



Article

Identification of Heat Shock Transcription Factor Genes Involved in Thermotolerance of Octoploid Cultivated Strawberry

Wan-Yu Liao ^{1,†}, Lee-Fong Lin ^{2,†}, Jing-Lian Jheng ³, Chun-Chung Wang ^{4,5}, Jui-Hung Yang ⁵ and Ming-Lun Chou ^{2,*}

¹ Institute of Medical Sciences, Tzu-Chi University, Hualien 97004, Taiwan; 98751101@gms.tcu.edu.tw

² Department of Life Sciences, Tzu-Chi University, Hualien 97004, Taiwan; leelin@gms.tcu.edu.tw

³ Department of Molecular Biology and Human Genetics, Tzu-Chi University, Hualien 97004, Taiwan; pudding30167@yahoo.com.tw

⁴ Institute of Molecular Medicine, National Tsing-Hua University, Hsinchu 30013, Taiwan; aricwang@itri.org.tw

⁵ Biomedical Technology and Device Research Laboratories, Industrial Technology Research Institute, Hsinchu 30011, Taiwan; JHYen@itri.org.tw

* Correspondence: mlchou1015@gms.tcu.edu.tw; Tel.: +886-3-846-5615; Fax: +886-3-857-2526

† These authors contributed equally to this study.

Academic Editor: Jianhua Zhu

Received: 2 September 2016; Accepted: 13 December 2016; Published: 17 December 2016

Abstract: Heat shock transcription factors (HSFs) are mainly involved in the activation of genes in response to heat stress as well as other abiotic and biotic stresses. The growth, development, reproduction, and yield of strawberry are strongly limited by extreme temperatures and droughts. In this study, we used Illumina sequencing and obtained transcriptome data set from *Fragaria × ananassa* Duchesne cv. Toyonoka. Six contigs and three unigenes were confirmed to encode HSF proteins (FaTHSFs). Subsequently, we characterized the biological functions of two particularly selected unigenes, *FaTHSFA2a* and *FaTHSFB1a*, which were classified into class A2 and B HSFs, respectively. Expression assays revealed that *FaTHSFA2a* and *FaTHSFB1a* expression was induced by heat shock and correlated well with elevated ambient temperatures. Overexpression of *FaTHSFA2a* and *FaTHSFB1a* resulted in the activation of their downstream stress-associated genes, and notably enhanced the thermotolerance of transgenic *Arabidopsis* plants. Besides, both *FaTHSFA2a* and *FaTHSFB1a* fusion proteins localized in the nucleus, indicating their similar subcellular distributions as transcription factors. Our yeast one-hybrid assay suggested that *FaTHSFA2a* has trans-activation activity, whereas *FaTHSFB1a* expresses trans-repression function. Altogether, our annotated transcriptome sequences provide a beneficial resource for identifying most genes expressed in octoploid strawberry. Furthermore, HSF studies revealed the possible insights into the molecular mechanisms of thermotolerance, thus rendering valuable molecular breeding to improve the tolerance of strawberry in response to high-temperature stress.

Keywords: octoploid cultivated strawberry (*Fragaria × ananassa* Duch. cv. Toyonoka); heat shock transcription factor; heat stress; Illumina sequencing; transcriptome; thermotolerance

1. Introduction

High temperature is one of the most crucial abiotic stresses in fields worldwide because it can considerably affect plant growth and crop production [1,2]. When the temperature increases beyond the optimal growth condition, it causes heat-stress responses in higher plants, leads to the inhibition of photosynthesis, and results in chlorophyll degradation [3]. Because of previous exposure

to high temperatures, plants develop a series of defense mechanisms to acquire thermotolerance against unavoidable high temperatures, which includes markedly increased transcript levels of heat-stress-responsive genes encoding heat-shock proteins (HSPs) and small HSPs, which function as molecular chaperones for protein quality control under heat stress [4,5]. In addition to HSPs, other regulatory proteins are involved in heat-stress responses, such as dehydration-responsive element-binding transcription factor 2 (DREB2), galactinol synthase 1 (GolS1), and ascorbate peroxidase 2 (APX2) that function to facilitate plant survival under stressful conditions [6–9].

Heat shock transcription factors (HSFs), the central regulators of heat shock stress response, regulate the expression of many heat-stress-inducible genes by recognizing the conserved binding motifs (heat stress element, HSE) that exist in their promoters in all eukaryotic organisms [10,11]. The HSF family, similar to many transcription factors, has a conserved modular structure with a N-terminal DNA-binding domain (DBD) characterized by a central helix-turn-helix motif; a hydrophobic coiled-coil region (HR-A/B) composed of hydrophobic heptad repeats essential for oligomerization; short peptide motifs required for nuclear import (nuclear localization signals, NLS) and export; and a C-terminal activation domain (CTAD) rich in aromatic, hydrophobic, and acidic amino acids, the so-called AHA motifs [12–17]. HSFs utilize their oligomerization domains to form trimers and function as sequence-specific trimeric DNA-binding proteins via the signal transduction pathway to activate the expression of the *HSP* genes [18].

Only a few *HSF* genes were identified in yeast, fruit fly and vertebrates, whereas genomes of Arabidopsis, rice, tomato, and soybean have been reported to contain 21, 25, 18 and 34 *HSF* genes, respectively [11,14,19–22]. According to structural characteristics and phylogenetic comparisons, HSFs of those plants have been categorized into three major classes (classes A, B, and C), which revealed the difference in their flexible linkers between the A and B parts of the HR-A/B regions [21,23]. Most class A HSFs contain the AHA motifs and activate the transcription of HSPs through trans-activation by binding some basic transcription protein complexes, whereas class B and C HSFs exhibit no trans-activation activity because of the lack of the AHA motif and function as repressors or co-activators [16,21,24]. It has been recently reported that the structure of class B HSFs (except HSF5) comprises a characteristic tetrapeptide (LFGV) in the C-terminal domain, acting as a repressor domain (RD) [25–27]. In addition to their role in heat stress, previous studies have reported that HSFs may play vital roles in plants under abiotic stress conditions: for instance, cold, salt, drought, and oxidative conditions [28–30]. Thus, the multiplicity of plant HSFs suggests their functional diversity and complexity in plants.

Recent studies have identified a high number of plant *HSF* genes from more than 20 plant species, including monocots and eudicots, on the basis of the latest development of next-generation sequencing (NGS) technology and availability of the growing number of complete plant genomic and transcriptome sequences resources. Furthermore, 15–56 HSF members were found in any given plant species, including 25 HSF-encoding genes in rice [17,31], 21 in Arabidopsis [17], 30 in maize [17], 24 in tomato [17,27], 25 in pepper [32], 29 in Chinese white pear [33], 17 in Chinese plum [33], 33 in European pear [33], 17 in peach [33], 52 in soybean [17], and at least 56 in wheat [34]. A diploid woodland strawberry species, *Fragaria vesca* Coville, whose haploid genome size was approximately 240 Mb, is the smallest among the rosaceous species [35]. Its complete genome sequence resources have been available since 2010 [36] and referenced to its octoploid cultivated relatives. Furthermore, another well-known octoploid strawberry cultivar, “Strawberry Festival”, was subjected to transcriptome analysis by using the Roche-454 GS-FLX system. To increase the transcript diversity and discovery in their analysis, Folta et al. [37] examined whole plants, and their detached tissues were subjected to at least one of the 30 specific treatments, including various growth regulators, light conditions, pharmacological agents, or stresses. The cDNA pools isolated from various tissues were sequenced using the Roche-454 GS-FLX system and assembled into more than 32,000 contigs [37]. Comparative genetic mapping analysis of genomes of diploid and octoploid strawberries showed that they shared a high level of macrosynteny and collinearity [38]. Because of its small herbaceous stature,

ease of propagation, short reproductive cycle, efficient regeneration and facile transformation, and available genomic and transcriptome resources, strawberry emerges as an advantageous and powerful functional genomics system for testing gene functions and representing many plants of the Rosaceae family [35,36,39–42].

Garden strawberry (pineapple strawberry), *F. × ananassa* Duchesne, is a hybrid cultivar of two wild *Fragaria* species (*F. virginiana* and *F. chiloensis*). It is a perennial herbaceous plant of the Rosaceae family, has gradually become a crucial and valuable fruit crop for many regional economies worldwide, and offers high levels of healthful compounds. In Taiwan, the most well-known and major cultivated strawberry in fields is *F. × ananassa* Duch. cv. Toyonoka because of its widely appreciated characteristic aroma, bright red color, juicy texture, and sweetness. Genomically, *F. × ananassa* is the most complex form among fruit tree crops and ornamental plants, harboring eight sets of chromosomes ($2n = 8 \times = 56$) originally derived from as many as four diploid ancestors [39]. In this study, we performed large-scale transcriptome sequencing of *F. × ananassa* Duch. cv. Toyonoka through high-throughput RNA-sequencing and described the sequences of more than 65,000 generated non-redundant (Nr) transcribed unigenes from six tissues. These unigenes were functionally annotated and analyzed considering their gene ontology (GO) and relationship to metabolic pathways. These unigene data sets from our studies will provide researchers with useful foundation for researching gene expression, genomics and functional genomics of not only *F. × ananassa* but also its closely related species.

In this study, we identified nine HSF genes with full-length cDNA sequences (*F. × ananassa* Duch. cv. Toyonoka Heat Shock Transcription Factor Protein, *FaTHSF*), determined from 31 contigs and 15 unigenes annotated as HSF proteins according to our results of *de novo* transcriptome sequencing. Two *FaTHSF* genes, *FaTHSFA2a* and *FaTHSFB1a*, were particularly selected not only because they represent class A2 and B1 HSFs, respectively, by multiple alignment and phylogenetic analysis, but also previous studies have revealed their crucial activities as strong heat-associated HSF proteins [43–48]. Reverse transcription polymerase chain reaction (RT-PCR) results showed that *FaTHSFA2a* and *FaTHSFB1a* transcripts were induced by high temperature, and their induced expression patterns correlated well with elevated ambient temperatures. We also found that the overexpression of *FaTHSFA2a* and *FaTHSFB1a* proteins caused the activation of downstream stress-associated genes, namely *Hsp101*, *MBF1C* (*multiprotein bridging factor 1C*), and *ELIP* (*early light-induced protein 1*), under normal conditions, thus enhancing the thermotolerance of transgenic Arabidopsis. Data from a subcellular localization assay revealed green fluorescent protein (GFP)-fused *FaTHSFA2a* and *FaTHSFB1a* proteins situated in the nucleus, suggesting their putative function as transcription factors. In addition, a yeast one-hybrid assay revealed that *FaTHSFA2a*-encoded proteins can be involved in transcriptional activation, whereas *FaTHSFB1a* functions as a transcriptional repressor. Altogether, our results provide vital information for the further investigation of strawberry as well as its future molecular breeding.

2. Results and Discussion

2.1. Illumina Sequencing, De Novo Assembly, and Functional Annotation of the Transcriptome

To obtain an overview of the strawberry transcriptome and gene activity at the nucleotide resolution level, we first extracted the total RNA from five strawberry tissues, namely the leaves, floral buds, fully bloomed flowers, immature fruits, and roots, for setting up the OF transcriptome data set. The second preparation involved extracting the total RNA from two distinct tissues containing vegetative and inflorescence meristems to set up the DU transcriptome data set. Equal amounts of total RNA from each sample were pooled, the mRNA was isolated, enriched, sheared into smaller fragments, and further reverse-transcribed into cDNA. The separated cDNA pools from the two preparations were subjected to Illumina HiSeq™ 2000 sequencing, and the resulting data were examined using bioinformatics analysis. The flowchart of our transcriptome analysis is shown in Figure 1.

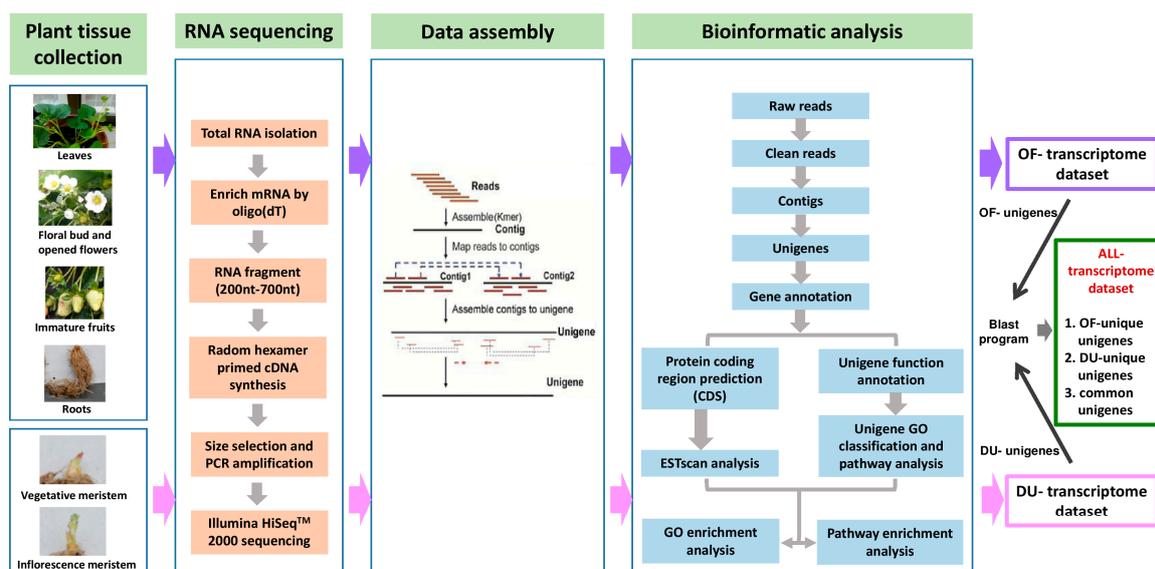


Figure 1. Flowchart of the transcriptome sequencing analysis (plant tissue collection, RNA sequencing, data assembly, and bioinformatics analysis). The total RNA was extracted from five strawberry tissues, namely leaves, floral buds, fully bloomed flowers, immature fruits, and roots. Following cDNA preparation of the extracted RNA and sequencing analysis, the retrieved data were used to set up the OF transcriptome data set. Simultaneously, the total RNA extracted from two other tissues comprising vegetative and inflorescence meristems was utilized to establish the DU transcriptome data set. Sequences of the unigenes obtained either from the OF or DU transcriptome data set were further analyzed, and common genes and unigenes unique to either OF or DU transcriptome data sets (OF- and DU-unique unigenes) were determined. To specify common unigenes, the sequence identity between the two unigenes identified separately from the data sets must be greater than 50%. In addition, its matched length must be higher than 50% of the length of both unigenes. The remaining unigenes appeared as unique to each data set. GO: gene ontology; CDS: consensus coding sequences.

After removing adaptor sequences, ambiguous reads, and low-quality reads, we obtained two transcriptome data sets. The data sets contain a total of 4,279,638,060 (4.28 Gb) and 4,412,888,820 (4.41 Gb) nucleotides with high-quality clean reads for OF and DU transcriptomes, respectively (Table 1). All high-quality reads were assembled de novo by using the Trinity program [49]. Consequently, 117,813 contigs, with an N50 of 740 bp representing 50% of the assembled bases were incorporated into contigs of 740 bp or longer in the OF transcriptome data set, and 122,261 contigs, with an N50 of 805 bp in the DU transcriptome data set. An overview of the sequencing and assembly data is outlined in Table 2.

Table 1. Output statistics of sequencing.

Samples	Total Clean Reads ¹	Total Clean Nucleotides (nt) ¹	Q20 Percentage ²	N Percentage ³	GC Percentage ⁴
OF	47,551,534	4,279,638,060	95.98%	0.00%	48.16%
DU	49,032,098	4,412,888,820	96.99%	0.00%	47.45%

¹ Total reads and total nucleotides are actually clean reads and clean nucleotides. The total nucleotides should be greater than contract provision; ² The Q20 percentage is the proportion of nucleotides with a quality value higher than 20; ³ The N percentage is the proportion of unknown nucleotides in clean read; ⁴ The GC percentage is the proportion of guanine and cytosine nucleotides among total nucleotides.

Table 2. Statistics of the assembly quality.

Classification	Sample	Total Number	Total Length (nt)	Mean Length (nt)	N50 ¹
Contig	OF	117,813	41,433,873	352	740
	DU	122,261	43,572,190	356	805
Unigene	OF	65,164	45,910,155	705	1218
	DU	72,914	50,183,039	688	1155
	All	65,768	57,422,693	873	1387

¹ N50: length of the smallest transcripts in the set that contain the fewest (largest).

Overview of the size distribution of contigs, unigenes, consensus coding sequences (CDS), and expressed sequence tags (ESTs) from OF, DU, and ALL transcriptome data sets was shown in Figure S1. Briefly, the mean contig size in the OF transcriptome data set was 352 bp (Table 2, Figure S1A). By contrast, the mean contig size in the DU transcriptome data set was 356 bp (Table 2, Figure S1B). In the OF transcriptome data set, a total of 65,164 unigenes were assembled, with an average unigene length of 705 bp and N50 of 1218 bp (Table 2, Figure S1C). Alternatively, a total of 72,914 unigenes were assembled in the DU transcriptome data set. Similarly, the average unigene length was 688 bp with an N50 of 1155 bp (Table 2, Figure S1D). An additional analysis of these unigenes from the two transcriptome data sets revealed common unigenes as well as unigenes unique to each data set. Following integration assays of these two data sets, the ALL transcriptome data set was generated and named. A total of 65,768 unigenes were identified in this data set after combining data from the OF and DU transcriptome data sets (Table 2). These unigenes were with an average length of 873 bp and N50 of 1387 bp (Table 2, Figure S1E). The distribution pattern of unigenes and contigs is similar, except that the unigene size is typically longer than the contig length. Furthermore, bioinformatics analysis was performed using the ALL transcriptome data set. In the Nr annotation, 67.8% of total 65,768 sequences matched perfectly with *E*-values of less than 10^{-45} (Figure S2A), and 61.1% of the matches showed more than 95% similarity (Figure S2B). The species distribution showed that 88.0% had top matches to *F. vesca* subsp. *vesca* genes (Figure S2C), followed by matches to *Prunus persica* (5.9%; Figure S2C). These results indicated that our transcriptome data sets can accurately predict the unigenes potentially useful for further analysis of strawberry species.

In addition, we observed that 72.7% (47,794/65,768, Table 2) consensus sequences from the ALL transcriptome data set showed homology with sequences in the Nr database, whereas 45.8% (30,114/65,768, Table 2) unigenes were similar to proteins in the SWISS-PROT database. A total of 54,813 unigenes [approximately 83.3% (54,813/65,768) of all assembled unigenes] were annotated by using combined Nr and SWISS-PROT databases, suggesting that these unigenes have relatively well-conserved functions in our strawberry transcriptome analysis.

GO, an international standardized gene-functional classification system, provides a standardized data by using a strictly defined concept to comprehensively describe the properties of genes and their products. GO was typically used to assign functions to uncharacterized sequences isolated from other organisms. According to GO classifications, a total of 65,768 strawberry unigenes with putative functions assigned to 33,064 unique sequences were categorized into three main GO categories and 55 sub-categories (functional groups; Figure S3A). The Clusters of Orthologous Groups of proteins (COG) database contains classifications of orthologous gene products that can predict and classify possible functions isolated from other species. In this study, the observed strawberry unigenes were searched against the COG database to predict and classify their possible functions. In total, out of 47,794 Nr hits, 17,441 sequences had COG classifications distributed into 25 COG categories (Figure S3B). Pathway-based analyses can further facilitate understanding the potential biological functions of genes. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database comprises information on a systematic analysis of inner-cell metabolic pathways, functions of gene products and their organism-specific variations [50]. To identify the biological pathways in strawberry, we evaluated 26,776 sequences assigned to 128 KEGG pathways (Table S1). These results highlight

the considerable potential of Illumina sequencing to discover metabolic pathway genes by using our strawberry transcriptome data set.

2.2. Identifying Nine FaTHSF Genes with Full-Length cDNA Sequences in *F. × ananassa* Duch. cv. Toyonoka

Plant HSFs play a central regulatory role in thermotolerance, which subsequently modulates genes encoding HSPs. Plant HSFs are regulated in association mainly with responses to abiotic stress and are associated with primed defense gene activation or pathogen-induced systemic acquired resistance (SAR) [43,51]. We evaluated the conceivable HSF family of *F. × ananassa* Duch. cv. Toyonoka; thus, we not only investigated the transcriptome databases generated in this study, but also the acquired annotation of the identified unigenes. As shown in Table S2, we discovered 31 contigs and 15 unigenes interpreted as HSF proteins, and they appeared to have high homology with *HSF* genes isolated from other plant species, including *Arabidopsis*, rice, and maize. Among these, we further identified nine *FaTHSF* genes containing full-length open reading frames (ORFs; Table S3). The online Multiple EM for Motif Elicitation (MEME) motif search tool was subsequently used and 23 corresponding conserved motifs in *FaTHSF*s were determined (Figure 2A). The number of motifs in the different *FaTHSF* proteins showed a high degree of variability. Considering the DBD, Motifs 1, 2, and 3 were basically situated at the N-terminal end of all nine *FaTHSF*s genes (Figure 2A,B). In addition, the presence of different oligomerization domains (the HR-A/B region) determines class A (Motif 4 or 5) or class B (Motif 8) of the *FaTHSF* family (Figure 2A,B). The class A *FaTHSF* motif contains other domains, such as CTAD or AHA (Motif 6), NLS (Motifs 7, 10, and 19), and NES (Motif 13; Figure 2A,B). Despite diverse length and sequences of amino acids in each *FaTHSF* protein, they were all predicted to encompass the DBD and oligomerization domain. Therefore, these and other specific domains result in functional diversity among different *FaTHSF* proteins.

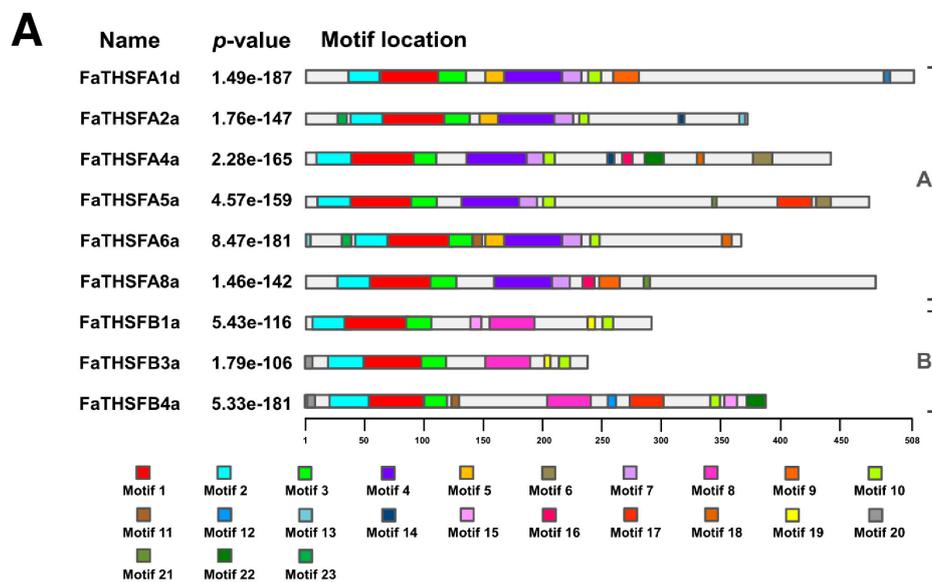


Figure 2. Cont.

B

Motif names	Best possible match protein sequences	Length	Function
1	·FVW ₁₂₃₄₅₆₇₈₉ ·F ₁₀ a ₁₁ e ₁₂ l ₁₃ p ₁₄ ·V ₁₅ F ₁₆ K ₁₇ ·W ₁₈ N ₁₉ F ₂₀ S ₂₁ S ₂₂ F ₂₃ Y ₂₄ R ₂₅ Q ₂₆ L ₂₇ N ₂₈ T ₂₉ G ₃₀ F ₃₁ R ₃₂ K ₃₃ Y ₃₄ S ₃₅ P ₃₆ Q ₃₇ W ₃₈ E ₃₉ F ₄₀ A ₄₁ N ₄₂ G ₄₃ F ₄₄	50	DNA-binding domain (DBD)
2	·P ₁ ·P ₂ F ₃ L ₄ a ₅ K ₆ T ₇ ·e ₈ b ₉ Y ₁₀ D ₁₁ C ₁₂ A ₁₃ T ₁₄ ·Q ₁₅ J ₁₆ Y ₁₇ S ₁₈ W ₁₉ ·e ₂₀ e ₂₁ e ₂₂ e ₂₃ e ₂₄ e ₂₅ e ₂₆ e ₂₇ e ₂₈ e ₂₉ e ₃₀ e ₃₁ e ₃₂ e ₃₃ e ₃₄ e ₃₅ e ₃₆ e ₃₇ e ₃₈ e ₃₉ e ₄₀ e ₄₁ e ₄₂ e ₄₃ e ₄₄ e ₄₅ e ₄₆ e ₄₇ e ₄₈ e ₄₉ e ₅₀	29	DNA-binding domain (DBD)
3	·e ₁ e ₂ e ₃ e ₄ e ₅ e ₆ e ₇ e ₈ e ₉ e ₁₀ e ₁₁ e ₁₂ e ₁₃ e ₁₄ e ₁₅ e ₁₆ e ₁₇ e ₁₈ e ₁₉ e ₂₀ e ₂₁ e ₂₂ e ₂₃ e ₂₄ e ₂₅ e ₂₆ e ₂₇ e ₂₈ e ₂₉ e ₃₀ e ₃₁ e ₃₂ e ₃₃ e ₃₄ e ₃₅ e ₃₆ e ₃₇ e ₃₈ e ₃₉ e ₄₀ e ₄₁ e ₄₂ e ₄₃ e ₄₄ e ₄₅ e ₄₆ e ₄₇ e ₄₈ e ₄₉ e ₅₀	21	DNA-binding domain (DBD)
4	·L ₁ L ₂ L ₃ L ₄ L ₅ L ₆ L ₇ L ₈ L ₉ L ₁₀ L ₁₁ L ₁₂ L ₁₃ L ₁₄ L ₁₅ L ₁₆ L ₁₇ L ₁₈ L ₁₉ L ₂₀ L ₂₁ L ₂₂ L ₂₃ L ₂₄ L ₂₅ L ₂₆ L ₂₇ L ₂₈ L ₂₉ L ₃₀ L ₃₁ L ₃₂ L ₃₃ L ₃₄ L ₃₅ L ₃₆ L ₃₇ L ₃₈ L ₃₉ L ₄₀ L ₄₁ L ₄₂ L ₄₃ L ₄₄ L ₄₅ L ₄₆ L ₄₇ L ₄₈ L ₄₉ L ₅₀	50	oligomerization domain
5	·E ₁ V ₂ C ₃ Y ₄ E ₅ F ₆ L ₇ E ₈ F ₉ F ₁₀ F ₁₁ F ₁₂ F ₁₃ F ₁₄ F ₁₅ F ₁₆ F ₁₇ F ₁₈ F ₁₉ F ₂₀ F ₂₁ F ₂₂ F ₂₃ F ₂₄ F ₂₅ F ₂₆ F ₂₇ F ₂₈ F ₂₉ F ₃₀ F ₃₁ F ₃₂ F ₃₃ F ₃₄ F ₃₅ F ₃₆ F ₃₇ F ₃₈ F ₃₉ F ₄₀ F ₄₁ F ₄₂ F ₄₃ F ₄₄ F ₄₅ F ₄₆ F ₄₇ F ₄₈ F ₄₉ F ₅₀	15	oligomerization domain
6	·D ₁ V ₂ F ₃ R ₄ E ₅ G ₆ L ₇ T ₈ E ₉ F ₁₀	15	AHA
7	·Q ₁ N ₂ R ₃ F ₄ L ₅ V ₆ L ₇ Y ₈ Q ₉ E ₁₀	15	nuclear localization signal (NLS)
8	·L ₁ L ₂ L ₃ L ₄ L ₅ L ₆ L ₇ L ₈ L ₉ L ₁₀ L ₁₁ L ₁₂ L ₁₃ L ₁₄ L ₁₅ L ₁₆ L ₁₇ L ₁₈ L ₁₉ L ₂₀ L ₂₁ L ₂₂ L ₂₃ L ₂₄ L ₂₅ L ₂₆ L ₂₇ L ₂₈ L ₂₉ L ₃₀ L ₃₁ L ₃₂ L ₃₃ L ₃₄ L ₃₅ L ₃₆ L ₃₇ L ₃₈ L ₃₉ L ₄₀ L ₄₁ L ₄₂ L ₄₃ L ₄₄ L ₄₅ L ₄₆ L ₄₇ L ₄₈ L ₄₉ L ₅₀	38	oligomerization domain
9	·D ₁ G ₂ R ₃ T ₄ V ₅ R ₆ P ₇ P ₈ M ₉ E ₁₀ A ₁₁ K ₁₂ K ₁₃	18	unknown
10	·e ₁ e ₂ e ₃ e ₄ e ₅ e ₆ e ₇ e ₈ e ₉ e ₁₀ e ₁₁ e ₁₂ e ₁₃ e ₁₄ e ₁₅ e ₁₆ e ₁₇ e ₁₈ e ₁₉ e ₂₀ e ₂₁ e ₂₂ e ₂₃ e ₂₄ e ₂₅ e ₂₆ e ₂₇ e ₂₈ e ₂₉ e ₃₀ e ₃₁ e ₃₂ e ₃₃ e ₃₄ e ₃₅ e ₃₆ e ₃₇ e ₃₈ e ₃₉ e ₄₀ e ₄₁ e ₄₂ e ₄₃ e ₄₄ e ₄₅ e ₄₆ e ₄₇ e ₄₈ e ₄₉ e ₅₀	8	nuclear localization signal (NLS)
11	·P ₁ L ₂ R ₃ Q ₄ E ₅	8	unknown
12	·C ₁ N ₂ K ₃ T ₄ N ₅	6	unknown
13	·M ₁ Y ₂ L ₃ T ₄ R ₅	6	nuclear export signal (NES)
14	·W ₁ E ₂ D ₃ E ₄ Y ₅	6	unknown
15	·N ₁ N ₂ G ₃ S ₄ N ₅ E ₆ T ₇ A ₈ N ₉ A ₁₀	12	unknown
16	·M ₁ L ₂ E ₃ E ₄ V ₅ D ₆ S ₇ P ₈ M ₉	11	unknown
17	·L ₁ L ₂ L ₃ L ₄ L ₅ L ₆ L ₇ L ₈ L ₉ L ₁₀ L ₁₁ L ₁₂ L ₁₃ L ₁₄ L ₁₅ L ₁₆ L ₁₇ L ₁₈ L ₁₉ L ₂₀ L ₂₁ L ₂₂ L ₂₃ L ₂₄ L ₂₅ L ₂₆ L ₂₇ L ₂₈ L ₂₉ L ₃₀ L ₃₁ L ₃₂ L ₃₃ L ₃₄ L ₃₅ L ₃₆ L ₃₇ L ₃₈ L ₃₉ L ₄₀ L ₄₁ L ₄₂ L ₄₃ L ₄₄ L ₄₅ L ₄₆ L ₄₇ L ₄₈ L ₄₉ L ₅₀	29	unknown
18	·I ₁ L ₂ A ₃ S ₄ R ₅ P ₆ M ₇	8	unknown
19	·K ₁ L ₂ F ₃ G ₄ V ₅ R ₆	6	nuclear localization signal (NLS)
20	·M ₁ L ₂ S ₃ K ₄ C ₅ S ₆	6	unknown
21	·G ₁ Y ₂ I ₃ X ₄ E ₅ Y ₆	6	unknown
22	·I ₁ Y ₂ S ₃ K ₄ I ₅ D ₆ L ₇ C ₈ M ₉ N ₁₀ S ₁₁ R ₁₂ P ₁₃ C ₁₄ S ₁₅ T ₁₆ Q ₁₇	18	unknown
23	·P ₁ a ₂ M ₃ E ₄ G ₅ L ₆ D ₇	9	unknown

Figure 2. Motifs identified using MEME tools in FaTHSFs. (A) The motif location and combined *p*-value of FaTHSFs are shown and denoted by rectangles with different colors. The black closed bracket on the right of the figure shows genes clustering into classes A and B; (B) Possible amino acid sequences of Motifs 1–23 identified using MEME tools for FaTHSFs.

We specifically selected *FaTHSFA2a* and *FaTHSFB1a* for further investigation not only because they represent class A2 and B1 HSFs, respectively, but also because previous studies have shown that HsfA2 responded strongly to long-term heat stress, accumulating to very high levels in tomato [44], Arabidopsis [45], or rice [46]. In addition, several lines of evidence have revealed that HsfA2 is associated with the expression of Hsps or stress-related non-chaperone encoding genes [44,45]. Thus, these data suggested that HsfA2 probably is one of the strongest heat-associated HSFs. Furthermore, HsfB1a and HsfB2a/b were demonstrated to be related to pathogen resistance in Arabidopsis [43,47,48]. Both *FaTHSFA2a* and *FaTHSFB1a* comprise the N-terminal highly conserved DBD and an adjacent oligomerization domain (the HR-A/B region), and their relationship with other HSF genes from *Arabidopsis*, rice (*Oryza sativa*), and tomato (*Solanum lycopersicum*) was examined. The *FaTHSFA2a* transcript possesses an ORF of 1119 nucleotides, encoding a protein of 372 amino acids with a predicted molecular weight (*Mw*) of 41.95 kDa and an isoelectric point (*pI*) of 4.93 (Table S3). Amino acid sequence analysis and multiple alignments showed that the deduced amino acid sequences of *FaTHSFA2a* encompass all critical domains and motifs as those in most HSF families, including the DBD, the HR-A/B region, nuclear localization signal (NLS), nuclear export signal (NES), and a CTAD, which contains AHA1 (ESLFAAAAALDN) and AHA2 (SEWGEDLQD). Thus, *FaTHSFA2a* is similar to atHSFA2 (Figure 3A). By contrast, the *FaTHSFB1a* transcript has an ORF of 873 nucleotides, encoding a protein composed of 290 amino acids with a predicted *Mw* of 32.10 kDa and *pI* of 6.4 (Table S3). Multiple alignment analysis revealed that the deduced amino acid sequences of *FaTHSFB1a* enclose the DBD, the HR-A/B region, NLS, and a B3 repression domain (BRD; Figure 3B). These domains or motifs were shown to be highly conserved in class B1 HSFs, such as atHSFB1. Therefore, our data suggested that *FaTHSFA2a* and *FaTHSFB1a* belong to class A2 and B1 HSFs, respectively.

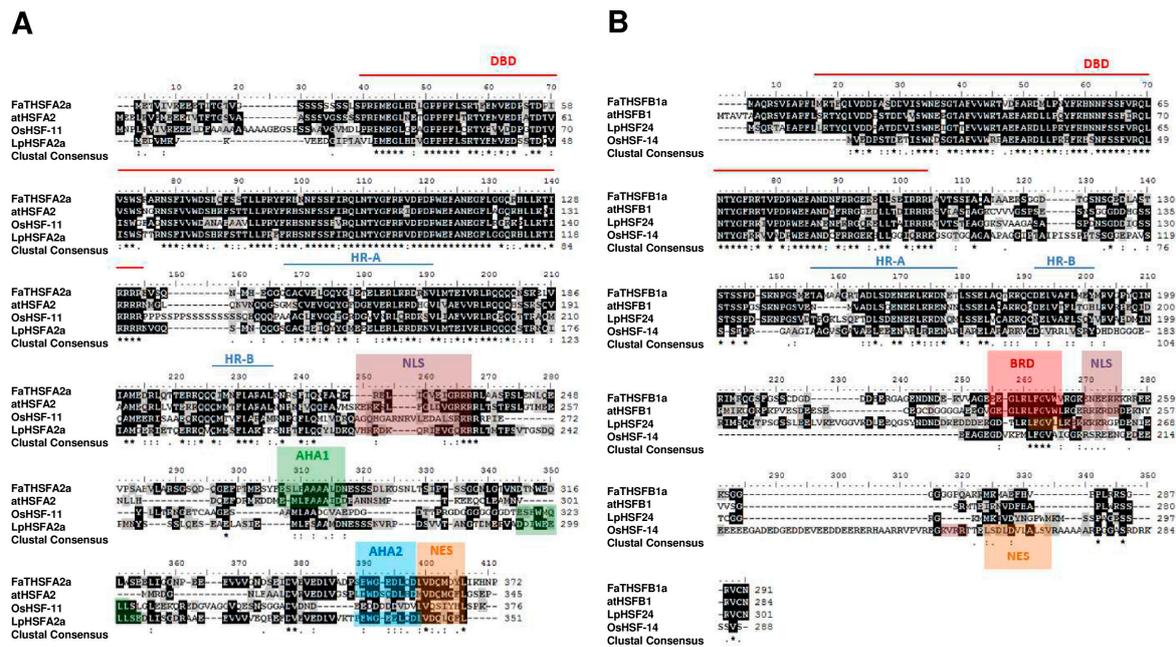


Figure 3. Sequence comparisons of FaTHSFA2a, FaTHSFB1a, and the related HSF proteins. (A) Amino acid sequence alignment of the full-length FaTHSFA2a and related HSF proteins, including atHSFA2 (At2g26150, *A. thaliana*), OsHSF-11 (AAQ23058, *O. sativa*), and LpHSFA2a (CAA47870, *L. peruvianum*); (B) Amino acid sequence alignment of the full-length FaTHSFB1a and its associated HSF proteins, such as atHSFB1 (At4g36990, *A. thaliana*), OsHSF-14 (AAQ23055, *O. sativa*), and LpHSF24 (P22335, *L. peruvianum*). This sequence alignment was performed by ClustalX 1.8 and BioEdit 7.0 software. Completely and partly conserved amino acids in proteins are shaded in black and gray, respectively. The letters and marks in the alignment are represented as follows: DBD (red line), HR-A and -B (blue line), NLS (highlighted in brown), AHA (highlighted in light green for AHA1 and light blue for AHA2), NES (highlighted in orange), and BRD (highlighted in red).

2.3. Classifying FaTHSFA2a and FaTHSFB1a into A2 and B1 HSFs, Respectively, by Phylogenetic Analysis

To determine the phylogenetic relationships among FaTHSFA2a, FaTHSFB1a, 22 Arabidopsis HSFs, 25 rice HSFs, and five tomato HSFs, a neighbor-joining phylogenetic tree was constructed using MEGA6 [52]. This method basically determines genetic distances according to the sequence identity and similarity among the full-length amino acid sequences of samples. Our phylogenetic analysis results revealed that FaTHSFA2a belongs to class A2 HSFs, with the nearest relative to LpHSFA2a (58.9% identity and 82.8% similarity) and atHSFA2 (55.2% identity and 79.0% similarity; Figure 4). However, FaTHSFB1a belongs to class B1 HSFs, and its nearest relatives are LpHSF24 (58.2% identity and 79.1% similarity) and atHSFB1 (58.3% identity and 80.0% similarity; Figure 4). According to the potential function of class A and B HSFs in heat shock responses and our data, we suggest that FaTHSFA2a acts as a transcriptional activator. By contrast, the putative function of FaTHSFB1a may be involved in the transcriptional repression. Seven class A or B HSFs, containing full-length genes were previously discovered: A1 HSF (*FaTHSFA1d*), A4 HSF (*FaTHSFA4a*), A5 HSF (*FaTHSFA5a*), A6 HSF (*FaTHSFA6a*), A8 HSF (*FaTHSFA8a*), B3 HSF (*FaTHSFB3a*), and B4 HSF (*FaTHSFB4a*; Figure 4). Among the wild-type diploid woodland strawberry *F. vesca*, which was composed of 17 classified *FvHSF* genes, namely 11 class A, five class B, and one class C genes, only 14 genes were isolated, including *FvHSFA4b* and *FvHSFA7a* (class A). Hu et al. (2015) could not isolate *FvHSFA8a* from class A HSFs [53]. In our study, other class A subgroup (A3, A7, and A9), class B subgroup (B2), and class C HSF genes could not be identified either because our collected tissues were not subjected to heat stress before transcriptome analysis, resulting in lower expression of those genes, or complete lack of those genes in *F. × ananassa*

Furthermore, phylogenetic analysis from comparison among nine *FaTHSFs*, 14 *FvHSFs*, *Arabidopsis*, rice, and tomato *HSF* genes revealed that *FaTHSFA1d* and *FvHSFA1d* belong to A1 *HSFs* and share 98.8% identity; *FaTHSFA2a* and *FvHSFA2a* are A2 *HSFs* and share 92.9% identity; *FaTHSFA4a* and *FvHSFA4a* are A4 *HSFs* and share 37.5% identity; *FaTHSFA6a* and *FvHSFA6a* are A6 *HSFs* and share 97.0% identity, suggesting they are homologous relatives. Notably, *FaTHSFB1a* and *FvHSFB1a* are B1 *HSFs* and share 96.6% identity, while *FaTHSFB3a* and *FvHSFB3a* belong to B3 *HSFs* and share 97.5% identity. Lastly, *FaTHSFB4a* and *FvHSFB4a* are B4 *HSFs* and share 99.3% identity. Altogether, these data indicate that they are indeed closely-related orthologous homologs. In addition, further overexpression of *FaTHSFA2a* and *FaTHSFB1a* in transgenic *Arabidopsis* enhanced plant resistance to heat stress, indicating that these genes play crucial roles in plants' thermotolerance by using "gain of function" strategies. Thus, this is the major difference between our studies and those reported by Hu et al. [53]. We mainly focused on functional assays, whereas they intended to discover whether all the *HSF*-related genes exhibited differential changes at the transcriptional level following heat treatment.

2.4. Enhanced Thermotolerance in *FaTHSFA2a*- and *FaTHSFB1a*-Overexpressed Transgenic *Arabidopsis* Plants

Semi-quantitative RT-PCR revealed that *FaTHSFA2a* and *FaTHSFB1a* expression was markedly upregulated, to approximately 2.4-fold for *FaTHSFA2a* and 5.7-fold for *FaTHSFB1a* compared with the control, when strawberry leaf tissues were subjected to heat shock at 37 °C for 10 min. By contrast, *FaTHSFA2a* and *FaTHSFB1a* expression was low when leaf tissues were subjected to a control temperature of 24 °C (Figure 5A). In addition, strawberry plants were grown in the field, and the temperature variation was recorded at different time periods, once every hour, particularly at 6 am (29.3 °C), 2 pm (32.2 °C), and 8 pm (30.5 °C). Strawberry leaf tissues were subsequently collected at these time points. The total RNA was extracted, and the samples were subjected to real-time quantitative RT (qRT)-PCR analysis. As shown in Figure 5B, *FaTHSFA2a* expression was increased by approximately 8.2-fold from time points 1 to 2 when the temperature was increased from 29.3 °C to 32.2 °C. However, the expression decreased to the original level from time points 2 to 3 when the temperature was decreased to 30.5 °C. Similarly, *FaTHSFB1a* was approximately 4.6-fold augmented from time points 1 to 2, thereafter declining at time point 3. Thus, these data indicated that *FaTHSFA2a* and *FaTHSFB1a* expression was induced by heat stress in a temperature-dependent manner. Notably, the induction decreased at a later stage of heat shock treatment.

To explore the *in vivo* function of *FaTHSFA2a* and *FaTHSFB1a*, transgenic plants were generated to overexpress full-length *FaTHSFA2a* and *FaTHSFB1a* genes under the control of cauliflower mosaic virus (CaMV) 35S promoter. Among T2 selection lines, six independent lines (#1–#6) for *35S::FaTHSFA2a* transgenic plants with a higher expression of *FaTHSFA2a* transcripts and six independent lines (#1–#6) for *35S::FaTHSFB1a* transgenic plants with an increased expression of *FaTHSFB1a* transcripts were selected for further observation (Figure S4). Semi-quantitative RT-PCR analysis was performed for transgenic plants grown under non-heat stress (22 °C) and heat stress (37 °C) conditions. Several HSPs, HSFs, and stress-associated genes, including *Hsp101*, *MBF1C*, and *ELIP* were upregulated and detected in *35S::FaTHSFA2a* and *35S::FaTHSFB1a* transgenic plants without heat stress treatment (22 °C), whereas they were barely detected in wild-type plants grown under the same conditions (Figure 6A). We also obtain similar results confirmed by real-time qPCR analysis (Figure 6B). As expected, these stress-associated genes exhibited a heat-inducible expression pattern in the wild-type plants grown at 37 °C (Figure 6A). Moreover, our data revealed that these stress-associated genes were upregulated in *35S::FaTHSFA2a* and *35S::FaTHSFB1a* transgenic plants during heat stress treatment, probably because of the increased expression of endogenous *HSFA2* and other members of the *HSF* family (*HSFB1* and *HSFB2a*) under heat stress (Figure 6A,B).

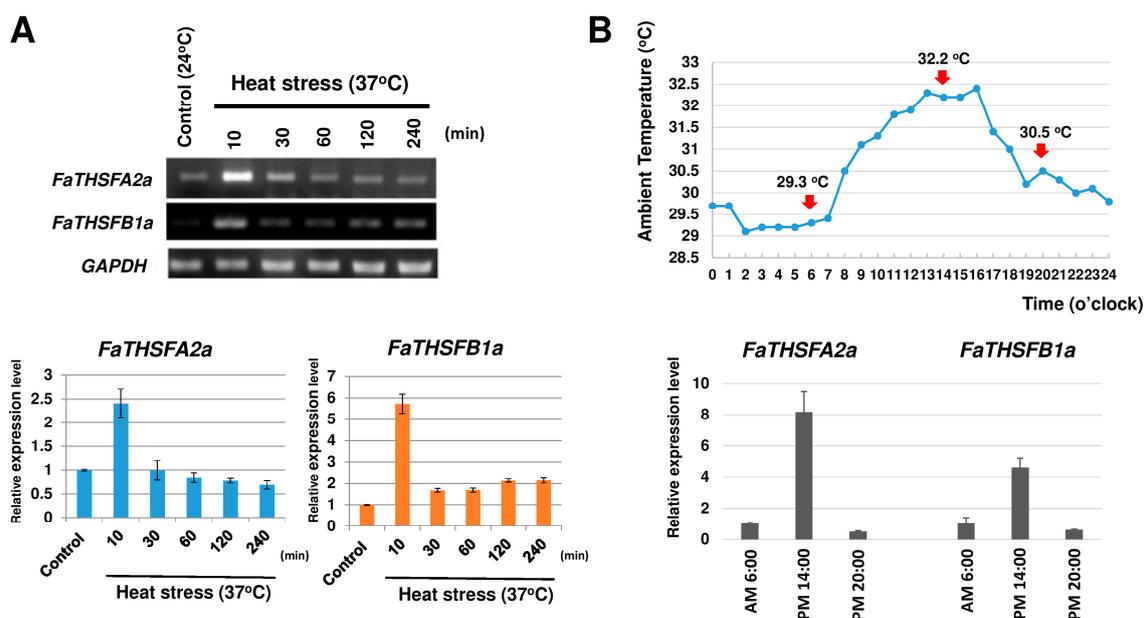


Figure 5. Induction of *FaTHSFA2a* and *FaTHSFB1a* expression under heat stress and the variation in these 2 gene expression were tightly monitored in response to various ambient temperatures within 24 h. (A) Induction of *FaTHSFA2a* and *FaTHSFB1a* expression under heat stress (37 °C). The garden strawberry was grown at 24 °C as a control or at 37 °C to subject it to heat stress for different time periods. The leaf tissues of the control and heat stress-treated plants were subsequently collected, and the total RNA was extracted thereafter at 10, 30, 60, 120, and 240 min. Semi-quantitative RT-PCR analysis was performed for each sample collected at different time points and normalized with GAPDH (=1). Error bars indicate standard deviation ($n = 3$); (B) Temperature variation in the environment where the garden strawberry was cultivated was recorded once every hour within 24 h on the same day the leaf samples were collected. The leaf tissues were collected and the total RNA was extracted at 6 am (29.3 °C), 2 pm (32.2 °C), and 8 pm (30.5 °C), respectively, and real-time qRT-PCR analysis was performed and normalized with GAPDH (=1). Error bars represent standard deviation ($n = 3$).

To examine whether the biological function of the C-terminal trans-activation domain in *FaTHSFA2a* and trans-repression domain in *FaTHSFB1a* were required for the thermotolerance of plants, we generated two transgenic plants overexpressing *FaTHSFA2a* Δ AD (containing 1-233aa of *FaTHSFA2a*, with a deleted trans-activation domain, named 35S::*FaTHSFA2a* Δ AD) and *FaTHSFB1a* Δ RD (containing 1-232aa of *FaTHSFB1a*, with a deleted trans-repression domain, named 35S::*FaTHSFB1a* Δ RD). RT-PCR confirmed a higher expression of these truncated transcripts among T2 selection lines compared with that in wild-type plants (Figure S4). The basal thermotolerance of 35S::*FaTHSFA2a*, 35S::*FaTHSFA2a* Δ AD, 35S::*FaTHSFB1a*, and 35S::*FaTHSFB1a* Δ RD plants were subsequently compared with that of wild-type plants. Under heat stress (45 °C), most 35S::*FaTHSFA2a* Δ AD, 35S::*FaTHSFB1a* Δ RD, and *hsp101* mutant-like wild-type plants died after 7 days, whereas 35S::*FaTHSFA2a* and 35S::*FaTHSFB1a* plants, such as 35S::*atHSF2A* survived under the same condition (Figure 6C). Thus, our results support the notion that the C-terminal functional domain of both *FaTHSFA2a*- and *FaTHSFB1a*-containing plants is essential for their increased thermotolerance in response to heat stress. Furthermore, our data revealed that the overexpression of *FaTHSFA2a* and *FaTHSFB1a* led to the constitutive expression of their downstream heat-stress-responsive genes in *Arabidopsis* transgenic plants, which may also contribute to the enhanced thermotolerance observed in the transgenic plants compared with wild-type plants.

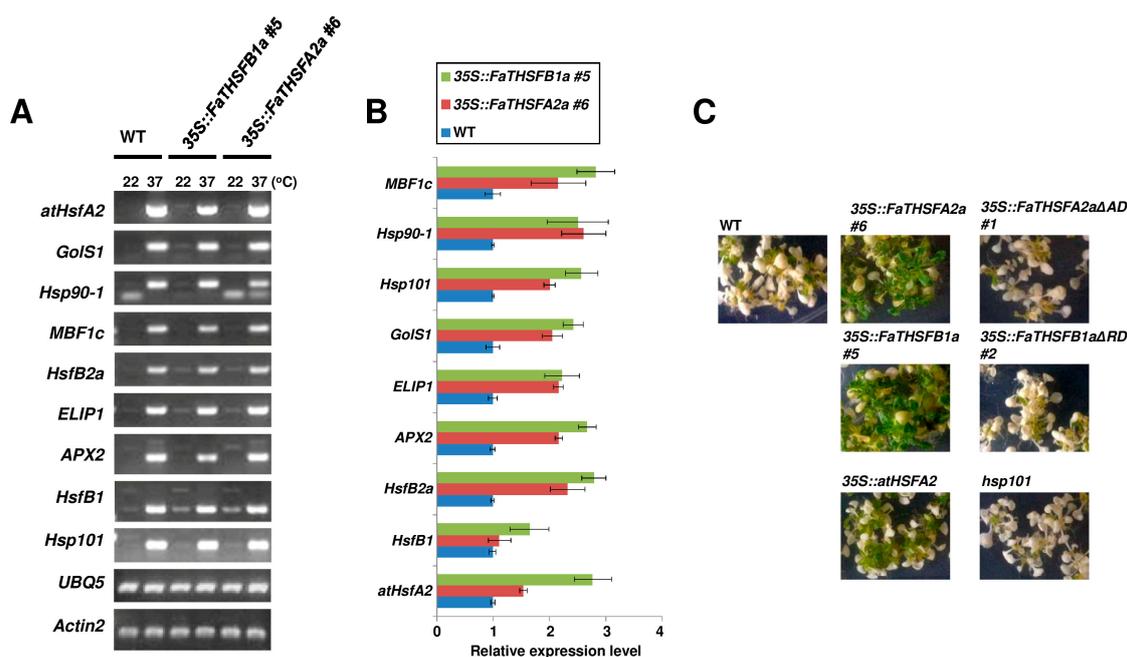


Figure 6. Roles of *FaTHSFA2a* and *FaTHSFB1a* in basal thermotolerance. **(A)** Up-regulated gene expression in *35S::FaTHSFA2a#6* and *35S::FaTHSFB1a#5* transgenic plants were observed after heat stress. Total RNAs were extracted from 10-day-old plants grown in MS medium at 22 °C and subsequently were either exposed to 22 °C as a control or moved to a 37 °C incubator for 1 h as heat stress treatment. The extracted RNAs were further used in semi-quantitative RT-PCR assays for wild-type (WT), *35S::FaTHSFA2a#6*, and *35S::FaTHSFB1a#5* plants by using gene-specific primer sets (Table S4). In this assay, *Ubiquitin 5 (UBQ5)* and *Actin2 (ACT2)* were used as the loading controls. Similar results were observed for *35S::FaTHSFA2a#4* and *#5* and *35S::FaTHSFB1a#3* and *#4* transgenic plants; **(B)** Analysis of mRNA levels in 10-days-old WT, *35S::FaTHSFA2a#6*, *35S::FaTHSFA2aΔAD#1*, *35S::FaTHSFB1a#5*, and *35S::FaTHSFB1aΔRD#2* at 22 °C by using gene-specific primer sets (Table S4). The levels of mRNA were determined by real-time qRT-PCR, and normalized with *Actin2/8* (=1). Error bars indicate the standard deviation ($n = 3$); **(C)** Thermotolerance evaluation of positive transgenic *Arabidopsis* lines with the overexpression of *FaTHSFA2a#6*, *FaTHSFA2aΔAD#1*, *FaTHSFB1a#5*, and *FaTHSFB1aΔRD#2*. Ten-day-old WT and transgenic plants exposed to heat stress at 45 °C for 1 h and subsequently recovered at 22 °C for 7 days were evaluated for their greening cultivars and photographed. More than 20 plants were analyzed in each experimental set ($n = 3$). Similar results were observed from triplicate experiments.

2.5. Nuclear Localization of *FaTHSFA2a* and *FaTHSFB1a* Suggesting Their Roles as Transcription Factors

According to our study, the unique functions of *FaTHSFA2a* and *FaTHSFB1a* proteins can act as HSFs. Therefore, these proteins were predicted to be located in the nucleus. To investigate the subcellular localization of *FaTHSFA2a* and *FaTHSFB1a* proteins, protoplasts prepared from flower lips of orchids were used for the transformation of *GFP-FaTHSFA2a* (*35S::GFP-FaTHSFA2a*) and *GFP-FaTHSFB1a* (*35S::GFP-FaTHSFB1a*) fusion genes, respectively, by particle bombardment. Our results detected the fluorescent signals of the *GFP-FaTHSFA2a* and *GFP-FaTHSFB1a* proteins mainly in the nucleus compared with the control GFP whose fluorescence was located in both nucleus and cytosol (Figure 7). These data revealed that *FaTHSFA2a* and *FaTHSFB1a* proteins were not confined in the nucleus; they can be transferred into the cytoplasm where they might conceivably perform other functions. Our results are in concordance with previous data reported by Hu et al. (2015) [53]; they clearly detected most FvHSF-GFP proteins in the nucleus. However, some FvHSF-GFP proteins, including FvHSF2a, FvHSF3a, FvHSF4a, FvHSF5a, FvHSF2b, and FvHSFC1a, remained detectable

in the cytosol [53]. Thus, our results revealed that FaTHSFA2a and FaTHSFB1a proteins have the capability to enter the nucleus and putatively function as transcriptional regulators.

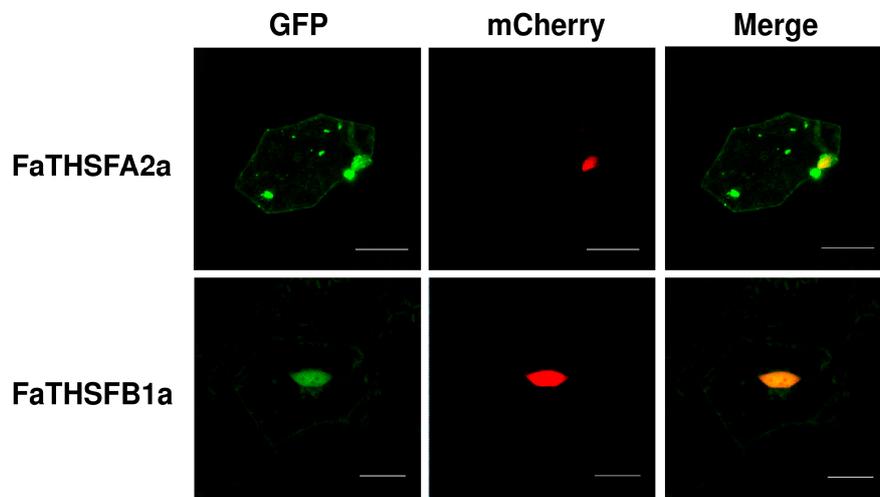


Figure 7. Subcellular localization of fluorescent GFP–FaTHSFA2a and GFP–FaTHSFB1a proteins. Recombinant plasmids harboring GFP–FaTHSFA2a and GFP–FaTHSFB1a, whose gene expression is driven by the CaMV35S promoter was transiently expressed in flower lips of orchids. These two recombinant plasmids (35S::GFP–FaTHSFA2a and 35S::GFP–FaTHSFB1a) were delivered into flower lips of orchids by using the particle bombardment method. The NLS domain of VirD2 fused with mCherry was used as the nuclear marker in this study. Overlays (merge) are shown in the extreme right column. The bar in each figure represents 20 μm .

2.6. *FaTHSFA2a* Exhibited Trans-Activation Function, Whereas *FaTHSFB1a* Showed Trans-Repression Activity

According to the deduced amino acid sequence and phylogenetic analysis, FaTHSFA2a belongs to the Class A2 HSF subgroup, whereas FaTHSFB1a is classified into the Class B1 HSF subgroup. To examine whether FaTHSFA2a possesses a CTAD similar to other Class A HSFs and FaTHSFB1a contains a BRD similar to other Class B HSFs in their gene structure, we performed a yeast one-hybrid assay. The full-length or C-terminal truncated fragments of FaTHSFA2a and FaTHSFB1a were fused to the GAL4 DBD and transformed into the yeast reporter strain AH109 harboring *HIS3* reporter genes. The activity of *HIS3* reporter genes was confirmed by a viability test with a selective medium lacking histidine. In the absence of histidine, the yeast cells transformed with the construct containing the full-length ORF of FaTHSFA2a survived. By contrast, yeast cells transformed with the C-terminal deletion of the *FaTHSFA2a* construct (no CTAD) could not survive under the same condition (Figure 8). By contrast, yeast cells transformed with either the full-length ORF or C-terminal deletion construct containing no BRD of *FaTHSFB1a*, could not survive in the selective medium without histidine (Figure 8). Subsequently, we co-transformed yeast cells containing the full-length ORF of *FaTHSFA2a* with either *FaTHSFB1a* or *FaTHSFB1a* Δ RD. In contrast to the survival of yeast cells harboring *FaTHSFA2a* in the medium without histidine, yeast cells harboring both full-length *FaTHSFA2a* and *FaTHSFB1a* could not survive. However, yeast cells harboring *FaTHSFA2a* and subsequently transformed with *FaTHSFB1a* Δ RD survived in the aforementioned medium. These results revealed that *FaTHSFA2a* functions as a transcriptional activator, exhibiting trans-activation activity similar to that of other class A HSFs, such as *HsfA1a*, *HsfA1b*, *HsfA1d*, *HsfA1e*, and *HsfA2* [45,54–58]. By contrast, *FaTHSFB1a* acts as a transcriptional repressor, which expresses trans-repression activity similar to that of other *HsfB1* identified in *Arabidopsis*, tomato, and soybean [25,26,48,59].

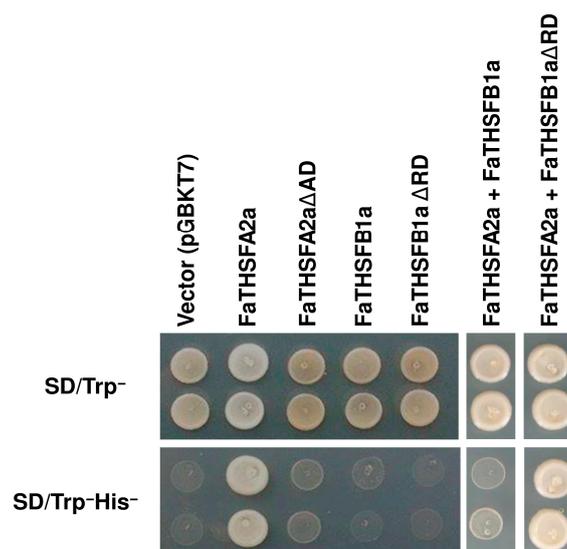


Figure 8. Analysis for measuring trans-activation activity in yeast cells. The full-length constructs of FaTHSFA2a and FaTHSFB1a were fused to GAL4 DBD and expressed in the yeast strain AH109. The transformed yeast cells were grown in non-selective media with histidine (SD/Trp⁻) or selective media without histidine (SD/Trp⁻His⁻), followed by incubation at 30 °C for 3 days. Two independent transformants were selected from each construct in the presence of histidine (SD/Trp⁻). Subsequently, yeast cells containing either single construct or cotransformed constructs were spotted on the media with (SD/Trp⁻) or without histidine (SD/Trp⁻His⁻). The pGBKT7 vector transformed into yeast cells was used as the negative control.

3. Materials and Methods

3.1. Plant Materials and RNA Extraction

In this study, we used the octoploid strawberry *F. × ananassa* Duch. cv. *Toyonoka*, which needs short-day and low temperature (cold) conditions to accelerate flower bud initiation [60,61]. Strawberry seedlings were grown in the field and widely planted in Industrial Technology Research Institute. All plants were watered daily and fertilized weekly. We selected plants with similar height, and their crown diameter was moved, and cultivated in the growth chamber. To reduce the microclimate effects in the growth chamber, all plants placed in each layer inside the chamber were rotated once weekly. The total RNA was extracted from strawberry tissues, namely leaves, floral buds, fully bloomed flowers, immature fruits, roots, and vegetative, and inflorescence meristems that were separately harvested and immediately frozen in liquid nitrogen, followed by storage at -80 °C until use. To perform Semi-quantitative RT-PCR for *FaTHSFA2a* and *FaTHSFB1a* expression, strawberry leaf tissues were subjected to heat shock at 37 °C for 10 min. All sample RNA were extracted with cetyl trimethylammonium bromide based buffer, as previously reported [62]. The quality and quantity of RNA extracts were evaluated using the Nanodrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and visualized through 1% agarose gel electrophoresis under denaturing conditions.

3.2. cDNA Preparation, Sequencing and De Novo Assembly

Equal amounts of total RNA from each strawberry tissue were mixed for the subsequent steps of the following experiments. Poly-A-containing mRNAs were purified from the mixed total RNA samples. The paired-end cDNA library was synthesized using the Genomic Sample Prep kit (Illumina, San Diego, CA, USA), according to the manufacturer's instructions. The sequencing cDNA library was constructed through PCR amplification and directly sequenced using the Illumina Genome Analyzer

according to the manufacturer's instructions (Genomics BioSci & Tech., Taipei, Taiwan). The raw sequencing data were subsequently filtered to remove low-quality sequences, including ambiguous nucleotides, adaptor sequences, and repeat sequences. The following de novo transcriptome assemblies of these short reads were performed using the SOAP de novo program [63]. To determine the consensus coding sequence (CDS) and sequence direction of the unigenes, these unigenes were analyzed using the BLASTX program (NCBI, Bethesda, MD, USA) with Nr, SWISS-PROT, KEGG, and COG databases [63]. The CDS of unigenes which were aligned to none of the aforementioned databases, were subsequently determined from the coding regions and sequence direction prediction by using ESTScan software [64]. Our transcriptome data sets are available at the NCBI Sequence Read Archive (SRA), under the accession numbers SRX1895539 and SRX1895540 for OF and DU transcriptome data sets, respectively.

3.3. Functional Annotation and Classification

Unigene annotations provide functional information, including protein sequence similarities, GO, Clusters of Orthologous Groups of proteins (COG) clusters, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway data. A sequence similarity search was conducted using the NCBI Nr, SWISS-PROT, COG, and KEGG pathway databases by using the BLASTX algorithm specifying *E*-values of less than 10^{-5} . Thus, unigene annotations may identify genes with their potential expression patterns and functional annotations. From the NCBI Nr protein database, we obtained GO annotations of these assembled unigenes by using the BLAST2GO program (<https://www.blast2go.com/>) [65]. Thereafter, we utilized WEGO software (Genomics BioSci & Tech., Taipei, Taiwan) to conduct GO functional classifications for all identified unigenes, and explored the macro-distribution of gene functions for this strawberry species [65]. Here, these strawberry unigenes with Nr annotation obtained using the BLAST2GO program [65], followed by WEGO software [66], were subjected to GO functional classifications, and the major distribution of the gene functions of this species was determined. Furthermore, we performed metabolic pathway analysis by using the KEGG database and related software applications (<http://www.genome.jp/kegg/kegg4.html>) [20]. Annotation with COGs and KEGG pathways was performed by searching with the BLASTX program against the COG [67] and KEGG [68] databases, with an *E*-value threshold of 10^{-5} .

3.4. Reverse Transcription Polymerase Chain Reaction (RT-PCR) and Real-Time RT-PCR Analysis

First strand cDNA was synthesized using 2 μ g of total RNA, an oligo dT primer, and an ImProm-IITM reverse transcription system (Promega, Madison, WI, USA) according to the manufacturer's instructions. Samples from each reaction (1 μ L) were subsequently used in a 20- μ L premix PCR mixture containing Platinum[®] Taq polymerase (Invitrogen, Carlsbad, CA, USA) and gene-specific primer sets (Table S4). For RT-PCR, amplification was continuously performed for 26 cycles, with each cycle at 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s. The PCR product was collected and analyzed through agarose gel electrophoresis, followed by ethidium bromide staining. By contrast, real-time PCR was conducted for 40 cycles by using 1 μ L of cDNA as the template and the CFX-96_Real Time system with SYBR Green Master Mix (Toyobo, Osaka, Japan). Subsequently, real-time RT-PCR data were analyzed using CFX Manager V2.1 software (Bio-Rad, Hercules, CA, USA). All sample data were normalized to tubulin or glyceraldehyde-3-phosphate dehydrogenase mRNA levels. The primer sets used for RT-PCR analysis are listed in Table S4.

3.5. Construction of 35S::FaTHSFA2a, 35S::FaTHSFA2a Δ AD, 35S::FaTHSFB1a and 35S::FaTHSFB1a Δ RD Recombinant Plasmids

Both full-length and C-terminal-truncated fragments of FaTHSFA2a and FaTHSFB1a coding regions (FaTHSFA2a, FaTHSFB1a and FaTHSFA2a Δ AD, and FaTHSFB1a Δ RD) were amplified from Toyonoka strawberry cDNAs through PCR with the gene-specific primer sets containing 5'-BamHI or 3'-SpeI recognition sites (Table S4). These primer sets were specifically designed to facilitate the cloning of these cDNAs by using Platinum[®] Taq DNA polymerase (Invitrogen). The PCR product

was subsequently purified using a purification column (BIOKIT, Miaoli, Taiwan) according to the manufacturer's instructions, cloned into the pGEM-T easy vector (Promega, Madison, WI, USA), and further sequenced to confirm the DNA sequences of the selected clones. T-vector with an appropriate insertion orientation was selected through the release of the respective full-length and C-terminal truncated cDNA fragment through BamHI and SpeI digestion and further subcloned from pGEM-T into the binary vector pCambia1390-35S vector (CAMBIA, Brisbane, Australia) between the BamHI and SpeI sites downstream of the CaMV 35S promoter. Gene-specific PCR combined with restriction enzyme digestion analysis was used to confirm the pCambia1390-35S vector harboring 35S::FaTHSFA2a, 35S::FaTHSFA2a Δ AD, 35S::FaTHSFB1a, or 35S::FaTHSFB1a Δ RD. *Agrobacterium tumefaciens* Strain GV3101 was used as the host for the resultant constructs.

3.6. *Arabidopsis thaliana* Transformation and Phenotypic Analysis

The aforementioned constructs harboring 35S::FaTHSFA2a, 35S::FaTHSFA2a Δ AD, 35S::FaTHSFB1a, or 35S::FaTHSFB1a Δ RD were transformed into wild-type *Arabidopsis* plants (ecotype: Col-0) by using the floral-dipping method [69]. The seeds were harvested and surface-sterilized with 25% sodium hypochlorite solution. These transgenic plants were selected on solid half MS medium containing 30 μ g/mL hygromycin. The plates were sealed with 3 M micropore tape and placed vertically in a growth chamber (CH-202, CHIN-HSIN, Taipei, Taiwan), with a long-day condition (16-h light/8-h dark cycle), and 120 $\text{Lmol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ photo flux density at 22 ± 2 °C for 12 days before transplanting to soil. The ectopic gene expression of FaTHSFA2a, FaTHSFA2a Δ AD, FaTHSFB1a, and FaTHSFB1a Δ RD was confirmed through RT-PCR by using gene-specific primer sets (Table S4). Each transgenic whole plant was photographed using a digital camera (Canon, PowerShot A640, Tokyo, Japan). Young seedlings of wild-type and all transgenic plants were viewed under a dissecting microscope (Leica, MS5, Heerbrugg, Switzerland), and all image data were collected using a digital camera (Canon, PowerShot S80, Tokyo, Japan) under a light-field.

3.7. Heat Stress Assays

For both basal thermotolerance and acquired thermotolerance assays, wild-type and transgenic plants were cultivated on a half MS medium in plastic Petri dishes and incubated in growth chambers (CH-202, CHIN-HSIN, Taipei, Taiwan). In basal thermotolerance assays, 5-day-old plants were subjected to 45 °C for 90 min, and the temperature was subsequently reduced to 22 °C [70]. Alternatively, 10-day-old plants were grown at 45 °C for 60 min and transferred to a 22 °C incubator. Following continuous growth at 22 °C for 5 days, these plants were photographed [70]. The number of survived plants was counted and recorded daily after the heat stress treatment.

3.8. Subcellular Localization

Both full-length FaTHSFA2a and FaTHSFB1a cDNAs were PCR-amplified with a pair of primers with attB1/B2 sites (Table S4) and cloned into the pDONR221 vector by using Gateway BP Clonase II Enzyme Mix (Invitrogen, Carlsbad, CA, USA). Each construct was subsequently recombined as an N-terminal fusion of GFP into the Gateway destination binary vector pK7WGF2 (Functional Genomics Division of the Department of Plant Systems Biology, Gent, Belgium), yielding 35S::GFP–FaTHSFA2a and 35S::GFP–FaTHSFB1a by an attL \times attR recombination reaction (Invitrogen, Carlsbad, CA, USA). These GFP fusion constructs were isolated and transformed into floral lips of orchids by bombardment transformation [71]. A Zeiss LSM 510 META laser-scanning confocal microscope using an LD C-Apochromat 409/1.1 W objective lens was subsequently used to observe the fluorescence emitted in the transformed cells, as previously described [71].

3.9. Transcriptional Activation or Repression Activity in Yeast

For measuring trans-activation or trans-repression activity, both full-length coding regions and C-terminal truncated fragments of FaTHSFA2a and FaTHSFB1a were amplified through PCR with

a set of primers containing attB1/B2 sites and cloned into the pDONR221 vector by using Gateway BP Clonase II Enzyme Mix (Invitrogen, Carlsbad, CA, USA). Subsequently, the cloned cDNAs were transferred into the pGBKT7 BD vector by using Gateway LR Clonase II Enzyme Mix (Invitrogen, Carlsbad, CA, USA) [72]. Each BD-construct was introduced into *Saccharomyces cerevisiae* strain AH109 and tested for its ability to activate or repress the transcription of HIS3 reporter genes by using a small-scale yeast transformation protocol (Clontech, Palo Alto, CA, USA). The transformed yeast cells were spotted on SD/Trp⁻ and SD/His⁻/Trp⁻ media. The plates were subsequently incubated at 30 °C for 3 days, and HIS3 activity was confirmed on histidine-lacking solid medium.

4. Conclusions

The assembled transcriptome database in the present study provides a foundation for the molecular genetics and functional genomics insights required to manipulate desirable agronomic traits or molecular breeding of garden strawberry. We used Illumina second-generation sequencing technology, and surveyed the transcriptome of *F. × ananassa* Duch. cv. Toyonoka. We combined two transcriptome data sets, and further assembled 65,768 unigenes and annotated 54,813 of these unigenes. This comprehensive coverage indicates the identification of almost all known genes from the major metabolic pathways responsible for strawberry development during different stages. Through bioinformatics analysis, we identified nine *FaTHSF* genes with full-length cDNAs (five contigs and four unigenes) from 31 contigs and 15 unigenes interpreted as HSF proteins. Classification of these genes was determined according to the sequence alignments, phylogenetic relationships, and expression analysis. *FaTHSFA2a* and *FaTHSFB1a* were specifically selected due to their representative roles as class A2 and B1 HSFs, respectively, and previous studies have revealed their crucial activities as strong heat-associated HSF proteins [46–51]. So, they were further isolated, characterized, and found up-regulated at the transcription level under heat stress. Notably, when functioning as transcription factors, *FaTHSFA2a* and *FaTHSFB1a* activated downstream heat stress-associated genes. In addition, enhanced thermotolerance was observed in *FaTHSFA2a*- and *FaTHSFB1a*-overexpressed transgenic *Arabidopsis* plants. Notably, yeast one-hybrid analysis revealed that *FaTHSFA2a* performs its trans-activation function by using the CTAD, and *FaTHSFB1a* exerts trans-repression activity through BRD. Thus, our data suggest that *FaTHSFA2a* and *FaTHSFB1a* function as crucial transcriptional activator and repressor, respectively, in the heat signaling pathways in garden strawberry. Consequently, it would be beneficial to employ the transformation of genes encoding *FaTHSFA2a* and *FaTHSFB1a* in plants to increase the thermotolerance of cultivated strawberry or other Rosaceae species in molecular-assisted breeding. This study also demonstrates that NGS technology is a rather rapid and cost-effective method for de novo transcriptome analysis of non-model plants, for which genomic information remains unavailable.

Supplementary Materials: Supplementary materials can be found at www.mdpi.com/1422-0067/17/12/2130/s1.

Acknowledgments: This study was supported in part by grants from the Tzu-Chi University Research Projects (TCMRC-P-102008, TCMRC-P-103016, and TCMRC-P-104013), Tzu-Chi University Industry-University Cooperative Research Projects (TCU101198A and TCU101198B), and The Industrial Technology Foresight Research Program Grant from the Industrial Technology Research Institute (C301ARB151). We are grateful to Choun-Sea Lin for biotechnical assistance with transient transformation and subcellular localization, and thank Yee-Yung Charng (Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan) for generously providing us with the seeds of *Arabidopsis* transgenic line (*35S::atHSFA2*) and mutant line (SALK_066374; hereafter labeled as *hsp101*). We also thank Ms. Amy Adams for editing and providing comments on our manuscript.

Author Contributions: Ming-Lun Chou, Jui-Hung Yang, and Wan-Yu Liao conceived and designed the experiments. Wan-Yu Liao, Jing-Lian Jheng, Chun-Chung Wang, and Jui-Hung Yang performed the experiments. Wan-Yu Liao, Lee-Fong Lin, Chun-Chung Wang, and Jui-Hung Yang analyzed the data. Lee-Fong Lin and Ming-Lun Chou drafted the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schlenker, W.; Roberts, M.J. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15594–15598. [[CrossRef](#)] [[PubMed](#)]
2. Lobell, D.B.; Schlenker, W.; Costa-Roberts, J. Climate trends and global crop production since 1980. *Science* **2011**, *333*, 616–620. [[CrossRef](#)] [[PubMed](#)]
3. Allakhverdiev, S.I.; Kreslavski, V.D.; Klimov, V.V.; Los, D.A.; Carpentier, R.; Mohanty, P. Heat stress: An overview of molecular responses in photosynthesis. *Photosynth. Res.* **2008**, *98*, 541–550. [[CrossRef](#)] [[PubMed](#)]
4. Lim, C.J.; Yang, K.A.; Hong, J.K.; Choi, J.S.; Yun, D.-J.; Hong, J.C.; Chung, W.S.; Lee, S.Y.; Cho, M.J.; Lim, C.O. Gene expression profiles during heat acclimation in *Arabidopsis thaliana* suspension-culture cells. *J. Plant Res.* **2006**, *119*, 373–383. [[CrossRef](#)] [[PubMed](#)]
5. Arkindale, J.; Vierling, E. Core genome responses involved in acclimation to high temperature. *Plant Physiol.* **2008**, *146*, 748–761. [[CrossRef](#)] [[PubMed](#)]
6. Panchuk, I.I.; Volkov, R.A.; Schöffl, F. Heat stress- and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in *Arabidopsis*. *Plant Physiol.* **2002**, *129*, 838–853. [[CrossRef](#)] [[PubMed](#)]
7. Panikulangara, T.J.; Eggers-Schumacher, G.; Wunderlich, M.; Stransky, H.; Schöffl, F. Galactinol synthase1: A novel heat shock factor target gene responsible for heat-induced synthesis of raffinose family oligosaccharides in *Arabidopsis*. *Plant Physiol.* **2004**, *136*, 3148–3158. [[CrossRef](#)] [[PubMed](#)]
8. Sakuma, Y.; Maruyama, K.; Qin, F.; Osakabe, Y.; Shinozaki, K.; Yamaguchi-Shinozaki, K. Dual function of an *Arabidopsis* transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 18822–18827. [[CrossRef](#)] [[PubMed](#)]
9. Nishizawa, A.; Yabuta, Y.; Shigeoka, S. Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. *Plant Physiol.* **2008**, *147*, 1251–1263. [[CrossRef](#)] [[PubMed](#)]
10. Bienz, M.; Pelham, H.R. Mechanisms of heat-shock gene activation in higher eukaryotes. *Adv. Genet.* **1987**, *24*, 31–72. [[PubMed](#)]
11. Baniwal, S.K.; Bharti, K.; Chan, K.Y.; Fauth, M.; Ganguli, A.; Kotak, S.; Mishra, S.K.; Nover, L.; Port, M.; Scharf, K.D. Heat stress response in plants: A complex game with chaperons and more. *J. Biosci.* **2004**, *29*, 471–487. [[CrossRef](#)] [[PubMed](#)]
12. Clos, J.; Westwood, J.T.; Becker, P.B.; Wilson, S.; Lambert, K.; Wu, C. Molecular cloning and expression of a hexameric Drosophila heat shock factor subject to negative regulation. *Cell* **1990**, *63*, 1085–1097. [[CrossRef](#)]
13. Treuter, E.; Nover, L.; Ohme, K.; Scharf, K.D. Promoter specificity and deletion analysis of three heat stress transcription factors of tomato. *Mol. Gen. Genet.* **1993**, *240*, 113–125. [[CrossRef](#)] [[PubMed](#)]
14. Nover, L.; Scharf, K.D.; Gagliardi, D.; Vergne, P.; Czarnecka-Verner, E.; Gurley, W.B. The Hsf world: Classification and properties of plant heat stress transcription factors. *Cell Stress Chaperones* **1996**, *1*, 215–223. [[CrossRef](#)]
15. Bailey, T.L.; Gribskov, M. Combining evidence using *p*-values: Application to sequence homology searches. *Bioinformatics* **1998**, *14*, 48–54. [[CrossRef](#)] [[PubMed](#)]
16. Döring, P.; Treuter, E.; Kistner, C.; Lyck, R.; Chen, A.; Nover, L. The role of AHA motifs in the activator function of tomato heat stress transcription factors HsfA1 and HsfA2. *Plant Cell* **2000**, *12*, 265–278. [[CrossRef](#)] [[PubMed](#)]
17. Scharf, K. D.; Berberich, T.; Ebersberger, I.; Nover, L. The plant heat stress transcription factor (Hsf) family: Structure, function and evolution. *Biochim. Biophys. Acta* **2012**, *1819*, 104–119. [[CrossRef](#)] [[PubMed](#)]
18. Chen, X.J.; Ye, C.J.; Lu, H.Y. Cloning of *GmHsFA1* gene and its overexpression leading to enhancement of heat tolerance in transgenic soybean. *Acta Genet. Sin.* **2006**, *28*, 1411–1420.
19. Morimoto, R.I. Regulation of the heat shock transcriptional response: Cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. *Genes Dev.* **1998**, *12*, 3788–3796. [[CrossRef](#)] [[PubMed](#)]
20. Nakai, A. New aspects in the vertebrate heat shock factor system: Hsf3 and Hsf4. *Cell Stress Chaperones* **1999**, *4*, 86–93. [[CrossRef](#)]

21. Kotak, S.; Port, M.; Ganguli, A.; Bicker, F.; von Koskull-Doring, P. 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.* **2004**, *39*, 98–112. [[CrossRef](#)] [[PubMed](#)]
22. Xing, H.; Wilkerson, D.C.; Mayhew, C.N.; Lubert, E.J.; Skaggs, H.S.; Goodson, M.L.; Hong, Y.; Park-Sarge, O.-K.; Sarge, K.D. Mechanism of *hsp70i* gene bookmarking. *Science* **2005**, *307*, 421–423. [[CrossRef](#)] [[PubMed](#)]
23. Nover, L.; Bharti, K.; Doring, P.; Mishra, S.K.; Ganguli, A.; Scharf, K.D. Arabidopsis and the heat stress transcription factor world: How many heat stress transcription factors do we need? *Cell Stress Chaperones* **2001**, *6*, 177–189. [[CrossRef](#)]
24. Schultheiss, J.; Kunert, O.; Gase, U.; Scharf, K.D.; Nover, L.; Ruterjans, H. Solution structure of the DNA-binding domain of the tomato heat-stress transcription factor HSF24. *Eur. J. Biochem.* **1996**, *236*, 911–921. [[CrossRef](#)] [[PubMed](#)]
25. Czarnecka-Verner, E.; Yuan, C.X.; Scharf, K.D.; Englich, G.; Gurley, W.B. Plants contain a novel multi-member class of heat shock factors without transcriptional activator potential. *Plant Mol. Biol.* **2000**, *43*, 459–471. [[CrossRef](#)] [[PubMed](#)]
26. Ikeda, M.; Ohme-Takagi, M. A novel group of transcriptional repressors in *Arabidopsis*. *Plant Cell Physiol.* **2009**, *50*, 970–975. [[CrossRef](#)] [[PubMed](#)]
27. Fragkostefanakis, S.; Röth, S.; Schleiff, E.; Scharf, K.D. Prospects of engineering thermotolerance in crops through modulation of heat stress transcription factor and heat shock protein networks. *Plant Cell Environ.* **2015**, *38*, 1881–1895. [[CrossRef](#)] [[PubMed](#)]
28. Schramm, F.; Larkindale, J.; Kiehlmann, E.; Ganguli, A.; Englich, G.; Vierling, E.; von Koskull-Döring, P. A cascade of transcription factor DREB2A and heat stress transcription factor HsfA3 regulates the heat stress response of Arabidopsis. *Plant J.* **2008**, *53*, 264–274. [[CrossRef](#)] [[PubMed](#)]
29. Yoshida, T.; Fujita, Y.; Sayama, H.; Kidokoro, S.; Maruyama, K.; Mizoi, J.; Shinozaki, K.; Yamaguchi-Shinozaki, K. AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. *Plant J.* **2010**, *61*, 672–685. [[CrossRef](#)] [[PubMed](#)]
30. Hwang, S.M.; Kim, D.W.; Woo, M.S.; Jeong, H.S.; Son, Y.S.; Akhter, S.; Choi, G.J.; Bahk, J.D. Functional characterization of *Arabidopsis* HsfA6a as a heat-shock transcription factor under high salinity and dehydration conditions. *Plant Cell Environ.* **2014**, *37*, 1202–1222. [[CrossRef](#)] [[PubMed](#)]
31. Guo, J.K.; Wu, J.; Ji, Q.; Wang, C.; Luo, L.; Yuan, Y.; Wang, Y.; Wang, J. Genome-wide analysis of heat shock transcription factor families in rice and *Arabidopsis*. *J. Genet. Genom.* **2008**, *35*, 105–118. [[CrossRef](#)]
32. Guo, M.; Lu, J.P.; Zhai, Y.F.; Chai, W.G.; Gong, Z.H.; Lu, M.H. Genome-wide analysis, expression profile of heat shock factor gene family (CaHsfs) and characterisation of CaHsfA2 in pepper (*Capsicum annuum* L.). *BMC Plant Biol.* **2015**, *15*, 151–170. [[CrossRef](#)] [[PubMed](#)]
33. Qiao, X.; Li, M.; Li, L.; Yin, H.; Wu, J.; Zhang, S. Genome-wide identification and comparative analysis of the heat shock transcription factor family in Chinese white pear (*Pyrus bretschneideri*) and five other Rosaceae species. *BMC Plant Biol.* **2015**, *15*, 12–27. [[CrossRef](#)] [[PubMed](#)]
34. Xue, G.P.; Sadat, S.; Drenth, J.; McIntyre, C.L. 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. *J. Exp. Bot.* **2014**, *65*, 539–557. [[CrossRef](#)] [[PubMed](#)]
35. Shulaev, V.; Korban, S.S.; Sosinski, B.; Abbott, A.G.; Aldwinckle, H.S.; Folta, K.M.; Iezzoni, A.; Main, D.; Aru's, P.; Dandekar, A.M.; et al. Multiple models for Rosaceae genomics. *Plant Physiol.* **2008**, *147*, 985–1003. [[CrossRef](#)] [[PubMed](#)]
36. Shulaev, V.; Sargent, D.J.; Crowhurst, R.N.; Mockler, T.C.; Folkerts, O.; Delcher, A.L.; Jaiswal, P.; Mockaitis, K.; Liston, A.; Mane, S.P.; et al. The genome of woodland strawberry (*Fragaria vesca*). *Nat. Genet.* **2011**, *43*, 109–116. [[CrossRef](#)] [[PubMed](#)]
37. Folta, K.M.; Clancy, M.A.; Chamala, S.; Dhingra, A.; Brunings, A.M.; Gomide, L.; Kulathinal, R.J.; Peres, N.A.; Davis, T.M.; Barbazuk, W.B. A Transcript accounting from diverse tissues of a cultivated strawberry. *Plant Genome* **2010**, *3*, 90–105. [[CrossRef](#)]

38. Rousseau-Gueutin, M.; Lerceteau-Köhler, E.; Barrot, L.; Sargent, D.J.; Monfort, A.; Simpson, D.; Arús, P.; Guérin, G.; Denoyes-Rothan, B. Comparative genetic mapping between octoploid and diploid *Fragaria* species reveals a high level of colinearity between their genomes and the essentially disomic behavior of the cultivated octoploid strawberry. *Genetics* **2008**, *179*, 2045–2060. [[CrossRef](#)] [[PubMed](#)]
39. Folta, K.M.; Davis, T.M. Strawberry genes and genomics. *Crit. Rev. Plant Sci.* **2006**, *25*, 399–415. [[CrossRef](#)]
40. Mezzetti, B.; Costantini, E. Strawberry (*Fragaria x ananassa*). *Methods Mol. Biol.* **2006**, *344*, 287–295. [[PubMed](#)]
41. Oosumi, T.; Gruszewski, H.A.; Blischak, L.A.; Baxter, A.J.; Wadl, P.A.; Shuman, J.L.; Veilleux, R.E.; Shulaev, V. High-efficiency transformation of the diploid strawberry (*Fragaria vesca*) for functional genomics. *Planta* **2006**, *223*, 1219–1230. [[CrossRef](#)] [[PubMed](#)]
42. Liu, H.; Xie, W.F.; Zhang, L.; Valpuesta, V.; Ye, Z.W.; Gao, Q.H.; Duan, K. Auxin biosynthesis by the YUCCA6 flavin monooxygenase gene in woodland strawberry. *J. Integr. Plant Biol.* **2014**, *56*, 350–363. [[CrossRef](#)] [[PubMed](#)]
43. Pick, T.; Jaskiewicz, M.; Peterhänsel, C.; Conrath, U. Heat shock factor HsfB1 primes gene transcription and systemic acquired resistance in *Arabidopsis*. *Plant Physiol.* **2012**, *159*, 52–55. [[CrossRef](#)] [[PubMed](#)]
44. Chan-Schamnet, K.Y.; Baniwal, S.K.; Bublak, D.; Nover, L.; Scharf, K.D. Specific interaction between tomato HsfA1 and HsfA2 creates hetero-oligomeric superactivator complexes for synergistic activation of heat stress gene expression. *J. Biol. Chem.* **2009**, *284*, 20848–20857. [[CrossRef](#)] [[PubMed](#)]
45. Schramm, F.; Ganguli, A.; Kiehlmann, E.; Englich, G.; Walch, D.; von Koskull-Döring, P. The heat stress transcription factor HsfA2 serves as a regulatory amplifier of a subset of genes in the heat stress response in *Arabidopsis*. *Plant Mol. Biol.* **2006**, *60*, 759–772. [[CrossRef](#)] [[PubMed](#)]
46. Mittal, D.; Chakrabarti, S.; Sarkar, A.; Singh, A.; Grover, A. Heat shock factor gene family in rice: Genomic organization and transcript expression profiling in response to high temperature, low temperature and oxidative stresses. *Plant Physiol. Biochem.* **2009**, *47*, 785–795. [[CrossRef](#)] [[PubMed](#)]
47. Kumar, M.; Busch, W.; Birke, H.; Kemmerling, B.; Nürnberger, T.; Schöffl, F. Heat shock factors HsfB1 and HsfB2b are involved in the regulation of Pdf1.2 expression and pathogen resistance in *Arabidopsis*. *Mol. Plant* **2009**, *2*, 152–165. [[CrossRef](#)] [[PubMed](#)]
48. Ikeda, M.; Mitsuda, N.; Ohme-Takagi, M. *Arabidopsis* HsfB1 and HsfB2b act as repressors of the expression of heat-inducible Hsfs but positively regulate the acquired thermotolerance. *Plant Physiol.* **2011**, *157*, 1243–1254. [[CrossRef](#)] [[PubMed](#)]
49. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)] [[PubMed](#)]
50. Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **2008**, *36*, D480–D484. [[CrossRef](#)] [[PubMed](#)]
51. Von Koskull-Döring, P.; Scharf, K.D.; Nover, L. The diversity of plant heat stress transcription factors. *Trends Plant Sci.* **2007**, *12*, 452–457. [[CrossRef](#)] [[PubMed](#)]
52. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [[CrossRef](#)] [[PubMed](#)]
53. Hu, Y.; Han, Y.-T.; Wei, W.; Li, Y.-J.; Zhang, K.; Gao, Y.-R.; Zhao, F.-L.; Feng, J.-Y. Identification, isolation, and expression analysis of heat shock transcription factors in the diploid woodland strawberry *Fragaria vesca*. *Front. Plant Sci.* **2015**, *6*, 736–751. [[CrossRef](#)] [[PubMed](#)]
54. Busch, W.; Wunderlich, M.; Schöffl, F. Identification of novel heat shock factor-dependent genes and biochemical pathways in *Arabidopsis thaliana*. *Plant J.* **2005**, *41*, 1–14. [[CrossRef](#)] [[PubMed](#)]
55. Nishizawa, A.; Yabuta, Y.; Yoshida, E.; Maruta, T.; Yoshimura, K.; Shigeoka, S. *Arabidopsis* heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *Plant J.* **2006**, *48*, 535–547. [[CrossRef](#)] [[PubMed](#)]
56. Charng, Y.Y.; Liu, H.C.; Liu, N.Y.; Chi, W.T.; Wang, C.N.; Chang, S.H.; Wang, T.T. A heat-inducible transcription factor, HsfA2, is required for extension of acquired thermotolerance in *Arabidopsis*. *Plant Physiol.* **2007**, *143*, 251–262. [[CrossRef](#)] [[PubMed](#)]
57. Liu, H.C.; Liao, H.T.; Charng, Y.Y. The role of class A1 heat shock factors (HSFA1s) in response to heat and other stresses in *Arabidopsis*. *Plant Cell Environ.* **2011**, *34*, 738–751. [[CrossRef](#)] [[PubMed](#)]

58. Nishizawa-Yokoi, A.; Nosaka, R.; Hayashi, H.; Tainaka, H.; Maruta, T.; Tamoi, M.; Ikeda, M.; Ohme-Takagi, M.; Yoshimura, K.; Yabuta, Y.; et al. HsfA1d and HsfA1e involved in the transcriptional regulation of HsfA2 function as key regulators for the Hsf signaling network in response to environmental stress. *Plant Cell Physiol.* **2011**, *52*, 933–945. [[CrossRef](#)] [[PubMed](#)]
59. Czarnecka-Verner, E.; Pan, S.; Salem, T.; Gurley, W.B. Plant class B HSFs inhibit transcription and exhibit affinity for TFIIB and TBP. *Plant Mol. Biol.* **2004**, *56*, 57–75. [[CrossRef](#)] [[PubMed](#)]
60. Morishita, M.; Mochizuki, T.; Yamakawa, O. Flower induction and selection on earliness of strawberry seedlings by short-day and low night temperature treatment. *J. Jpn. Soc. Hortic. Sci.* **1993**, *61*, 857–864. [[CrossRef](#)]
61. Ledesma, N.A.; Nakata, M.; Sugiyama, N. Effect of high temperature stress on the reproductive growth of strawberry cvs. “Nyoho” and “Toyonoka”. *Sci. Hortic-Amsterdam* **2008**, *116*, 186–193. [[CrossRef](#)]
62. Chang, S.; Puryear, J.; Cairney, J. A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* **1993**, *11*, 113–116. [[CrossRef](#)]
63. Li, R.; Zhu, H.; Ruan, J.; Qian, W.; Fang, X.; Shi, Z.; Li, Y.; Li, S.; Shan, G.; Kristiansen, K.; et al. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **2010**, *20*, 265–272. [[CrossRef](#)] [[PubMed](#)]
64. Iseli, C.; Jongeneel, C.V.; Bucher, P. ESTScan: A program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1999**, 138–148.
65. Götz, S.; García-Gómez, J.M.; Terol, J.; Williams, T.D.; Nagaraj, S.H.; Nueda, M.J.; Robles, M.; Talón, M.; Dopazo, J.; Conesa, A. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* **2008**, *36*, 3420–3435. [[CrossRef](#)] [[PubMed](#)]
66. Ye, J.; Fang, L.; Zheng, H.; Zhang, Y.; Chen, J.; Zhang, Z.; Wang, J.; Li, S.; Li, R.; Bolund, L.; et al. WEGO: A web tool for plotting GO annotations. *Nucleic Acids Res.* **2006**, *34*, W293–W297. [[CrossRef](#)] [[PubMed](#)]
67. Tatusov, R.L.; Natale, D.A.; Garkavtsev, I.V.; Tatusova, T.A.; Shankavaram, U.T.; Rao, B.S.; Kiryutin, B.; Galperin, M.Y.; Fedorova, N.D.; Koonin, E.V. The COG database: New developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* **2001**, *29*, 22–28. [[CrossRef](#)] [[PubMed](#)]
68. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [[CrossRef](#)] [[PubMed](#)]
69. Clough, S.J.; Bent, A.F. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **1998**, *16*, 735–743. [[CrossRef](#)] [[PubMed](#)]
70. Ogawa, D.; Yamaguchi, K.; Nishiuchi, T. 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.* **2007**, *58*, 3373–3383. [[CrossRef](#)] [[PubMed](#)]
71. Hsu, C.T.; Liao, D.C.; Wu, F.H.; Liu, N.T.; Shen, S.C.; Chou, S.J.; Tung, S.Y.; Yang, C.H.; Chan, M.T.; Lin, C.S. Integration of molecular biology tools for identifying promoters and genes abundantly expressed in flowers of *Oncidium Gower Ramsey*. *BMC Plant Biol.* **2011**, *11*, 60–74. [[CrossRef](#)] [[PubMed](#)]
72. Karimi, M.; Inze, D.; Depicker, A. GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* **2002**, *7*, 193–195. [[CrossRef](#)]

