

MOLECULAR EVOLUTION AND DIVERSITY OF SMALL HEAT SHOCK PROTEINS GENES IN PLANTS

WASEEM SAFDAR^{1*}, HAMID MAJEED², BARKAT ALI³ AND ISHRAT NAVEED⁴

¹Department of Biochemistry, PMAS Arid Agriculture University, Rawalpindi, Pakistan,

²Food Science Department, Jiangnan University, Wuxi, China,

³National Agriculture Research Council, Islamabad, Pakistan,

⁴Department of Biotechnology, Quaid-i-Azam University, Islamabad, Pakistan

*Corresponding author E-mail: waseem.safdar@yahoo.com; +92-0331-5223268

Abstract

Small heat shock proteins (sHSPs) are the most abundant proteins and considered as *Cinderella* of molecular chaperon world. The present study is to understand the evolutionary process that led to the diversification of sHSPs specific to plants because of dramatic daily fluctuation in the temperature and other environmental factors which may prompt more efficient chaperon activity of sHSPs. For this purpose mRNA and protein sequences of 62 different plant sHSPs was mined from different databases and analyzed with Clustal W and MEGA 5 Beta # 7 software. Two Neighbor Joining (NJ) and two Dendrogram large congruent trees were obtained from the phylogenetic analysis of mRNA and amino acids. These analyses reveal that sHSPs encoded by one gene family are similar to each other even in different plant species. However, sHSPs belonging to different gene families show very low sequence similarity even in same plant species. These analyses also suggest that gene duplication and mutation play an important role in the evolution and diversification of sHSP.

Introduction

Every living organism responds to temperatures greater than their optimal growth temperature with increased expression of a defined class of functionally related proteins called heat shock proteins (HSPs). There are six structurally conserved distinct families of HSPs, HSP100, HSP90, HSP70, HSP60 (Chaperonins), small HSPs (sHSPs) (~17-30 kDa) and ubiquitin (8.5 kDa) (Waters *et al.*, 1996; Vierling, 1997). sHSPs are the most abundant proteins and considered as *Cinderella* of HSP world (Van-den-Ussel *et al.*, 1999, Mahmood *et al.*, 2010) because sHSPs are also expressed under normal conditions in several organisms such as vertebrates (Liopsis *et al.*, 1998), *Drosophila* (McKenzie and Meselson 1977) and plants (Carranco *et al.*, 1999). sHSPs are divided into 6 classes; three classes (CI, CII and CIII) are in the nucleus or cytosole and the other three (CIV, CV and CVI) are in the mitochondria chloroplasts and endoplasmic reticulum (Mahmood *et al.*, 2010). This diverse family contains a α -crystalline structural domain. The significance of this α -crystalline domain is its 90 amino acid conserved region found in the C-terminal end of the sHSPs while N-terminal side contains unconserved variable length amino acid sequences (Basha *et al.*, 2010). The α -crystalline domain contains several beta-strands organized into two beta-sheets responsible for dimer formation, the basic building block of most sHSPs. The amino-terminal extension modulates oligomerization, subunit dynamics and substrate binding, whereas the flexible carboxy-terminal extension promotes solubility, chaperoning and oligomerization, the latter by inter-subunit linkage (Sun & MacRae, 2005).

sHSPs organized into large, sphere-like structures commonly consisting of 12 or 24 subunits forming two distinct types of octahedral oligomers. During the stress, some structural changes occurs resulting in increased chaperon activity (Haslbeck *et al.*, 2008). sHSP-substrate complexes with varying stability and composition can protect all substrate equally, and substrate protection is

not correlated with sHSP oligomeric stability (Basha *et al.*, 2006).

sHSPs recognized damaged proteins and sort them into repaired form, degraded or transported across the membranes by helping them to stabilize (Nakamoto & Vigh, 2007; Toth *et al.*, 2010). sHSPs specifically interact with the membrane lipid and control fluidity and permeability (Porta *et al.*, 2010; Horváth & Vigh, 2010.). High temperature treatment increased the expression of HSP which inhibit tobacco mosaic viral RNA synthesis and resume when temperature decreased (Arif *et al.*, 2005). Iqbal *et al.*, (2010) reported that Rice sHSPs not only protect the high molecular weight proteins from heat induced but also enhance their impact in chilling tolerance. These proteins also play a fundamental role in the pathology of human diseases like cancer, neurodegenerative diseases, diabetes, prevention of apoptosis after ischemic injury, cardiac myocyte function, platelet aggregation and skeletal muscle function (Fan *et al.*, 2005; Nakamoto & Vigh, 2007). In aged-muscle sHSPs shows a dramatic increase in expression for essential cellular response to fiber aging and might therefore be a novel therapeutic option to treat sarcopenia of old age (Doran *et al.*, 2007).

sHSPs are also expressed during normal developmental stages like fruit maturation, pollen growth, germination and embryogenesis for the recovery of damaged and newly synthesized proteins (Mehmood *et al.*, 2010). In *Pisum sativum*, *Arabidopsis thaliana*, *Zea mays*, *Triticum aestivum*, sunflower, alfalfa, tobacco and tomato sHSPs are produced in a stage-specific fashion suggesting that certain sHSP genes may play specific roles in early, others during later stages of development (waters *et al.*, 1996; Waters & Rioflorida, 2007).

Gene duplication and mutation play an important role in the evolutionary process and diversification of species (Wright and Gaut, 2005). Advantageous mutation in duplicated genes may derive the gene to a new function

and the original function remains preserved in another copy (Wu *et al.*, 1995). Genetic changes and diversification in sHSPs specific to plants is of particular interest because dramatic daily fluctuation in the temperature and other environmental factors may prompt more efficient chaperon activity of sHSPs (Kriehuber *et al.*, 2010). The sHSPs encoded by one gene family are similar to each other even in different plant species. The sequence similarity can be up to 93% and identity up to 85% (Vierling, 1991). However, sHSPs of one plant species belonging to different families show very low sequence similarity (50-75%), and identity usually below 50%. This applies not only for the comparisons of sHSPs between divergent species, but also for comparisons between different classes of plant sHSPs (Waters *et al.*, 1996). Even under same heat shock conditions different sHSPs of are accumulated in different variety of the same species (Iqbal *et al.*, 2010).

Naz *et al.*, (2006) reported that the most of the high and low molecular weight HSPs remain conserved in wild and hybrid rice. The conserved region of mitochondrial sHSP (MT-sHSP) genes among diploid genome of cotton species contain one single nucleotide polymorphism per 14 bp indicating higher degree of evolution among MT-sHSP of different cotton species (Shaheen *et al.*, 2009). Phylogenetic analysis of the chloroplast localized small heat shock proteins (CP sHSPs) from angiosperms, with other plant CP sHSPs and eukaryotic, archaeal, and bacterial sHSPs shows that the CP sHSPs are not closely related to the cyanobacterial sHSPs (Waters & Vierling, 1999). Yildiz & Terzi (2008) reported that heat shock treatment may be helpful to determine the genetic variability in chlorophyll accumulation and thermotolerance of cereals. So molecular and genetic basis of heat tolerance is strongly required to elucidate in cereals for identification of beneficial genes and alleles. These genes may be utilized in molecular breeding programs to produce superior cereal cultivars (Yildiz & Terzi, 2008). In the present study phylogenetic analysis of different sHSPs was carried out to analyze their evolutionary relationship among different plant species on the bases of mRNA and protein sequences.

Materials and Methods

Data collection: mRNA sequences of 62 different plants sHSPs were downloaded from the literature and NCBI Genbank database. Their species name and accession numbers are given in the Table 1.

Evolutionary analysis: Phylogenetic analysis of these mRNA sequences was conducted using different bioinformatic tools. Sequence alignment and Dendrogram tree was dragged by using online available program Clustal W (<http://align.genome.jp/>). Statistical selection pairing, by applying Tajima's test (Tajima, 1989) and Neighbour Joining (NJ) (Saitou & Nei 1987) methods were used for phylogenetic reconstruction and the NJ *p-distance* model was used for distance analysis (Nei & Kumar 2000). Base statistical robustness was performed

by using 500 bootstrap repeats for the validity of results and the whole process was developed by MEGA 5 Beta # 7 software (Kumar *et al.*, 2004). Conceptual translated amino acid sequences of plants sHSPs were also subjected to Clustal W and MEGA 5 Beta # 7 software for phylogenetic analysis.

Results and Discussions

sHSP are considered to be the most conserved among protein families. To understand the molecular evolution and diversity of sHSP genes, mRNA and protein sequences were mined from different data bases including GenBank. Sequence alignment and Dendrogram tree was dragged by using online available program Clustal W. Sequence alignment of mRNA (Fig. 1) and protein sequences (Fig. 2). HSP 16.9 that belong to different species of genus *Triticum*, *Pennisetum* and *Oryza* showed highly conserved sequences of mRNA and Proteins. Similarly HSP 17.6C-II, HSP 20.2 and HSP 17.4 that belong to same genus *Lycopersicon* also showed highly conserved sequences of mRNA and Proteins. This indicates that different sHSPs that belong to same species showed high sequence similarity and sHSP that belongs to different species remain conserved during evolution. The dragged dendrograms (Fig. 3) were the rooted trees; with tree topologies indicated the bifurcating internodes with asymmetrical branching structures. Dendrogram of the studied accessions was divided into III clusters. All clusters evolved almost at the same time with branch topologies indicated a slow and stable evolution; one possible reason might be the gradual increase in temperature of the environment. In cluster I and II sHSPs belongs to genus *Triticum* form a distinct evolutionary conserved group.

The overall mean distance with standard error 0.02 among all studies mRNA sequences are 0.74. This showed the mean pairwise distance and standard error for the set of sequences under study. When the values above were recalculated using the deduced protein data, the overall mean distance became higher to 0.90 (with Std. Err. 0.02) among protein sequences. Standard error estimate(s) were obtained by a bootstrap procedure (500 replicates). Mean diversity of entire studied population from mRNA sequences is 3.98 and from protein sequences is 0.83. For the entire population, the mean diversity is calculated by the formula:

$$\pi T = \frac{q}{q-1} \sum_{i=1}^q \sum_{j=1}^q X_i X_j d_{ij}$$

Where X_i is the estimate of average frequency of the i -th allele in the entire population, and q is the number of different sequences in the entire sample. Tajima's Neutrality Test was also performed (Table 2). The analysis involved 62 nucleotide and amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 151 positions in the final dataset of mRNA and 130 positions in the final dataset of proteins. Evolutionary analyses were

Table 1. List of plants sHSPs along with their mRNA and protein accession numbers and species name used in this study.

S. No.	Species	Proteins	Accession Numbers	
			Nucleotides	Proteins
1.	<i>Agrostis stolonifera</i>	HSP16.5	AF007762	AAC01560
2.	<i>Arabidopsis thaliana</i>	HSP17.6-II	X63443	CAA45039
3.	<i>Arabidopsis thaliana</i>	HSP21	X54102	CAA38036
4.	<i>Arabidopsis thaliana</i>	HSP17.6	X89504	CAA61675
5.	<i>Arabidopsis thaliana</i>	HSP23.5	NM_124523	NP_199957
6.	<i>Carica papaya</i>	HSP17.5	AY387588	AAR25848
7.	<i>Carica papaya</i>	HSP17.7	AY242075	AAP73794
8.	<i>Daucus carota</i>	HSP18.0	X53852	CAA37848
9.	<i>Daucus carota</i>	HSP17.8	X53851	CAA37847
10.	<i>Funaria hygrometrica</i>	HSP18.3	AF089846	AAD09185
11.	<i>Funaria hygrometrica</i>	HSP16.4	AF089845	AAD09184
12.	<i>Glycine max</i>	HSP23.9	U21722	AAB03096
13.	<i>Glycine max</i>	HSP22.0	X07188	CAA30168
14.	<i>Helianthus annuus</i>	HSP17.9	Z29554	CAA82653
15.	<i>Hordeum vulgare</i>	HSP18.0	X64561	CAA45862
16.	<i>Hordeum vulgare</i>	HSP17.0	Y07844	CAA69172
17.	<i>Lycopersicon esculentum</i>	HSP21.0	LEU66300	AAB07023
18.	<i>Lycopersicon esculentum</i>	HSP17.6CI	AF123257	AAD30454
19.	<i>Lycopersicon esculentum</i>	HSP17.8	AF123256	AAD30453
20.	<i>Lycopersicon esculentum</i>	HSP17.7	AF123255	AAD30452
21.	<i>Lycopersicon esculentum</i>	HSP17.6CII	LEU72396	LEU72396
22.	<i>Lycopersicon peruvianum</i>	HSP17.4	AY608694	AAT36481
23.	<i>Lycopersicon peruvianum</i>	HSP20.2	AJ225049	CAA12390
24.	<i>Medicago sativa</i>	HSP17.0	X98617	CAA67206
25.	<i>Medicago sativa</i>	HSP18.1	X58710	CAA41546
26.	<i>Medicago sativa</i>	HSP18.2	X58711	CAA41547
27.	<i>Nicotiana tabacum</i>	HSP26.0	D88584	BAA29064
28.	<i>Oryza sativa</i>	HSP16.9	X60820	CAA43210
29.	<i>Oryza sativa</i>	HSP CII	DQ180746	ABA29610
30.	<i>Oryza sativa</i>	HSP17.8	EU715987	ACH72824
31.	<i>Oryza sativa</i>	HSP26.0	AB020973	BAA78385
32.	<i>Pennisetum glaucum</i>	HSP16.9	X94192	CAA63902
33.	<i>Pennisetum glaucum</i>	HSP17.9	GQ121016	ACR78191
34.	<i>Petroselinum crispum</i>	HSP17.9	X95716	CAA65020
35.	<i>Petunia hybrid</i>	HSP21.0	X54103	CAA38037
36.	<i>Picea glauca</i>	HSP17.0	L47717	AAB01561
37.	<i>Picea glauca</i>	HSP23.5	L47741	AAB01557
38.	<i>Pisum sativum</i>	HSP22.0	X86222	CAA60120
39.	<i>Pisum sativum</i>	HSP17.7	M33901	AAA33670
40.	<i>Pisum sativum</i>	HSP21.0	X07187	CAA30167
41.	<i>Rosa hybrid</i>	HSP17.5	EF157600	ABO84842
42.	<i>Triticum aestivum</i>	HSP26.6B	X67328	CAA47745
43.	<i>Triticum aestivum</i>	HSP17.3	X58279	CAA41218
44.	<i>Triticum aestivum</i>	HSP26.6	X58280	CAA41219
45.	<i>Triticum aestivum</i>	HSP16.9B	X64618	CAA45902
46.	<i>Triticum aestivum</i>	HSP17.8	AF350423	AAK51797
47.	<i>Triticum aestivum</i>	HSP16.9A	EU649679	ACD03605
48.	<i>Triticum aestivum</i>	HSP16.9C	L14444	AAA34294
49.	<i>Triticum durum</i>	HSP26.5	AJ971373	CAI96515
50.	<i>Triticum durum</i>	HSP17.6	AJ971359	CAI96501
51.	<i>Triticum monococcum</i>	HSP16.8	AM709755	CAM96546
52.	<i>Triticum monococcum</i>	HSP17.0	AM709756	CAM96547
53.	<i>Triticum monococcum</i>	HSP16.9A	AM709757	CAM96548
54.	<i>Triticum monococcum</i>	HSP16.9B	AM709758	CAM96549
55.	<i>Triticum monococcum</i>	HSP26.6	AJ971374	CAI96516
56.	<i>Triticum monococcum</i>	HSP23.5	AJ971365	CAI96507
57.	<i>Triticum turgidum</i>	HSP26.8	AJ971372	CAI96514
58.	<i>Vitis vinifera</i>	HSP17.4	GU169701	ACZ48684
59.	<i>Vitis vinifera</i>	HSP17.3	GU169700	ACZ48683
60.	<i>Zea mays</i>	HSP18.0	X54075	CAA38012
61.	<i>Zea mays</i>	HSP17.5	NM_001154982	NP_001148454
62.	<i>Zea mays</i>	HSP26.0	L28712	AAA33477

Table 2. Tajima's neutrality Test for mRNA and proteins.

Sequences	m	S	p_s	Θ	π	D
mRNA	62	151	1.000000	0.212935	0.737320	8.624317
Proteins	62	130	1.000000	0.212935	0.896111	11.196385

m = number of sites, S = Number of segregating sites, $p_s = S/m$, $\Theta = p_s/a_1$, π = nucleotide diversity, and D is the Tajima test statistic

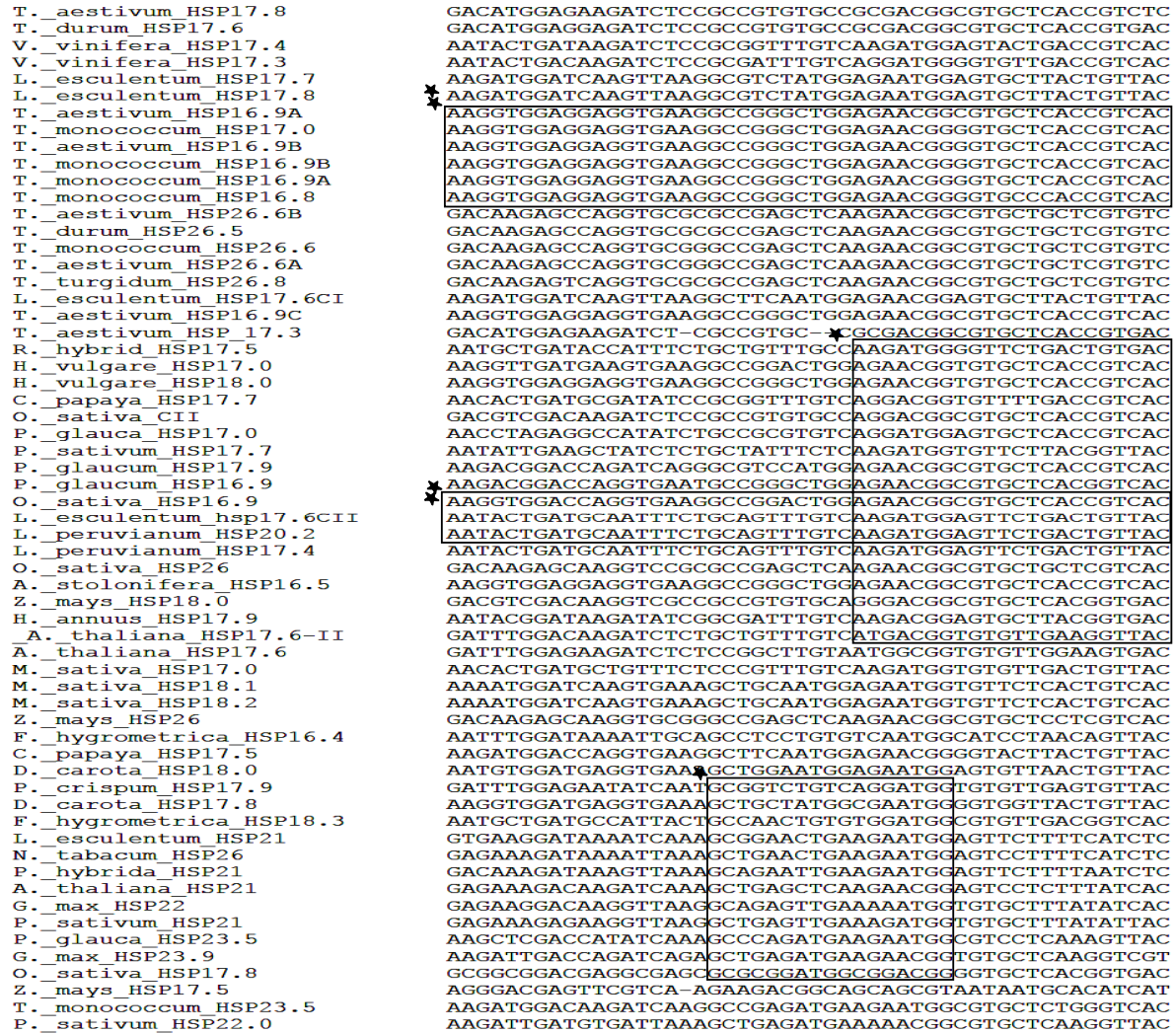


Fig. 1. mRNA sequence alignment of 62 sHSP from different plant species. Box mark conserved region.

*Highly conserved region; **Completely conserved region.

conducted in MEGA 5 (Saitou & Nei, 1987; Nei & Kumar, 2000; Tamura *et al.*, 2007). Maximum Likelihood fits of 24 different nucleotide substitution models was also sketched.

The evolutionary history was inferred using the Neighbor-Joining method (Figs. 4 and 5). The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were computed using the p-distance method and are in the units of the number of base and amino acid differences per site. The analysis involved 62 nucleotide and amino acid sequences. All positions containing gaps and missing data were eliminated. There

were a total of 151 positions in the final dataset of mRNA and 130 positions in the final dataset of proteins. Evolutionary analyses were conducted in MEGA 5 (Saitou and Nei, 1987; Nei and Kumar, 2000; Tamura *et al.*, 2007). The evolutionary history was also inferred using the Maximum Parsimony method. For mRNA and Proteins, the consistency index is 0.107743 & 0.465045, the retention index is 0.330649 & 0.541383 and the composite index is 0.035625 & 0.251767 for all sites respectively. All Evolutionary analysis revealed that different sHSPs evolve prior to the divergence of the plant species. The branch topology of cytosolic, mitochondrial, chloroplast and Endoplasmic reticulum families are highly supported by Bootstrap analysis of mRNA and Protein sequences (Figs. 4 and 5). The molecular evolution

```

N. tabacum_HSP26      KDDENEIKMRFDMPLGSLKDEVKVSVEDDLLVIKGEY-----K
P. hybrida_HSP21     KDDENEIKMRFDMPLGSLKDEVKVSVEDDVLVIKGEH-----K
L. esculentum_HSP21  HDDENEIKMRFDMPLGSLKDEVKVSVENDMLVIKGEH-----K
G. max_HSP22         KDEEHEIRMRFDMPGLAKEDVKVSVEDDMLVIKGGH-----K
P. sativum_HSP21     KDEEHEIRMRFDMPGVSKEDEVKSVEDDVLVIKSDH-----R
A. thaliana_HSP21    *KEEHEIKMRFDMPLGSLKDEVKISVEDNVLVIKGEQ-----K
T. durum_HSP26.5     *MEDEKEVKMRFDMPLGSLREEVKVMVEDDALVIRGEH-----K
T. monococcum_HSP26.6 MEDDKEVKMRFDMPLGSLREEVKVMVEDDALVIRGEH-----K
T. aestivum_HSP26.6A MEDEKEVKMRFDMPLGSLREEVKVMVEDDALVIRGEH-----K
T. aestivum_HSP26.6A MEDEKEVKMRFDMPLGSLREEVKVMVEDDALVIRGEH-----K
T. turgidum_HSP26.8 *MEDDKEVKMRFDMPLGSLREEVKVMVEGDALVIRGEH-----K
Z. mays_HSP26        VEDEKEVKMRFDMPLGLARDEVKVMVEDDTLVIRGEH-----K
O. sativa_HSP26      MEDDKEVKMRFDMPLGSLREEVKVMVEDDALVIRGEH-----K
G. max_HSP23.9       RETEDALHLRVDMPGLAKEDVKISVEQNTLTIKGE-----G
P. sativum_HSP22     RETEDALFLRLDMPGLKEDVKISVEQNTLTIKGE-----G
A. thaliana_HSP23.5 KEKDDALHLRLDMPGLSREDEVKLALEQNTLVIIRGEG-----E
P. glauca_HSP23.5    *VEDEKALHLRVDMPGLGKEDIKVYAEENALVIKGE-----
T. aestivum_HSP17.3 KEIPGAYAFVVDMPGLGSGDIKVQVEDERVLVVISG-----
T. durum_HSP17.6     KEIPGAYAFVVDMPGLGSGDIKVQVEDERVLVVISG-----
T. aestivum_HSP17.8 KEIPGAYAFVVDMPGLGSGDIKVQVEDERVLVVISG-----
O. sativa_CII        KDIPGAYAFVVDMPGLKSSDIKVQVEERLVLVISG-----
Z. mays_HSP18.0      KEIPGAYAFVVDMPGLGTGDIRVQVEDERVLVVISG-----
L. esculentum_hsp17.6CII KEYPNSYVFVVDMPGLKSGDIKVQVEDNVLVISG-----
L. peruvianum_HSP20.2 KEYPNSYVFVVDMPGLKSGDIKVQVEDNVLVISG-----
L. peruvianum_HSP17.4 KEYPDSYVFVVDMPGLKSGDIKVQVEDNVLVISG-----
C. papaya_HSP17.7    KEYPNSYVFVVDMPGLKSGDIKVQVEDNVLQISG-----
P. sativum_HSP17.7   KEHPNSYVFVVDMPGVKSGDIKVQVEDENVLLISG-----
V. vinifera_HSP17.4 KEYPNSYTFIVDMPGLKSGDIKVQVEDDNVLVISG-----
V. vinifera_HSP17.3 KEYPNSYAFIIVDMPGLKSGDIKVQVEDDNVLVISG-----
R. hybrid_HSP17.5    KEIPNSYVFVVDMPGLKSGDIKVQVEDDNVLVISG-----
M. sativa_HSP17.0    KENPNSYVFVVDMPGLKSGDIKVQVEDDNVLVISG-----
P. crispum_HSP17.9   KEYPNSYVFVVDMPGLKSGDIKVQVEDNVLVVISG-----
H. annuus_HSP17.9    KECPNNSYVFIVDMPGLKSGDIKVQVEDRNVLVISG-----
A. thaliana_HSP17.6-II IEHPNAYAFVVDMPGIKGDIEKVQVEDNVLVVISG-----
A. thaliana_HSP17.6 IEHPDAYVFVVDMPGIKGDIEKVQVIENENVLVVISG-----
F. hygrometrica_HSP16.4 KELPDAYIFVADMPGLKSADVKVQLENDNVLVIGG-----
P. glauca_HSP17.0    KEYPNSYVFIIIDMPGLKSNDIKVQVEDENVLISG-----
F. hygrometrica_HSP18.3 KEIKDAYLFFVADVPGLQKTDIEVQVENENVLTMRG-----
T. monococcum_HSP16.9B *KETPEAHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
T. aestivum_HSP16.9B KETPEAHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
T. monococcum_HSP16.8 KEAPEAHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
T. monococcum_HSP17.0 KETPEAHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
T. monococcum_HSP16.9A *KETPEAHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
T. aestivum_HSP16.9A KETPEAHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
T. aestivum_HSP16.9C KETPEAHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
H. vulgare_HSP17.0   GRRLEAHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
A. stolonifera_HSP16.5 KETPEAHVFKADLPGVKKEEVKVEVEGGNVLVVISG-----
P. glaucum_HSP16.9   KETPEVHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
O. sativa_HSP16.9    KETPESHVFKADLPGVKKEEVKVEVEEGNVLVVISG-----
H. vulgare_HSP18.0   GRRVAHVFKADLPGVKKEEVKVEVEDGNVLVVISG-----
L. esculentum_HSP17.7 KETPEAHVFKVDLPLGKKEEVKVEVEEDRVLQISG-----
L. esculentum_HSP17.8 KETPEAHVFKVDLPLGKKEEVKVEVEEDRVLQISG-----
L. esculentum_HSP17.6CI KETPEAHVFKADLPLGKKEEVKVEVEEDRVLQISG-----
M. sativa_HSP18.1    KETPEAHVFKADLPLGKKEEVKVEVEDDRVLQISG-----
M. sativa_HSP18.2    KETPEAHVFKADLPGMKKEEVKVEVEDDRVLQISG-----
C. papaya_HSP17.5    EETPEAHVFRADLPLGKKEEVKVELEDDRVLQISG-----
D. carota_HSP18.0    KETPQAHVFKADLPLGKKEEVKVEVEEGKVLQISG-----
D. carota_HSP17.8    KETPQAHVFKADLPLGKKEEVKVELEEGKVLQISG-----
P. glaucum_HSP17.9   KETPEAHVFKADVPGLKKEEVKVEVEDGNVLQISG-----
O. sativa_HSP17.8    RETPVAHVFEMLPLGLAKDQVAVEVVDGHIILRVRAGGEHEDANNAKAGK
T. monococcum_HSP23.5 KEDDDAVYLVKVPMPGLTKEHVEVRADKNILVIKGE-----
Z. mays_HSP17.5     VESPREYAFVLDVPLGSLKSDIQVTLEEDRVLVVKGG-----

```

Fig. 2. Protein sequence alignment of 62 sHSP from different plant species. Box mark conserved region.

*Highly conserved region; **Completely conserved region.

of the sHSPs in plants can be considered as the best example of gene evolution. Darwinian selection reflects different selective constraints for sequence evolution rate among members of a gene family (Waters *et al.*, 1996). Variation in branch length indicates difference in rate of evolution. The phylogenetic relationships reveal that sHSPs are evolved due to gene duplication, sequence divergence and gene conversion at different levels (Waters & Rioflorido, 2007). Phylogenetic analysis of a representative 62 sHSPs from plants reveals a close relationship between plant species, which appears to exhibit polydispersity (Fu *et al.*, 2006). Phylogenetic analysis of the chloroplast localized small heat shock proteins (CP sHSPs) from angiosperms, with other plant

CP sHSPs and eukaryotic, archaeal, and bacterial sHSPs shows that the CP sHSPs are not closely related to the cyanobacterial sHSPs (Waters & Vierling, 1999). Thus, they most likely evolved via gene duplication from a nuclear-encoded cytosolic sHSP and not via gene transfer from the CP endosymbiont (Vierling, 1991). The evolutionary relationships among all of the plant sHSPs identified and homologues from bacteria and other eukaryotes are consistent with the hypothesis that the land plant chloroplast and mitochondrion sHSPs did not originate from the endosymbionts of the chloroplast and mitochondria (Waters & Rioflorido, 2007). This tremendous diversification of small heat shock proteins in plants may reflect adaptations to stresses unique to plants

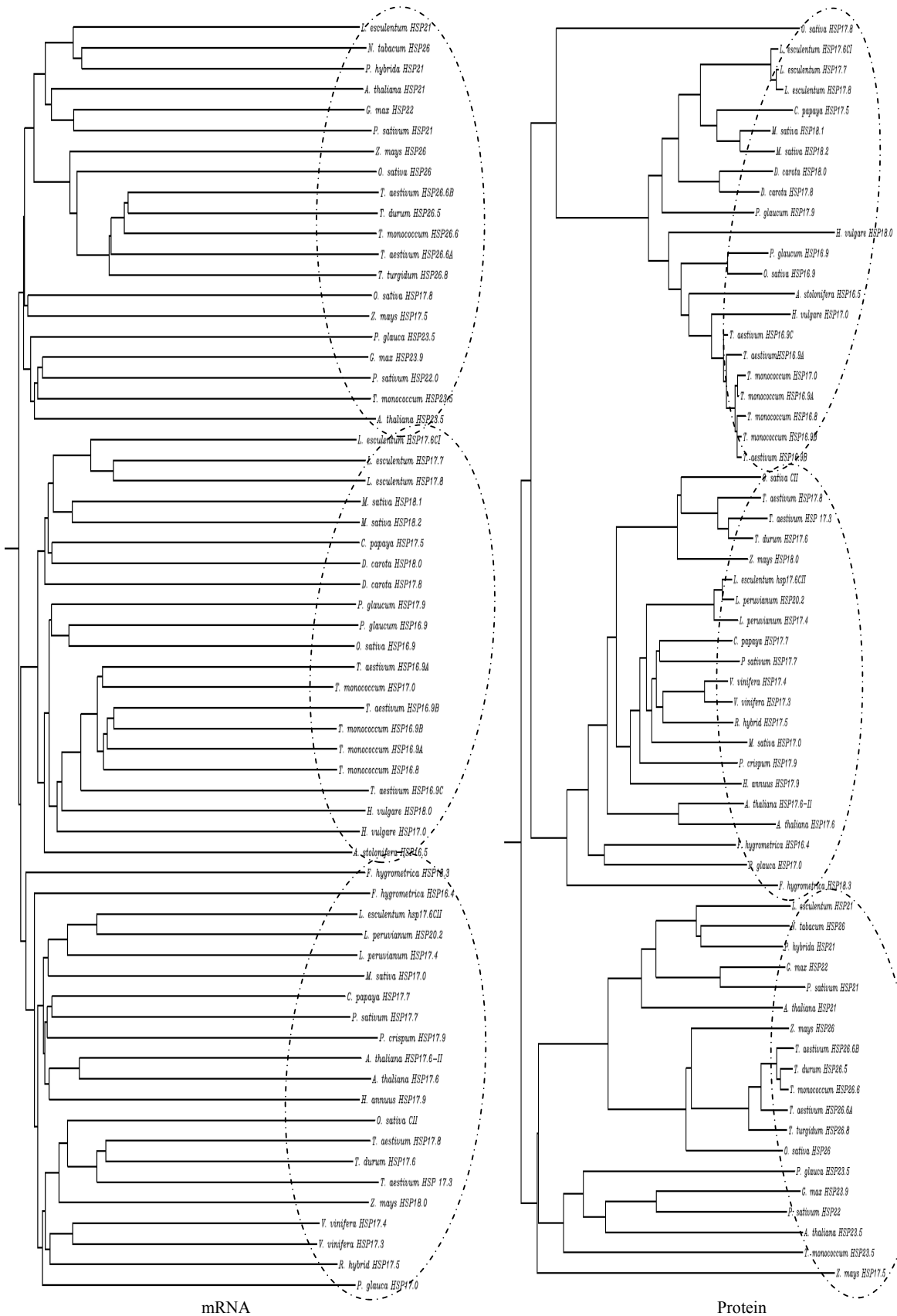


Fig. 3. Dendrogram tree of 62 mRNA and protein sequences with branch length.

(Vierling, 1991). The gene duplication that gave rise to the sHSP genes families certainly occurred before the divergence of the monocots and dicots, a minimum of 150 million years ago (Doyle & Donoghue, 1993).

References

- Arif, M., M. Ibrahim, A. Ahmad and S. Hassan. 2005. Elimination of *Citrus tristeza* Closterovirus from Citrus bud-wood through thermotherapy. *Pak. J. Bot.*, 37(2): 423-430.
- Basha, E., C. Jones, V. Wysocki, E. Vierling. 2010. Mechanistic differences between two conserved classes of small heat shock proteins found in the plant cytosol. *J. Biol. Chem.* 285: 11489-11497.
- Basha, E., K.L. Friedrich and E. Vierling. 2006. The N-terminal arm of small heat shock proteins is important for both chaperone activity and substrate specificity. *J. Biol. Chem.*, 281(52): 39943-39952.
- Carranco, R., C. Almoquera and J. Jordano. 1999. An imperfect heat shock element and different upstream sequences are required for the seed-specific expression of a small heat shock protein gene. *Plant Physiol.*, 121: 723-730.
- Doran, P., J. Gannon, O. K. Connell and K. Ohlendieck. 2007. Aging skeletal muscle shows a drastic increase in the small heat shock proteins α -crystallin/HspB5 and cvHsp/HspB7. *Eur. J. Cell Biol.*, 86(10): 629-640.
- Doyle, J.A. and M.J. Donoghue. 1993. Phylogenies and angiosperm diversification. *Paleobiol.*, 19: 141-67.
- Fan, G.C., X. Ren, J. Qian, Q. Yuan, P. Nicolaou, Y. Wang, W.K. Jones, G. Chu and E.G. Kranias. 2005. Novel cardioprotective role of a small heat-shock protein, Hsp20, against ischemia/reperfusion injury. *Circulation*, 111(14): 1792-9.
- Fu, X., W. Jiao and Z. Chang. 2006. Phylogenetic and biochemical studies reveal a potential evolutionary origin of small heat shock proteins of animals from bacterial class A. *J. Mol. Evol.*, 62: 257-266.
- Haslbeck, M., A. Kastenmuller, J. Buchner, S. Weinkauff and N. Braun. 2008. Structural Dynamics of Archaeal Small Heat Shock Proteins. *J. Mol. Biol.*, 378(2): 362-374.
- Horváth, I. and L. Vigh. 2010. Stability in times of stress. *Nature*, 463(7280): 436-438
- Iqbal, N., S. Farooq, R. Arshad and A. Hameed. 2010. Differential accumulation of high and low molecular weight heat shock proteins in Basmati rice (*Oriza sativa* L.) cultivars. *Genet. Resour. Crop. Evol.*, 57: 65-70.
- Kriehuber, T., T. Rattei, T. Weinmaier, A. Bepperling, M. Haslbeck and J. Buchner. 2010. Independent evolution of the core domain and its flanking sequences in small heat shock proteins. *FASEB J.*, 24(10): 3633-3642.
- Kumar, S., K. Tamura and M. Nei. 2004. Mega-3, integrated software for molecular evolutionary genetics analysis and sequence alignment. *Bri. Bioinform.*, 5: 150-163.
- Liossis, S.N., C.S. Via and G.C. Tsokos. 1998. The alter ego of heat shock proteins. *Clin. Immunol. Immunopathol.*, 86: 235-236.
- Mahmood, T., W. Safdar, B.H. Abbasi and S.M.S. Naqvi. 2010. An overview on the small heat shock proteins. *Afr. J. Biotechnol.*, 9(7): 727-949.
- McKenzie, S.L. and M. Meselson. 1977. Translation *in vitro* of *Drosophila* heat-shock messages. *J. Mol. Biol.*, 117: 279-283.
- Nakamoto, H. and L. Vigh. 2007. The small heat shock proteins and their clients. *Cell. Mol. Life Sci.*, 64: 294-306.
- Naz, F., S. Farooq, R. Arshad, M. Afzaal and M. Akram. 2006. Discriminating upland and lowland rice genotypes through proteomic approach. *Pak. J. Bot.*, 38(5): 1731-1738.
- Nei M. and S. Kumar. 2000. Molecular Evolution and Phylogenetics. *Oxford University Press, New York*, pp. 241-287.
- Porta, A., Z. Török, I. Horvath, S. Franceschelli, L. Vigh and B. Maresca. 2010. Genetic modification of the Salmonella membrane physical state alters the pattern of heat shock response. *J. Bacteriol.*, 192(7): 1988-1998.
- Saitou, N. and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406-425.
- Shaheen, T., M. Asif, Y. Zafar and Mehboob-ur-rahman. 2009. Single nucleotide polymorphism analysis of MT-sHSP Gene of *Gossypium arboreum* and its relationship with other diploid cotton genomes, *G. hirsutum* and *Arabidopsis thaliana*. *Pak. J. Bot.*, 41(1): 177-183.
- Sun, Y. and T.H. MacRae. 2005. Small heat shock proteins: Molecular structure and chaperone function. *Cell Mol. Life Sci.*, 62(21): 2460-2476.
- Tajima, F. 1989. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics*, 123: 585-595.
- Tamura, K., J. Dudley, M. Nei and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, 24: 1596-1599.
- Toth, M.E., S. Gonda, L. Vigh and M. Santha. 2010. Neuroprotective effect of small heat shock protein, Hsp27, after acute and chronic alcohol administration. *Cell Str. Chaperon.*, 15(6): 807-817.
- Van-den-Ussell, P., D.G. Norman and R.A. Quinlan. 1999. Molecular chaperones: Small heat shock proteins in the limelight. *Current Biol.*, 9(3): 103-105.
- Vierling, E. 1991. The heat shock response in plants. *Ann. Rev. Plant Physiol Plant Mol. Biol.*, 4: 579-620.
- Vierling, E. 1997. The small heat shock proteins in plants are members of an ancient family of heat induced proteins. *Acta Physiol. Plan.*, 19(4): 539-547.
- Waters, E.R., G.J. Lee and E. Vierling. 1996. Evolution, structure and function of the small heat shock proteins in plants *J. Exp. Bot.*, 47(296): 325-338.
- Waters, E. R. and E. Vierling. 1999. Chloroplast small heat shock proteins: Evidence for atypical evolution of an organelle-localized protein. Proceedings of the National Academy of Sciences of the United States of America, 96(25): 14394-14399.
- Waters, E.R. and I. Rioflorida. 2007. Evolutionary analysis of the small heat shock proteins in five complete algal genomes. *J. Mol. Evol.*, 65: 162-174.
- Waters, E.R., G.J. Lee and E. Vierling. 1996. Evolution, structure and function of the small heat shock proteins in plants. *J. Exp. Bot.*, 47(296): 325-338.
- Wright, S.I. and B.S. Gaut. 2005. Molecular population genetics and the search for adaptive evolution in plants. *Mol. Biol. Evol.*, 22: 506-519.
- Wu, C. 1995. Heat shock transcription factors: structure and regulation. *Ann. Rev. Cell Develop. Biol.*, 11: 441-69.
- Yildiz, M. and H. Terzi. 2008. Evaluation of acquired thermotolerance in wheat (*triticum aestivum* and *T. Durum*) cultivars grown in turkey *Pak. J. Bot.*, 40(1): 317-327.