



OPEN

## Genome-wide identification and comprehensive analysis heat shock transcription factor (Hsf) members in asparagus (*Asparagus officinalis*) at the seeding stage under abiotic stresses

Caihua Li<sup>1,3</sup>, Yuhuan Li<sup>1,3</sup>, Zeng Zhou<sup>2</sup>, Yudi Huang<sup>2</sup>, Zunzun Tu<sup>2</sup>, Xin Zhuo<sup>2</sup>, Dingyuan Tian<sup>2</sup>, Yibo Liu<sup>2</sup>, Hongli Di<sup>2</sup>, Ze Lin<sup>2</sup>, Mingxin Shi<sup>2</sup>, Xue He<sup>2</sup>, Haiyu Xu<sup>1</sup>, Yi Zheng<sup>1</sup> & Zhongsheng Mu<sup>1</sup>✉

Heat shock transcription factors (Hsf) are pivotal as essential transcription factors. They function as direct transcriptional activators of genes regulated by thermal stress and are closely associated with various abiotic stresses. Asparagus (*Asparagus officinalis*) is a vegetable of considerable economic and nutritional significance, abundant in essential vitamins, minerals, and dietary fiber. Nevertheless, asparagus is sensitive to environmental stresses, and specific abiotic stresses harm its yield and quality. In this context, Hsf members have been discerned through the reference genome, and a comprehensive analysis encompassing physical and chemical attributes, evolutionary aspects, motifs, gene structure, *cis*-acting elements, collinearity, and expression patterns under abiotic stresses has been conducted. The findings identified 18 members, categorized into five distinct subgroups. Members within each subgroup exhibited analogous motifs, gene structures, and *cis*-acting elements. Collinearity analysis unveiled a noteworthy pattern, revealing that Hsf members within asparagus shared one, two, and three pairs with counterparts in *Arabidopsis*, *Oryza sativa*, and *Glycine max*, respectively. Furthermore, members displayed tissue-specific expression during the seedling stage, with roots emerging as viable target tissue. Notably, the expression levels of certain members underwent modification under the influence of abiotic stresses. This study establishes a foundational framework for understanding Hsf members and offers valuable insights into the potential application of molecular breeding in the context of asparagus cultivation.

Heat shock transcription factors (Hsf) constitute a family of pivotal transcription factors primarily responsible for regulating the heat shock response and preserving protein homeostasis within cells<sup>1</sup>. Hsf members exhibit a remarkably conserved structure, with the DNA binding domain (DBD) situated at the N-terminus displaying significant conservation. This DBD employs four reverse equilibria to precisely locate and recognize the  $\beta$ -Fold (comprising  $\beta 1$ ,  $\beta 2$ ,  $\beta 3$ , and  $\beta 4$ ) as well as three  $\alpha$ -structures ( $\alpha 1$ ,  $\alpha 2$ , and  $\alpha 3$ )<sup>2</sup>. In addition to the conserved DNA binding domain located at the N-terminus, the oligomerization domain (OD), nuclear localization signals (NLS), nuclear export signals (NES), and the aromatic large hydrophobic and acid amino residues domain (AHA) at the C-terminus constitute the fundamental structural components of Hsf members<sup>3</sup>. The Hsf members can be categorized into subgroups based on variations in the length of amino acid residues between the DNA binding domain (DBD) and the oligomerization domain (OD)<sup>3</sup>. Hsfs can be activated to exert their functions in response to various stress conditions<sup>4</sup>. Hsf proteins function as molecular chaperones, engaging in vital cellular processes. They play a role in preventing the aggregation of misfolded proteins, rectifying misfolded protein structures, and facilitating the degradation of irreparably damaged proteins under stressful conditions<sup>5</sup>. Hsf

<sup>1</sup>Jilin Academy of Agricultural Sciences, Changchun, Jilin, China. <sup>2</sup>Heilongjiang Bayi Agricultural University, Daqing, Heilongjiang, China. <sup>3</sup>These authors contributed equally: Caihua Li and Yuhuan Li. ✉email: muzs@163.com

members have been reported to participate in various aspects of plant growth and development, particularly emphasizing their involvement in responding to both abiotic and biotic stresses<sup>6</sup>. HSF-HSP pathway was considered as a classical thermal regulatory mechanisms in plants<sup>7</sup>, which heat signaling is transmitted through reactive oxygen and calcium ions<sup>8</sup>. *HsFA1* as a regulator of the HSF-HSP pathway, which activates the expression of downstream HSP genes and interacts with HSF members such as *HsFA1* and *HsFB1*. Also, *HsFA1* induces the expression of *DREB2A* directly, which *DREB2A* is a related osmotic and heat stress and positively controls osmotic- and heat-inducible gene and regulates the expression of the *HsFA3* to enhance tolerance of plants<sup>9,10</sup>; In *Arabidopsis*, *HsFA1b* has been documented to play a role in enhancing plant resistance against *Pseudomonas aeruginosa*. Notably, the *fc/a* mutant displays heightened sensitivity to this pathogen<sup>11</sup>. In tomato (*Solanum lycopersicum*), the Hsf member known as *SIHsFA1a* plays a crucial role in generating extracellular H<sub>2</sub>O<sub>2</sub> upon nematode infection. This process further facilitates the accumulation of heat shock proteins and augments the foundational defense mechanisms<sup>12</sup>; *HsFA4a*'s function is closely intertwined with cell death processes during pathogen infections. It upregulates the expression of pathogen-responsive genes, contributing to the intricate interplay between plants and pathogens<sup>13</sup>; *HsFB4d*, a member of the Hsf family in rice, is pivotal in enhancing resistance against bacterial wilt disease. It achieves this by inducing the expression of various genes involved in the plant's defense mechanisms<sup>14</sup>.

Additionally, Hsf members have been reported to play a significant role in responding to abiotic stresses<sup>6</sup>: *HsFA1* is an important Hsf member in *Arabidopsis* known for its role in regulating heat tolerance. Mutants of *HsFA1*, such as *hsfa1abcd* mutants, show severe impairment in responding to heat stress<sup>15</sup>; Overexpression of *HsFA2* has been shown to restore the heat-sensitive phenotype observed in *hsfa1abcd* mutants<sup>16</sup>; ABA has been demonstrated to induce the expression of several HSF members, including *HsFA6a*, *HsFA6b*, *HsFA8*, *HsFB2a*, *HsFB2b*, and *HsFC1*<sup>17</sup>; Cold stress can cause the expression of *HsFA4a*, *HsFA4b*, *HsFA8*, and *HsFC1*. Additionally, overexpression of *HsFA1*, *HsFA2*, *HsFA3*, and *HsFAa* has been demonstrated to enhance heat tolerance in plants and improve tolerance to various other stresses, including drought, salt, hypoxia, and osmotic stress<sup>18–21</sup>.

Hsf members have been identified in various plant species, including *Arabidopsis*<sup>4</sup>, rice (*Oryza sativa*)<sup>22</sup>, maize (*Zea Mays*)<sup>23</sup>, soybean (*Glycine max*)<sup>24</sup>, wheat (*Triticum aestivum*)<sup>25</sup> and common bean (*Phaseolus vulgaris*)<sup>26</sup>. However, comprehensive research on Hsf members in asparagus (*Asparagus officinalis*) is lacking. Asparagus is a vegetable of significant nutritional and health value, primarily harvested for its tender stems, which are rich in essential vitamins, minerals, and medicinal properties<sup>27</sup>. Asparagus is known for its abundant nutrients, including vitamins, carbohydrates, proteins, diverse amino acids, and various bioactive compounds such as polyphenols, flavonoids, saponins, and asparaginic acid<sup>28</sup>. These compounds contribute to different dietary effects, such as immune regulation, moistening of the lungs and relieving cough, cardiovascular health support, and potential anti-tumor properties<sup>27</sup>. Renowned as the “king of vegetables” in the global market, asparagus holds a cherished position among consumers worldwide<sup>29</sup>. Nevertheless, asparagus stands vulnerable to various environmental stressors, particularly abiotic stresses, leading to substantial reductions in both yield and quality.

Consequently, enhancing the abiotic stress tolerance of asparagus through molecular breeding has become imperative. In this context, a thorough genome-wide exploration of Heat shock transcription factor (Hsf) members within the reference genome was undertaken. This comprehensive study encompassed an exhaustive analysis of these Hsf members, encompassing their physical and chemical attributes, evolutionary patterns, underlying motifs, gene structural compositions, *cis*-acting elements, collinearity, and expression profiles in response to abiotic stressors. The outcomes of this investigation not only unraveled the unique attributes of Hsf members but also furnished a vital theoretical foundation for the practical application of molecular breeding strategies in asparagus improvement.

## Materials and methods

### The identification of Hsf members in asparagus

Genomic data, encompassing gene, cDNA, and protein information, were sourced from the National Center for Biotechnology Information (NCBI) database (<https://www.ncbi.nlm.nih.gov/>), with the accession number PRJNA376608 (<https://www.ncbi.nlm.nih.gov/genome/10978>). The identification of Heat shock transcription factor (Hsf) members was facilitated by utilizing HMMER software, which focused on detecting the domain (PF00447). This approach enabled the comprehensive identification and subsequent exploration of Hsf members within the provided genomic dataset<sup>30</sup> while 1e<sup>-15</sup> was set as a filter threshold. Also, the Hsf members in *Arabidopsis* and rice (*Oryza sativa*) were came from the reference of Busch's<sup>31</sup> and Guo's<sup>22</sup> results for comprehensive analysis; The roster of identified Hsf members was subjected to a screening process facilitated by ExpASY Proteomics Server and Plant Protein Phosphorylation Database software. This meticulous scrutiny allowed for a more refined selection and analysis of the Hsf members within the dataset<sup>32</sup> by default parameters. All of the HSF members identified in asparagus after screened was used for subsequent analysis. The selected candidates, resulting from the screening process, were definitively identified as members of the Hsf family, marking the commencement of subsequent in-depth analysis.

### The analysis of Hsf members in asparagus

Various databases, including NCBI and Phytozome, furnished essential details regarding Hsf members, encompassing attributes such as location and coding sequence length, enabling a thorough examination of these members' physical and chemical properties. In parallel, the evolutionary relationship among Hsf members was elucidated through MEGA X software. The optimal model for this analysis was predicted using MEGA X<sup>30</sup> while while maximum likelihood method and 1000 bootstrap was used in this study, which the suitable model was predicted by MEGA X. Moreover, the reference sources for Hsf members in *Arabidopsis* and rice (*Oryza sativa*) were derived from Busch's work<sup>31</sup> and Guo's<sup>33</sup> results. For motif analysis of Hsf members, the MEME software

was harnessed, with the criterion that each motif's value was below  $1e^{-20}$ . Indeed, the length of each motif encompassed a range of 10 to 50 amino acids<sup>34</sup>. Indeed, the Gene Structure Display Server (GSDS) software was employed to analyze the gene structure of the Hsf members. Additionally, the Gene-wise software facilitated the analysis of DNA coordinates, including exonic and intronic regions, as well as coding sequence and protein sequences<sup>35</sup>, which filter according to the default parameters of the software; Indeed, the PlantCARE software was utilized to predict the *cis*-acting elements of the Hsf members, providing a comprehensive analysis of these elements<sup>36</sup>, which 0.01 was set as a filter default parameters; The collinear relationship of the Hsf members was analyzed using the Circos and MCScanX software tools<sup>37</sup>, which the location information and cds sequences of two species' members, and filtering the collinear regions, all collinear gene pairs are ultimately obtained by McScanX.

### The expression pattern analysis of Hsf members in asparagus

The expression patterns of Hsf members in asparagus were studied using Jinglvlu III, a locally grown variety provided by the Jilin Academy of Agricultural Sciences. Four-month-old seedlings with normal growth were selected as the plant materials. These seedlings were divided into different tissues, including roots, stems, and leaves, for the analysis of tissue-specific expression using the quantitative real-time PCR (qRT-PCR) method<sup>38</sup>. Subsequently, the roots were chosen as the target tissue to analyze the expression of Hsf members under abiotic stresses. This analysis designated water treatment as the control (CK) treatment. At the same time, 70 mmol/L NaCl was employed as the salt stress (S) treatment<sup>39</sup>. A temperature of 43 °C was designated as the heat stress (H) treatment, and a temperature of 4 °C was selected as the cold stress (C) treatment<sup>39</sup>; After subjecting the samples to stress treatments for 48 h, the target tissues were collected for subsequent expression pattern analysis. The RNA from each sample was extracted using the MolPure® RNA Kit (19,291, YEASEN, Shanghai, China). The extracted RNA was then reverse transcribed into cDNA using the Evo M-MLV reverse transcription reagent premix (AG11706, AG, Hunan) after its quality was assessed by 1% agarose gel electrophoresis and a NanoDrop instrument (OneC, Thermo Fisher, Waltham, MA, USA).

To determine the expression levels of Hsf members, qRT-PCR analysis was conducted. The primers for amplifying Hsf members were designed, and 2 × SYBR qPCR Master Mix (M4QS02, bioeast, Hangzhou, China) was utilized. The LightCycler® system (Roche 480II, Roche, Switzerland) was employed to analyze the expression levels of Hsf members, and the experiments were carried out with three biological replicates. The relative expression level was calculated using the  $2^{-\Delta\Delta Ct}$  method<sup>40</sup>.

### Statement

The plant materials used in this experiment are local cultivated varieties, which comply with national guidelines and legislation. All the methods were carried out in accordance with relevant Institutional guidelines and regulations.

## Results

### The identification of Hsf members in asparagus (AoHsfs)

In this study, 21 members were initially identified through the reference genome. However, three redundant members were found after removing the duplicates, leaving a final count of 18 unique members. These members were then designated with names based on their chromosomal arrangement, resulting in the nomenclature AoHsf01 to AoHsf08, respectively. The analysis revealed that these AoHsf members were distributed across eight different chromosomes. Notably, AoHsf08 was not associated with any specific chromosome (NW\_017972489.1). Among these, four members, namely AoHsf11, AoHsf12, AoHsf13, and AoHsf14, were located on chromosome NC\_033803.1 (Fig. 1).

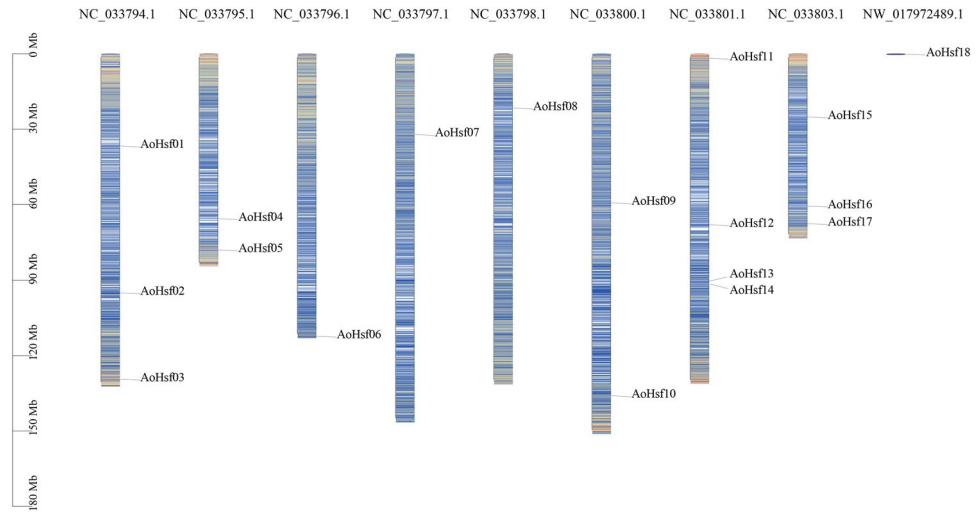
Detailed information regarding the AoHsf members was retrieved from the NCBI and Phytozome databases and is provided in Table S2. The length of the AoHsf proteins ranged from 170 to 504 amino acids. Specifically, AoHsf02 exhibited the largest protein length, while *AoHsf06* had the shortest. The instability index of the AoHsf members varied between 42.04 (*AoHsf02*) and 67.62 (*AoHsf06*). Furthermore, the aliphatic index of these members ranged from 54.61 to 80.96. The comprehensive information compiled here serves as a foundational resource for understanding the characteristics of the AoHsf members.

### The evolution of AoHsfs

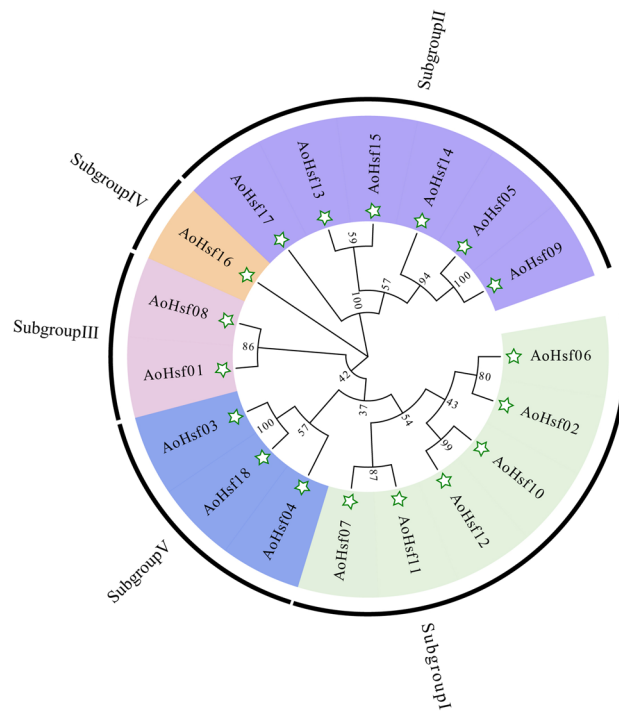
The protein sequences of AoHsfs were subjected to detection and analysis using MEGA X software. For optimization, the jtt+g+i model was selected based on predictions by MEGA X. The outcome revealed the classification of the 18 AoHsf members into five distinct clusters, designated as subgroup I, II, III, IV, and V (Fig. 2). Among these clusters, subgroup IV contained only one AoHsf member, representing the subgroup with the fewest members. In contrast, subgroup I and II comprised six AoHsf members each, signifying the subgroups with the highest number of members in this classification.

### The motifs and gene structure of AoHsfs

The motifs and gene structure analysis of AoHsfs was conducted using MEME and GSDS software (Fig. 3). The findings revealed that members within the same subgroup shared similar motifs (Fig. 3A,B and Table S3). Specifically, only subgroup IV lacked motif 1, and subgroup II lacked motifs 3, 5, 6, and 8. Additionally, motif four was exclusively identified in subgroup II. The gene structure analysis results indicated that AoHsfs within the same subgroup also exhibited a similar gene structure. Notably, the longest member was identified in subgroup I, while members in subgroup III were consistently shorter than those in other subgroups (Fig. 3A,C).



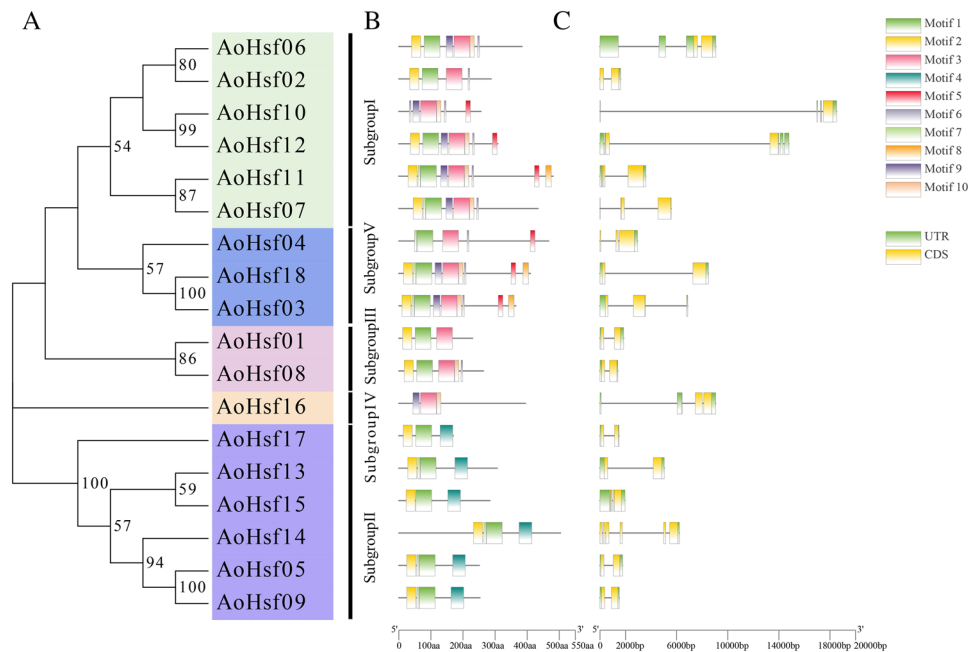
**Figure 1.** The spatial distribution of AoHsf members was illustrated in the analysis. The vertical number axis on the left represented the chromosomal ruler, indicating the positions of chromosomes. The members were depicted by black lines, signifying their locations along the chromosomes. To denote gene density on the chromosomes, the lines were colored differently. The color gradient ranged from red to blue, symbolizing the gene density from high to low respectively.



**Figure 2.** The evolutionary relationship of AoHsfs. The green stars were used to indicate AoHsf members and the background was color-coded, with green, purple, pink, orange, and blue representing subgroup I, II, III, IV, and V, respectively. The numbers displayed on the tree corresponded to the evolutionary values of the AoHsfs.

**The evolution and motifs of Hsf members**

This comprehensive study examined Hsf members across three plant species: *Arabidopsis*, *Oryza sativa*, and *Asparagus officinalis*. A total of 64 Hsf members were identified from the reference genomes of these species. The evolutionary relationship analysis revealed their classification into five subgroups, with subgroup IV having the lowest number (3 members) and subgroup I containing the most significant number of Hsf members. In the motif analysis, members within the same subgroup displayed similar motif profiles, with specific motifs, such as motif 6 and 15, unique to subgroup II (Fig. 4 and Table S4).



**Figure 3.** The motifs and gene structure of AoHsfs were intricately analyzed. **(A)** The evolutionary trajectory of AoHsfs, where the five distinct colors signify subgroups I through V. **(B)** Motif representation of AoHsfs. The ten differently colored squares denote the distinct motifs, numbered 1 through 10. **(C)** Illustrates the gene structure of AoHsfs. Here, the green boxes depict the untranslated region (UTR), the yellow boxes indicate the coding sequence (CDS) region and the black lines represent the intron regions.

### The *cis*-acting elements of AoHsfs

In this analysis, the *cis*-acting elements of AoHsfs were thoroughly examined using PlantCare software. The predicted functions of these *cis*-acting elements are detailed in Table 1. These elements were categorized into three types and visualized with different colors in Fig. 5: stress-related elements (blue), hormone-related elements (red), and sprout-related elements (yellow). Nearly all *AoHsf* members possessed hormone-related and stress-related characteristics, except *AoHsf03* (lacking hormone-related elements) and *AoHsf17* (lacking stress-related aspects). This analysis underscores the close association between AoHsfs and stress, hormones, and sprouting, revealing their potential roles in these processes. The analysis of *cis*-acting elements in *AoHsfs* revealed three distinct types, represented by different colors in Fig. 5: stress-related elements (blue), hormone-related elements (red), and sprout-related elements (yellow). Almost all *AoHsf* members were found to have both hormone-related and stress-related features, except for *AoHsf03* and *AoHsf17*. This observation highlights the close association between *AoHsfs* and hormonal responses and their involvement in stress-related processes. Notably, *AoHsf06* was identified to contain an RY element, suggesting its potential role during sprouting. These findings underscore the multifaceted participation of *AoHsfs* in stress, hormonal regulation, and sprouting processes.

### The collinearity of AoHsfs

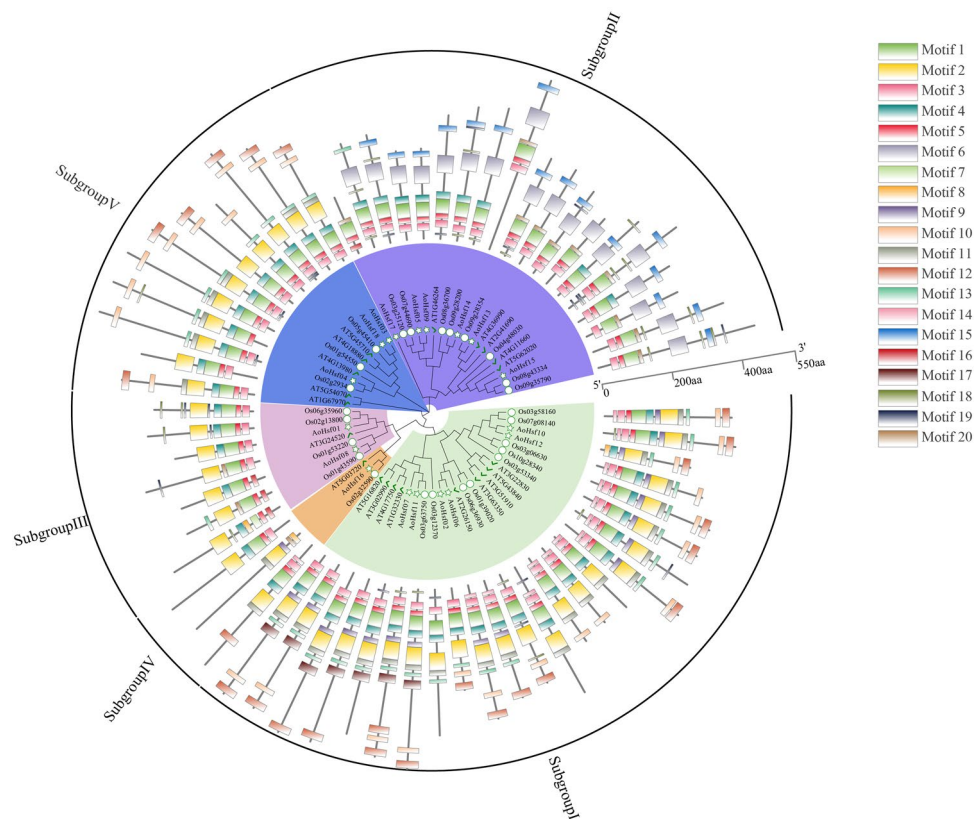
Also, the collinear relationship was analyzed. The result of collinearity showed that there were three pairs of collinear genes with soybean (*Glycine max*), which were *AoHsf10* with *GLYMA\_03G175300*, *AoHsf05* with *GLYMA\_02G278400*, and *AoHsf14* with *GLYMA\_02G278400*. Two pairs of collinear genes with rice (*Oryza sativa*) and only one pair had a collinear relationship, *AoHsf11* with *AT1G32330*. All these results revealed that these members might have a close relationship with these collinear genes (Fig. 6 and Table S5).

### The expression patterns of AoHsfs

The expression patterns of AoHsfs in different tissues at the seeding stage were tested by qRT-PCR analysis. The plant was divided into four tissues: leaves, stems, buds, and roots (Fig. 7A). The results showed that almost all AoHsfs had higher expression levels in roots (10) than in leaves, stems, and buds, which suggests that roots might be a target tissue for studying AoHsfs at the seeding stage (Fig. 7B).

### The expression level under abiotic stresses

Nine *AoHsfs*, which exhibited higher expression levels in roots, were selected as target genes to investigate their expression changes under abiotic stresses. The results showed that nearly all of these members exhibited increased expression levels under abiotic stresses, including heat, cold, and salt, except for *AoHsf12*, which displayed differential expression patterns under different stresses (Fig. 8). Notably, *AoHsf05*, *AoHsf13*, and *AoHsf15* showed significantly higher expression levels under all three abiotic stress conditions than the control (CK) treatment.



**Figure 4.** The evolutionary trajectory and motif patterns of Hsf members from *Arabidopsis*, rice (*Oryza sativa*), and asparagus (*Asparagus officinalis*) were extensively analyzed. The inner circle delineates the evolutionary relationships among the Hsf members, while the outer circle showcases the motif distribution across these members. Each of the 20 differently colored boxes signifies a unique motif, ranging from motif 1 to motif 20.

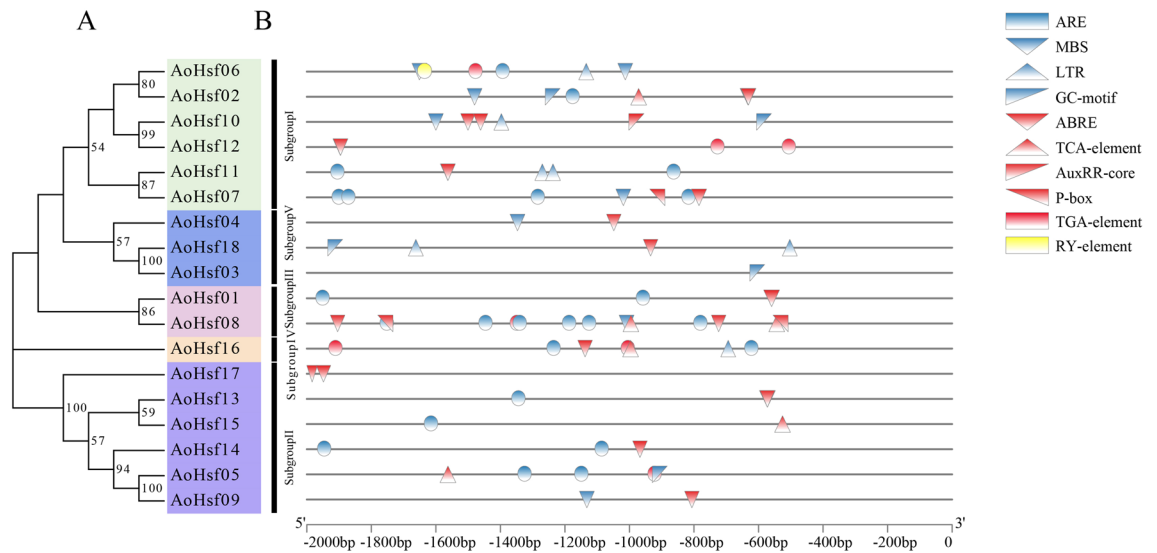
Element name	The function of <i>cis</i> -elements
ARE	<i>Cis</i> -acting regulatory element essential for the anaerobic induction
LTR	<i>Cis</i> -acting element involved in low-temperature responsiveness
ABRE	<i>Cis</i> -acting element involved in the abscisic acid responsiveness
TGA-element	Auxin-responsive element
TCA-element	<i>Cis</i> -acting element involved in salicylic acid responsiveness
GC-motif	Enhancer-like element involved in anoxic specific inducibility
MBS	MYB binding site involved in drought-inducibility
RY-element	<i>Cis</i> -acting regulatory element involved in seed-specific regulation
P-box	Gibberellin-responsive element
AuxRR-core	<i>Cis</i> -acting regulatory element involved in auxin responsiveness

**Table 1.** The prediction of *cis*-acting elements of AoHsfs.

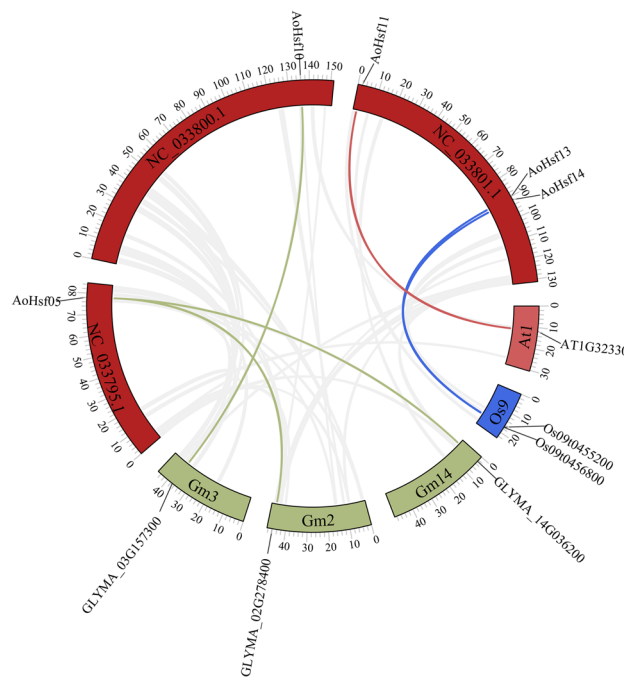
These genes hold promise as candidate *AoHsfs* for breeding programs to enhance stress tolerance. Interestingly, most of the AoHSF members with significant changes in differential expression levels had the stress-related elements while the expression of *AoHsf12* had no such great change under stress (Figs. 5, 8), these information also indirectly revealed the potential features of AoHSF members in response to abiotic stresses.

## Discussion

Hsf members have been identified in different species, and the count of Hsf members varies between species. For example, *Arabidopsis* possesses 21 Hsf members<sup>3</sup>; however, tomato (*Solanum lycopersicum*) has 18 identified Hsf members<sup>4</sup>. From the alfalfa (*Medicago sativa*) genome, 16 Hsf members were successfully identified<sup>41</sup> while 30 members were found in common bean (*Phaseolus vulgaris*)<sup>26</sup>, 29 members were identified in buckwheat (*Fagopyrum tataricum*)<sup>42</sup> Poplar (*Populus*) revealed a total of 28 Hsf members within its genome<sup>41</sup>. In this study,



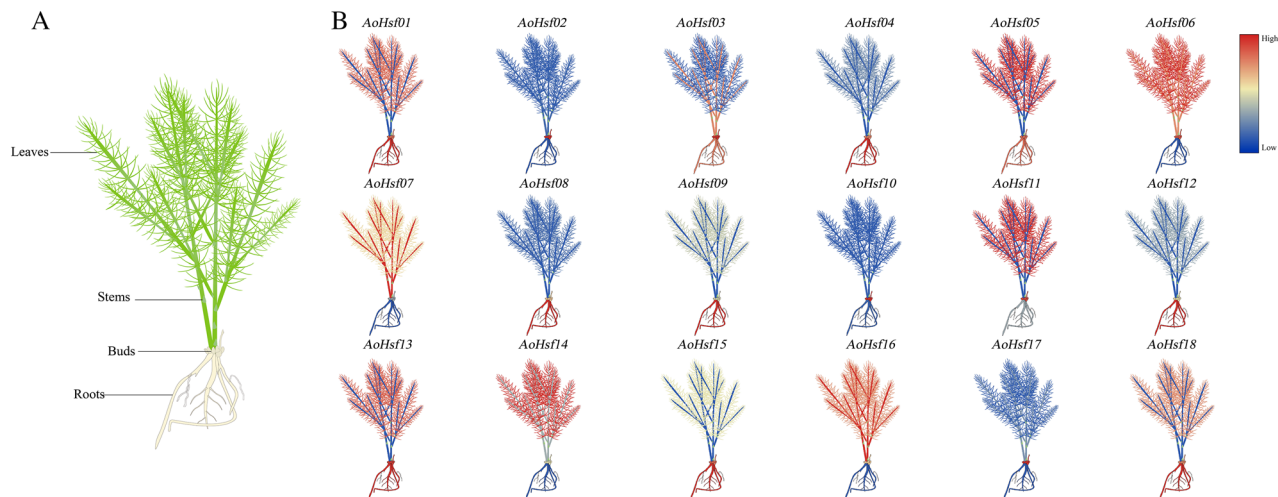
**Figure 5.** The *cis*-acting elements analysis of AoHsfs. (A) The evolution of AoHsfs. Five different colors represented subgroup I–V; (B) The *cis*-acting elements of AoHsfs. Blue boxes were the stress-related elements; red boxes were hormone-related elements; Yellow boxes were the elements with sprout-related functions.



**Figure 6.** The collinearity analysis of AoHsfs. The red squares represented the chromosomes of asparagus (*Asparagus officinalis*), while the green squares represented the chromosomes of soybean (*Glycine max*); The blue squares represented the chromosomes of rice (*Oryza sativa*), while pink squares represented the chromosomes of Arabidopsis. The bright lines were the collinear relationship with members, while the gray lines were the collinearity background.

we identified 18 Hsf members in asparagus through the reference genome in the database. The relatively lower number of Hsf members in asparagus compared to some other species can be attributed to two main factors. First, the reference genome of asparagus is relatively tiny compared to other species. Second, the lower number could result from gene duplication events during genomic replication and evolution<sup>43</sup>.

Furthermore, the Hsf members exhibited diverse classifications in their evolutionary relationships, ranging from 3 to 9 subgroups. This diversity reflects the complexity and differentiation within the Hsf gene family in asparagus<sup>44–46</sup>. In this study, these 18 members were classified into five subgroups, which aligns with the number of subgroups observed in the evolutionary analysis. The evolutionary and motif analyses involving three species (*Arabidopsis*, *Oryza sativa*, and *Asparagus officinalis*) revealed that these 64 members were grouped into five



**Figure 7.** Analysis of AoHsfs' Expression Patterns at the Seeding Stage. **(A)** Schematic representation of seedlings in various tissues, including leaves, stems, buds, and roots; **(B)** Expression patterns of AoHsfs across different tissues. Colors range from blue to red, indicating expression levels from low to high.

subgroups. Remarkably, each subgroup contained Hsf members from *Arabidopsis*, *Oryza sativa*, and *Asparagus officinalis*. This observation suggests that Hsf members are conserved across monocotyledonous and dicotyledonous plants throughout evolution.

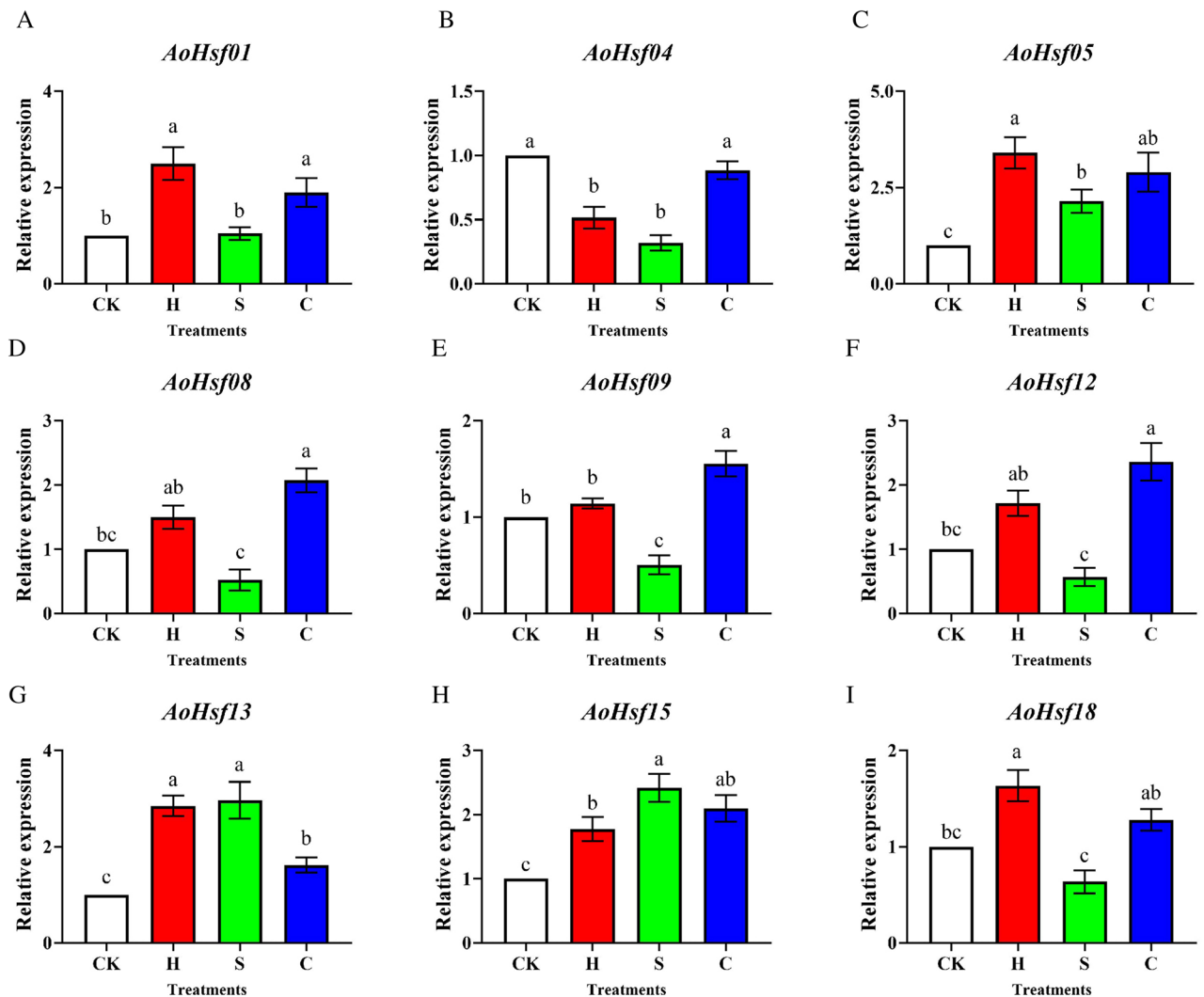
Each subfamily of Hsf members exhibited distinct characteristics based on motifs, gene structure, and *cis*-acting elements. These characteristics were similar to those of Hsf members in various other plant species<sup>46</sup>. Motifs 1 and 2 contained the DBD domain and were present in all AoHsf members. Furthermore, members within the same subfamily shared similar motifs. It's worth noting that Hsf members have also been identified in bamboo (*Phyllostachys edulis*)<sup>47</sup>. Similar results have been observed in tea<sup>48</sup> (*Camellia sinensis*) and maize<sup>23</sup> (*Zea Mays*). Gene structures, including coding sequences (CDS) and introns, play crucial roles in regulating gene expression levels and can provide insights into gene functions<sup>49</sup>.

The gene structure analysis of AoHsfs indicated that members within each subgroup had a similar structure, suggesting potential functional similarities among members within the same subgroup. Similar findings have been reported in common bean and cotton as well<sup>46</sup>; *Cis*-elements located in the promoter regions of genes play a critical role in regulating gene expression levels, which can significantly impact the secondary metabolism of plants and influence various aspects of their biology<sup>36</sup>. The analysis of *cis*-acting elements in AoHsf members revealed the presence of different hormone-related elements (including ARBE, TCA-element, AuxRR-core, P-box, and TGA-element) as well as stress-related elements (including ARE, MBS, LTR, and GC-motif). These findings suggest that AoHsfs are closely associated with hormonal regulation and stress responses. Similar *cis*-elements, such as ABRE, TGA-element, AuxRR-core, and P-box, have been identified in Hsf members from other plant species like *Dianthus caryophyllus* and *Hypericum perforatum*. This suggests that these elements are conserved across different plant species and likely play a significant role in regulating Hsf genes<sup>50</sup>.

Furthermore, certain *cis*-acting elements, including LTR, ARE, and MBS, were also identified in Hsf genes from other plant species like *Brassica juncea* and *Camellia sinensis*<sup>48</sup>. Similar *cis*-acting elements in the promoter regions of Hsf genes from different plant species suggest that these elements are characteristic features of Hsfs and play a role in their regulation. Additionally, identifying collinear genes, such as ATIG32330 as a collinear gene of AoHsf11, suggests that AoHsf11 may have a role in responding to abiotic stress. The fact that AoHsfs show more collinear relationships with soybean than with rice and Arabidopsis could be attributed to differences in genome size and gene duplication events within the Hsf family. Also, GLYMA\_02G278400, GLYMA\_14G036200 and GLYMA\_03G175300 had a connection with stresses through genome-wide association study as the collinearity gene of AoHsfs<sup>51,52</sup>.

Seeding is a critical phase in a plant's life cycle, as it significantly influences plant development and ultimately affects yield. Moreover, this stage is susceptible to various abiotic stresses, which can profoundly impact plant growth and productivity<sup>53</sup>. Identifying the target tissue for understanding the function of AoHsfs is crucial in unraveling their roles in plant growth and stress responses<sup>54</sup>. In this study, the roots have been chosen as the primary target tissues for further investigation due to their consistently elevated expression levels of AoHsfs both during the seeding stage and in response to abiotic stress conditions. Similarly, in the case of PvHsfs, the root also emerged as the most suitable tissue for studying their functions and roles due to their significant expression levels in this organ<sup>46</sup>. The expression levels of AoHsfs significantly changed under abiotic stresses, indicating their involvement in responding to these stresses. However, it's worth noting that not all Hsf members necessarily respond to every type of abiotic stress. Therefore, it is crucial to identify broad-spectrum genes that can react to multiple abiotic stresses, providing a more comprehensive defense mechanism for the plant<sup>46</sup>. HSF members had an important role under heat stress in lilies (*Lilium longiflorum*), which HSF members were considered as differentially expressed genes while HSF-HSP pathway was considered as a candidate pathways; PsnHSF21





**Figure 8.** The expression patterns analysis of *AoHsfs* under abiotic stresses. The white column represented CK treatment, the red column represented heat stress treatment, the green column represented salt stress treatment, and the blue column described cold stress treatment. Lowercase letter(s) indicated significant differences ( $\alpha = 0.05$ ). (A) The expression of *AoHsf01* exhibited variations across different abiotic stress treatments; (B) The expression level of *AoHsf04* under abiotic stresses; (C) The expression level of *AoHsf05* under abiotic stresses; (D) The expression level of *AoHsf08* under abiotic stresses; (E) The expression level of *AoHsf09* under abiotic stresses; (F) The expression level of *AoHsf12* under abiotic stresses; (G) The expression level of *AoHsf13* under abiotic stresses; (H) The expression level of *AoHsf15* under abiotic stresses; (I) The expression level of *AoHsf18* under abiotic stresses.

improved salt tolerance by specifically binding to the stress-related *cis*-acting element HSE while the transgenic poplar overexpressing plant had a better growth state under salt stress. .

The study demonstrates that the expression levels of many *AoHsf* members changed under abiotic stress conditions, suggesting their potential roles in responding to such stress. Members like *AoHsf05*, *AoHsf13*, and *AoHsf15* exhibited significantly higher expression levels under all three tested abiotic stresses, which these three members all had hormone and stress related *cis*-acting elements (Fig. 5). This finding suggests that these particular *AoHsfs* could be promising candidates for breeding programs to enhance stress tolerance in asparagus.

## Conclusions

In this study, we identified 18 *AoHsfs* through the reference genome. Subsequently, we conducted a comprehensive analysis to gain insights into the evolutionary relationships, motifs, gene structures, collinearity, and expression patterns of these *AoHsfs*. Our findings revealed that these *AoHsfs* could be classified into five distinct clades, each showing similarities in motifs, gene structure, and *cis*-acting elements. Moreover, due to their higher expression levels, we identified roots as a suitable target tissue for studying *AoHsfs* during the seeding stage. Additionally, *AoHsf05*, *AoHsf13*, and *AoHsf15* exhibited significant upregulation under various abiotic stress conditions, making them promising candidates for enhancing stress tolerance through breeding programs.

This study provides valuable insights into the characterization of *AoHsfs* and highlights potential candidates for future breeding efforts.

## Data availability

**Datasets:** The datasets analyzed during this study could be available from the corresponding author on reasonable request.

Received: 31 May 2023; Accepted: 18 October 2023

Published online: 23 October 2023

## References

1. Westerheide, D. S., Raynes, R., Powell, C., Xue, B. & Uversky, N. V. HSF transcription factor family, heat shock response, and protein intrinsic disorder. *Curr. Prot. Pept. Sci.* **13**, 86–103 (2012).
2. Harrison, C. J., Bohm, A. A. & Nelson, H. C. Crystal structure of the DNA binding domain of the heat shock transcription factor. *Science* **263**, 224–227 (1994).
3. Nover, L. *et al.* Arabidopsis and the heat stress transcription factor world: How many heat stress transcription factors do we need?. *Cell Stress Chaperones* **6**, 177–189 (2001).
4. Scharf, K.-D., Berberich, T., Ebersberger, I. & Nover, L. The plant heat stress transcription factor (Hsf) family: Structure, function and evolution. *Biochimica et Biophysica Acta (BBA) Gene Regul. Mech.* **1819**, 104–119 (2012).
5. Zhang, N., Wang, Y., Wang, Z., Yue, Z. & Niu, Y. Heat shock transcription factor family in plants: a review. *Sheng Wu Gong Cheng Xue Bao* **37**, 1155–1167 (2021).
6. Guo, M. *et al.* The plant heat stress transcription factors (HSFs): Structure, regulation, and function in response to abiotic stresses. *Front. Plant Sci.* **7**, 114 (2016).
7. Gao, Z., Zhou, Y. & He, Y. Molecular epigenetic mechanisms for the memory of temperature stresses in plants. *J. Genet. Genomics* **49**, 991–1001 (2022).
8. Song, G. *et al.* Transcriptome analysis of Cd-treated switchgrass root revealed novel transcripts and the importance of HSF/HSP network in switchgrass Cd tolerance. *Plant Cell Rep.* **37**, 1485–1497 (2018).
9. Sakuma, Y. *et al.* Dual function of an *Arabidopsis* transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 18822–18827 (2006).
10. Zhou, Y., Xu, F., Shao, Y. & He, J. Regulatory mechanisms of heat stress response and thermomorphogenesis in plants. *Plants* **11**, 3410 (2022).
11. Bechtold, U. *et al.* Arabidopsis HEAT SHOCK TRANSCRIPTION FACTOR1b overexpression enhances water productivity, resistance to drought, and infection. *J. Exp. Bot.* **64**, 3467–3481 (2013).
12. Zhou, J. *et al.* Heat shock factor HsfA1a is essential for R gene-mediated nematode resistance and triggers H<sub>2</sub>O<sub>2</sub> production. *Plant Physiol.* **176**, 2456–2471 (2018).
13. Baniwal, S. K., Chan, K. Y., Scharf, K.-D. & Nover, L. Role of heat stress transcription factor HsfA5 as specific repressor of HsfA4. *J. Biol. Chem.* **282**, 3605–3613 (2007).
14. Yang, W. *et al.* OsHsfB4d binds the promoter and regulates the expression of OsHsp18.0-CI to resistant against *Xanthomonas oryzae*. *Rice* **13**, 28 (2020).
15. Yoshida, T. *et al.* Arabidopsis HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. *Mol. Genet. Genomics* **286**, 321–332 (2011).
16. Liu, H. & Charng, Y. Common and distinct functions of Arabidopsis class A1 and A2 heat shock factors in diverse abiotic stress responses and development. *Plant Physiol.* **163**, 276–290 (2013).
17. Fujii, H. & Zhu, J.-K. Arabidopsis mutant deficient in 3 abscisic acid-activated protein kinases reveals critical roles in growth, reproduction, and stress. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8380–8385 (2009).
18. 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.* **58**, 3373–3383 (2007).
19. Banti, V., Mafessoni, F., Loreti, E., Alpi, A. & Perata, P. The heat-inducible transcription factor HsfA2 enhances anoxia tolerance in Arabidopsis. *Plant Physiol.* **152**, 1471–1483 (2010).
20. Pérez-Salamó, I. *et al.* The heat shock factor A4A confers salt tolerance and is regulated by oxidative stress and the mitogen-activated protein kinases MPK3 and MPK6. *Plant Physiol.* **165**, 319–334 (2014).
21. Li, C. AtHsfA2 modulates expression of stress responsive genes and enhances tolerance to heat and oxidative stress in Arabidopsis. *Sci. China Ser. C* **48**, 540 (2005).
22. Guo, J. *et al.* Genome-wide analysis of heat shock transcription factor families in rice and *Arabidopsis*. *J. Genet. Genomics* **35**, 105–118 (2008).
23. Zhang, H. *et al.* Genome-wide identification, transcriptome analysis and alternative splicing events of Hsf family genes in maize. *Sci. Rep.* **10**, 8073 (2020).
24. Li, P.-S. *et al.* Genome-wide analysis of the Hsf family in soybean and functional identification of GmHsf-34 involvement in drought and heat stresses. *BMC Genomics* **15**, 1009 (2014).
25. Yun, L. *et al.* Phylogenetic and expression analyses of HSF gene families in wheat (*Triticum aestivum* L.) and characterization of TaHSFB4-2B under abiotic stress. *Front. Plant Sci.* **13**, 1047400 (2023).
26. Zhang, Q. *et al.* Heat shock transcription factor (Hsf) gene family in common bean (*Phaseolus vulgaris*): genome-wide identification, phylogeny, evolutionary expansion and expression analyses at the sprout stage under abiotic stress. *BMC Plant Biol.* **22**, 33 (2022).
27. Chitrakar, B., Zhang, M. & Adhikari, B. Asparagus (*Asparagus officinalis*): Processing effect on nutritional and phytochemical composition of spear and hard-stem byproducts. *Trends Food Sci. Technol.* **93**, 1–11 (2019).
28. Huang, X. & Kong, L. Steroidal saponins from roots of *Asparagus officinalis*. *Steroids* **71**, 171–176 (2006).
29. Albanese, D., Russo, L., Cinquanta, L., Brasiello, A. & Di Matteo, M. Physical and chemical changes in minimally processed green asparagus during cold-storage. *Food Chem.* **101**, 274–280 (2007).
30. Finn, R. D. *et al.* HMMER web server: 2015 update. *Nucleic Acids Res.* **43**, W30–W38 (2015).
31. Busch, W., Wunderlich, M. & Schöffl, F. Identification of novel heat shock factor-dependent genes and biochemical pathways in *Arabidopsis thaliana*. *Plant J.* **41**, 1–14 (2005).
32. Hoogland, C., Mostaguir, K., Appel, R. D. & Lisacek, F. The world-2DPAGE constellation to promote and publish gel-based proteomics data through the ExpASY server. *J. Proteomics* **71**, 245–248 (2008).
33. Guo, J. *et al.* Genome-wide analysis of heat shock transcription factor families in rice and *Arabidopsis*. *J. Genet. Genomics* **35**, 105–118 (2008).
34. Bailey, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).
35. Hu, B. *et al.* GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* **31**, 1296–1297 (2015).

36. Lescot, M. *et al.* PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **30**, 325–327 (2002).
37. Krzywinski, M. *et al.* Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).
38. Zhang, Q. *et al.* Transcriptome-based analysis of salt-related genes during the sprout stage of common bean (*Phaseolus vulgaris*) under salt stress conditions. *Biotechnol. Biotechnol. Equip.* **35**, 1086–1098 (2021).
39. Zhang, Q. *et al.* Genome-wide identification reveals the potential functions of the bZIP gene family in common bean (*Phaseolus vulgaris*) in response to salt stress during the sprouting stage. *J. Plant Growth Regul.* **41**, 3075–3090 (2022).
40. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-Delta Delta C(T))</sup> method. *Methods* **25**, 402–408 (2001).
41. Wang, F. *et al.* Genome-wide analysis of the heat shock transcription factors in *Populus trichocarpa* and *Medicago truncatula*. *Mol. Biol. Rep.* **39**, 1877–1886 (2012).
42. Liu, M. *et al.* Genome-wide investigation of the heat shock transcription factor (*Hsf*) gene family in Tartary buckwheat (*Fagopyrum tataricum*). *BMC Genomics* **20**, 871 (2019).
43. Liu, M. *et al.* The potential role of auxin and abscisic acid balance and FtARF2 in the final size determination of Tartary buckwheat fruit. *IJMS* **19**, 2755 (2018).
44. Scharf, K.-D., Berberich, T., Ebersberger, I. & Nover, L. The plant heat stress transcription factor (*Hsf*) family: Structure, function and evolution. *Biochimica et Biophysica Acta (BBA) Gene Regul. Mech.* **1819**, 104–119 (2012).
45. Nover, L. *et al.* Arabidopsis and the heat stress transcription factor world: How many heat stress transcription factors do we need?. *Cell Stress Chaperones* **6**, 177–189 (2001).
46. Zhang, Q. *et al.* Heat shock transcription factor (*Hsf*) gene family in common bean (*Phaseolus vulgaris*): Genome-wide identification, phylogeny, evolutionary expansion and expression analyses at the sprout stage under abiotic stress. *BMC Plant Biol.* **22**, 33 (2022).
47. Huang, B. *et al.* Genome-wide identification and analysis of the heat shock transcription factor family in moso bamboo (*Phyllostachys edulis*). *Sci. Rep.* **11**, 16492 (2021).
48. Zhang, X., Xu, W., Ni, D., Wang, M. & Guo, G. Genome-wide characterization of tea plant (*Camellia sinensis*) *Hsf* transcription factor family and role of CsHsfA2 in heat tolerance. *BMC Plant Biol.* **20**, 244 (2020).
49. Shaul, O. How introns enhance gene expression. *Int. J. Biochem. Cell Biol.* **91**, 145–155 (2017).
50. Li, W., Wan, X.-L., Yu, J.-Y., Wang, K.-L. & Zhang, J. Genome-wide identification, classification, and expression analysis of the *Hsf* gene family in carnation (*Dianthus caryophyllus*). *IJMS* **20**, 5233 (2019).
51. Zhang, Y., Wang, Y., Zhou, W., Zheng, S. & Ye, R. Detection of candidate gene networks involved in resistance to *Sclerotinia sclerotiorum* in soybean. *J. Appl. Genet.* **63**, 1–14 (2022).
52. Xiong, H. *et al.* A genome-wide association study and genomic prediction for *Phakopsora pachyrhizi* resistance in soybean. *Front. Plant Sci.* **14**, 1179357 (2023).
53. Ibrahim, E. A. Seed priming to alleviate salinity stress in germinating seeds. *J. Plant Physiol.* **192**, 38–46 (2016).
54. Fahad, S. *et al.* Crop production under drought and heat stress: Plant responses and management options. *Front. Plant Sci.* **8**, 1147 (2017).

## Author contributions

C.L., Y.L., Z.M. and Y.H.: experiments and manuscript writing; Z.Z., Z.T., and X.Z.: Experimental processing and data collection; D.T., Y.L., H.D. and Z.L.: Analysis and plotting of data; M.S., X.H., H.X. and Y.Z.: Processing of Plant Materials and RNA Extraction. All authors have agreed to the publication of the manuscript.

## Funding

This study was supported by Jilin Provincial Department of Science and Technology (20230505019ZP and 20210202091NC).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-45322-w>.

**Correspondence** and requests for materials should be addressed to Z.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023