Analysis and prediction of leucine-rich nuclear export signals. Protein Eng Des Sel 17(6):527–536. https://doi.org/10.1093/protein/gzh062

Lamers J, van der Meer T, Testerink C (2020) How plants sense and respond to stressful environments. Plant Physiol 182(4):1624–1635. https://doi.org/10.1104/pp.19.01464

Le Hir H, Nott A, Moore MJ (2003) How introns influence and enhance eukaryotic gene expression. Trends Biochem Sci 28(4):215–220. https://doi.org/10.1016/S0968-0004(03)00052-5

Lee U, Rioflorido I, Hong SW, Larkindale J, Waters ER, Vierling E (2007) The Arabidopsis ClpB/Hsp100 family of proteins: chaperones for stress and chloroplast development. Plant J 49(1):115–127. https://doi.org/10.1111/j.1365-313X.2006.02940.x

Lescot M, Dehais P, Gloor G, Zhang H, Luo J, Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31(8):1296–1297. https://doi.org/10.1093/bioinformatics/btu817
analysis of promoter sequences. Nucleic Acids Res 30(1):325–327. https://doi.org/10.1093/nar/30.1.325
Li M, Ji L, Yang X, Meng Q, Guo S (2012) The protective mechanisms of CaHSP26 in transgenic tobacco to alleviate photoinhibition of PSII during chilling stress. Plant Cell Rep 31(11):1969–1979. https://doi.org/10.1007/s00299-012-1309-x
Li H, Liu SS, Yi CY, Wang F, Zhou J, Xiao XJ, Shi K, Zhou YH, Yu JQ (2014a) Hydrogen peroxide mediated abscisic acid-induced HSP70 accumulation and heat tolerance in grafted cucumber plants. Plant Cell Environ 37(12):2768–2780. https://doi.org/10.1111/pce.12360
Li M, Guo S, Xu Y, Meng Q, Li G, Yang X (2014b) Glycine betaine-mediated potentiating of HSP gene expression involves calcium signaling pathways in tobacco exposed to NaCl stress. Physiol Plant 150(1):63–75. https://doi.org/10.1111/ppl.12067
Li M, Ji L, Jia Z, Yang X, Meng Q, Guo S (2018) Constitutive expression of CaHSP22.5 enhances chilling tolerance in transgenic tobacco by promoting the activity of antioxidative enzymes. Funct Plant Biol 45(5):575–585. https://doi.org/10.1071/FP17226
Li Q, Li H, Huang W, Xu Y, Zhou Q, Wang S, Ruan J, Huang S, Zhang Z (2019) A chromosome-scale genome assembly of cucumber (Cucumis sativus L.). Gigascience 8(6):giz072. https://doi.org/10.1093/gigascience/giz072
Lin BL, Wang JS, Liu HC, Chen RW, Meyer Y, Barakat A, Delseny M (2001) Genomic analysis of the Hsp70 superfamily in Arabidopsis thaliana. Cell Stress Chaperones 6(3):201–208. https://doi.org/10.1379/jhsp.2001.00600.0201:ga0ths62.2.0.c:2
Liu HC, Chang YY (2012) Acquired therotolerance independent of heat shock factor A1 (HsfA1), the master regulator of the heat stress response. Plant Signal Behav 7(5):547–550. https://doi.org/10.4161/psb.19803
Liu HC, Chang YY (2013) Common and distinct functions of Arabidopsis class A1 and A2 heat shock factors in diverse abiotic stress responses and development. Plant Physiol 163(1):276–290. https://doi.org/10.1104/pp.113.221168
Liu Y, Wang L, Zhang D, Li D (2012) Expression analysis of segmentally duplicated ZmMPK3-1 and ZmMPK3-2 genes in maize. Plant Mol Biol Report 31(2):457–463. https://doi.org/10.1007/s11105-012-0489-4
Liu Y, Wan H, Yang Y, Wei Y, Li Z, Ye Q, Wang R, Ruan M, Yao Z, Zhou G (2014) Genome-wide identification and analysis of heat shock protein 90 in tomato. Yi Chuan 36(10):1043–1052. https://doi.org/10.13724/SPJ1.2014.1043
Liu C, Xie T, Chen C, Luan A, Long J, Li C, Ding Y, He Y (2017) Genome-wide organization and expression profiling of the R2R3-MYB transcription factor family in pineapple (Ananas comosus). BMC Genomics 18(1):503. https://doi.org/10.1186/s12864-017-3896-y
Luan Q, Chen C, Liu M, Li Q, Wang L, Ren Z (2019) CsWRKY50 mediates defense responses to Pseudoperonospora cubensis infection in Cucumis sativus. Plant Sci 279:59–69. https://doi.org/10.1016/j.plantsci.2018.11.002
Ma W, Guan X, Li J, Pan R, Wang L, Liu F, Ma H, Zhu S, Hu J, Ruan YL, Chen Y, Zhang T (2019) Mitochondrial small heat shock protein mediates seed germination via thermal sensing. Proc Natl Acad Sci U S A 116(10):4716–4721. https://doi.org/10.1073/pnas.1815790116
Mishra SK, Tripp J, Winkelhaus S, Tschiersch B, Theres K, Nover L, Scharf KD (2002) In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. Genes Dev 16(12):1555–1567. https://doi.org/10.1101/gad.228802
Nejat N, Mantri N (2017) Plant immune system: crosstalk between responses to biotic and abiotic stresses the missing link in understanding plant defence. Curr Issues Mol Biol 23:1–16. https://doi.org/10.21775/cimb.023.001
Nelson CJ, Li L, Millar AH (2014) Quantitative analysis of protein turnover in plants. Proteomics 14(4–5):579–592. https://doi.org/10.1002/pmic.201300240
Ogawa D, Yamaguchi K, Nishiuichi T (2007) High-level overexpression of the Arabidopsis HsfA2 gene confers not only increased thermotolerance but also salt/osmotic stress tolerance and enhanced callus growth. J Exp Bot 58(12):3373–3383. https://doi.org/10.1093/jxb/erm184
Ohama N, Sato H, Shinozaki K, Yamaguchi-Shinozaki K (2017) Transcriptional regulatory network of plant heat stress response. Trends Plant Sci 22(1):53–65. https://doi.org/10.1016/j.tplants.2016.08.015
Pei HC, Sun QX, Hao QQ, Lv B, Wu JJ, Fu DL (2015) The HSPP90-RARI-SGT1 based protein interactome in barley and stripe rust. Physiol Mol Plant Pathol 91:11–19. https://doi.org/10.1016/j.pmpp.2015.02.004
Rikhvanov EG, Gamborg KZ, Varakina RN, Rusaleva TM, Fedoseeva IV, Tauson EL, Stupnikova IV, Stepanov AV, Borovskii GB, Voinikov VKJTP (2007) Nuclear-mitochondrial cross-talk during heat shock in Arabidopsis cell culture. Plant J 52(4):763–778. https://doi.org/10.1111/j.1365-313X.2007.03275.x
Sarkar NK, Kim YK, Grover A (2009) Rice hsp genes: genomic organization and expression profiling under stress and development. BMC Genomics 10(1):393. https://doi.org/10.1186/1471-2164-10-393
Sarkar NK, Kundnani P, Grover A (2013) Functional analysis of Hsp70 superfamily proteins of rice (Oryza sativa). Cell Stress Chaperones 18(4):427–437. https://doi.org/10.1007/s12192-012-0395-6
Scharf K-D, Siddique M, Vierling E (2001) The expanding family of Arabidopsis thaliana small heat stress proteins and a new family of proteins containing α-crystallin domains (Acd proteins). Cell Stress Chaperones 6(3):225. https://doi.org/10.1379/1466-1268(2001)0062.0.CO;2
Scharf KD, Berberich T, Ebersberger I, Nover L (2012) The plant heat stress transcription factor (Hsf) family: structure, function and evolution. Biochim Biophys Acta 1819(2):104–119. https://doi.org/10.1016/j.bbagrm.2011.10.002
Scharf KD, Rose S, Zott W, Schoffl F, Nover L (1990) Three tomato genes code for heat stress transcription factors with a region of remarkable homology to the DNA-binding domain of the yeast HSF. EMBO J 9(13):4495–4501. https://doi.org/10.1002/1460-2075.1990.tb07900.x
Singh A, Singh U, Mittal D, Grover A (2010) Genome-wide analysis of rice CbP/HSP100, CbPc and CbPd genes. BMC Genomics 11(1):95. https://doi.org/10.1186/1471-2164-11-95
Su H, Zhang S, Yuan X, Chen C, Wang X-F, Hao Y-J (2013) Genome-wide analysis and identification of stress-responsive genes of the NAM–ATAF1, 2–CUC2 transcription factor family in apple. Plant Physiol Biochem 71:11–21. https://doi.org/10.1016/j.plaphy.2013.06.022
Sun L, Liu Y, Kong X, Zhang D, Pan J, Zhou Y, Wang L, Li D, Yang X (2012) ZmHSP16.9, a cytosolic class I small heat shock protein in maize (Zea mays), confers heat tolerance in transgenic tobacco. Plant Cell Rep 31(8):1473–1484. https://doi.org/10.1007/s00201-012-1262-8
Swindell WR, Huebner M, Weber AP (2007) Transcriptional profiling of Arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. BMC Genomics 8(1):125. https://doi.org/10.1186/1471-2164-8-125
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.
