

Analysis of Heat Shock Transcription Factors (HSFs) Gene Family in *Coffea arabica*

Waki'atil Rosida¹, Fuad Bahrul Ulum¹, Dwi Indahning Rohmah¹, Mukhamad Su'udi^{1*}

¹Department of Biology, Faculty of Mathematics and Natural Sciences, University of Jember, Jember, East Java, Indonesia

*Corresponding author: msuudi.fmipa@unej.ac.id

Editor: Samuel Pereira de Carvalho

Received in November 5, 2025 and approved in May 7, 2026

ABSTRACT

Coffee (*Coffea arabica*) is an important commodity with global trade reaching US\$90 billion. Coffee production is greatly shaped by environmental factors to obtain high yield and quality. The growth of coffee plants is strongly influenced by how they respond to various stresses. Heat shock factors (HSF) play a role in plant growth and response to stress. This study aims to analyze the role and characteristics of heat shock factors (HSF) in *C. arabica* plants (CaHSF). The analysis method was carried out bioinformatically, consisting of HSF identification in the *C. arabica* genome, analysis of physical properties, structure, motifs, subcellular localization, phylogenetics, chromosomal localization, and protein interactions. *C. arabica* contains 54 HSF genes that exhibit diverse intron–exon arrangements, and motifs 1 and 3 are conserved across all CaHSF. The analysis of the CaHSF gene promoters reveals several cis-acting elements that play a role in plant stress responses, such as ARE, ABRE, LTR, and MBS. Most CaHSF proteins exhibit a strong degree of subcellular localization in the nucleus, and the phylogenetic results of these proteins are grouped into a single clade according to their respective species. CaHSF protein interactions are also involved with WD-repeats proteins that function in stress regulation and plant development.

Key words: Abiotic stress; coffee plant; HSFs; molecular regulation; plant stress.

1 INTRODUCTION

Coffee is one of the commodities that is traded internationally in many countries. Global coffee trade sales reached US\$90 billion with two coffee species dominating, namely *Coffea arabica* L. (Arabica coffee) and *Coffea canephora* L. (Robusta coffee). Coffee production and productivity are strongly influenced by climate to achieve high yields and quality (Damatta et al., 2008). The coffee products produced are not only used to meet domestic coffee consumption needs but are also distributed in the international market. Indonesian coffee has more than 50 export countries, such as the United States, Germany, Malaysia, Russia, and Italy. International trade is based on demand and supply between countries, driven by differences in natural factors. Climate variability is the main factor driving reductions in coffee yields worldwide (Apriani et al., 2022). The growth of coffee plants is strongly influenced by how they respond to various stresses.

Plants have protective responses to environmental stresses. One of those involved is the Heat shock transcription factors (HSFs), which regulate transduction signals related to heat stress in plants. HSFs play an important role in responding to stress in plants caused by climate or environmental change and in increasing resistance to abiotic stress (Jagadish et al., 2021).

HSF are interconnected with Heat Shock Protein (HSP) in activating transcription in response to heat stress that occurs in plants. HSFs bind to the Heat Shock Element (HSE) in the promoter region of the HSP gene, thereby activating transcription of heat-resistant proteins. The function of HSFs in plants are very broad, including the molecular defense of

plants that are under stress. Stress that occurs in plants can cause proteins to be denatured, as a molecular response by plants, HSFs signals will be activated to induce transcription of HSP genes which can help plant tolerance to stress (Guo et al., 2016). HSFs have several functional domains such as the N-terminal DNA-binding domain (DBD), nuclear export signal (NES), oligomerization domain (OD), nuclear localization signal (NLS), and C-terminal transcription activation domain (CTAD). The most conserved structures and domains are DBD and OD. DBD is located at the N-terminus and recognizes and binds the heat shock element (HSE) motif on the promoter of the target gene. HSFs regulate gene expression to maintain cell homeostasis and to control responses to heat stress (HS) and other environmental factors. Plants under stress will activate HSFs transcription and function as an aid to damaged folding and distributing it intracellularly and degraded under environmental stress (Wang et al., 2024). Research on HSF proteins in coffee has never been conducted, so it is necessary to conduct genome and protein analysis related to the mechanism and role of HSFs in *C. arabica* plants. This study aims to analyze the role and characteristics of HSF proteins in *C. arabica* plants through an *in silico* approach.

2 MATERIAL AND METHODS

2.1 Identification of HSF genes in *C. arabica*

The HSFs of *Arabidopsis thaliana* protein sequence was obtained from the NCBI (National Centre for Biotechnology Information) (NCBI, <https://www.ncbi.nlm.nih.gov/>).

.nih.gov) database. In the next, analyze the domain or motif of the AtHSFs protein sequence (AEK67477) using MOTIF tools (<https://www.genome.jp/tools/motif/>). Identification of the HSF genes in *C. arabica* was carried out through BLAST of AtHSFs motif against the coffee genome in the Phytozome database (Goodstein et al., 2012) ;<https://phytozome-next.jgi.doe.gov/>) and recorded transcript ID, location, CDS length, genome and protein sequences.

2.2 Physical properties of HSF proteins in *C. arabica*

Analysis of the physical properties of *C. arabica* HSF(CaHSF) proteins, including molecular weight (kDa), isoelectric point (pI), and Grand Average of Hydropathicity (GRAVY), was carried out using the ExPASy ProtParam tool software (Gasteiger et al., 2005); <https://web.expasy.org/protparam/>).

2.3 Gene structure and cis-elements analysis in the promoter of coffee HSF genes

Exon and intron distribution patterns in *C. arabica* HSF genes were analyzed using Gene Structure Display Server 2.0 (Hu et al., 2015; <https://gsds.gao-lab.org/>) based on the collection of coding sequences (CDS) and genomic sequences. Analysis of cis-acting elements in the promoter regions of the HSF gene of coffee was carried out using the plantCARE database (Lescot et al., 2002; <https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) at 2000 bp in the upstream part before the start of the codon (ATG), which was then visualized using the TBtools-II software (Chen et al., 2023).

2.4 Sub-cellular localization

The sub-cellular localization prediction of the *C. arabica* HSF proteins were performed using CELLO v.2.5 software (Yu et al., 2004; Yu et al., 2006;<http://cello.life.nctu.edu.tw/>) and WoLF PSORT (Horton et al., 2007;<https://wolfpsort.hgc.jp/>), and then visualized with TBtools-II software (Chen et al., 2023).

2.5 Interaction Protein Analysis

Protein interaction analysis was carried out using STRING (<https://string-db.org/>) software by inserting full-length protein sequences from *C. arabica*. This analysis was carried out to determine the interaction of proteins from coffee with other proteins.

2.6 Phylogenetic analysis of coffee HSF

Analysis of phylogenetic tree construction was carried out by collecting protein sequences of *C. arabica*, *A. thaliana*, *Oryza sativa*, *Manihot esculenta*, and *Zea mays* in the phytozome database (<https://phytozome-next.jgi.doe.gov/>) and alignment was carried out using ClustalW

(Thompson et al., 1994) with MEGA11 software (Tamura et al., 2021). The phylogenetic tree was constructed using the maximum likelihood method with partial deletion parameters and the bootstrap method with 1000 bootstrap replications.

3 RESULT

3.1 Analysis characteristic

Arabica coffee (*C. arabica*) has 54 HSF genes, named *CaHSF-1* to *CaHSF-54*, based on their chromosomal locations (Table 1). The length of the genes in *CaHSF* varies by an average of 1200 bp. CaHSF proteins have lengths ranging from 117 aa to 1033 aa, with molecular weights of 3.55 kDa to 116.85 kDa. The *isoelectric point* (pI) value ranges from 4.72 to 9.51. Based on the pI value, there are 35 CaHSF which is acidic (pI < 7) and 19 others are alkaline (pI > 7). The pI value shows that HSF proteins that have a value of 4.83-6.25 or are acidic are included in the HSF class A group while pI that has a value of more than 7 is included in HSF class B (Li et al., 2023). CaHSF proteins are located inside the nucleus and have a negative gravity value, thus all CaHSF proteins are hydrophilic (Jaspard & Hunault, 2014).

3.2 Gene structure and motif analysis of CaHSF

Gene structure analysis was used to see the structural linkage of the HSF gene family in *Arabica coffee*. Differences in gene structure relate to gene function in subcellular (Shamshad et al., 2023). The structure of the *CaHSF* genes consists of introns and exons with varying lengths (Figure 1). The longest structure is *CaHSF-21* at 21kb in length, and the shortest is *CaHSF-17* with a length of 300 bp. *CaHSF-21* has the most exon number of 26 and the least is found in *CaHSF-26* with only 1 exon. The most number of introns is found in *CaHSF-21* as many as 25 while in *CaHSF-26* there are no intron. The different gene structures in *CaHSF* affect the transcription process and gene expression.

Analysis motif showed that *CaHSF* genes have 10 motifs in their gene products. Motifs number 1, 2, and 3 are the most preserved motifs, so they exist in all *CaHSF* genes. The same motif is shared by *CaHSF 1, 2, 21-30, 40-43, 53, and 54* on all CaHSF proteins. The general structure of HSF is indicated by the presence of DNA binding domains (DBD) which are characterized by red and green motifs (Figure 2). This motif is an important regulator of gene expression in response to environmental stress. DBD motif is highly structured and located near the N-terminal of all HSF genes in *Arabica coffee*. This domain plays a role in the process of recognizing and binding heat shock elements (HSEs) to DNA (Feng et al., 2021).

Table 1: Characteristic of HSF in *Coffea arabica*.

Gene ID	Name	Chr	No. CDS (bp)	No. Amino acids (AA)	Molecular Weight (kDa)	pI	GRAVY	Protein Localization Prediction
CAG016072.t1	<i>CaHSF-1</i>	1	1065	355	40.33	5.57	-0.931	Nuclear
CAG019868.t1	<i>CaHSF-2</i>	1	1065	355	40.37	5.56	-0.912	Nuclear
CAG012822.t1	<i>CaHSF-3</i>	1	1080	360	41.04	7.67	-0.992	Nuclear
CAG013422.t1	<i>CaHSF-4</i>	1	624	208	23.74	9.08	-0.604	Nuclear
CAG022526.t1	<i>CaHSF-5</i>	2	1137	379	43.01	4.83	-0.543	Nuclear
CAG028785.t1	<i>CaHSF-6</i>	2	525	175	20.75	9.51	-0.956	Nuclear
CAG020780.t1	<i>CaHSF-7</i>	2	1041	347	37.65	5.43	-0.631	Nuclear
CAG027124.t1	<i>CaHSF-8</i>	2	1041	347	37.72	5.25	-0.616	Nuclear
CAG028979.t1	<i>CaHSF-9</i>	2	1032	344	39.57	8.98	-0.724	Nuclear
CAG025622.t1	<i>CaHSF-10</i>	2	1023	341	39.12	7.70	-0.713	Nuclear
CAG022725.t1	<i>CaHSF-11</i>	2	1992	664	75.20	5.37	-0.474	Nuclear, cytoplasmic
CAG028268.t1	<i>CaHSF-12</i>	2	1032	344	39.42	8.38	-0.72	Nuclear
CAG028309.t1	<i>CaHSF-13</i>	2	1191	397	44.55	5.60	-0.610	Nuclear
CAG022038.t1	<i>CaHSF-14</i>	2	1071	357	40.11	6.37	-0.634	Nuclear
CAG028980.t1	<i>CaHSF-15</i>	2	1080	360	40.62	5.61	-0.723	Nuclear
CAG028772.t1	<i>CaHSF-16</i>	2	1290	430	49.69	5.61	-0.906	Nuclear
CAG028308.t1	<i>CaHSF-17</i>	2	831	277	31.87	9.33	-0.71	Nuclear
CAG022156.t1	<i>CaHSF-18</i>	2	633	211	24.86	9.39	-0.914	Nuclear
CAG028416.t1	<i>CaHSF-19</i>	2	633	211	24.87	9.47	-0.891	Nuclear
CAG022514.t1	<i>CaHSF-20</i>	2	351	117	12.94	9.51	-0.607	Nuclear
CAG023567.t1	<i>CaHSF-21</i>	2	3099	1033	116.85	8.16	-0.214	Nuclear
CAG029837.t1	<i>CaHSF-22</i>	2	2799	933	106.07	8.38	-0.139	Nuclear
CAG033200.t1	<i>CaHSF-23</i>	3	1089	363	41.54	4.93	-0.796	Nuclear
CAG037049.t1	<i>CaHSF-24</i>	3	1467	489	54.18	5.40	-0.745	Nuclear
CAG034205.t1	<i>CaHSF-25</i>	3	1467	489	54.26	5.46	-0.724	Nuclear
CAG050375.t1	<i>CaHSF-26</i>	5	1332	444	49.81	5.32	-0.668	Nuclear
CAG047099.t1	<i>CaHSF-27</i>	5	1332	444	49.80	5.45	-0.648	Nuclear
CAG057045.t1	<i>CaHSF-28</i>	6	933	311	3.55	9.20	-0.407	Nuclear
CAG056450.t1	<i>CaHSF-29</i>	6	1437	479	53.39	5.06	-0.507	Cytoplasmic
CAG052168.t1	<i>CaHSF-30</i>	6	1587	529	59.03	5.15	-0.550	Nuclear
CAG052925.t1	<i>CaHSF-31</i>	6	1277	409	46.38	5.74	-0.736	Nuclear
CAG057295.t1	<i>CaHSF-32</i>	6	1227	409	46.59	5.95	-0.719	Nuclear
CAG053848.t1	<i>CaHSF-33</i>	6	786	262	30.03	7.06	-0.801	Nuclear
CAG054123.t1	<i>CaHSF-34</i>	6	1215	405	46.37	8.69	-0.500	Nuclear
CAG058422.t1	<i>CaHSF-35</i>	6	1215	405	46.32	8.16	-0.510	Nuclear
CAG054134.t1	<i>CaHSF-36</i>	6	1215	405	46.31	8.50	-0.519	Nuclear
CAG054124.t1	<i>CaHSF-37</i>	6	1089	363	40.65	5.98	-0.564	Nuclear
CAG058423.t1	<i>CaHSF-38</i>	6	1158	386	43.06	6.00	-0.546	Nuclear
CAG054135.t1	<i>CaHSF-39</i>	6	789	263	28.59	4.87	-0.538	Nuclear
CAG060570.t1	<i>CaHSF-40</i>	7	915	305	33.69	5.11	-0.721	Nuclear
CAG063979.t1	<i>CaHSF-41</i>	7	912	304	33.62	5.20	-0.732	Nuclear
CAG067295.t1	<i>CaHSF-42</i>	8	1560	520	56.75	4.93	-0.651	Nuclear

Continue...

Table 1: Continuation.

Gene ID	Name	Chr	No. CDS (bp)	No. Amino acids (AA)	Molecular Weight (kDa)	pI	GRAVY	Protein Localization Prediction
CAG070348.t1	CaHSF-43	8	1560	520	56.77	4.89	-0.660	Nuclear
CAG069016.t1	CaHSF-44	8	1173	391	43.88	4.72	-0.603	Nuclear
CAG072132.t1	CaHSF-45	8	1170	390	43.87	4.72	-0.597	Nuclear
CAG069486.t1	CaHSF-46	8	1035	345	37.56	5.04	-0.636	Nuclear
CAG072589.t1	CaHSF-47	8	999	333	36.14	5.41	-0.505	Nuclear
CAG067601.t1	CaHSF-48	8	1878	626	72.44	6.72	-0.590	Nuclear, mitochondrial, cytoplasmic
CAG070582.t1	CaHSF-49	8	1683	561	63.67	6.52	-0.667	Nuclear
CAG067602.t1	CaHSF-50	8	1176	392	43.90	5.73	-0.636	Nuclear
CAG001204.t1	CaHSF-51	10	1134	378	42.45	7.79	-0.594	Nuclear
CAG004341.t1	CaHSF-52	10	1134	378	42.48	7.82	-0.619	Nuclear
CAG008266.t1	CaHSF-53	11	1149	383	43.46	4.94	-0.646	Nuclear
CAG011255.t1	CaHSF-54	11	1149	383	43.48	4.94	-0.631	Nuclear

Chr: chromosome; CDS: coding sequence; pI: isoelectric point; GRAVY: grand average of hydropathicity.

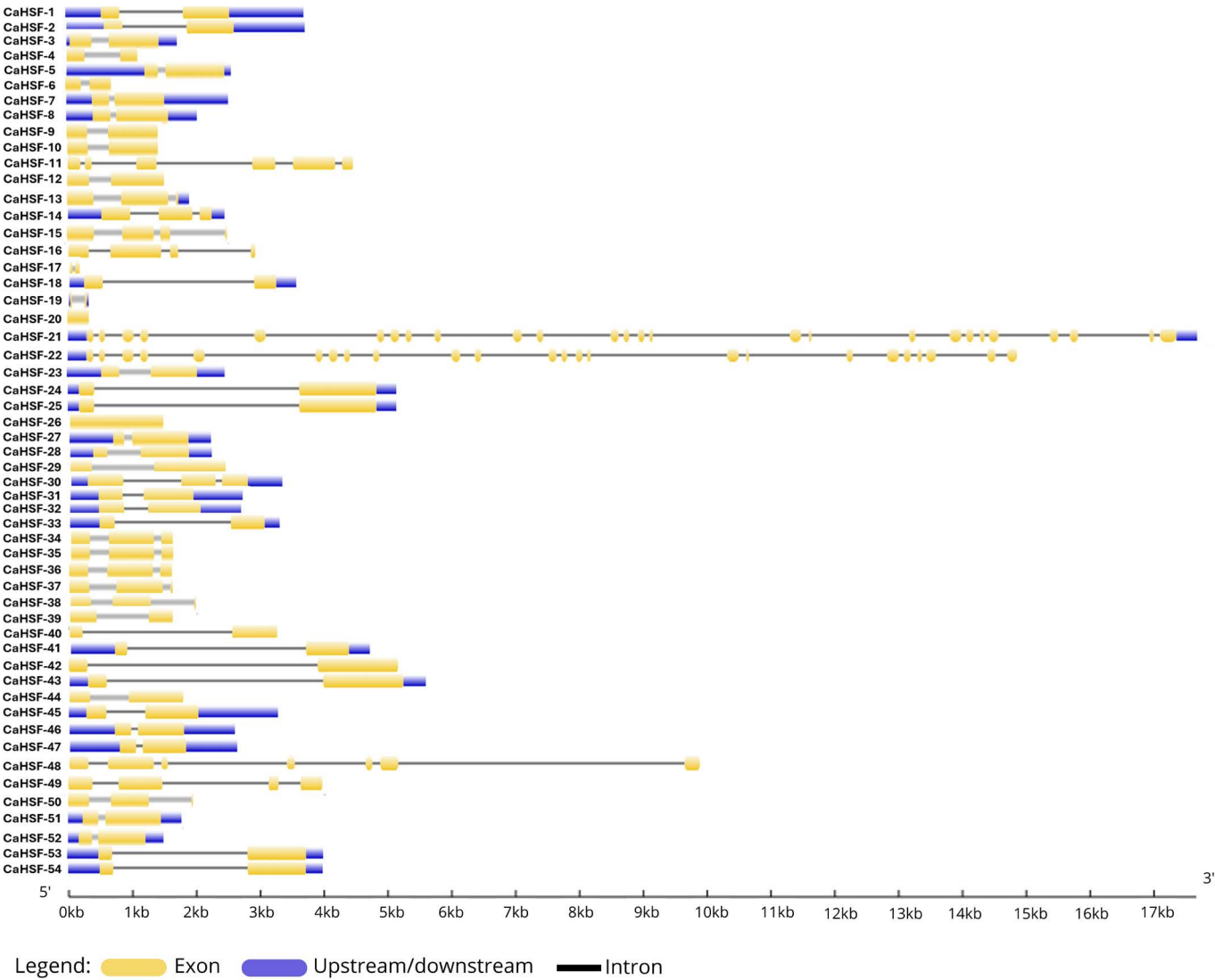


Figure 1: Gene Structure of HSF in *C. arabica*.

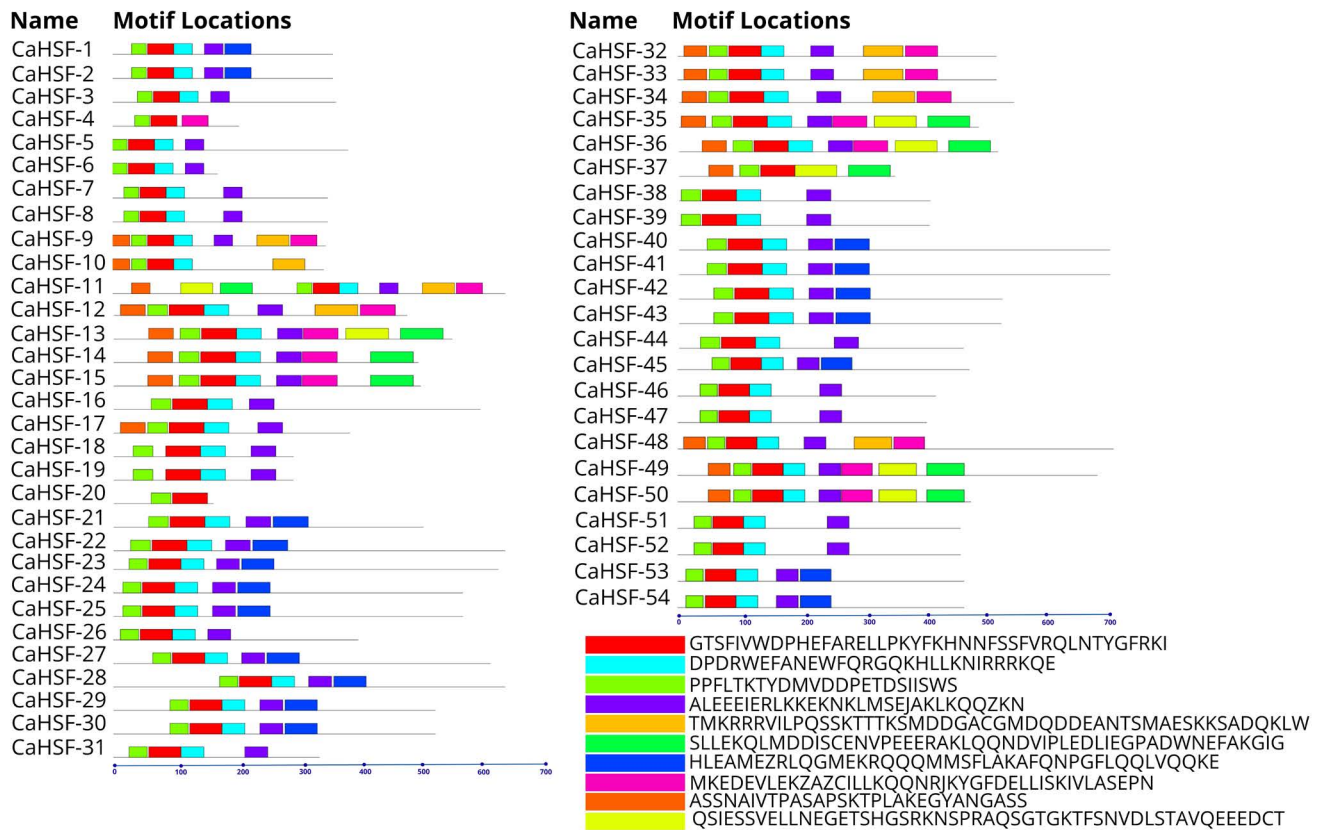


Figure 2: Motif of CaHSF.

3.3 Cis acting element analysis

Several cis-acting elements have been identified in *CaHSF* promoters that respond to heat stress (Li et al., 2023). Cis-acting elements that have been predicted to have a response to heat stress include MYB binding site (MBS), Low Temperature Response (LTR), Anaerobic Responsive Element (ARE), and ABA-responsive element (ABRE). The ARE and ABRE elements are found in almost all *CaHSF* promoters except *CaHSF-4*, *-7*, *-10*, *-28*, *-35*, *-38*, *-40*, and *-41* (Figure 3). ABRE is an element that regulates gene expression in response to abiotic stress. This element is activated by the hormone abscisic acid (ABA) that accumulates in plants when experiencing environmental stress conditions such as drought, thus it help plant tolerance to heat stress (Li et al., 2023). The LTR element is only found in some *CaHSF* promoters consisting of *CaHSF-1*, *-7*, *-8*, *-10*, *-11*, *-13*, *-15*, *-18*, *-19*, *-22*, *-24*, *-25*, *-29*, *-30*, *-36-40*, *-42-46*, and *-50-52*. LTR and MBS are also part of the plant stress response. MBS is a promoter element involved in drought induction. This element binds to the MYB domain when the plant is experiencing water deficit and activates the expression of genes associated with drought adaptation (Shariatipour & Heidari, 2020).

3.4 Sub Cellular Localization

Subcellular localization of proteins serves to predict the *CaHSF* gene family in *Arabica coffee*. HSF is in the nucleus

as a regulatory protein. The location and distribution pattern of cellular in *CaHSF* indicate dominant localization in the nucleus (Figure 4). This is in accordance with the function of HSF as a transcription factor that can regulate the expression of genes that will bind in the Heat Shock element (HSE) region of the Heat Shock Protein (HSP) to repair damaged proteins and restore the cell's condition to a normal state again. Some proteins have localization in the cytoplasm and mitochondria. Important domains involved in bringing proteins into the nucleus are NLS and NES. NLS will help the process of translocating HSF proteins into the nucleus when plants are under stress and NES will help the process of exporting proteins out of the nucleus and will be in the cytoplasm in an inactive state (Zhang et al., 2020). The selection of *CaHSF* for further analysis is *CaHSF-1*, *CaHSF-2*, *CaHSF-5*, *CaHSF-6*, *CaHSF-7*, *CaHSF-24*, *CaHSF-25*, *CaHSF-26*, *CaHSF-27*, *CaHSF-51* and *CaHSF-52*.

3.5 Phylogenetic Analysis and Protein Interaction

Phylogenetic analysis was used to determine the HSF protein classification of *C. arabica* against several other plants, such as *A. thaliana*, *O. sativa*, *Z. mays*, and *M. esculenta* (Figure 5). Phylogenetic results show that *CaHSF* is divided into 3 groups, namely HSFA, HSFB and HSFC. HSFA has the most members, which is 35 *CaHSF*. Meanwhile, HSFB has 13 members of the

CaHSF and HSFC has 6 members of the CaHSF. The bootstrap value in phylogenetic results varies, ranging from 13-100%. High bootstrap value indicates close evolutionary relationships between species. Some CaHSF groups are also in the same group with

other species, such as *A. thaliana* and *M. esculenta*. This shows the similarity of the sequence between the two. Phylogenetic analysis of CaHSF showed that the evolutionary distribution of *CaHSF-26* is similar to that of Manes.02G087400 (*MeHSF-5*)

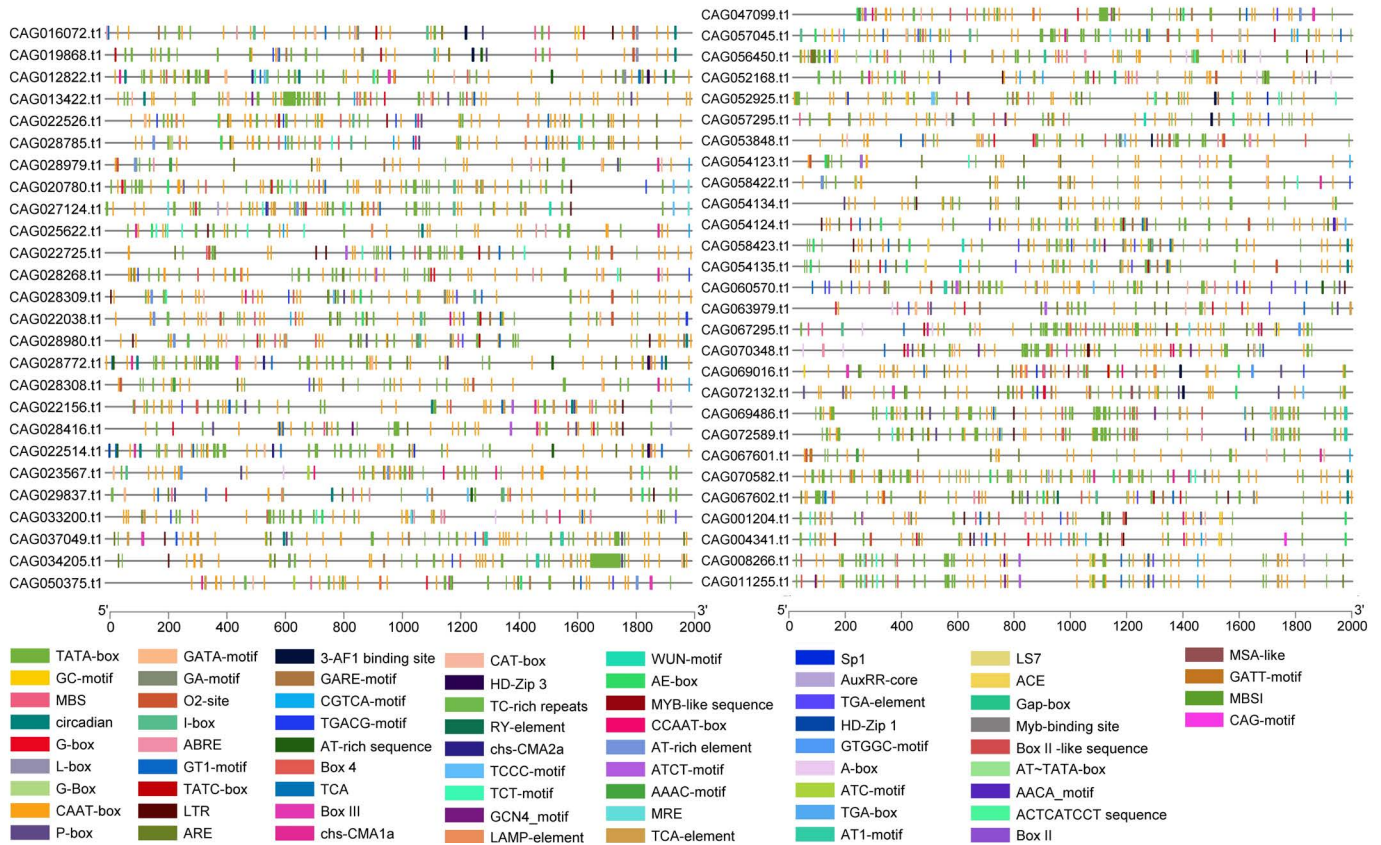


Figure 3: Cis acting element analysis of *CaHSF*.

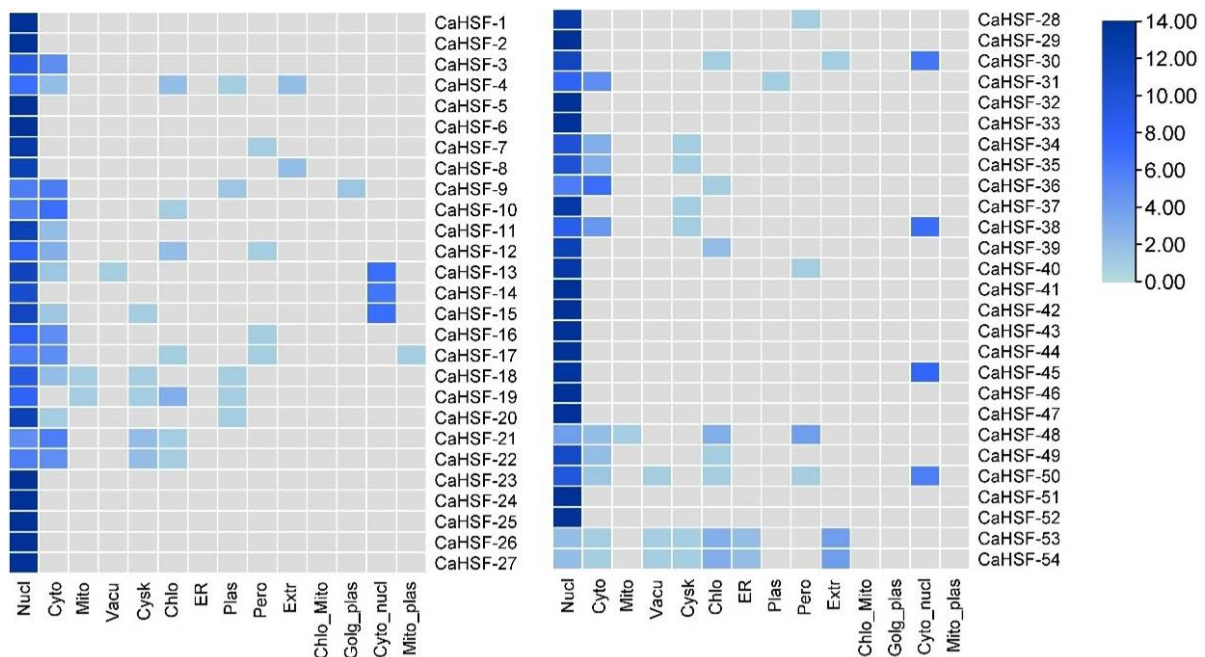


Figure 4: Subcellular localization.

which is known to have a sensitive response to drought (Zeng et al., 2020). *CaHSF-51* and *CaHSF-52* show similarities to *At1G46264* (*AtHSF-2*) which belongs to HSFB and is known to be associated with drought and heat stress (Guo et al., 2008).

Coffea arabica HSF proteins interactions were analyzed using STRING software (Figure 6). These proteins interact with each other, which is indicated by lines with different types of colors to represent these interactions. The HSF protein works in conjunction with other proteins in response to stress.

4 DISCUSSION

Coffea arabica is a large commodity crop whose global demand continues to increase every year, so that high-quality materials are needed. Environmental factors are important factors that can improve the quality of *C. arabica* seeds. Tolerance to extreme, such as drought, can increase *C. arabica* production to meet market demand and reduce pressure on agricultural production. Understanding the genes involved in plant stress in *C. arabica* is important to discuss and study as illustrated in our results.

Heat shock transcription factors play a major role in not only in plant development, but also in biotic and abiotic

stresses. The number of HSF genes in *C. arabica* is 54, which is much different from *A. thaliana*, which has 21 HSF genes. This number is relatively large, and the amount is almost similar to that of soybean and wheat plants. This can be related to selective pressure during plant evolution (Li et al., 2023).

The different structures of CaHSF affect the transcription process and gene expression. A smaller number of introns can speed up transcription and mRNA processing, allowing genes that respond to stress to be expressed more quickly (Jeffares et al., 2008). Some genes that do not have introns will be directly expressed without the splicing process, thereby accelerating the process of gene expression in response to stress.

Phylogenetic results in CaHSF show that each clade is closely related based on the grouping of closed-seeded plants, namely dicot in *C. arabica*, *A. thaliana*, and *M. esculenta* and monocots in *Z. mays* and *O. sativa*. The HSF grouping in *C. arabica* is divided into three groups, namely HSFA, HSFB and HSFC. HSFA has the most CaHSF members. The HSF grouping is based on the existence of the HRA/B domain. HSFA has an HRA/B domain while HSFB and HSFC do not have such domains. HSFA acts as an activator that can activate the transcription process. HSFB and HSFC play the role of regulator (Feng et al., 2021).

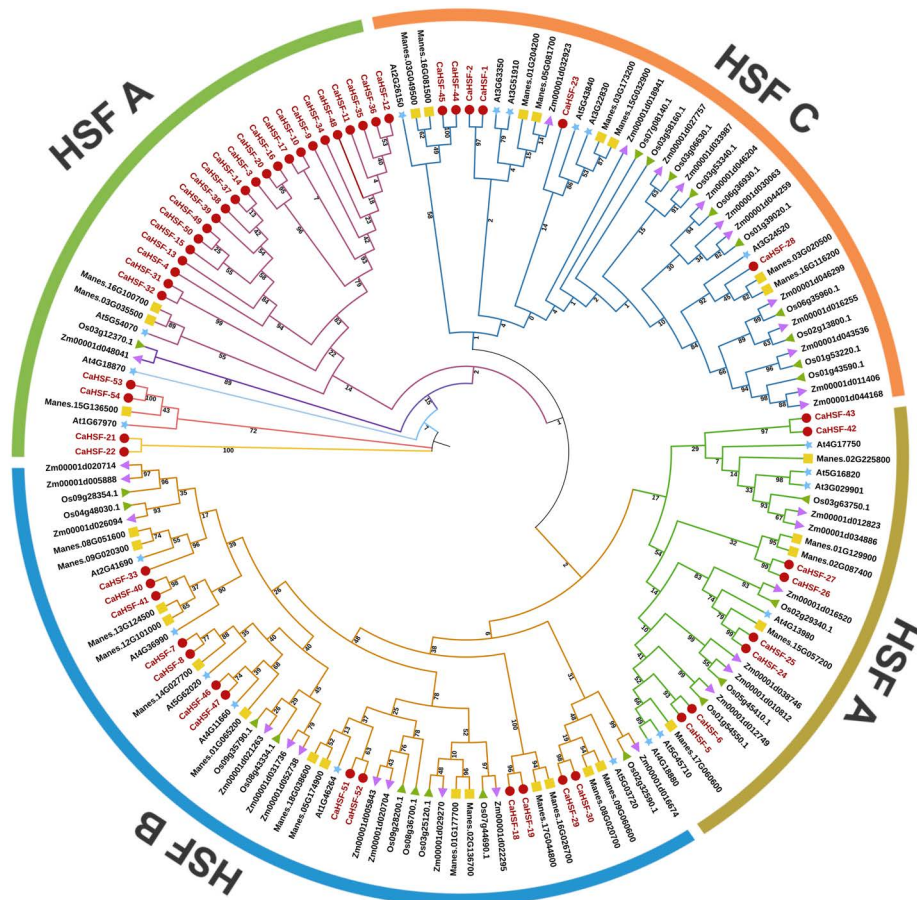


Figure 5: Phylogenetic HSF protein in *Coffea arabica* (CaHSF), *Arabidopsis thaliana* (At), *Manihot esculenta* (Manes), *Oryza sativa* (Os) and *Zea mays* (Zm).

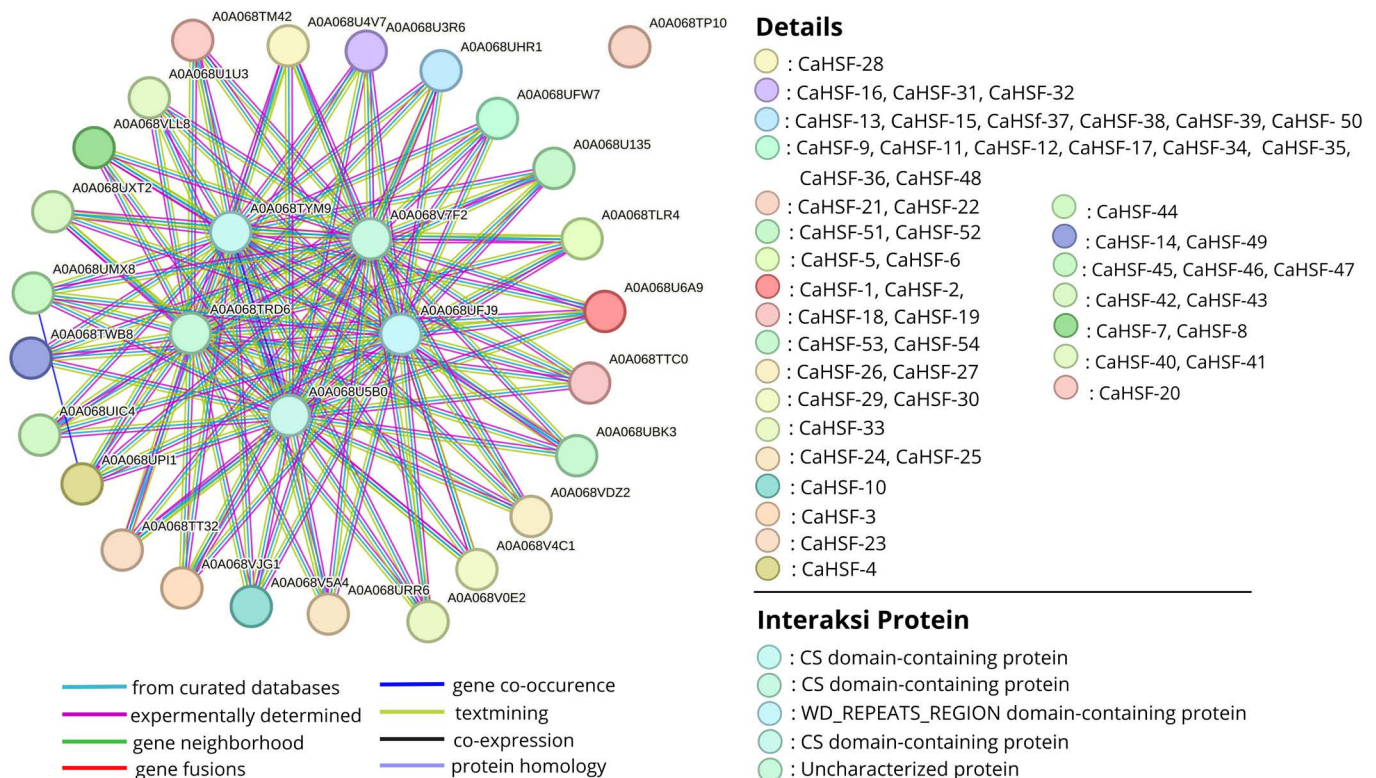


Figure 6: Protein Interaction of Heat shock factors (HSFs).

The HSF proteins *C. arabica* also have other interactions with the CS domain containing protein that functions in HSF dimerization and stabilization of protein structure before binding to the target promoter cis element. Activation of HSFs and its interaction with other proteins will help plants maintain cellular homeostasis during heat stress and drought stress (Guo et al., 2016; Driedonks et al., 2015). CaHSF proteins interactions are also involved with WD-repeats proteins that function not only in abiotic responses, but also in the regulation of plant growth and development. WD-repeat proteins are widely found in plants and play important roles in abiotic stress and physiological processes (Mu et al., 2025).

The presence of HSF plays a very important role in the regulation of plants under abiotic stress. HSF analysis has been performed on several plants such as tomatoes and paddy suggests that the *Solanum lycopersicum* HSF (*SlyHSF*) and *Oryza sativa* HSF (*Oshsf*) gene plays a role in the heat stress (Yang et al., 2024; Shamshad et al., 2023). In-silico analysis of HSF *Coffea* is needed as an analytical study for the selection of *CaHSF* gene candidates related to the adaptability of coffee to heat stress conditions so that it can be used as information for the development of new coffee cultivars that are resistant to heat stress. The selection of *CaHSF* gene candidates was carried out by selecting the 56 most potential *CaHSF* genes for further analysis. Gene selection includes the presence of DNA-binding Domain (DBD) motifs, the number of introns, the presence of elements in the promoter region, subcellular localization in

the nucleus and the results of phylogenetic analysis. Based on these criteria, there are three potential *CaHSF* gene candidates, namely *CaHSF-26*, *CaHSF-51* and *CaHSF-52*.

5 CONCLUSIONS

Heat shock factors (HSFs) are important transcription factors that regulate stress responses in plants. A total of 54 HSF genes were identified in *C. arabica* in this study. Most CaHSF genes' products are predicted to nuclear-localized resident, consistent with their function as regulators of the expression of genes encoding heat shock proteins (HSPs) and other stress-related proteins. HSFs also act as a regulator that activates the expression of protective genes when plants experience environmental stress.

6 AUTHOR CONTRIBUTION

Conceptual idea: Su'udi, M.; Methodology design: Rosida, W., Su'udi, M.; Data collection: Rosida, W., Su'udi, M.; Data analysis and interpretation: Rosida, W., Su'udi, M., Ulum, F.B.; Writing and editing: Rosida, W., Su'udi, M., Rohmah, D.I.

7 DATA AVAILABILITY STATEMENT

Data available upon request to authors.

8 REFERENCES

- Apriani, D. et al. (2022). Indonesian coffee at the international market. *Jurnal Paradigma Ekonomika*, 17(2):261-272.
- Chen, C. et al. (2023). TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Molecular Plant*, 16(11):1733-1742.
- Damatta, F. M. et al. (2008). Ecophysiology of coffee growth and production. *Brazilian Journal of Plant Physiology*, 19(4):485-510.
- Driedonks, N. et al. (2015). Multi-level interactions between heat shock factors, heat shock proteins, and the redox system regulate acclimation to heat. *Frontiers in Plant Science*, 6:999.
- Feng, N. et al. (2021). Structures of heat shock factor trimers bound to DNA. *ISCIENCE*, 24(9):102951.
- Gasteiger, E. et al. (2005). The Proteomics Protocols Handbook. The Proteomics Protocols Handbook, 571-608. <https://doi.org/10.1385/1592598900>
- Goodstein, D. M. et al. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1):1178-1186.
- Guo, J. et al. (2008). Genome wide analysis of heat shock transcription factor families in rice and Arabidopsis. *Journal of Genetics and Genomics*, 35(2):105-118.
- Guo, M. et al. (2016). The plant heat stress transcription factors (HSFs): Structure, regulation, and function in response to abiotic stresses. *Frontiers in Plant Science*, 7:114.
- Horton, P. et al. (2007). WoLF PSORT: Protein localization predictor. *Nucleic Acids Research*, 35:585-587.
- Hu, B. et al. (2015). GSDB 2.0: An upgraded gene feature visualization server. *Bioinformatics*, 31(8):1296-1297.
- Jagadish, S. V. K. et al. (2021). Plant heat stress: Concepts directing future research. *Plant Cell and Environment*, 44(7):1992-2005.
- Jaspard, E., & Hunault, G. (2014). Comparison of amino acids physico-chemical properties and usage of late embryogenesis abundant proteins, hydrophilins and WHY domain. *PLoS ONE*, 9(10):e109570.
- Jeffares, D. C. et al. (2008). Rapidly regulated genes are intron poor. *Trends in Genetics*, 24(8):375-378.
- Lescot, M. et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, 30(1):325-327.
- Li, L. et al. (2023). Genome-wide analysis of the heat shock transcription factor family reveals saline-alkali stress responses in *Xanthoceras sorbifolium*. *PeerJ*, 11:e15929.
- Mu, X. et al. (2025). Whole-genome identification of HSF family genes in *Cerasus humilis* and expression analysis under high-temperature stress. *Frontiers in Plant Science*, 16:1553187.
- Shamshad, A. et al. (2023). In-silico analysis of heat shock transcription factor (OsHSF) gene family in rice (*Oryza sativa* L.). *BMC Plant Biology*, 23:395.
- Shariatipour, N., & Heidari, B. (2020). Meta-analysis of expression of the stress tolerance associated genes and uncover their cis-regulatory elements in rice (*Oryza sativa* L.). *The Open Bioinformatics Journal*, 13(1):39-49.
- Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38(7):3022-3027.
- Thompson, J. D. et al. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22):4673-4680.
- Wang, N. et al. (2024). Characterization of the heat shock transcription factor family in *lycoris radiata* and its potential roles in response to abiotic stresses. *Plants*, 13(2):271.
- Yang, X. et al. (2024). Genome-Wide Identification and Expression Analysis of the HSF Gene Family in *Ammopiptanthus mongolicus*. *BMC Genomics*, 25(729):1-15.
- Yu, C.-S. et al. (2006). Prediction of protein subcellular localization. *PROTEINS: Structure, Function, and Bioinformatics*, 64:643-651.
- Yu, C. et al. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Science*, 13(5):1402-1406.
- Zeng, J. et al. (2020). Genomic analyses of heat stress transcription factors (HSFs) in simulated drought stress response and storage root deterioration after harvest in cassava. *Molecular Biology Reports*, 47:5997-6007.
- Zhang, X. et al. (2020). Genome-wide characterization of tea plant (*Camellia sinensis*) Hsf transcription factor family and role of CsHsfA2 in heat tolerance. *BMC Plant Biology*, 20:244.