JPMB | Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 2

Genome-wide analysis of the *HSP90* gene family and their roles in soybean growth and development

Original paper

Samira Mohammadi*, Firouzeh Sohrevardi, Ghorbanali Nematzadeh

Genetics and Agricultural Biotechnology Institute of Tabarestan (GABIT), Sari Agricultural Sciences and Natural Resources University (SANRU). Sari. Iran.

Abstract: Heat shock protein of 90 kDa or HSP90 plays an important dynamic role in regulating biotic and abiotic stresses through multiple functional mechanisms. The present study aimed to perform a comprehensive analysis of the *HSP90* gene family in soybean. In total, 20 *HSP90* genes from soybean were identified and showed unequal distribution on the 13 chromosomes. The evolutionary tree divided these genes into three main groups based on their subcellular localization. In Group I, nearly all of the *HSP90* genes are distributed in the nucleus or cytoplasm. In Group II, the *HSP90s* were mostly classified in the endoplasmic reticulum. *HSP90* genes were exclusively found in the mitochondria or chloroplast in Group III. Phylogenetic relationships have shown that genes in similar subgroups have the same exon-intron structure and number of introns. Glyma14g219700, Glyma17g258700, and Glyma07G207600 were identified as hub proteins based on their high degrees of interaction. In addition, *Glyma02g302500*, *Glyma08g332900*, *Glyma14g219700*, *Glyma17g258700*, and *Glyma18g074100* genes displayed high expression levels in all of the tissues at different developmental stages. These findings provide a complete overview of the *GmHSP90* gene family classification and evolution, which can help to identify the functional properties of the *HSP90* genes in soybean growth and development.

Keywords: gene expression profiling, gene structure, heat shock protein, phylogenetic analysis, protein-protein interaction network, subcellular localization.

OPEN ACCESS

Edited by

Dr. Parviz Heidari,
Shahrood University of Technology,
Iran

Date

Received date: 06 June 2022
Accepted date: 16 May 2024
Published: 26 May 2024

Correspondence

Dr. Samira Mohammadi
s.mohamadi@stu.sanru.ac.ir

Citation

Mohammadi, S., Sohrevardi, F. and Nematzadeh, G.A. (2023). Genome-wide analysis of the *HSP90* gene family and their roles in soybean growth and development . *J Plant Mol Breed 11* (2): 119-132.
doi: 10.22058/JPMB.2024.555256.1256.

Introduction

Plants encounter inevitable environmental challenges such as heat shock, cold, salt, submergence, drought, nutrient deficiency, and chemical pollution, which severely limit their normal growth and development (Vaughan et al., 2018). When plants are faced with these stresses, specific defense responses or mechanisms activate at both molecular and physiological levels. One of these responses is the induction of heat shock proteins (HSPs) for the survival of stressed cells (Sangster and Queitsch, 2005). HSPs are divided into five main categories based on their molecular weight, namely, HSP100, HSP90, HSP70, HSP60, and small HSP (sHSP, HSP20) (Wang et al., 2004).

HSP90s are molecular chaperones and highly conserved in prokaryotes to eukaryotes, which helps to regulate and maintain the compatibility of various proteins and also protects normal cells against stress stimuli. In eukaryotes, HSP90s play a wide variety of roles in stress signaling, including stabilizing and, or correctly folding their client proteins, and assisting in protein degradation (Hoter et al., 2018).

HSP90 chaperones are involved in transcription factors and protein kinase folding, along with activating a substrate for initiating stress signaling (Song et al., 2019). The *HSP90* gene family under normal or stressful conditions is responsible for preventing protein aggregation and facilitating denatured proteins refolding, which helps protein folding with the participation of other chaperones and the formation of a mechanism (Picard, 2002; Wang et al., 2004).

HSP90 protein includes an ATP-binding domain at the N-terminal, an intermediate domain, and a dimerization domain at the C-terminal (Pearl and Prodromou, 2006). The expression patterns of the *HSP90* gene family have been broadly studied in various plant tissues and organs, and it has been found that HSP90s affect the plant's tolerance to biotic and abiotic stresses (Chen et al., 2018; di Donato and Geisler, 2019; Song et al., 2019; Bettaieb et al., 2020).

The *HSP90* genes were examined in *Aeluropus littoralis* (Hashemipetroudi et al., 2019), *Cucumis sativus* L. (Zhang et al., 2021), pepper (Jing et al., 2020), barley (Chaudhary et al., 2019), *Nicotiana tabacum* (Song et al., 2019), and a variety of lower plants to higher plants (Li et al., 2020). Cucumber *HSP90* genes were differentially expressed under gibberellin, photoperiod, and temperature, powdery, and downy mildew stimuli (Zhang et al., 2021). *HSP90* genes of *Brassica napus* played crucial roles in seed germination and development (Reddy et al., 1998). It has been found that the *HSP90* family is involved in the differentiation and development of cotton fiber through the homeostasis maintenance of cellular (Sable et al., 2018).

Soybean (*Glycine max* L.) is a nutritionally and economically important crop, the world's most important edible oilseed crop due to its oil content, vegetable protein, and nutritional value. This plant also provides essential amino acids for animals and humans. In addition, it is used as a food supplement and pharmaceutical source (Shelke et al., 2023). However, the risk of damage from abiotic and biotic stresses threatens soybean production)Zhang et al., 2015(.

In rice, *OsHsp90-2* and *OsHsp90-4* genes were upregulated to heat, cold, salt, and drought stresses (Hu et al., 2009). It has been shown that in Arabidopsis, expression of *HSP90* genes increased by salt and heat stresses (Prasad et al., 2010; Mishra and Grover, 2016). According to the available information about the functional diversity of *HSP90* genes in plants, HSP90 proteins can be considered ideal targets for improving the development of soybean cultivars tolerant to a wide range of stresses. Thus, we investigated *HSP90* genes and estimated their evolutionary relationships, functional protein domains analysis, physicochemical features, gene structure, conserved motifs analysis, subcellular and chromosomal localization, and expression analysis during soybean growth and development. The findings of this study have revealed a better comprehension of *GmHSP90* genes in soybean, while providing a basis for the complete investigation of candidate genes in the future.

Materials and Methods

Identification of soybean HSP90 proteins

The *G. max* genome sequences were downloaded from the Phytozome v13 database. To collect GmHSP90 protein sequences, profile hidden

Markov models (HMMs) of the HSP90 domain (Pfam: PF00183) were retrieved from the Pfam database and sent as a query for the search of BLASTP in the whole *G. max* genome sequence (Finn et al., 2015). Then, identified sequences were checked for the existence of the HSP90 domain using the InterProScan (Jones et al., 2014) and the SMART program (Letunic et al., 2014).

Protein properties analysis

Information of the GmHSP90 proteins, including amino acid number, molecular weight, instability index, aliphatic index, isoelectric point, and grand average of hydropathicity index (GRAVY), were calculated using compute pI/Mw tool from ExPASy (Gasteiger et al., 2005).

Subcellular and chromosomal localization

The online program WoLF PSORT server was used to predict the subcellular locations of the GmHSP90 proteins (Horton et al., 2007). The chromosomal distribution of *GmHSP60* genes was determined using the MapChart program (Voorrips, 2002).

Phylogenetic analysis

In this study, we construct two phylogenetic trees so that we could first classify the HSP90 protein sequences in soybean and then compare them with the HSP90 proteins of other plants including Arabidopsis, maize, and rice. MEGA 7.0 software was used to align the HSP90 sequences and then constructed phylogenetic trees using the Neighbor-Joining algorithm with 1000 bootstrapping tests to ensure clustering accuracy (Kumar et al., 2016).

Conserved motif and gene structure analysis

The Gene Structure Display Server GSDS 2.0 tool was used to identify the exon-intron structure of GmHSP90 members (Hu et al., 2014). The conserved motifs in the GmHSP90 proteins were identified with the MEME tool using the following parameters: maximum number of motifs, 20; minimum motif width, 6; and maximum motif width, 100 (Bailey et al., 2009). Then, conserved motifs of GmHSP90s were annotated using the InterProScan tool.

Proteins interaction network prediction

Functional and physical interactions between soybean HSP90 proteins were predicted using the STRING database (http://string-db.org/) (Szklarczyk et al., 2019) and visualized by the Cytoscape v3.10.1 software (http://cytoscape.org/) (Shannon et al., 2003).

RNA-Seq analysis

To determine *HSP90* gene family expression during soybean growth and development, RNA-seq data from 14 tissues, including flower, nodule, young leaf, root, one cm pod, pod shell 10 DAF, pod shell 14 DAF, seed 10 DAF, seed 14 DAF, seed 21 DAF, seed 25 DAF, seed 28 DAF, seed 35 DAF, and seed 42 DAF, were downloaded from SoyBase (Brown et al., 2021). The normalized expression data were transformed by log10, and the heat map was drawn using the CIMminer program (Scherf et al., 2000).

Results

Identification of the HSP90 proteins in Soybean

In total, 20 *HSP90* genes were identified in the soybean genome (Table 1). Domain analysis revealed the presence of HSP90 (Pfam: PF00183) and Histidine kinase-like ATPase (HATPase_c) (Pfam: PF02518) domains. Although all GmHSP90 proteins have at least one conserved HSP90 domain, the HATPase_c domain is present in 13 GmHSP90s. Noteworthy, Glyma10G098300, and Glyma07G207600 have two HSP90 domains. The number of amino acids of the GmHSP90 proteins ranged from 281 (Glyma10G098300) to 847 (Glyma14G219700) and had a molecular weight of 33.10 – 97.38 kDa. The aliphatic indexes varied from 77.09 (Glyma14G007700) to 94.93 (Glyma07G207600). The values of the isoelectric point were between 4.86 (Glyma17G258700) and 9.60 (Glyma19G098200), which indicates the change of protein nature from acidic to basic. Investigating the protein instability index showed that only 13 GmHSP90 proteins have indexes below 40, which can be considered stable proteins. The GRAVY values of GmHSP90 proteins were negative, which can be concluded that all of them are hydrophilic (Table 1).

Subcellular localization

Subcellular localization of the 20 GmHSP90s revealed that 11 GmHSP90 proteins are active in the cytoplasm, 5 proteins in the chloroplast, 2 proteins in the nucleus, and 2 proteins in the endoplasmic reticulum (Table 1).

Table 1. The physicochemical characteristics and subcellular location of the GmHSP90 proteins.

Proposed name	Animo acids	Molecular weight (kDa)	pI	GRAVY	Instability index	Aliphatic index	Subcellular localization
Glyma10G098300	281	33.10	5.37	-0.741	40.63	83.20	Cytoplasm
Glyma17G182500	360	42.92	5.70	-0.750	39.07	79.03	Cytoplasm
Glyma17G220000	323	37.71	6.04	-0.627	40.44	74.24	Nucleus
Glyma17G258700	814	93.29	4.86	-0.717	31.37	79.64	Endoplasmic reticulum
Glyma09G131500	699	80.39	4.95	-0.646	40.01	82.56	Cytoplasm
Glyma07G207600	304	34.44	5.88	-0.356	25.93	94.93	Cytoplasm
Glyma02G305600	791	89.72	5.25	-0.550	37.42	79.86	Chloroplast
Glyma02G302500	702	80.33	4.96	-0.584	35.98	82.51	Cytoplasm
Glyma02G124500	794	90.09	4.91	-0.560	50.01	80.31	Chloroplast
Glyma19G098200	311	36.89	9.60	-0.511	47.69	85.24	Nucleus
Glyma03G114400	740	85.55	8.73	-0.354	30.98	88.96	Cytoplasm
Glyma11G227000	406	46.37	5.53	-0.225	44.23	87.36	Cytoplasm
Glyma14G011600	700	80.19	4.98	-0.590	36.29	82.60	Cytoplasm
Glyma14G219700	847	97.38	4.91	-0.693	31.38	80.57	Endoplasmic reticulum
Glyma14G007700	797	90.50	5.13	-0.575	36.20	77.09	Chloroplast
Glyma08G332900	699	80.18	4.97	-0.594	36.49	83.96	Cytoplasm
Glyma08G032900	655	75.88	6.28	-0.530	38.60	82.90	Chloroplast
Glyma01G068000	793	90.11	4.94	-0.567	52.70	79.81	Chloroplast
Glyma18G074100	702	80.38	4.94	-0.579	35.53	84.30	Cytoplasm
Glyma16G178800	699	80.29	4.97	-0.637	39.62	82.98	Cytoplasm

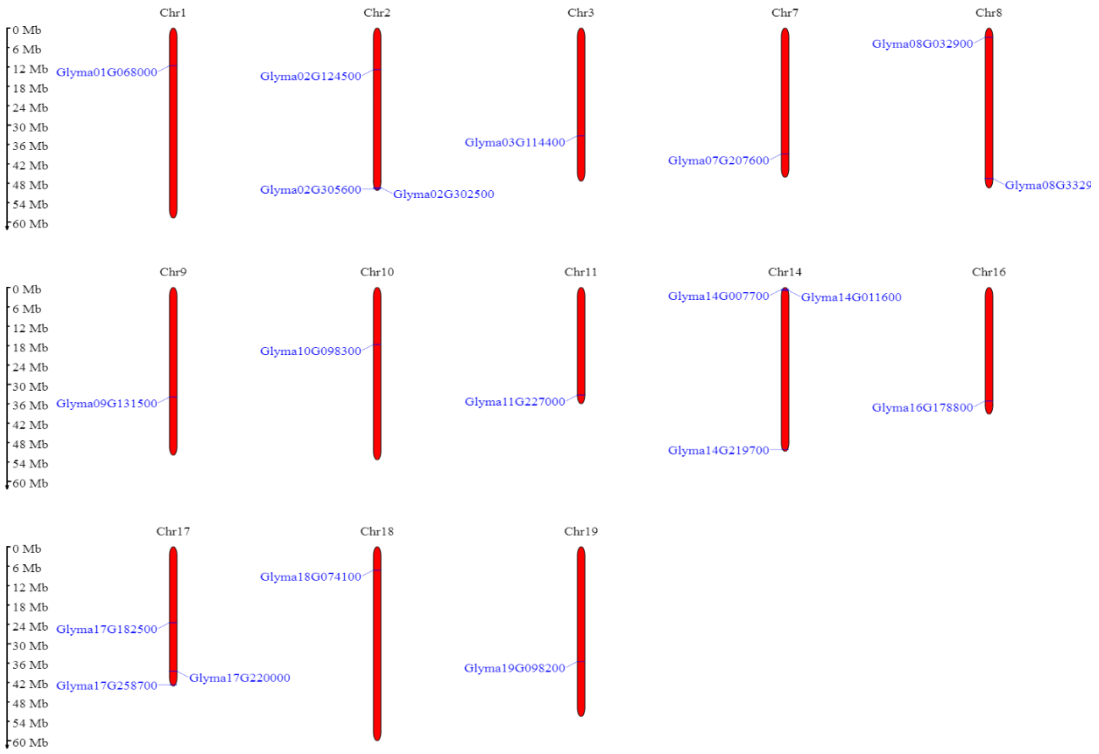

Figure 1. Chromosomal location of *GmHSP90* genes. The scale bar represents gene position (Mb) in soybean chromosomes.

Chromosomal location of GmHSP90 genes

The 20 soybean *HSP90* genes were distributed among 13 chromosomes, excluding chromosomes 4, 5, 6, 12, 13, 15 and 20 (Figure 1). The number of *GmHSP90* genes in each chromosome was significantly different. Chromosomes 2, 14, and 17 carried 3 *GmHSP90* genes; two genes were found in chromosome 8; only one gene was found in other chromosomes.

Motif analysis and gene structure of the GmHSP90s

Analysis of exon-intron structure showed that the number of introns in the soybean *HSP90* genes varied from 0 to 18 (Figure 2). However, no intron was observed in the Glyma19G098200, whereas 18 introns were observed in the Glyma02G305600, Glyma14G007700, Glyma02G124500, and Glyma01G068000 genes. The motif analysis identified 20 conserved motifs, and 13 motifs were detected as the domains associated with the HSP90 protein family (Figure 3 and Table 2). The results showed that related proteins shared the conserved motifs. Furthermore, the number and composition of motifs in the same subgroups were similar but were different from the proteins in other subgroups (Figure 3).

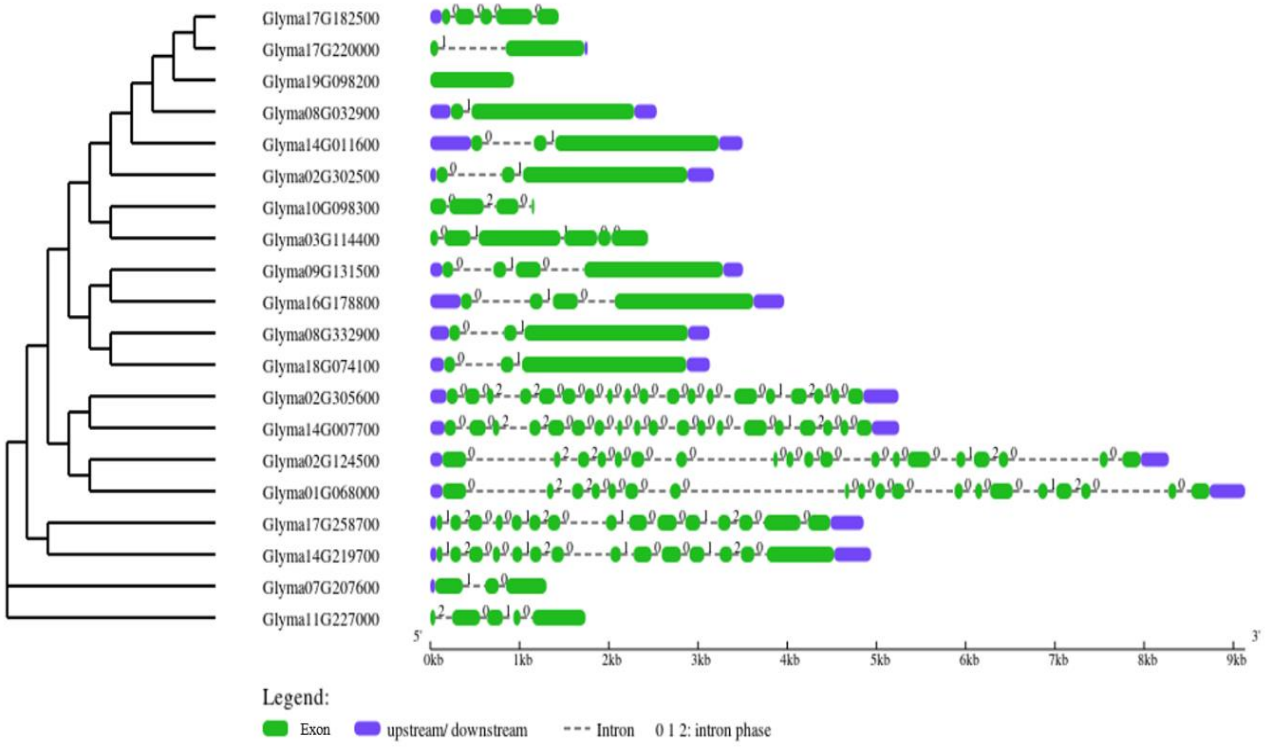

Figure 2. Phylogenetic tree and exon-intron distribution in the HSP90 genes in G. max. Phylogenetic tree was mapped using MEGA7.0 based on the Neighbor-Joining distance algorithm with 1000-replicates bootstrap. Blue and green boxes indicate the upstream/downstream UTR regions and exons, respectively, while dashed lines indicate introns. The splicing phases are indicated by 0, 1, and 2. Using the size at the bottom, it is possible to estimate the length of introns and exons.

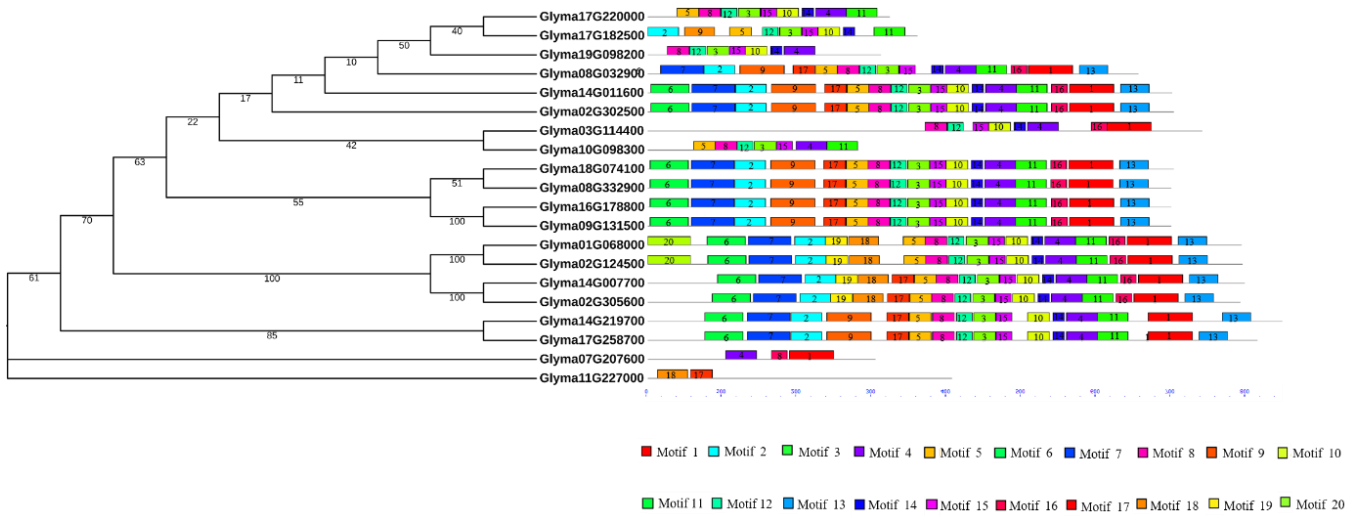

Figure 3. Phylogenetic tree and motif analysis of GmHSP90 proteins.. The colored boxes illustrate motifs and gray lines show the non-conserved sequences.

Table 2. List of the putative motifs and annotation in GmHSP90s . Numbers in the first column indicate the motifs represented in Figure 3.

Motif	Width	Best possible match	E-value	Annotation
1	59	LVDSPCCLVTGEYGWSANMERIMKAQALRDSSMAGYMSSKKTMEINPDNPIMEELRKRA	4.5e-549	HSP90 domain
2	41	MIGQFGVGFYSAYLVADKVIVTTKHNDDEQYVWESQAGGSF	2.2e-326	HSP90 domain
3	29	LMPEYLSFVKGIVDSEDLPLNISREMLQQ	2.5e-320	HSP90 domain
4	41	YVTRMKEGQKDIYYITGESKKAVENSPFLEKLKKKGYEVLY	9.1e-419	HSP90 domain
5	29	PIWMRKPEEITKEEYAAFYKSLTNDWEEH	2.7e-252	HSP90 domain
6	51	AETFEFQAEINRLLDLIINSLYSNKEIFLRELISNASDALDKJRFESLTDK	2.9e-355	HSP90 domain
7	57	AQPELFIHIKPDKDNNTLSIIDSGIGMTKADLVBNLGTIARSGTKEFMEALAAGADV	1.5e-337	HSP90 domain
8	29	LAVKHFSVEGQLEFKAILFVPKRAPFDLF	1.4e-273	HSP90 domain
9	59	GENLGRGTKITLFLKEDQLEYLEERRLKDLIKKHSEFISYPISLWIEKTTEKEISDDED	4.6e-256	HSP90 domain
10	29	AENKEDYNKFYEAFSKNLKLGIHEDSQNK	4.7e-238	HSP90 domain
11	41	MVDAIDEYAVGQLKEYEGKKLVSATKEGLKLDESEEEKKKK	3.4e-256	HSP90 domain
12	21	TRKKPNNIKLYVRRVFIMDNC	6.4e-218	HSP90 domain
13	38	KDLVLLLFETALLTSGFSLDDPNTFGNRIHRMLKLGLS	3.4e-217	HSP90 domain
14	15	LLRYHSTKSGDEMTS	3.3e-112	-
15	21	KILKVIRKNLVKKCIEMFFEI	4.2e-150	-
16	21	FDGLCKVIKDVLGDKVEKVVV	1.6e-109	-
17	29	DVDEDKEKEEKKKKTIKEVSHEWELVNKQ	4.7e-101	-
18	40	FSEPERIEGLVKNYSQFVSFPIYTWQEKSTTKEVEEDEDP	2.0e-061	HSP90 domain
19	29	SYVIKEETDPEKLJPRGTRJTLYLKEDDK	1.6e-024	-
20	57	MAPVPSRTMATASLASLPPSSPFARASLLRSAFLPPQIGRGRKCFSPAAGLRWTQRR	3.1e-020	

Evolutionary classification of HSP90 genes

To investigate the phylogenetic relationships of *HSP90* genes in soybean, Arabidopsis, maize, and rice species, the phylogenetic classification of the amino acid sequences of 47 identified proteins was constructed based on the Neighbor-Joining distance algorithm (Figure 4). The evolutionary tree divided these genes into three main groups based on their subcellular localization. The distribution of *HSP90* genes in different species showed unique patterns

and slight differences in each group. In Group I, nearly all of the *HSP90* genes are distributed in the nucleus or cytoplasm. In Group II, the *HSP90*s were mostly classified in the endoplasmic reticulum. *HSP90* genes were exclusively found in the mitochondria or chloroplast in Group III. Among the three groups, group I was the largest, with 28 members. After that, there was group II with 13 members, and group III with six members.

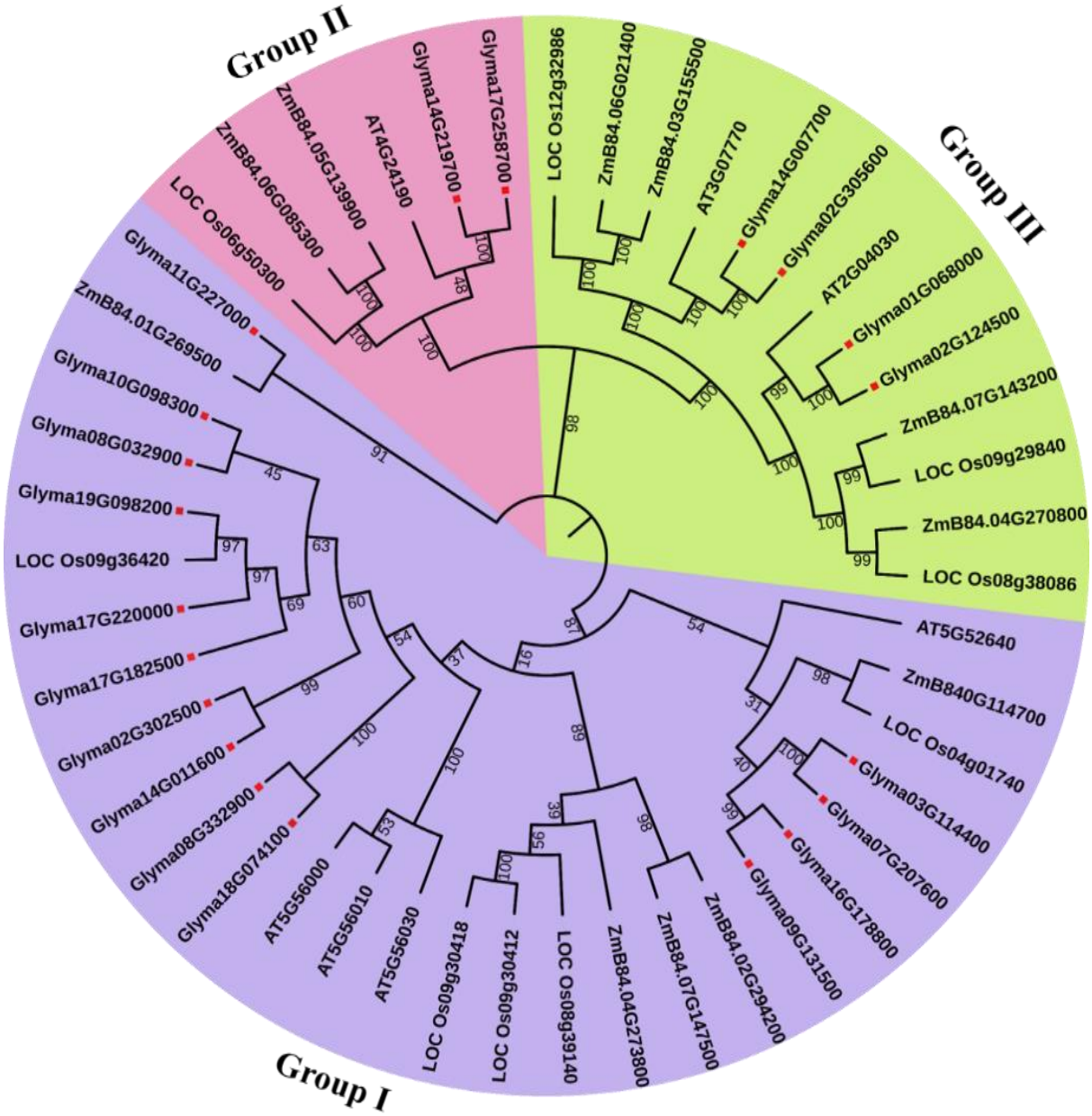

Figure 4. Phylogenetic tree of HSP90 family members from *Glycine max*, *Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays* species. The HSP90s were divided into three groups (groups I–III) based on the clustering of the protein sequence which were showed by different colors. GmHSP90 proteins are marked with red shapes Bootstrap values are indicated on each branch

Protein-protein interaction network of GmHSP90 proteins

The protein interaction network for GmHSP90 proteins was predicted using the STRING database and plotted by the Cytoscape software. Clustering of the network using MCODE algorithm showed that this protein network has three clusters, the first, second, and third clusters are composed of 8, 8, and 4 proteins (nodes), respectively (Figure 5A). We identified three key hub proteins, Glyma14g219700, Glyma17g258700, and Glyma07G207600, based on their high degrees of interactions, which are shown in red color (Figure 5B). This indicates the importance of these proteins in signaling responses to various environmental conditions.

Soybean HSP90 family gene expression patterns

Due to the lack of the expression patterns of two *GmHSP90* genes (*Glyma07g207600* and *Glyma03g114400*) in the soybean RNA-seq database, the expression profiles of 18/20 *GmHSP90* members were analyzed. A heat map showed that the majority of the *GmHSP90s* revealed preferential expression patterns, and only four *GmHSP90* genes (*Glyma10g098300*, *Glyma17g182500*, *Glyma17g220000*, and *Glyma19g098200*) had no expression in any tissues (Figure 6). The expression data analysis demonstrated that 14 genes were expressed in 14 tissues and organs at different developmental stages. *Glyma02g302500*, *Glyma08g332900*, *Glyma14g219700*, *Glyma17g258700*, and *Glyma18g074100* genes displayed high expression levels in all of the tissues. Therefore, it can be concluded that the *GmHSP90* genes played vital roles in different processes during soybean growth and development.

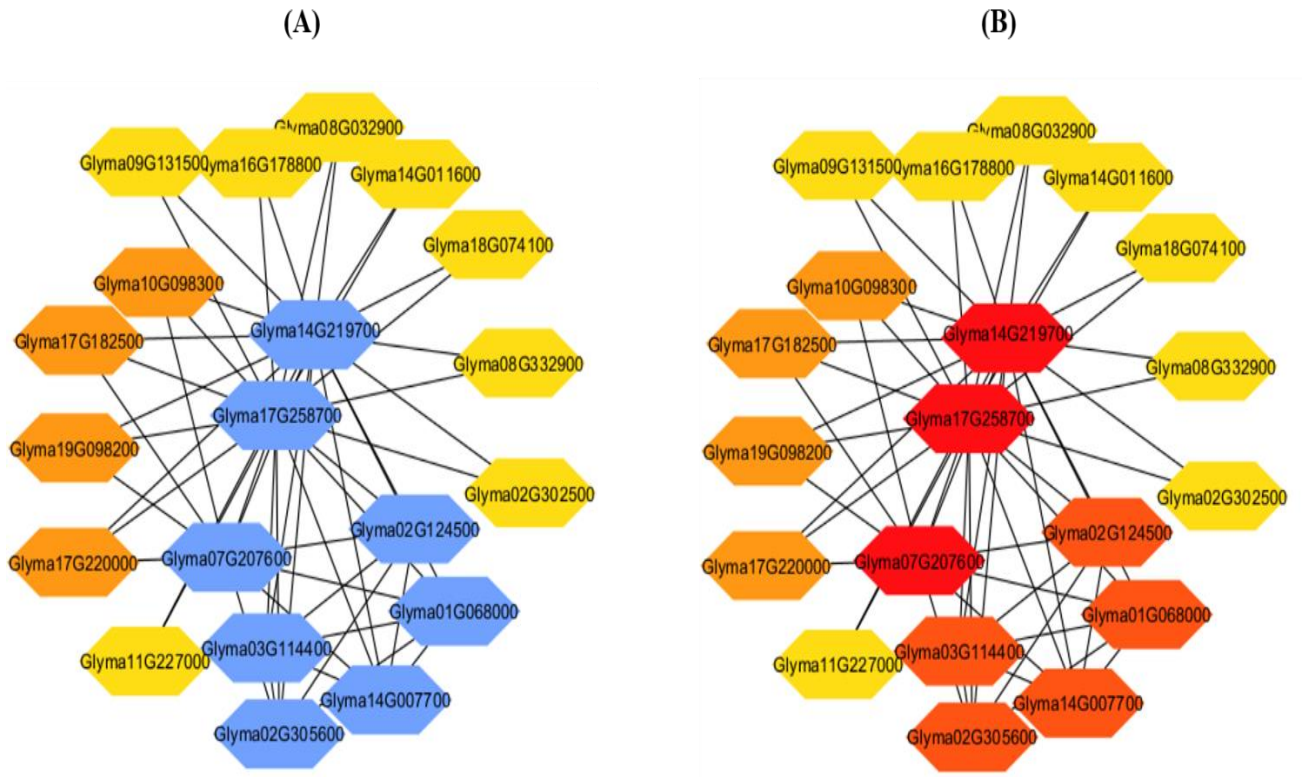

Figure 5. Protein-protein interaction network of GmHSP60 proteins. The protein-protein interaction network for GmHSP60 proteins was extracted using the STRING database and visualized by the Cytoscape software. Colored hexagons and lines represent nodes and edges, respectively. A) Network clustering: blue, yellow, and orange hexagons show the proteins located in the first, second, and third clusters, respectively. B) Network analysis: proteins located in red hexagons were identified as hubs.

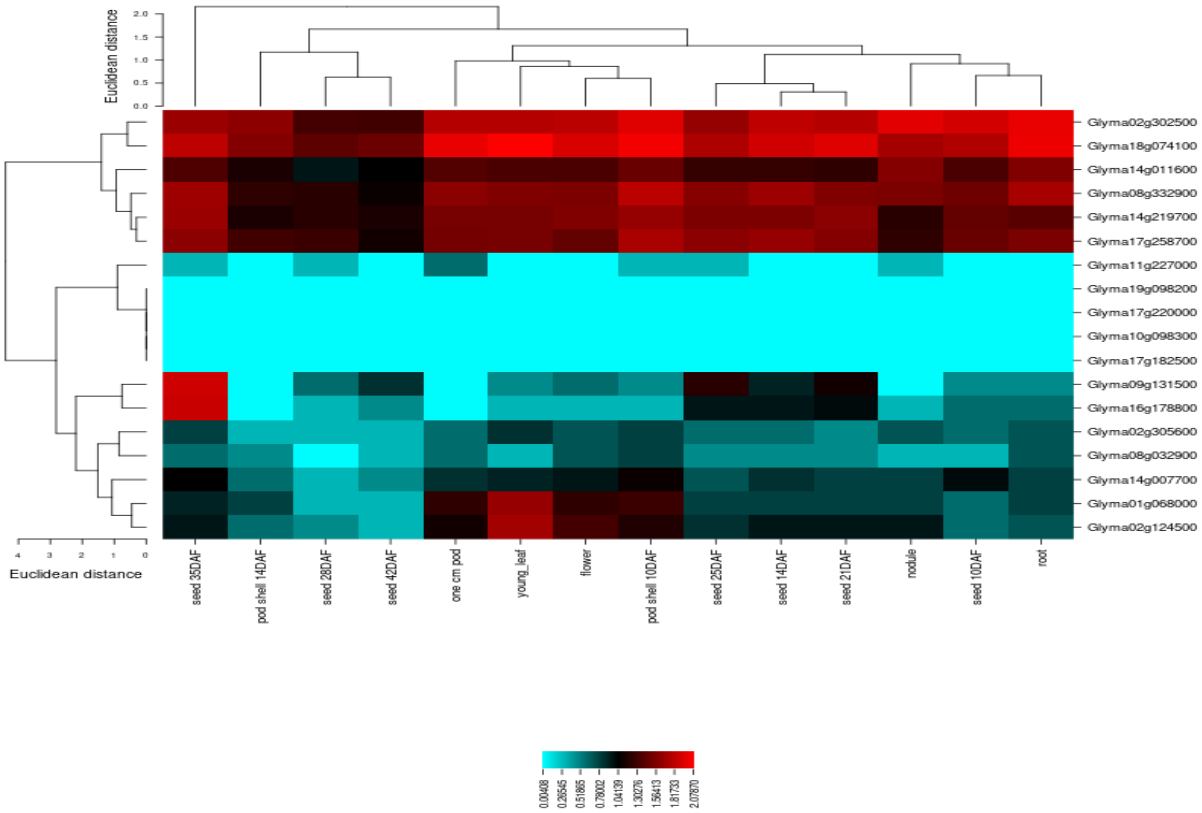

Figure 6. Heat map showing the levels of *GmHSP90* genes expression in multiple tissues, organs, and developmental stages. Blue and red colors respectively indicate low and high expressions.

Discussion

HSP90 proteins, which can be found in prokaryotes and eukaryotes, are highly conserved molecular chaperones. These proteins play a major role in the cell cycle, transduction of signals, and more essential biological processes (Pearl and Prodromou, 2006). A comprehensive analysis of the soybean *HSP90* gene family has been conducted in this study to assess evolutionary relationships, subcellular and chromosomal localization, conserved motifs, gene structure, and expression profiles.

The *GmHSP90* genes were mainly distributed at two ends of the chromosomes, similar to the *HSP90* genes distribution of *Malus sieversii*, rice, and *Nicotiana tabacum* (Hu et al., 2009; Song et al., 2019; Haxim et al., 2021). Most of the GmHSP90 proteins are located in the cytoplasm. This feature is shared with *Brachypodium distachyon* (Zhang et al., 2021) and Arabidopsis (Sarkar et al., 2009), and suggests that HSP90 proteins may have their primary site of action in the cytoplasm, where protein assembly occurs (Vabulas et al., 2010). Some proteins are active in the chloroplast, nucleus, and endoplasmic reticulum. The existence of spatial variation for GmHSP90 proteins probably indicates the diverse functional roles of these genes in different cellular processes.

To describe and understand the existing functional characteristics of examined genes, it may be important to consider these evolutional relationships. To determine the species' evolutional relationships and to estimate paralogs and orthologs, within and between species, phylogenetic analysis is frequently used (Bettaieb et al., 2020). High bootstrapping values have supported the phylogenetic tree, and have shown that *GmHSP90* genes have been classified into three main subgroups. This is consistent with the findings of prior studies (Zhang et al., 2013; Zhang et al., 2017). The results indicated that exon-intron structures and the number of introns were similar

within each subgroup, while genes from various subgroups have differing gene structures. Also, the distribution of conserved motifs in the same group shows a similar pattern. These analyses suggested that the evolutionary classification of *GmHSP90* genes is reliable. In addition, these findings revealed that the genes located in each group are highly conserved, which probably indicates the similar functions of the members of each group.

Analysis of subcellular location indicated that in Group I, nearly all of the *HSP90* genes are distributed in the nucleus or cytoplasm. In Group II, the HSP90s were mostly classified in the endoplasmic reticulum. *HSP90* genes were almost exclusively found in the mitochondria or chloroplasts in Group III.

The number of introns was similar on some sister pairs within the branches of a phylogenetic tree. However, some sister pairs continued to show changes in gene structure, as well as numbers. The results demonstrated intron loss/gain in *HSP90* encoding genes during the structural evolution. It is known that the protein structure determines its function (Bai et al., 2002). This study indicated that there was a different gene structure for GmHSP90 protein sequences. The number of introns mainly depends upon gene transcription regulation sensitivity, and it is more likely that plants can respond to various developmental and environmental stimuli with lower numbers of introns (Appiah et al., 2021).

Using the protein-protein interaction network, Glyma14g219700, Glyma17g258700, and Glyma07G207600 were identified as hub proteins. It has been found that hub genes play significant roles in various biological processes (Mishra et al., 2023). The heat map data shows that the majority of *GmHSP90* genes are expressed in a variety of soybean organs and tissues. So, it can be concluded that they may be involved in growth and development. Heat map revealed expression profiles of paralogous pairs from various subfamilies, and it was found that most paralogous pairs at the high level of sequence homology have identical expression profiles in the number of different tissues. In particular, the *Glyma02g302500* and *Glyma18g074100* pairs were strongly expressed in all types of tissues. However, in all seed stages, *Glyma08g032900* and *Glyma02g305600* genes showed very low expression.

Conclusion

Soybean has 20 genes encoding *HSP90*, and they are located on 13 chromosomes. The organization of conserved motifs and gene structure confirmed the phylogenetic classification. *GmHSP90* gene family expression profiles in various tissues indicate that these genes play essential roles in soybean growth and development. According to the general information about these genes and their possible participation in plant growth and development, further study of the *HSP90* gene family will be facilitated, especially about their biological functions and evolutionary history.

Author contributions

Conceptualization, S. M. and F. S.; methodology, S. M.; software, S. M.; validation, S. M., F. S. and G. A.N.; formal analysis, S. M.; investigation, S. M.; resources, S. M.; data curation, S. M.; writing—original draft preparation, S. M.; writing—review and editing, S. M.; visualization, S. M.; supervision, S. M.; project administration, S. M.. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgments:

Conflict of interest statement

The authors declare no conflict of interest

References

Appiah, C., Yang, Z.-F., He, J., Wang, Y., Zhou, J., Xu, W.-Z., Nie, G., and Zhu, Y.-Q. (2021). Genome-Wide identification of *Hsp90* gene family in *Perennial Ryegrass* and expression analysis under various abiotic stresses. *Plants* 10(11): 2509.

Bai, J., Pennill, L.A., Ning, J., Lee, S.W., Ramalingam, J., Webb, C.A., Zhao, B., Sun, Q., Nelson, J.C., and Leach, J.E. (2002). Diversity in nucleotide binding site–leucine-rich repeat genes in cereals. *Genome Res* 12(12): 1871-1884.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 37: 202-208.

Bettaieb, I., Hamdi, J., and Bouktila, D. (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. *Physiol Mol Biol Plants* 26(11): 2301-2318.

Brown, A.V., Conners, S.I., Huang, W., Wilkey, A.P., Grant, D., Weeks, N.T., Cannon, S.B., Graham, M.A., and Nelson, R.T. (2021). A new decade and new data at SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res* 49(D1): D1496-D1501.

Chaudhary, R., Baranwal, V.K., Kumar, R., Sircar, D., and Chauhan, H. (2019). Genome-wide identification and expression analysis of *Hsp70*, *Hsp90*, and *Hsp100* heat shock protein genes in barley under stress conditions and reproductive development. *Funct Integr Genomics* 19(6): 1007-1022.

Chen, J., Gao, T., Wan, S., Zhang, Y., Yang, J., Yu, Y., and Wang, W. (2018). Genome-wide identification, classification and expression analysis of the *HSP* gene superfamily in tea plant (*Camellia sinensis*). *Int J Mol Sci* 19(9): 2633.

di Donato, M., and Geisler, M. (2019). *HSP 90* and co‐chaperones: a multitaskers' view on plant hormone biology. *FEBS letters* 593(13): 1415-1430.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., and Sangrador-Vegas, A. (2015). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44: 279-285.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). "Protein identification and analysis tools on the ExPASy server," in *The proteomics protocols handbook,* ed. J.M. Walker. (New York City, New York, United States: Humana Press), 571-607.

Hashemipetroudi, S.H., Mohammadi, S., and Kuhlmann, M. (2019). Analysis of expression pattern of genome and analysis of *HSP90* gene family in *Aeluropus littoralis* under salinity stress. *Crop Breed J* 11(31): 134-143.

Haxim, Y., Si, Y., Liu, X., Wen, X., Kahar, G., Ding, Y., Li, X., and Zhang, D. (2021). Genome-wide characterization of *HSP90* gene family in *Malus sieversii* and their potential roles in response to Valsa mali infection. *Forests* 12(9): 1232.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C., and Nakai, K. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res* 35: 585-587.

Hoter, A., El-Sabban, M.E., and Naim, H.Y. (2018). The HSP90 family: structure, regulation, function, and implications in health and disease. *Int J Mol Sci* 19(9): 2560.

Hu, B., Jin, J., Guo, A.-Y., Zhang, H., Luo, J., and Gao, G. (2014). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31(8): 1296-1297.

Hu, W., Hu, G., and Han, B. (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 Sci* 176(4): 583-590.

Jing, W., Fangjun, T., Chengliang, L., Xilu, Z., Lijun, O., Juntawong, N., Fei, W., Chunhai, J., Xuexiao, Z., and Wenchao, C. (2020). Genome-wide identification and analysis of *HSP90* gene family in pepper. *Acta Horticulturae Sinica* 47(4): 665.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., and Nuka, G. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9): 1236-1240.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7): 1870-1874.

Letunic, I., Doerks, T., and Bork, P. (2014). SMART: recent updates, new developments and status in 2015. *Nucleic acids res* 43(D1): D257-D260.

Li, W., Chen, Y., Ye, M., Wang, D., and Chen, Q. (2020). Evolutionary history of the heat shock protein 90 (*Hsp90*) family of 43 plants and characterization of Hsp90s in *Solanum tuberosum*. *Mol Biol Rep* 47(9): 6679-6691.

Mishra, R.C., and Grover, A. (2016). Constitutive over-expression of rice ClpD1 protein enhances tolerance to salt and desiccation stresses in transgenic *Arabidopsis* plants. *Plant Sci* 250: 69-78.

Mishra, S., Chaudhary, R., Pandey, B., Singh, G., and Sharma, P. (2023). Genome-wide identification and expression analysis of the GRAS gene family under abiotic stresses in wheat (*Triticum aestivum* L.). *Sci Rep* 13(1): 18705.

Pearl, L.H., and Prodromou, C. (2006). Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu Rev Biochem* 75(1): 271-294.

Picard, D. (2002). Heat-shock protein 90, a chaperone for folding and regulation. *Cell Mol Life* 59(10): 1640-1648.

Prasad, B.D., Goel, S., and Krishna, P. (2010). In silico identification of carboxylate clamp type tetratricopeptide repeat proteins in *Arabidopsis* and rice as putative co-chaperones of Hsp90/Hsp70. *Plos One* 5(9): e12761.

Reddy, R.K., Chaudhary, S., Patil, P., and Krishna, P. (1998). The 90 kDa heat shock protein (hsp90) is expressed throughout *Brassica napus* seed development and germination. *Plant Sci* 131(2): 131-137.

Sable, A., Rai, K.M., Choudhary, A., Yadav, V.K., Agarwal, S.K., and Sawant, S.V. (2018). Inhibition of heat shock proteins HSP90 and HSP70 induce oxidative stress, suppressing cotton fiber development. *Sci Rep* 8(1): 1-17.

Sangster, T.A., and Queitsch, C. (2005). The HSP90 chaperone complex, an emerging force in plant development and phenotypic plasticity. *Curr Opin Plant Biol* 8(1): 86-92.

Sarkar, N.K., Kim, Y.-K., and Grover, A. (2009). Rice sHsp genes: genomic organization and expression profiling under stress and development. *BMC Genom* 10(1): 1-18.

Scherf, U., Ross, D.T., Waltham, M., Smith, L.H., Lee, J.K., Tanabe, L., Kohn, K.W., Reinhold, W.C., Myers, T.G., and Andrews, D.T. (2000). A gene expression database for the molecular pharmacology of cancer. *Nat Genet* 24(3): 236-244.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11): 2498-2504.

Shelke, D.B., Chambhare, M.R., Nikalje, G.C., and Nikam, T. (2023). Improvement of soybean crop for yield, stress tolerance, and value-added products using a transgenic approach. *Adv Agric* 2023. doi: 10.1155/2023/8166928.

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

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., and Bork, P. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47(D1): D607-D613.

Vabulas, R.M., Raychaudhuri, S., Hayer-Hartl, M., and Hartl, F.U. (2010). Protein folding in the cytoplasm and the heat shock response. *Cold Spring Harb Perspect Biol* 2(12): a004390.

Vaughan, M.M., Block, A., Christensen, S.A., Allen, L.H., and Schmelz, E.A. (2018). The effects of climate change associated abiotic stresses on maize phytochemical defenses. *Phytochem Rev* 17: 37-49.

Voorrips, R. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93(1): 77-78.

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci* 9(5): 244-252.

Zhang, J., Li, J., Liu, B., Zhang, L., Chen, J., and Lu, M. (2013). Genome-wide analysis of the Populus *Hsp90* gene family reveals differential expression patterns, localization, and heat stress responses. *BMC Genom* 14(1): 532.

Zhang, K., He, S., Sui, Y., Gao, Q., Jia, S., Lu, X., and Jia, L. (2021). Genome-wide characterization of *HSP90* gene family in cucumber and their potential roles in response to abiotic and biotic stresses. *Front Genet* 12: 584886.

Zhang, L., Zhao, H.-K., Dong, Q.-l., Zhang, Y.-Y., Wang, Y.-M., Li, H.-Y., Xing, G.-J., Li, Q.-Y., and Dong, Y.-S. (2015). Genome-wide analysis and expression profiling under heat and drought treatments of *HSP70* gene family in soybean (*Glycine max* L.). *Front Plant Sci* 6: 773.

Zhang, M., Shen, Z., Meng, G., Lu, Y., and Wang, Y. (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(12): e0189187.

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2023 | Volume 11 | Issue 2

مقاله پژوهشی

تجزیه و تحلیل گسترده ژنومی خانواده ژنی *HSP90* و نقش آن‌ها در رشد و نمو سویا

سمیرا محمدی*، فیروزه سهروردی، قربانعلی نعمت‌زاده

پژوهشکده ژنتیک و زیست‌فناوری کشاورزی طبرستان، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران

چکیده: پروتئین شوک حرارتی ۹۰ کیلو دالتونی یا HSP90 از طریق چندین مکانیسم عملکردی، نقش مهمی در تنظیم تنش‌های زیستی و غیرزیستی بر عهده دارد. مطالعه حاضر با هدف انجام تجزیه و تحلیل جامعی از خانواده ژنی *HSP90* در سویا انجام شد. در مجموع، ۲۰ ژن *HSP90* در سویا شناسایی شدند و توزیع غیر یکنواختی روی ۱۳ کروموزوم نشان دادند. روابط فیلوژنتیکی نشان داد که ژن‌های واقع در زیرگروه‌های یکسان دارای ساختار اگزون – اینترون و تعداد اینترون‌های مشابه هستند. تجزیه و تحلیل مکان‌یابی سلولی نشان داد که تقریباً تمام ژن‌های *HSP90* واقع در گروه اول، در هسته یا سیتوپلاسم توزیع شده‌اند. ژن‌ها در گروه دوم، بیشتر در شبکه آندوپلاسمی واقع شده‌اند. ژن‌های *HSP90* واقع در گروه سوم نیز تقریباً در میتوکندری یا کلروپلاست یافت شده‌اند. Glyma14g219700، Glyma17g258700 و Glyma07G207600 بر اساس بیشترین ارتباط با سایر پروتئین‌ها، به‌عنوان پروتئین‌های هاب شناسایی شدند. علاوه بر این، ژن‌های *Glyma02g302500*، *Glyma08g332900*، *Glyma14g219700*، *Glyma17g258700* و *Glyma18g074100* سطوح بیان بالایی را در تمام بافت‌ها و در مراحل نموی مختلف نشان دادند. این یافته‌ها با ارائه اطلاعات کلی از طبقه‌بندی و تکامل خانواده ژنی *GmHSP90*، می‌تواند به شناسایی خصوصیات کارکردی ژن‌های *HSP90* در رشد و نمو سویا کمک نماید.

کلمات کلیدی: پروتئین شوک حرارتی، پروفایل بیان ژن، ساختار ژنی، شبکه برهم‌کنش پروتئین – پروتئین، روابط فیلوژنتیکی، مکان‌یابی سلولی.

ویراستار علمی
دکتر پرویز حیدری،
دانشگاه صنعتی شاهرود، ایران

تاریخ
دریافت: ۱۶ خرداد ۱۴۰۱
پذیرش: ۲۷ اردیبهشت ۱۴۰۳
چاپ: ۶ خرداد ۱۴۰۳

نویسنده مسئول
دکتر سمیرا محمدی
s.mohamadi@stu.sanru.ac.ir

ارجاع به این مقاله
Mohammadi, S., Sohrevardi, F. and Nematzadeh, G.A. (2023). Genome-wide analysis of the *HSP90* gene family and their roles in soybean growth and development . *J Plant Mol Breed 11* (2): 119-132.
doi: 10.22058/JPMB.2024.555256.1256.