Pelita Perkebunan 41(2) 2025, 73–85
DOI: 10.22302/iccri.jur.pelitaperkebunan.v41i2.656 ISSN: 0215-0212 / e-ISSN: 2406-9574

Identification of Heat Shock Protein 90 (HSP90) Gene Family in Cocoa (*Theobroma cacao* L.) in Silico Study

Mukhamad Su'udi[1)], Qori'atul Mustafidah[1)], and Sulistyani Pancaningtyas[2)]

[1)]Department of Biology, Faculty of Mathematics and Natural Science, University of Jember,
Jl. Kalimantan No. 37, Jember, Indonesia
[2)] Indonesian Coffee and Cocoa Research Institute (ICCRI), Kaliwining, Jember, Indonesia
[*)]Corresponding author: msuudi.fmipa@unej.ac.id
Received: March 1, 2025 / Accepted: July 6, 2025

Abstract

Heat Shock Protein 90 (HSP90) has an important role in regulating heat stress response, plant development control, and defense reactions. The HSP90 gene family has been extensively reported in numerous plant species, but studies on the HSP90 gene in cocoa (*Theobroma cacao* L.) remain lacking. Continuous global warming and the use of low-grade quality planting materials are the reasons for low cocoa productivity in Indonesia. Based on this, it is necessary to have good adaptability in cocoa as well as the development of new cultivars so that there is no decrease in cocoa productivity. This study aims to identify the HSP90 gene family in cocoa to recognize the HSP90 functioning in cocoa and provide the necessary information to characterize HSP90 at the genome level. Identification of the HSP90 gene family in *T. cacao* was carried out through analysis of *TcHSP90* sequences obtained from various related databases. The results showed that a total of seven *TsHSP90* genes are distributed into ten *T. cacao* chromosomes. Analysis of the *TcHSP90* motif showed that HSP90 is conserved in each *T. cacao* gene. *Cis-acting* element analysis showed that *TcHSP90* has important roles in growth and development as well as abiotic and hormonal stress responses including defense against pathogens. Phylogenetic analysis showed that *TcHSP90* has a similar evolutionary distribution to *A. thaliana* and *S. lycopersicum*. The results of this study can be used as a scientific foundation and basic knowledge for the development of new *T. cacao* cultivars that are resistant to biotic and abiotic stresses by selecting suitable candidate genes for *T. cacao* breeding.

Keywords: Abiotic stress, genome-wide study, Heat Shock Protein (HSP90), *Theobroma cacao* L.

INTRODUCTION

Plant growth and development are influenced by various factors including interrelated biotic and abiotic stresses. Abiotic stresses such as temperature, salinity, high light intensity, heavy metal toxicity and other abiotic stresses can cause stress conditions in plants. Global warming that has occurred in recent decades is the main factor that causes extreme weather. Extreme weather conditions will have a detrimental impact that can affect plant growth and development and reduce crop quality and quantity (Guihur *et al.,* 2022). Extreme weather due to prolonged global warming has a major impact on plants because it can inhibit starch synthesis, carbon assimilation, and cause a decrease in crop yields which then presents a new challenge to food security (Oldroyd & Leyser, 2020; Tigchelaar *et al.,* 2018). These conditions allow plants to develop adaptation to high temperatures through regulatory mechanisms to overcome heat stress by producing stress proteins, namely Heat Shock Protein (HSP) (Peng *et al.,* 2024).

HSP is one of the proteins that has an important role in regulating the response to heat stress and is highly conserved at the cellular and organismal levels (Appiah *et al.,* 2021). The mechanism of HSPs in stress response in plants works by helping to refold unfolded or misfolded proteins due to the disruption of cell homeostasis after experiencing stress (Bettaieb *et al.,* 2020). HSPs can be divided into five families based on molecular weight and sequence homogeneity, namely HSP20, HSP60, HSP70/DnaK, HSP90, and HSP100/ClpB (di Donato & Geisler, 2019; Wasilah *et al.,* 2019).

One of the most abundant proteins in the prokaryotic and eukaryotic cytoplasm is the HSP90 family which makes up 1-2% of the cellular protein level. HSP90 is generally composed of three structural domains namely N-terminal ATP-binding domain, M domain, and C-terminal substrate-binding domain (Chiosis *et al.,* 2023). HSP90 is directly or indirectly involved in several physiological processes ranging from plant growth and development to abiotic and biotic stress responses (Peng *et al.,* 2024). HSP90 will be abundantly expressed in the plant cytoplasm under normal physiological conditions but will accumulate rapidly in the nucleus under heat stress conditions (Appiah *et al.,* 2021). The function of HSP90 is to regulate and maintain the conformation of various proteins, assist normal cell survival under stress, and act as a negative feedback regulator of the heat stress response (Peng *et al.,* 2024). HSP90 along with other molecular chaperones will function in protein folding, prevent protein aggregation, and facilitate the folding of inactive proteins to increase resistance to certain cells. HSP90 expression is regulated when plants experience stress by joining nonprotein substances to allow the repair of damaged proteins (Li *et al.,* 2020). Protein damage can be caused by heat stress because it can cause Rubisco inactivation and increased oxygen affinity for Rubisco. This condition can have more harmful effects, especially in C3 plants such as soybean, coffee, and cocoa because it can reduce photosynthetic efficiency (Arachchige *et al.,* 2024).

Cocoa (*Theobroma cacao* L.) is one of the leading plantation commodities in Indonesia that has high potential and has a large contribution to the country's economy through export activities. The number of people interested in processed cocoa products has increased the demand for cocoa to 380.72 thousand tons of cocoa per year (Larasati *et al.,* 2022). Indonesia is the largest cocoa producer in Asia and ranks seventh among the world's largest cocoa producers (ICCO, 2025). The area of cocoa plantations in Indonesia is known to continue to decline every year, resulting in cocoa productivity also continuing to decline. The cocoa plantation area in Indonesia in 2020 was recorded at around 1,5 million hectares but this condition continued to decline to 1,39 million hectares in 2023 (BPS, 2023). Cocoa productivity will be indirectly affected by global warming. The producing Heat Shock Protein (HSP) in response toheat stress is a mechanism that this condition enables cocoa to develop into more adaptable, preventing a decline in output and ensuring that the demand for cocoa exports is continuously fulfilled.

HSP production in cocoa plants which are C3 plants has an important role especially in supporting the photosynthesis process by maintaining and protecting the proteins involved in it. Heat stress in plants will respond to the thylakoid membrane protein complex in chloroplasts by producing a large number of companion proteins to protect

PSII which is the most sensitive target to heat stress and other abiotic stresses (Chauhan *et al.,* 2023; Hu *et al.,* 2020). One of these companion proteins is HSP90C, which is a member of HSP90 that has a function in maintaining protein homeostasis and chloroplast protein transport (Mu *et al.,* 2024).

HSP90 has been identified and found in several plants including *Arabidopsis thaliana* with seven HSP90 genes (Krishna & Gloor, 2001), *Solanum lycopersicum* with six HSP90 genes (Liu *et al.,* 2014), *Oryza sativa* with nine HSP90 genes (Hu *et al.,* 2009), and *Zea mays* with eleven HSP90 genes (Magnard & Vergne, 1996). The identification of HSP90 in *Theobroma cacao* related to adaptability to extreme weather conditions due to global warming has not yet been reported. These extreme weather conditions will certainly affect the growth and development of cocoa plants so that it will have an impact on the productivity produced. Therefore, identification of the HSP90 gene family in *Theobroma cacao* is very important to understand the HSP90 family that works in *Theobroma cacao* and provides the necessary information to characterize HSP90 at the whole genome level of *T. cacao*.

MATERIALS AND METHODS

Cocoa genome database

This study used the HSP90 protein sequence from *Arabidopsis thaliana* (AtHSP90) obtained from the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/protein/). The protein sequences were then used for domain or motif analysis of AtHSP90 protein sequences using MOTIF tools (https://www.genome.jp/tools/motif/). The cocoa genome to be identified was obtained from the Phytozome database (Goodstein *et al.,* 2012; https://phytozome-next.jgi.doe.gov/) with annotation version v1.1 or *Theobroma cacao* v1.1.

Identification of the HSP90 gene family in *T. cacao*

The identification of the *HSP90* gene in *T. cacao* was carried out by performing BLASTP on the *AtHSP90* motif against the *T. cacao* genome in the Phytozome database (Goodstein *et al.,* 2012; https://phytozome-next.jgi.doe.gov/). Data extraction of *T. cacao* HSP90 was carried out by searching for information including transcript ID, chromosome number, chromosome location data, strand, CDS (coding sequence) base pair length and protein length (AA), and phytozome annotation. The *T. cacao* HSP90 chromosome number was obtained from the NCBI protein database and then used to determine the name of the HSP90 gene in *T. cacao* based on the sequenced chromosome number. Based on the sequence of chromosome numbers, seven HSP90 genes in *T. cacao* were obtained, so they were named *TcHSP90-1* to *TcHSP90-7*. Data extraction was also carried out on the genomic sequence, CDS sequence, and peptide sequence of *TcHSP90-1* to *TcHSP90-7* in the Phytozome database. Analysis of physicochemical properties including molecular weight (kDa), isoelectric point (pI), and Grand Average of Hydropathicity (GRAVY) was carried out through the Expasy ProtParam tool software (Gasteiger *et al.,* 2005; https://web.expasy. org/protparam/) using the obtained *T. cacao* HSP90 protein sequence.

Analysis of gene structure

The exon-intron structures of *TcHSP90-1* to *TcHSP90-7* were analyzed through genomic sequence alignment with the corresponding CDS and genomic sequences using Genes Structure Display Server 2.0 software (Hu *et al.*, 2015;https://gsds.gao-lab.org/).

Promoter analysis of HSP90 gene families

Analysis of *cis-acting* elements from the promoter region of the *TcHSP90-1* to *TcHSP90-7* genes was carried out by searching 2000 bp genomic sequence before the start codon from the Phytozome database. The data obtained were then analyzed using the plantCARE software (Lescot *et al.,* 2002; https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) and visualized with TBtools-II software (Chen *et al.,* 2023)

Sub-cellular localization

Sub-cellular localization of proteins was predicted 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 the heat-map illustration were visualized with TBtools-II software (Chen *et al.,* 2023).

Phylogenetic tree

Analysis of the evolutionary relationships of HSP90 members of *T. cacao* was carried out through protein sequence alignment with *A. thaliana*, *S. lycopersicum*, *Z. mays*, and *O. sativa*. Sequence alignment was performed using ClustalW (Thompson *et al.,* 1994) with MEGA11 software (Tamura *et al.,* 2021) while the phylogenetic tree was constructed using the maximum likelihood method with partial deletion parameters and the bootstrap method with 1000 bootstrap replications.

RESULT AND DISCUSSION

Identification of *HSP90* genes in *Theobroma cacao*

BLAST results between seven *AtHSP90* genes and the *T. cacao* genome for the identification of HSP90 proteins in *T. cacao*

Table 1. List of Heat Shock Protein 90 gene family in *Theobroma cacao* L.

Transcript ID	Gene Name	Chromo-some	Location	Strand	CDS	AA	Molecular weight (kDa)	pI	GRAVY	Exon; Intron	Protein Localization Prediction	Phytozome Annotations
Thecc1EG002328t1	*TcHSP90-1*	1	13190310..13196723	reverse	2445	815	92.031	5.49	-0.530	20;19	cytoplasmic (1.589), nuclear (1.405)	Heat shock protein 90 // Heat shock protein 89.1
Thecc1EG011421t1	*TcHSP90-2*	2	40043872..40050867	reverse	2370	790	90.027	4.99	-0.564	19;18	cytoplasmic (2.670)	Heat shock protein 90 // Heat shock protein 89.1
Thecc1EG015252t1	*TcHSP90-3*	3	27005587..27009274	forward	2112	704	80.596	5.00	-0.610	4;3	cytoplasmic (3.936)	Molecular chaperone HtpG (htpG, HSP90A)
Thecc1EG016330t1	*TcHSP90-4*	3	32396252..32399517	forward	2118	706	81.166	4.99	-0.571	4;3	cytoplasmic (4.053)	Molecular chaperone HtpG (htpG, HSP90A)
Thecc1EG030579t1	*TcHSP90-5*	7	398958..402967	forward	2100	700	79.952	5.01	-0.583	3;2	cytoplasmic (3.812)	Heat shock protein 90 // Subfamily not named
Thecc1EG034239t1	*TcHSP90-6*	8	2054242..2063021	reverse	1878	626	69.739	4.85	-0,324	9;8	ER (2.327), cytoplasmic (2.085)	Heat shock protein 90
Thecc1EG034246t1	*TcHSP90-7*	8	2090294..2097175	reverse	2523	841	95.408	4.84	-0.687	15;14	ER (4.002)	Heat shock protein 90

showed that there were seven *HSP90* genes that were successfully selected from the *T. cacao* genome. HSP90 genes in *T. cacao* (*TcHSP90*) were renamed according to the order of chromosomal location, resulting in genes *TcHSP90-1* to *TcHSP90-7* (Table 1).

Data extraction of TcHSP90 showed that of the 10 chromosomes in the *T. cacao* genome, two genes were located on chromosomes 3 and 8, one gene on chromosomes 1, 2, and 7 and no genes were located on chromosomes 4, 5, 6, 9, and 10. The seven TcHSP90 genes have varying sequence lengths with Coding Sequence (CDS) ranging from 1878 bp (TcHSP90-6) to 2523 bp (TcHSP90-7) while amino acids range from 626 aa (TcHSP90-6) to 841 aa (TcHSP90-7).

The molecular weight of the seven TcHSP90s ranged from 69,739 kDa (TcHSP90-6) to 95,408 kDa (TcHSP90-7). The isoelectric point (pI) of the seven TcHSP90s showed values varying from 4.84 (TcHSP90-7) to 5.49 (TcHSP90-1). These pI values indicate that all HSP90 proteins of *T. cacao* are acidic ($pI < 7.0$) and these results are similar to those found in *A. thaliana*, tomato and some other plants (Sajad *et al.,* 2022). The Grand Average of Hydropathicity (GRAVY) also showed values varying between -0.687 (TcHSP90-7) to -0.324 (TcHSP90-6). The GRAVY values obtained indicate that all *T. cacao* HSP90 proteins are negative, so the *T. cacao* HSP90 proteins are hydrophilic (Zhang *et al.,* 2021). Protein localization prediction shows that most of the TcHSP90 are located in cytoplasmic with TcHSP90-1 located in cytoplasmic and nuclear, TcHSP90-2 to TcHSP90-4 and TcHSP90-6 located in cytoplasmic, TcHSP90-5 located in cytoplas-mic and ER (Endoplasmic Reticulum), and TcHSP90-7 located in ER (Endoplasmic Reticulum). This is in accordance with the research of Appiah *et al.,* 2021 which states that HSP90 is expressed in the cytoplasm and has an important role in regulating the response to heat stress.

Motif and gene structure analysis of TcHSP90

Based on motif analysis of seven HSP90 in *T. cacao*, TcHSP90-1 to TcHSP90-7 contain histidine kinase-like ATPase (HATPase_c) domain. The HATPase_c domain is an evolutionarily conserved protein domain found in several ATP-binding proteins including histidine kinase, DNA gyrase B, topoisomerase (Bellon *et al.,* 2004), HSP90 molecular chaperones (Immormino *et al.*, 2004), phytochrome-like ATPase and DNA mismatch repair protein (Bettaieb *et al.,* 2020).

The longest HSP90 motif is found in TcHSP90-7 with a length range from 252 to 835 aa while the shortest HSP90 motif is found in TcHSP90-6 with a length range from 266 to 525 (Figure 1A). Zhang *et al.,* 2021 stated that the protein sequence containing HSP90 and the domain structure of HATPase_c were identified as HSP90. HATPase_c together with HSP90 is known to have an important role in ATP binding and hydrolysis and as a binding site for ATP/ADP with ATPase activity (Bettaieb *et al.,* 2020).

Gene structure was used to further investigate the structural characteristics of the *HSP90* gene family in *T. cacao*. Differences in gene structure between HSP90 groups are related to differences in gene function in subcellular compartments (Bettaieb *et al.,* 2020). Based on seven *TcHSP90* genes (Figure 1B), it shows that from *TcHSP90-1* to *TcHSP90-7* have different exon-intron distribution patterns. The largest genome size is found in *TcHSP90-6* with a length of around 8.8 kb while the smallest genome size is found in *TcHSP90-4* with a length of around 3.3 kb with an average genome size between the seven *TcHSP90* genes are 5.76 kb. The highest number of exons is found in *TcHSP90-1* with 20 exons while the least is found in *TcHSP90-5* which only has 3 exons. Additionally, the highest

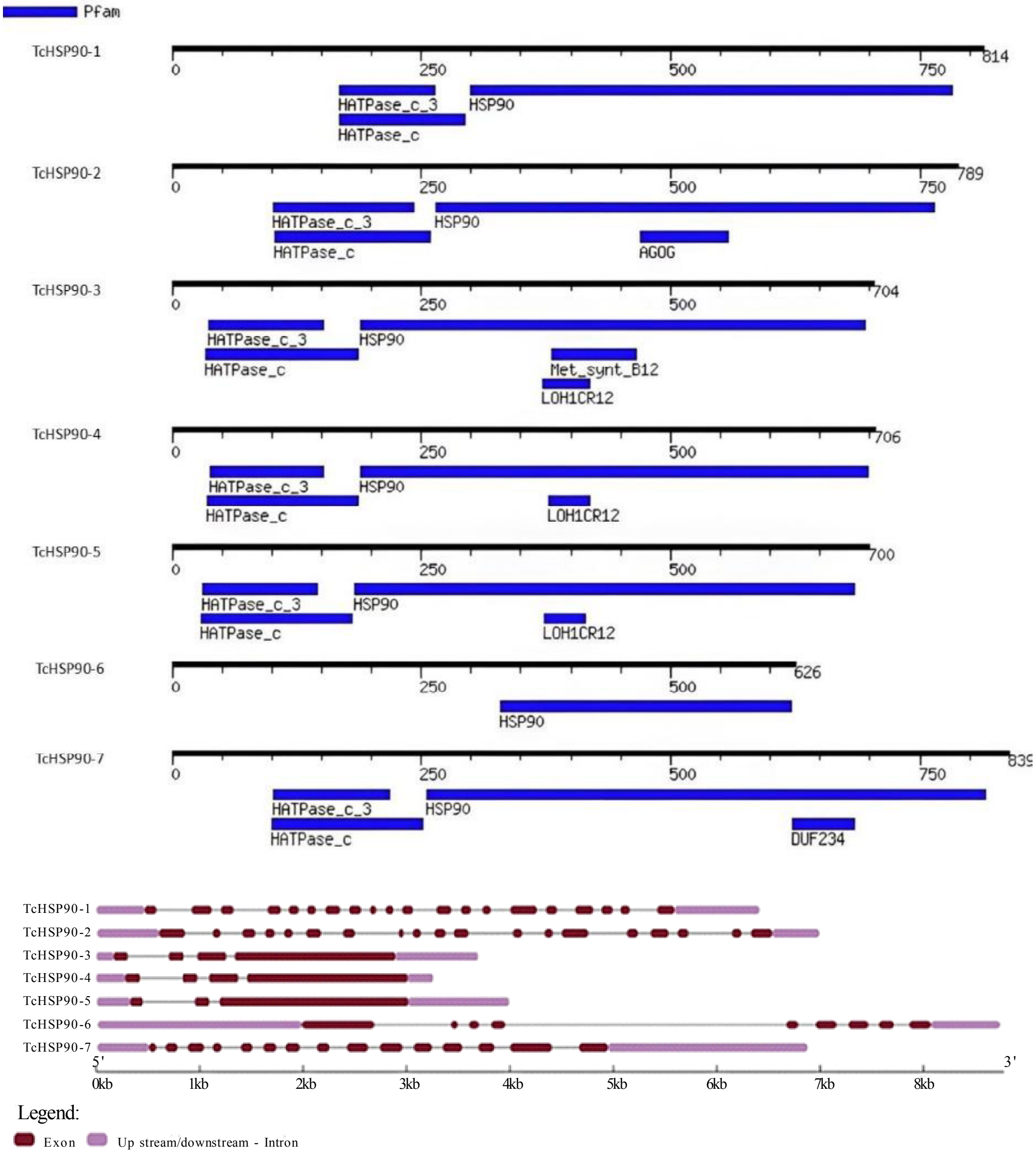

Figure 1. Motif and gene structure of *TcHSP90*

number of introns is found in *TcHSP90-1* with 19 introns while the least is found in *TcHSP90-5* which only has 2 introns. *TcHSP90-6* has the longest gene structure due to the length of introns although it does not have the highest number of introns. The gene structure also shows that all *TcHSP90* genes have an upstream that leads to the 5' end and a downstream that leads to the 3' end.

Analysis of the gene structure showing the exon-intron structure can be important evidence of the evolution of a gene (Wang *et al.,* 2022). Analysis of the exon-intron structure of the *HSP90* sequence from *T. cacao* will show information about the evolution of the *TcHSP90* gene family. The number of introns is largely related to the sensitivity of gene transcription regulation, so the fewer the number of introns, the more likely the plant will have the ability to respond to various environmental stimuli and developmental processes (Appiah *et al.,* 2021; Sajad *et al.,* 2022).

Cis-acting element analysis in the promoter of *TcHSP90* genes

Analysis of *cis-acting* elements in the promoter region was carried out to find out more about the regulation of *TcHSP90* gene expression as a reaction to stress conditions under abiotic pressure (Peng *et al.,* 2024). Based on this analysis, there are about 35 types of *cis-acting* elements that can be identified in the promoter sequence (Figure 2.). Some of the *cis-acting* elements found in *TcHSP90* are ACE, G-box, MRE, ATCT-motif, Box 4, GATA-motif, GT1-motif which are known to be responsive to light (Bettaieb *et al.,* 2020; Zhang *et al.,* 2021). The ACE *cis* element in the *TcHSP90* gene is only present in *TcHSP90-1* with locations around 820 bp and 1,320 bp before the start codon ATG. *Cis* element ATCT-motif and GATA-motif are also only found in one *TcHSP90* gene, namely in *TcHSP90-7* with the location of ATCT-motif around 1,180 bp and GATA-motif around 1,480 bp before the start codon. In addition, there is also Box 4 which is spread throughout the *TcHSP90* genes (Figure 2.).

Another *cis-acting* element found in *TcHSP90* is the LTR which is associated with responsiveness to drought, low temperature, and salt stress (Bettaieb *et al.,* 2020; Xue *et al.,* 2023). The gene associated with this response is found in *TcHSP90-2*, *TcHSP90-4*, and *TcHSP90-5* with the location being around 580 bp before the start codon in *TcHSP90-4*. The *TcHSP90* promoter also contains anaerobic responsive elements (ARE) which are scattered throughout the *TcHSP90* genes and are known to have an important role in responding to oxygen limitation. Other *cis-acting* elements found in *TcHSP90* are MYB Binding Site (MBS) involved in drought induction and TC-rich repeats involved in defense and stress responsiveness (Bettaieb *et al.,* 2020; Xue *et al.,* 2023).

Some *TcHSP90* promoters also contain elements induced by phytohormones, such as GARE-motif and P-box that are responsive to gibberellin, AuxRR-core and TGA-element that are responsive to auxin, and ABRE that is responsive to abscisic acid. In addition, there are CGTCA-motifs and TGACG-motifs that are responsive to Me-JA and TCA-elements that are responsive to salicylic acid (Bettaieb *et al.,* 2020). Other elements such as O2-site and GCN4-motif are known to be found only on specific promoters, suggesting that the corresponding genes will be controlled only under certain conditions (Xue *et al.,* 2023).

The analysis of the *cis-acting* element showed that abiotic stress and hormone regulation can regulate the expression of the *TcHSP90* genes. A total of 38 cis-acting elements that are randomly distributed in the *TcHSP90* genes have important roles in growth and development as well as abiotic and hormonal stress responses, thus showing the complex function of *TcHSP90*. Abiotic stress will increase the expression of the *TcHSP90* gene to maintain protein stability and function, thereby increasing its tolerance to environmental stress conditions. The HSP90 mechanism play a role in regulating and maintaining the conformation of various proteins and binds with non-protein substances to enable the repair of proteins damaged by stress (Li *et al.*, 2020). Abiotic stress will be related to the function of hormones in plants such as auxin, gibberellin, methyl-jasmonic, and abscisic acid. *Cis acting* ABRE related to abscisic acid responsiveness has been identified in *A. thaliana* and showed its function in the regulation of lignin deposition and

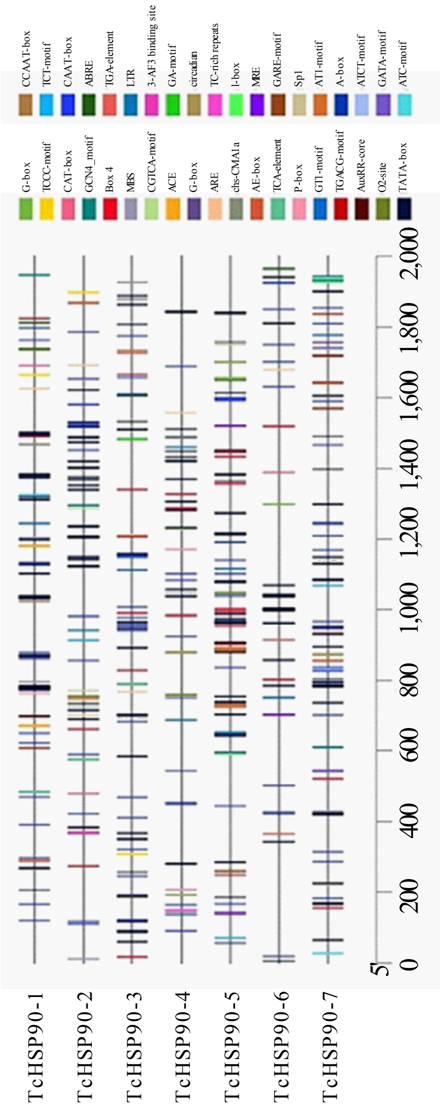

Figure 2. Cis-acting element of *TcHSP90*

secondary cell wall formation through phosphorylation of wall NST1 (Liu *et al.,* 2021). Auxin and gibberellin identified in cotton fiber cells function as promoters of cell elongation while jasmonic acid functions as a response to plant pathogen defense. The various *cis-acting* types found in the *TcHSP90* gene can show different and specific expression patterns in various tissues including the function and regulation of the genes contained therein. The gene expression pattern is involved in the growth of *T. cacao* and the development of various tissues and organs (Bettaieb *et al.,* 2020).

Subcellular localization of TcHSP90

The subcellular localization of proteins is used to predict TcHSP90 members in different cellular compartments indicating the involvement and function of each gene in various subcellular processes (Bettaieb *et al.,* 2020; Zhang *et al.,* 2021). The expression of ABA-responsive genes is known to be inhibited by overexpression of *AtHSP90-2* in the cytosol but not affected by overexpression of *AtHSP90-5* and *AtHSP90-7* under drought and salt stress conditions. This suggests that the cellular compartment that localizes *Arabidopsis thaliana* HSP90 protein respond to abiotic stresses with different functional mechanisms (Appiah *et al.,* 2021). The subcellular localization of HSP90 proteins in *T. cacao* (Figure 3.) showed that six of the seven proteins were predicted to be localized in the nucleus and none of the proteins were predicted to be localized in the cytoskeleton.

From the various predictions of HSP90 protein localization in *T. cacao*, it shows that the cytoplasm has the highest prediction value. This is in accordance with the research of Appiah *et al.,* 2021, that HSP90 is expressed in the cytoplasm. The highest predicted subcellular localization of *TcHSP90* protein is also in the mitochondria, namely in *TcHSP90-1*, chloroplast in *TcHSP90-2*, and cytoplasm

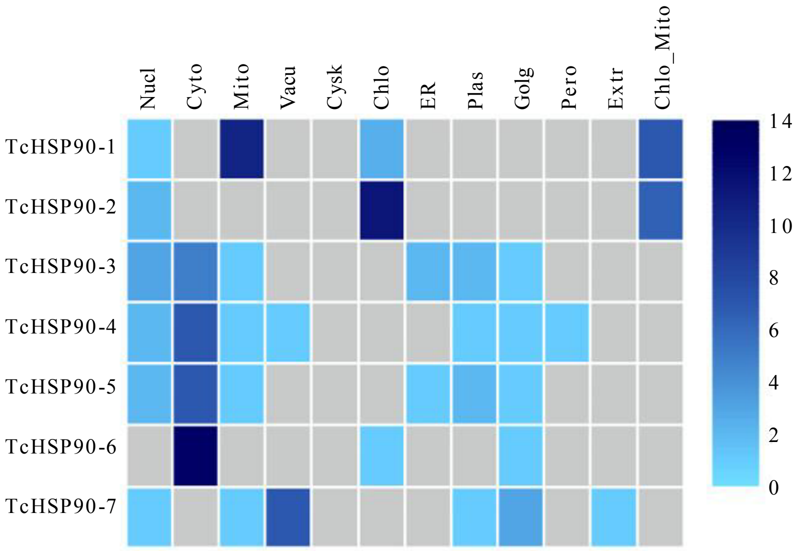

Figure 3. Heat map of *TcHSP90*

in *TcHSP90-6*. Between *TcHSP90-1* in the mitochondria, *TcHSP90-2* in the chloroplast, and *TcHSP90-6* in the cytoplasm shows that the highest subcellular localization is in *TcHSP90-6* in the cytoplasm. This was also shown in *A. thaliana* (Sarkar *et al.,* 2009) and *Brachypodium distachyon* (Zhang *et al.,* 2017) and indicates that the cytoplasm is the site of protein assembly which may be the main site of HSP90 protein activity (Bettaieb *et al.,* 2020).

Phylogenetic analysis

Phylogenetic analysis is used to obtain information on the evolutionary relationships of species and to help identify orthologs between species and paralogs within species. Construction of a phylogenetic tree on *T. cacao* HSP90 (Figure 4.) is aimed at determining the evolutionary relationships and classification of *T. cacao* HSP90 proteins. *T. cacao* HSP90 phylogenetic tree construction was formed through the alignment of HSP90 protein sequences from several plants such as *A. thaliana* with 7 sequences, *Solanum lycopersicum* with 6 sequences, *Oryza sativa* with 9 sequences, and *Zea mays* with 11 sequences. Based on the results of the phylogenetic tree construction obtained, 6 groups were formed which showed that group I was the largest group with 13 genes consisting of 4 genes from *Z. mays*, 3 genes from *O. sativa*, 2 genes from *S. lycopersicum*, 2 genes from *A. thaliana*, and 2 genes from *T. cacao* while group 5 was the smallest group because it only consisted of 2 genes from *Z. mays*. The phylogenetic tree construction shows that each clade has a close relationship based on the grouping of closed seed plants, namely dicots in *T. cacao*, *A. thaliana*, and *S. lycopersicum* and monocots in *Z. mays* and *O. sativa*. Different HSP90 protein families have different biophysical properties, suggesting that there is diversity among HSP90 members for further study.

The results of the phylogenetic tree construction showed that there were 3 orthologous genes among species, namely *TcHSP90-2* and *AtHSP90-5*, *TcHSP90-1* with *AtHSP90-6* and Solyc07g047790.2.1 (*SlHSP90-4*) and *TcHSP90-5* and Solyc12g 015880.1.1 (*SlHSP90-7*). The phylogenetic tree construction results also showed the presence of 9 pairs of paralogs within species,

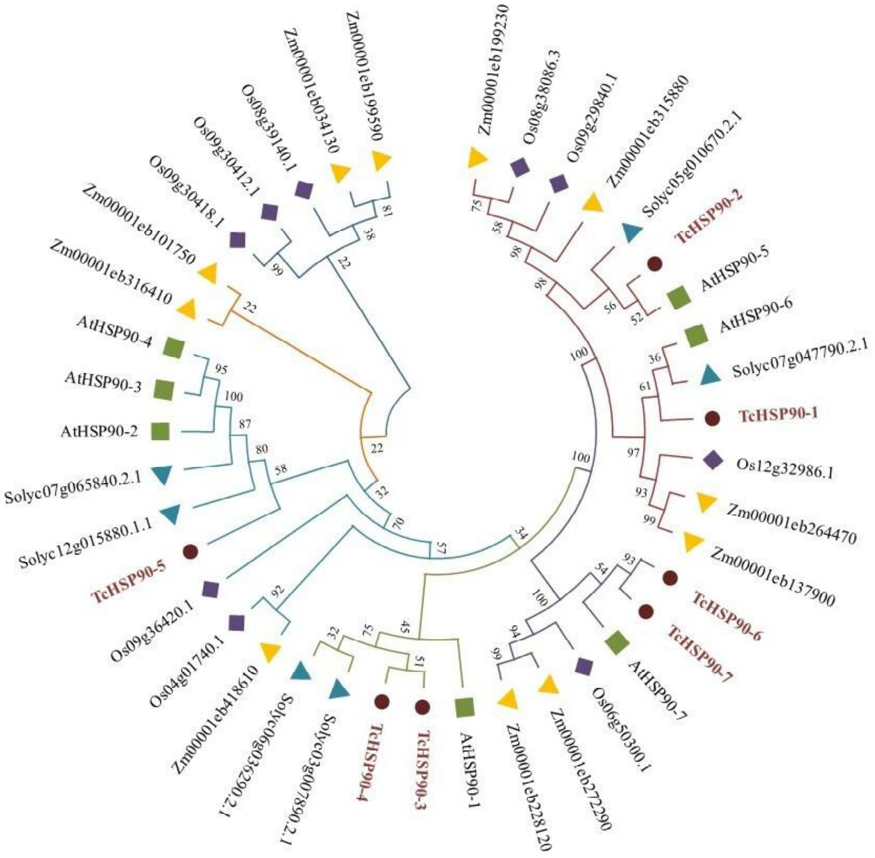

Figure 4. Phylogenetic tree of the HSP90 protein from *T. cacao, A. thaliana, S. lycopersicum, Z. mays,* and *O. sativa*

one of them from *A. thaliana* (*AtHSP90-3* and *AtHSP90-4*), 2 pairs from *T. cacao* (*TcHSP90-6* and *TcHSP907*, *TcHSP90-3* and *TcHSP90-4*), 2 pairs from *S. lycopersicum* (Solyc03g 007890.2.1 and Solyc06g036290. 2.1, Solyc07g 065840.2.1 and Solyc12g 015880.1.1), and 4 pairs from *Z. mays*. The nine pairs of paralogous genes in the phylogenetic tree construction indicated that most species evolved according to their own species-specific approaches during the evolution of the HSP90 gene family (Song *et al.,* 2019). The analysis of orthologous and paralogous genes in the phylogenetic tree can be used to predict the characteristics and functions of *TcHSP90* genes in evolutionary relationships with *A. thaliana*, *S. lycopersicum*, *Z. mays*, and *O. sativa* (Appiah *et al.,* 2021).

Based on the analysis of genes that are orthologous to *TcHSP90*, the evolutionary distribution of *TcHSP90-2* is similar to that of *AtHSP90-5*, indicating a degree of sensitivity to drought and salt stress (Song *et al.,* 2009). *TcHSP90-1* has an evolutionary distribution similar to *AtHSP90-6* and Solyc07g047790.2.1 (SlHSP90-4) which are associated with plant development (Luo *et al.,* 2019; Zai *et al.,* 2015). *TcHSP90-5* has a similar evolutionary distribution to Solyc12g 015880.1.1 (SlHSP90-7) which is also involved in plant development (Zai *et al.,* 2015). Based on the analysis of paralogous genes in *TcHSP90*, *TcHSP90-6* and *TcHSP90-7* are in a clade with *AtHSP90-7*, which shows a level of sensitivity to drought and salt stress (Song *et al.,* 2009). Other paralogs in *TcHSP90* are found in *TcHSP90-3* and *TcHSP90-4* which are in one clade with Solyc03g 007890.2.1 (*SlHSP90-1*) and Solyc06g 036290.2.1 (*SlHSP90-3*) and *AtHSP90-1* which are related to plant development and osmotic stress responses from ABA signaling (Zai *et al.,* 2015; Zhang *et al.,* 2017).

CONCLUSIONS

The identification and characterization of seven *HSP90* gene families found in *Theobroma cacao* were successfully analysed related to the motif and structure of the *TcHSP90* genes, the *cis* element in the promoter, the prediction of subcellular localization, and the phylogenetic tree. Different HSP90 protein families have different biophysical properties, indicating diversity among HSP90 members. Analysis of the expression pattern of *T. cacao* HSP90 family genes showed that *TcHSP90-1*, *TcHSP90-3*, and *TcHSP90-5* were expressed in response to biotic and abiotic stresses, which means that the three genes have an important role in *T. cacao* development. Analysis of the expression patterns of *TcHSP90-2*, *TcHSP90-4*, *TcHSP90-6*, and *TcHSP90-7* was related to the response to drought and salt stress. This study not only provides a scientific basis for HSP90 in *T. cacao* but can also be used as basic knowledge for the development of new *T. cacao* cultivars that are resistant to biotic and abiotic stresses and can help in the selection of candidate genes for *T. cacao* breeding.

REFERENCES

Arachchige, S.M.; A. Razzaq; H. Dai & J. Wang (2024). Confronting heat stress in crops amid global warming: impacts, defense mechanisms, and strategies for enhancing thermotolerance. *Crop Breed Genet Genom*, 6.

Appiah, C.; Z.F. Yang; J. He; Y. Wang; J. Zhou; W.Z. Xu; G. Nie & Y.Q. Zhu (2021). Genome-wide identification of Hsp90 gene family in perennial ryegrass and expression analysis under various abiotic stresses. *Plants*, 10, 1–14.

Bellon, S.; J.D. Parsons; Y. Wei; K. Hayakawa; L.L. Swenson; P.S. Charifson; J.A. Lippke; R. Aldape & C.H. Gross (2004). Crystal structures of *Escherichia coli* Topoisomerase IV ParE subunit (24 and 43 Kilodaltons): A single residue dictates differences in novobiocin potency against Topoisomerase IV and DNA Gyrase. *Antimicrobial Agents and Chemotherapy*, 48, 1856–1864.

Bettaieb, I.; J. Hamdi & D. Bouktila (2020). Genome-wide analysis of HSP90 gene family in the Mediterranean olive (*Olea europaea* subsp. europaea) provides insight into structural patterns, evolution and functional diversity. *Physiology and Molecular Biology of Plants*, 26, 2301–2318.

BPS (2023). Luas Perkebunan Kakao di Indonesia. Badan Pusat Statistik Indonesia. Jakarta, Indonesia.

Chauhan, J.; M.D. Prathibha; P. Singh; P. Choyal; U.N. Mishra; D. Saha; R. Kumar; H. Anuragi; S. Pandey; B. Bose; B. Mehta; P. Dey; K.K. Dwivedi; N.K. Gupta & R.K. Singhal (2023). Plant photosynthesis under abiotic stresses: damages, adaptive, and signaling mechanisms. *In Plant Stress* (Vol. 10). *Elsevier B.V.* 10, 1–13.

Chen, C.; Y. Wu; J. Li; X. Wang; Z. Zeng; J. Xu; Y. Liu; J. Feng; H. Chen; Y. He & R. Xia (2023). TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Molecular Plant*, 16, 1733–1742.

Chiosis, G.; C.S. Digwal; J.B. Trepel & L. Neckers (2023). Structural and functional complexity of HSP90 in cellular homeostasis and disease. In *Nature Reviews Molecular Cell Biology*, 24, 797–815.

di Donato, M. & M. Geisler (2019). HSP90 and co-chaperones: a multitaskers’ view on plant hormone biology. In *FEBS Letters*, 593, 1415M. 1430.

Gasteiger, E.; C. Hoogland; A. Gattiker; S. Duvaud; M.R. Wilkins; R.D. Appel & A. Bairoch (2005). *Protein Analysis Tools on the ExPASy Server 571 571 From: The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExPASy Server*. http://www.expasy.org/tools/.

Goodstein, D.M.; S. Shu; R. Howson; R. Neupane; R.D. Hayes; J. Fazo; T. Mitros; W. Dirks; U. Hellsten; N. Putnam & D.S. Rokhsar (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, 40.

Guihur, A.; M.E. Rebeaud & P. Goloubinoff (2022). How do plants feel the heat and survive?. In *Trends in Biochemical Sciences,* 47, 824–838.

Horton, P.; K.J. Park; T. Obayashi; N. Fujita; H. Harada; C.J. Adams-Collier & K. Nakai (2007). WoLF PSORT: Protein localization predictor. *Nucleic Acids Research*, 35.

Hu, B.; J. Jin; A.Y. Guo; H. Zhang; J. Luo & G. Gao (2015). GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics*, 31, 1296–1297.

Hu, S.; Y. Ding & C. Zhu (2020). Sensitivity and responses of chloroplasts to heat stress in plants. *Frontiers in Plant Science*, 11, 1–11.

Hu, W.; G. Hu & B. Han (2009). Genome-wide survey and expression profiling of heat shock proteins and heat shock factors revealed overlapped and stress specific response under abiotic stresses in rice. *Plant Science*, 176, 583–590.

ICCO (2025). Production of cocoa beans. ICCO Quarterly Bulletin of Cocoa Statistics, LI (1).

Immormino, R.M.; D.E. Dollins; P.L. Shaffer; K.L. Soldano; M.A. Walker & D.T. Gewirth (2004). Ligand-induced conformational shift in the N-terminal domain of GRP94, an Hsp90 chaperone. *Journal of Biological Chemistry*, 279, 46162–46171.

Krishna, P. & G. Gloor (2001). The Hsp90 family of proteins in *Arabidopsis thaliana*. *Cell Stress & Chaperones*, 6, 238–246.

Larasati, R.J.; R. Anindita & W. Widyawati (2022). Peningkatan ekspor kakao Indonesia di pasar internasional. *Jurnal Ekonomi Pertanian Dan Agribisnis*, 6, 1025–1037.

Lescot, M.; P. Déhais; G. Thijs; K. Marchal; Y. Moreau; Y. Van De Peer; P. Rouzé; & S. Rombauts (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, 325–327.

Li, W.; Y. Chen; M. Ye; D. Wang & Q. Chen (2020). Evolutionary history of the heat shock protein 90 (Hsp90) family of 43 plants and characterization of Hsp90s in *Solanum tuberosum*. *Molecular Biology Reports*, 47, 6679–6691.

Liu, C.; H. Yu; X. Rao; L. Li & R.A. Dixon (2021). Abscisic acid regulates secondary cell-wall formation and lignin deposition in *Arabidopsis thaliana* through phosphorylation of NST1. *Proceedings of the National Academy of Sciences,* 118, 1–11.

Luo, A.; X. Li; X. Zhang; H. Zhan; H. Du; Y. Zhang & X. Peng (2019). Identification of *AtHSP90*.6 involved in early embryogenesis and its structure prediction by molecular dynamics simulations. *Royal Society Open Science*, 6, 1–15.

Magnard, J.-L. & P. Vergne (1996). Complexity and genetic variability of heat-shock protein expression in isolated maize microspores'. *Plant Physiol*, 111, 1085–1096.

Mu, B.; A.M. Nair & R. Zhao (2024). Plastid HSP90C C-terminal extension region plays a regulatory role in chaperone activity and client binding. *Plant Journal*, 119, 2288–2302.

Oldroyd, G.E.D. & O. Leyser (2020). A plant's diet, surviving in a variable nutrient environment. *Science*, 368, 1–9.

Peng, J.; S. Liu; J. Wu; T. Liu; B. Liu; Y. Xiong; J. Zhao; M. You; X. Lei & X. Ma (2024).

Genome-wide analysis of the oat (*Avena sativa*) hsp90 gene family reveals its identification, evolution, and response to abiotic stress. *International Journal of Molecular Sciences*, 25, 1–17.

Sajad, S.; S. Jiang; M. Anwar; Q. Dai; Y. Luo; M.A. Hassan; C. Tetteh & J. Song (2022). Genome-wide study of hsp90 gene family in cabbage (*Brassica oleracea* var. capitata l.) and their imperative roles in response to cold stress. *Frontiers in Plant Science*, 13, 1–13.

Sarkar, N.K.; Y.K. Kim & A. Grover (2009). Rice sHsp genes: Genomic organization and expression profiling under stress and development. *BMC Genomics*, 10, 1–18.

Song, H.; P. Fan & Y. Li (2009). Overexpression of organellar and cytosolic *AtHSP90* in *Arabidopsis thaliana* impairs plant tolerance to oxidative stress. *Plant Molecular Biology Reporter*, 27, 342–349.

Song, Z.; F. Pan; C. Yang; H. Jia; H. Jiang; F. He; N. Li; X. Lu & H. Zhang (2019). Genome-wide identification and expression analysis of HSP90 gene family in *Nicotiana tabacum*. *BMC Genetics*, 20, 1–12.

Tamura, K.; G. Stecher & S. Kumar (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38, 3022–3027.

Thompson, J.D.; D.G. Higgins+ & T.J. Gibson (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. In *Nucleic Acids Research*, 22, 4673–4680.

Tigchelaar, M.; D.S. Battisti; R.L. Naylor & D.K. Ray (2018). Future warming increases probability of globally synchronized maize production shocks. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 6644–6649.

Wang, L.; F. Liu; L. Ju; B. Xue; Y. Wang; D. Wang & D. Hou (2022). Genome structures and evolution analysis of HSP90 gene family in brassica napus reveal the possible roles of members in response to salt stress and the infection of *Sclerotinia sclerotiorum*. *Frontiers in Plant Science*, 13, 1–13.

Wasilah, U.; D.A.G. Perwitasari & M. Su'udi (2019). Peran chaperone pada tumbuhan: mini review. *Jurnal Biologi Papua*, 11, 110–115.

Xue, P.; Y. Sun; D. Hu; J. Zhang & X. Wan (2023). Genome-wide characterization of DcHsp90 gene family in carnation (*Dianthus caryophyllus* L.) and functional analysis of DcHsp90-6 in heat tolerance. *Protoplasma*, 260, 807–819.

Yu, C.; C. Lin & J. Hwang (2004). Predicting subcellular localization of proteins for Gram negative bacteria by support vector machines based on n peptide compositions *Protein Science*, 13, 1402–1406.

Yu, C.S.; Y.C. Chen; C.H. Lu & J.K. Hwang (2006). Prediction of protein subcellular localization. *Proteins: Structure, Function and Genetics*, 64, 643–651.

Zai, W.S.; L.X. Miao; Z.L. Xiong; H.L. Zhang; Y.R. Ma; Y.L. Li; Y.B. Chen S.G.Ye (2015). Comprehensive identification and expression analysis of Hsp90s gene family in *Solanum lycopersicum*. *Genetics and Molecular Research*, 14, 7811–7820.

Zhang, K.; S. He; Y. Sui; Q. Gao; S. Jia; X. Lu & L. Jia (2021). Genome-wide characterization of hsp90 gene family in cucumber and their potential roles in response to abiotic and biotic stresses. *Frontiers in Genetics*, 12, 1–16.

Zhang, M.; Z. Shen; G. Meng; Y. Lu & Y. Wang (2017). Genome-wide analysis of the *Brachypodium distachyon* (L.) P. Beauv. Hsp90 gene family reveals molecular evolution and expression profiling under drought and salt stresses. *PLoS ONE*, 12, 1–18.

-o0o-