

Different evolutionary lineages of large and small heat shock proteins in eukaryotes

Li-Hua Huang¹, Hong-Sheng Wang¹, Le Kang¹

¹State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, B5 Datum Road, ChaoYang District, Beijing 100101, China

Cell Research (2008) **18**:1074-1076. doi: 10.1038/cr.2008.283; published online 2 September 2008

Dear Editor,

Heat shock proteins (Hsps) occur in nearly all organisms [1], and usually act as molecular chaperones, promoting correct refolding and preventing aggregation of denatured proteins [2]. Based on molecular weight, Hsps can be categorized into several families, including Hsp90, Hsp70, Hsp60, Hsp40 (so-called large Hsps) and small Hsps (sHsps, with the molecular weights varying from 12 to 43 kDa). Extensive studies have been carried out on the evolution of individual Hsp families, which indicate that the Hsps are organellar proteins [3, 4]. The endosymbiont theory (or gene transfer) is widely accepted to explain the origin of organelle-localized proteins [5]. However, some Hsps, such as chloroplast (CP) Hsp90s, do not conform to this rule, and they appear to arise from the endoplasmic reticulum (ER) version rather than from the cyanobacteria [3, 4]. Moreover, the origins of mitochondria (MT) versions of both Hsp90s and Hsp40s presented in our study do not agree with the gene transfer theory. In order to fully understand the evolutionary lineage of the Hsp superfamily, we retrieved the Hsp analogs from diverged eukaryotic organisms from the GenBank (Table S1-S2), and constructed phylogenetic trees for each Hsp family. Intriguingly, we found that the large and small Hsp families had different evolutionary patterns and lineages, and that gene replacement (i.e. intruding eubacterial genes replacing pre-existing nuclear counterparts) could be another evolutionary mechanism in the large Hsps.

Heat shock proteins are found to locate in cytosol (CY),

endoplasmic reticulum, mitochondria and chloroplast [4, 6]. However, no Hsp60 has been identified in endoplasmic reticulum, in which other Hsps have been demonstrated to be widely present. By searching the GenBank, three Hsp60s (accession numbers: NP_078961, XP_581684 and Q9DB12) were retrieved from *Homo sapiens*, *Bos taurus* and *Mus musculus*. They are the products of Bardet-Biedl syndrome 10 (BBS10) genes, and all have the cpn60 domain, which is the typical trait of Hsp60 (Supplementary information, Figure S1). In addition, their C- termini possess a conserved peptide "DEL", which usually acts as the signal for endoplasmic reticulum targeting [7]. The structural features suggest that the three BBS10s are members of ER-Hsp60s.

The phylogenetic trees of large Hsps showed that all the eukaryotic Hsps fell into four main groups according to their localizations in the cell: CY, ER, MT and CP (Figure 1A, Supplementary information, Figures S2-S4), and each group was strongly supported with high bootstrap value (73~100%), indicating that the evolution of these Hsps is highly related to the formation of eukaryotic cells. The CY- and ER- versions of the large Hsps are most closely related to one another, which is similar to the result that the CY- and ER- versions of both Hsp90 and Hsp70 may evolve from the same ancestor by gene duplication [4, 8]. The phylogenetic relationship of the MT- and CP- Hsps varies for different Hsp families (Figure 1A, Supplementary information, Figures S2-S4). The MT versions of both Hsp70 and Hsp60 formed a well supported cluster with the α -proteobacterial analogs, suggesting that they may have evolved by gene transfer from the endosymbiont. However, the MT versions of Hsp90s and Hsp40s are related to the analogs of firmicute and firmicute/actinobacteria, respectively. The CP versions of Hsp70s, Hsp60s and Hsp40s showed the closest relationship with the cyanobacteria analogs, which is consistent

Correspondence: Le Kang

Tel: 86-10-64807219; Fax: 86-10-64807099

E-mail: lkang@ioz.ac.cn

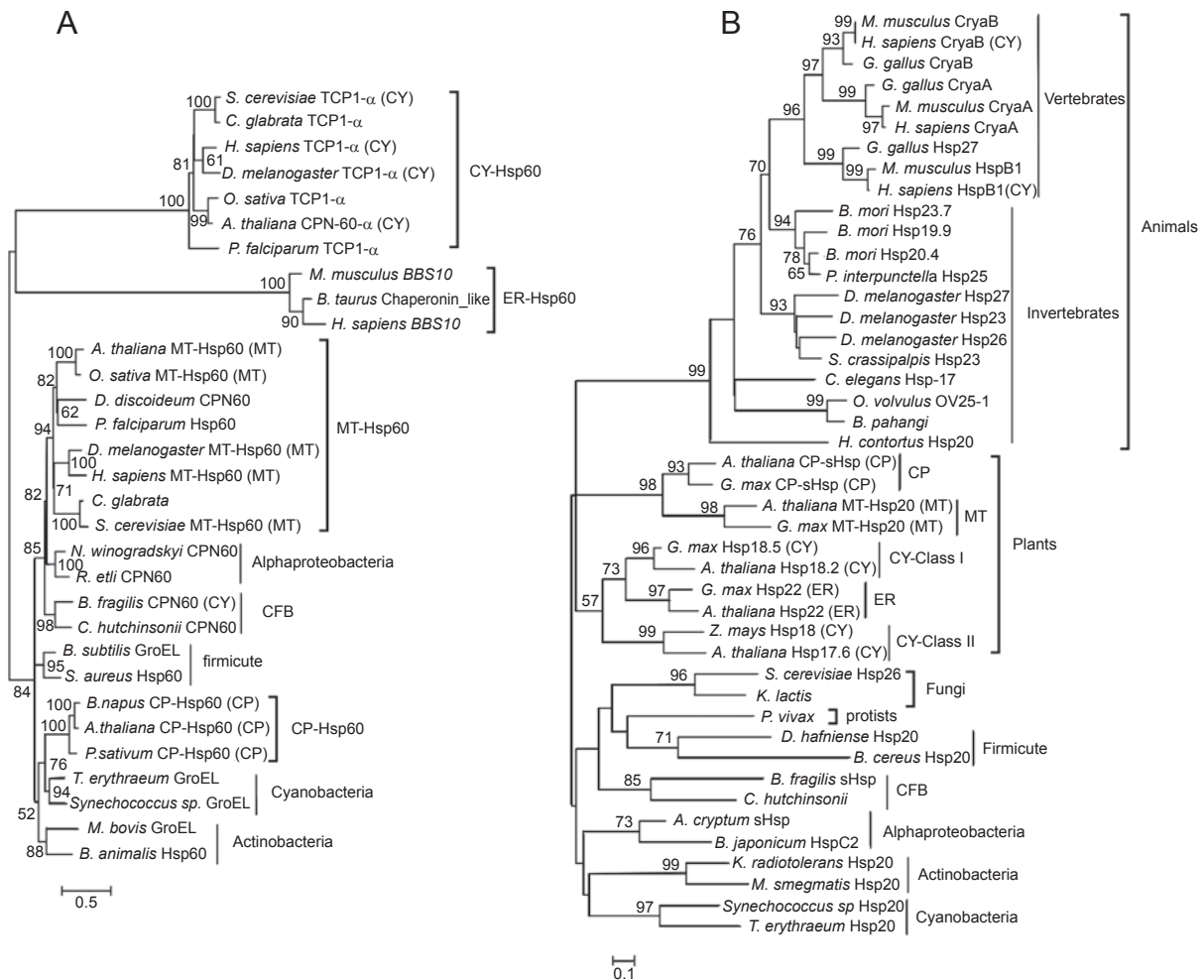


Figure 1 Neighbor-Joining trees of Hsp60s (A) and sHsps (B). The Hsp homologs were screened out from the GenBank database by Blast program using the representative Hsps (Supplementary information, Table S1) as queries. The sequences from various species that showed high scores (E value < 0.01) in the BLAST searches were retrieved. All the selected sequences were validated to contain typical Hsp domains by BLASTP. The full amino acid sequences were used to construct the NJ tree of Hsp60s, while only the α -crystalline domains (about 80 amino acids) were selected in the phylogenetic analysis of sHsps because of the variable terminals. All the trees were constructed using the NJ method with the Jones-Taylor-Thornton matrix-based model, and sites containing missing data or alignment gaps were removed in a pairwise fashion. For detailed procedure please refer to Supplementary Information, Data S1. The bootstrap values below 50% are not shown. Sequence labels are denoted by the names of species and proteins. The cellular localization of heat shock proteins is shown in parentheses. The full species names and Genbank accession numbers are listed in Supplementary information, Table S2. Abbreviation: CY, cytoplasm; ER, endoplasmic reticulum; MT, mitochondria; CP, chloroplast; CFB, Cytophaga-Fibrobacter-Bacteroides group.

with the classic endosymbiont theory that the chloroplast has originated from cyanobacteria [9]. However, the evolution of CP-Hsp90 appears to disagree with this pattern.

Quite different from those of large Hsps, the phylogenetic tree of sHsps (Figure 1B) revealed that the sHsps seemed to have evolved independently in the main eukaryotic lineages, including animals, plants and fungi. Moreover, the sHsps may have evolved at the very late stage when the animals and non-animals began

to diverge. On the other hand, all the vertebrate sHsps formed a well supported clade (96% bootstrap value), from which all the invertebrate sHsps were excluded, which indicates that the sHsps of vertebrates and invertebrates may have evolved independently. Interestingly, the sHsps of *Drosophila melanogaster* and *Bombyx mori* formed distinct groups (Figure 1B). Although most of the insect sHsps clustered according to insect order (Supplementary information, Figure S5), including those found

in the families of Lepidoptera, Diptera and Coleoptera, the orthologs of *D. melanogaster* CG14207 (Accession NO. NP_608326) apparently exist in other invertebrates, such as *B. mori* (BAD74197), *Tribolium castaneum* (XP_973685) and *Caenorhabditis elegans* (AAW57827). In other words, these sHsp genes may have duplicated at an early stage, followed by a more recent, order-specific duplication. The plant sHsps clustered according to the cellular localization (Figure 1B). They formed five groups, including CY-Class I, CY-Class II, ER, MT and CP. Each group was supported with high bootstrap value (>93%). The plant sHsps did not cluster with any prokaryotic homologs. Therefore, plant sHsps would not have originated from the prokaryotic version by gene transfer during the early evolutionary history.

Earlier studies have indicated that gene transfer is an important mechanism of the large Hsp evolution. Here, we found that gene replacement may be another evolutionary mechanism of the large Hsps. According to the endosymbiont theory, the MT- and CP- Hsps should have been derived respectively from the α -proteobacteria and cyanobacteria, such as the MT- versions of Hsp70 and Hsp60, CP- versions of Hsp70, Hsp60 and Hsp40. However, the CP-Hsp90s arose from the ER-versions rather than from cyanobacteria [4]. Gene replacement often takes place during the origin of organelles [10]. In the process, organelle and nuclear proteins are functionally equivalent or redundant in the early eukaryote. Thus, the cyanobacterial Hsp90 gene might have been replaced by a copy derived from the ER homolog [4]. The MT- versions of Hsp90s and Hsp40s might have undergone similar evolution. Based on the endosymbiont theory, the MT- organellar Hsps should have arisen from α -proteobacteria, but they came of different subdivisions of prokaryotes: firmicutes (Hsp90s), α -proteobacteria (Hsp70s and Hsp60s), actinobacteria/firmicute (Hsp40s), suggesting that the eukaryotic cell might have captured different endosymbionts during their evolutionary history, and that the α -proteobacterial versions of Hsp70s and Hsp60s had been kept in the eukaryotic genomes, whilst the versions of Hsp90s and Hsp40s might be functionally redundant with and subsequently replaced by analogs of the firmicutes or actinobacteria.

In conclusion, the eukaryotic large Hsps have evolved according to the cellular localizations (CY, ER, MT and

CP), and diverged at the very early stage of eukaryotic cell formation. However, the small Hsps probably have evolved independently in the main eukaryotic lineages, including animals, plants and fungi. Gene replacement may be another evolutionary mechanism for the large Hsps in addition to the endosymbiont occurrence or gene transfer.

Acknowledgments

We thank Dr Wen Wang (Kunming Institute of Zoology, Chinese Academy of Sciences), Drs Jun Yu and Jun Wang (Beijing Genomic Institute, Chinese Academy of Sciences) for valuable comments on early version of the manuscript. The research is supported by National Basic Research Program of China (2006CB102002).

References

- 1 Craig EA. The heat shock response. *CRC Crit Rev Biochem* 1985; **18**: 239-280.
- 2 Johnston JA, Ward CL, Kopito RR. Aggresomes: A cellular response to misfolded proteins. *J Cell Biol* 1998; **143**:1883-1898.
- 3 Krishna P, Gloor G. The Hsp90 family of proteins in *Arabidopsis thaliana*. *Cell Stress Chaperones* 2001; **6**:238-246.
- 4 Emelyanov VV. Phylogenetic relationships of organellar Hsp90 homologues reveal fundamental differences to organellar Hsp70 and Hsp60 evolution. *Gene* 2002; **299**:125-133.
- 5 Weeden NF. Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. *J Mol Evol* 1981; **17**:133-139.
- 6 Boorstein WR, Ziegelhoffer T, Craig EA. Molecular evolution of the Hsp70 multigene family. *J Mol Evol* 1994; **38**:1-17.
- 7 Rubin DM, Mehta AD, Zhu JS, et al. Genomic structure and sequence analysis of *Drosophila melanogaster* HSC70 genes. *Gene* 1993; **128**:155-163.
- 8 Gupta RS, Aitken K, Felah M, Singh B. Cloning of *Giardia lamblia* heat shock protein HSP70 homologs: Implications regarding origin of eukaryotic cells and of endoplasmic reticulum. *Proc Natl Acad Sci USA* 1994; **91**:2895-2899.
- 9 Martin W, Stoebe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* 1998; **393**:162-165.
- 10 Martin M, Schnarrenberger C. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: A case study of functional redundancy in ancient pathways through endosymbiosis. *Curr Genet* 1997; **32**:1-18.

(Supplementary information is linked to the online version of the paper on the Cell Research website.)