Article, Discoveries

The complex evolution of the HSP70 gene family

Er-meng Yu[1,2], Tatsuki Yoshinaga[3], Frank L. Jalufka[1], Hashimul Ehsan[1], David B. Mark Welch[4*], Gen Kaneko[1*]

[1]School of Arts & Sciences, University of Houston-Victoria, Victoria, TX.

[2]Key Laboratory of Tropical & Subtropical Fishery Resource Application & Cultivation, Key Laboratory of Aquatic Animal Immune Technology of Guangdong Province, Pearl River Fisheries Research Institute of CAFS, Guangzhou, China.

[3]School of Marine Biosciences, Kitasato University, Kanagawa, Japan.

[4]Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, MA.

***Corresponding author:**

Gen Kaneko, kanekog@uhv.edu

David Mark Welch, dmarkwelch@mbl.edu

Abstract

The 70-kDa heat shock protein (HSP70) family contains several isoforms localized in different subcellular compartments. The cytosolic isoforms have been classified into stress-inducible HSP70s and constitutive heat shock cognates (HSC70s), but occasional reports of "constitutive HSP70s" and the lack of cross-phylum comparisons have been a source of confusion in the evolution of the metazoan HSP70 family. Here we provide novel insights into the evolutionary history of this important molecular chaperone. We first cloned two HSP70 genes from the rotifer *Brachionus plicatilis,* an emerging model in evolutionary genetics, and confirmed their stress inducibility. Subsequent phylogenetic analyses of 100 full-length HSP70 family member genes revealed an ancient duplication that gave rise to two lineages from which all metazoan cytosolic HSP70s descend. One lineage contains a relatively small number of Lophotrochozoan and Ecdysozoan genes, none of which have been shown to be constitutively expressed. The second included both inducible and constitutive genes from diverse phyla. Species-specific duplications are present in both lineages, and in the second there are well-supported phylum-specific clades for Rotifera, Nematoda, and Chordata. Some genes in this lineage have likely independently acquired stress inducibility, which may explain the sporadic distribution of genes designated as "HSP70" or "HSC70" in previous analyses. Consistent with the history of diversification within each group, stress-inducible members of the second lineage show lower purifying selection pressure compared to constitutive members. These results illustrate the evolutionary history of the HSP70 family independent from their expression patterns, encouraging the development of new nomenclature based on evolutionary history.

Key words: heat shock cognates, molecular evolution, rotifer, stress inducibility, synonymous and nonsynonymous substitution.

Introduction

The 70-kDa heat shock protein (HSP70) family members play important roles in various cellular processes including heat shock response, folding of newly synthesized proteins, protein transport, and protein degradation. These apparently diverse functions are attributed to their chaperone activity (Daugaard et al. 2007; Hartl et al. 2011), by which they prevent the aggregation and misfolding of target proteins. Typically, HSP70 family members bind to denaturing or newly synthesized proteins by recognizing up to ten hydrophobic amino acid residues on the protein surface. The subsequent release of the HSP70 family members, which is triggered by ATP hydrolysis, facilitates the proper folding of the target proteins (Zuiderweg et al. 2017). The HSP70 family contains organelle-specific members localized in cytosol, endoplasmic reticulum (ER), mitochondria, and chloroplasts (Miernyk 1997), and these organelle-specific types are also known to contribute to protein transport across organelle membranes.

HSP70 family members are often upregulated upon various stresses that disrupt protein folding, such as heat treatment, exposure to toxic materials, ultraviolet irradiation, and pathogen attack (Sørensen et al. 2003; Baird et al. 2006). The upregulation is primarily regulated at the transcription level by transcription factors called heat shock factors (HSFs), particularly HSF1. Under unstressed conditions, HSF1 remains in a monomeric state. Various stresses lead to the formation of HSF1 trimers, which in turn bind to heat shock elements (HSEs) in the promoter region of HSP70 family member genes and promote their transcription. The HSF1-dependent transactivation system is generally conserved among eukaryotes, although species-specific differences in HSF1 function have been reported in many organisms such as the fruit fly *Drosophila melanogaster* (Jedlicka et al. 1997; Marchler and Wu 2001) and yeast *Saccharomyces cerevisiae* (Sorger and Pelham 1988). HSP70 family member genes have therefore been used as biomarkers of environmental stresses both in laboratory and field experiments (Sanders 1993; Ceyhun et al. 2010; Judge et al. 2011).

On the other hand, several cytosolic HSP70 family members show constitutive expression

71 patterns. These proteins are involved in the folding of newly synthesized proteins and are
72 traditionally called 70-kDa heat shock cognates (HSC70). A previous phylogenetic analysis
73 indicated that metazoan HSP70 family members can be classified into invertebrate HSP70s,
74 vertebrate HSP70s, and HSC70s from both vertebrates and invertebrates (Kourtidis et al. 2006).
75 However, occasional reports of "constitutive HSP70s," especially those from invertebrates
76 (Jayasena et al. 1999; Piano et al. 2002; Liu et al. 2017), have been a source of continued confusion
77 in the classification of HSP70 family members. Stress-induced upregulation of HSC70 has also
78 been reported in several animals including shrimps (Luan et al. 2010), snails (Zheng et al. 2012),
79 and fish (Yabu et al. 2011). In-depth phylogenetic analysis on HSP70 family members with a broad
80 sampling would offer further insight into the classification of this important group of molecular
81 chaperones.
82 In this study, we cloned two stress-inducible HSP70 family member genes from a monogonont
83 rotifer, *Brachionus plicatilis*, and performed phylogenetic analyses of these sequences and more
84 than 100 other HSP70 family members. The phylum Rotifera, composed of mostly microscopic
85 aquatic animals with about 1,000 cells and a characteristic ciliates head structure, is part of the
86 Gnathifera, a group of basal-branching phyla related to Lophotrochozoa, which contains Mollusca
87 and Annelida. Together these groups form a sister-clade to Ecdysozoa, which contains the
88 established invertebrate models *Caenorhabditis elegans* and *D. melanogaster* (Struck et al. 2014;
89 Fröbius and Funch 2017). Genetic information from Gnathifera remains limited, making rotifers
90 useful models in evolutionary studies. Our results added novel groups of stress-inducible HSP70,
91 providing important insight into the evolutionary history of the metazoan HSP70 family.
92
93 **Results**
94 Primary Structures of Rotifer HSP70s
95 A combination of 3' and 5' RACE identified two *B. plicatilis* HSP70 genes consisting of 2,121
96 and 2,248 bp, designated as HSP70-1 and HSP70-2, respectively. The nucleotide sequence of the

97 open reading frames, as well as the absence of introns, were ascertained by single PCR experiments
98 using cDNA or genomic DNA, respectively, followed by DNA sequencing. The deduced amino acid
99 sequences of HSP70-1 and HSP70-2 genes were more than 99% identical to each other
100 (supplementary fig. S1). Both contain the HSP70 protein family signatures, IDLGTTYS,
101 IFDLGGGTFDVSIL, and IVLVGGSTRIPKVQK (Prosite motifs PS00297, PS00329, and
102 PS01036, respectively), as well as the non-organelle stress protein motif RARFEEL found in
103 several cytosolic HSP70s (Cottin et al. 2008; Simoncelli et al. 2010; Zheng et al. 2012). The major
104 difference between *B. plicatilis* HSP70 genes was the number of GGMP repeats, which are involved
105 in binding to the HSP70-HSP90-organizing protein (Hop), a cofactor of HSP70 (Demand et al.
106 1998). The nucleotide sequences of *B. plicatilis* HSP70-1 and HSP70-2 genes were registered into
107 the DDBJ/EMBL/GenBank databases with accession numbers AB775784 and AB775785,
108 respectively.
109
110 Effects of Heat Stress on *B. plicatilis* HSP70 mRNA Levels
111 We next examined whether the mRNA levels of *B. plicatilis* HSP70s are increased by heat
112 stress. Due to the high sequence identity between *B. plicatilis* HSP70 genes, we employed semi-
113 quantitative RT-PCR to measure their respective mRNA levels. Primers were designed to amplify
114 DNA fragments of 112 and 124 bp from HSP70-1 and HSP70-2 cDNAs, respectively (fig. 1A), and
115 the PCR products were separated using a polyacrylamide gel. The mRNA levels of HSP70-1 and
116 HSP70-2 in heat-treated rotifers were 2.9 and 7.5 times higher, respectively, than those in control
117 rotifers (fig. 1B and 1C). These results indicate that both genes encode heat-inducible HSP70.
118 Since the two *B. plicatilis* HSP70 genes showed a similar expression pattern in response to
119 heat stress, we used quantitative real-time PCR primers that amplify both HSP70-1 and HSP70-2
120 genes in the subsequent time-course expression analyses (fig. 2A). The relative mRNA levels were
121 increased by heat stress at 40˚C for 10 min and reached at the maximum level 2 h after the heat
122 stress. The difference in mRNA levels between the control and 2 h groups was statistically

123 significant. The mRNA levels of *B. plicatilis* HSP70s remained twofold 8 h after the heat stress. We
124 then performed *in situ* hybridization using DIG-labeled RNA probes (fig. 2B). The probe region
125 contained only five base pair differences between the two HSP70 genes, and thus the RNA probes
126 were likely to hybridize with both HSP70-1 and HSP70-2 mRNAs. The hybridization signals
127 increased in all tissues of heat-treated rotifers. Altogether, these results indicate that the *B. plicatilis*
128 HSP70-1 and HSP70-2 genes are heat-inducible and have similar expression profiles.
129
130 Phylogenetic Analysis of HSP70 Family Members
131 To assess the relationship of the *B. plicatilis* HSP70 genes with those of other rotifers and
132 other metazoans, we constructed gene trees of more than 100 HSP70 family members, including
133 members associated with the mitochondria and endoplasmic reticulum and many sequences known
134 only from automated genome annotation as "HSP70." To assess the robustness of the results we
135 used alignments made by three different methods (Clustal-Omega, M-Coffee, and Expresso) to
136 construct consensus trees with maximum-likelihood (RAxML) and Bayesian (MrBayes)
137 approaches. The RAxML PROTGAMMAAUTO function found that the LG model of protein
138 evolution (Le and Gascuel 2008) was optimal for all three alignments and this model was used to
139 construct ML and Bayesian consensus trees. The six trees were in general agreement for major
140 nodes; fig. 3 shows the Bayesian consensus tree of the Expresso alignment and support for major
141 nodes from each approach is shown in table 1. In all analyses HSP70 family members known to be
142 mitochondria- and ER-specific formed two monophyletic groups (nodes 11 and 12 in fig. 3) in
143 agreement with previous reports (Boorstein et al. 1994; Miernyk 1997; Nikolaidis and Nei 2004).
144 All family members known to be cytosolic formed a third group (node 1; clades 1–6 in fig. 3), with
145 yeast cytosolic HSP70s (clade 6) basal to all metazoan cytosolic HSP70s. Within the metazoan
146 cytosolic HSP70s (node 2; clades 1–5) there were two lineages: Lineage I was well supported and
147 contained of genes from arthropods, molluscs, and rotifers (node 3; clade 1); Lineage II was more
148 poorly resolved and represented by many diverse vertebrate and invertebrate phyla (node 4; clades

149 2–5). There are two groups of vertebrate cytosolic HSP70s (nodes 8 and 9), which form a
150 monophyletic group (node 7; clade 5) in Bayesian and RAxML analyses of Expresso and Clustal
151 alignments. Unexpectedly, we found two additional clusters of cytosolic HSP70 family members in
152 all analyses: one composed of HSP70 genes from the phylum Rotifera (node 5; clade 2) and the
153 other from the phylum Nematoda (node 6; clade 3). This suggests the evolution of two unique
154 subgroups of HSP70, distinct from conventional vertebrate and invertebrate HSP70s.
155 The phylogenetic breadth of our analyses provided the opportunity to search for sequence
156 motifs characteristic of different groups within the HSP70 family. We identified a single region near
157 the N terminus (but after signal peptide motifs) that discriminates between metazoan mitochondrial,
158 ER, and cytosolic HSP70s (fig. 4). We also found the RARFEEL motif, previously identified as
159 possibly cytosolic-specific (Cottin et al. 2008; Simoncelli et al. 2010; Zheng et al. 2012), in several
160 invertebrate genes that are found with high confidence in the ER-associated node 12, suggesting
161 that this is not a reliable predictor of cytosol association.
162 Stress-inducible invertebrate HSP70s are reported to have an extra serine residue in their
163 ATPase domain (Kourtidis et al. 2006; Garbuz et al. 2011). In our analysis this insertion occurs only
164 in Lineage I irrespective of the stress-inducibility, where it is present in all sequences except *B.*
165 *ibericus* HSP70 and one of four *Lottia* HSP70s; in *Priapulus* HSP70-1 the serine has been replaced
166 by alanine (fig. 5). These results suggest that the extra serine residue is a characteristic of this
167 lineage and may be associated with the stress-inducibility of the ancestral HSP70. The alignment
168 result supports the idea that HSP70s from Nematoda and Rotifera evolved separately from the those
169 in Annelida, Mollusca and Arthropoda, which have been generally recognized as "invertebrate
170 HSP70s."
171
172 Synonymous and Nonsynonymous Substitution Rates
173 Lastly, we calculated the synonymous and nonsynonymous substitution rates of representative
174 cytosolic HSP70 member genes. We primarily selected genes with known expression patterns

175 (stress-inducible or constitutive) from each clade, and only well-conserved regions were used for
176 the calculation (supplementary fig. S2). We noted that *B. ibericus* HSP70 (GU574486) has an
177 insertion at nucleotide position 403 and a nucleotide deletion at position 563, causing a frame shift.
178 Deduced amino acid sequences encoded by nucleotides 403–563 were markedly different from
179 those of other HSP70 family members because of the frame shift, which resulted in the long genetic
180 distance in phylogenetic trees (fig. 3). Because the frame shift greatly affects the synonymous and
181 nonsynonymous substitution rates, this sequence was excluded from our analysis.
182 We calculated synonymous and nonsynonymous substitution rates (Ks and Ka) for all
183 combinations of representative cytosolic members of the HSP70 family from metazoans (fig. 6A).
184 Synonymous substitution rates (Ks) were saturated in many combinations and only a limited
185 number of combinations produced meaningful results (229 out of 465 combinations; supplementary
186 material 1). The Ka/Ks values calculated for the 229 combinations were averaged for each clade
187 (fig. 6B). Overall, Ka/Ks values were markedly lower than 1, indicating that cytosolic HSP70
188 family member genes have been under purifying selection. The purifying selection pressure appears
189 to have been less in Lineage I (clade 1) than in Lineage II (clades 2-5) and within Lineage II the
190 selection pressure was strongest in Nematode HSP70s (clade 3), invertebrate HSC70/HSP70s (clade
191 4), and vertebrate HSC70s (clade 5, node 8). In particular, vertebrate HSC70s (clade 5, node 8) and
192 HSP70s (clade 5, node 9) showed significant difference in the Ka/Ks despite their close
193 phylogenetic relationship.
194
195 **Discussion**
196 A previous phylogenetic analysis by Kourtidis et al. (2006) classified metazoan cytosolic
197 HSP70s into three groups, invertebrate HSP70, vertebrate HSP70, and HSC70 including both
198 vertebrate and invertebrate orthologues (Kourtidis et al. 2006). The monophyletic origin of HSC70s
199 was also proposed by Nikolaidis and Nei (2004) (Nikolaidis and Nei 2004). Surprisingly, to our
200 knowledge, there has been no significant update on the evolution of metazoan HSP70 family after

201 these reports. Here we show that there are at least two types of ancestral cytosolic HSP70 family
202 member genes in metazoans, one (node 3; clade 1) giving birth to a lineage of invertebrate HSP70s,
203 and a second (node 4; clade 5) to a lineage of both vertebrate and invertebrate HSP70s and all
204 HSC70 genes. This second lineage has further diversified within diverse phyla: Rotifera (node 5;
205 clade 2), Nematoda (node 6; clade 3), and Chordata (node 7; clade 5) (fig. 3).
206 Inter-clade comparison of the synonymous and nonsynonymous substitution rates shows that
207 while all HSP70 family members are under strong purifying selection, the pressure to conserve
208 amino acid sequence varies across clades. Most dramatically, the canonical HSC70 genes of
209 vertebrates (node 8; clade 5) are more conserved than their HSP70 cognates (node 9; clade 5). This
210 suggests that the role of HSC70 protein, the folding of newly synthesized proteins, is more
211 conserved and potentially ancestral. In addition, vertebrate HSC70, and not vertebrate HSP70, is
212 involved in chaperon-mediated autophagy (Kaushik and Cuervo 2012). Indeed, cells deficient in the
213 HSC70 gene are non-viable (Florin et al. 2004), whereas HSP70 knockout mice are mostly viable
214 and fertile (Daugaard et al. 2007).
215 Given the presence of inducible and constitutively expressed genes in many clades of
216 cytosolic HSP70, it is difficult to predict the ancestral state of this phenotype, and whether a gene is
217 inducible or constitutive expression does not predict its relationship to other family members.
218 Phylogenetic analyses, the novel signature sequences near the N terminus (fig. 4), and the extra
219 serine residue (fig. 5) would provide better clues for the classification of HSP70 family members
220 than their expression patterns under stress.
221 In conclusion, the present study identified two novel HSP70 family members from a
222 monogonont rotifer, an emerging model in evolutionary biology. The subsequent phylogenetic
223 analyses illustrated the evolutionary history of the metazoan HSP70 family, in which stress
224 inducibility does not reflect evolutionary history. The development of new nomenclature will help
225 us understand the true nature of the HSP70 family.
226

227 **Materials and Methods**

228 Culture

229 We used the *B. plicatilis* Ishikawa strain (also called *Brachionus* sp. ISKW), originally
230 isolated from a Japanese eel culture pond (Yoshinaga et al. 2004), for cDNA cloning, semi-
231 quantitative RT-PCR, quantitative real-time PCR, and *in situ* hybridization. The cytochrome oxidase
232 subunit I gene sequence has been registered for detailed identification of this strain (GenBank
233 accession number LC422762). Rotifers were cultured at 25 ˚C by a standard protocol (Kaneko et al.
234 2016) using half-diluted Brujewicz artificial seawater and fed the algae *Chlorella regularis* (Nikkai
235 Center, Tokyo, Japan) at a final concentration of approximately 7×10^6 cells/mL.

236

237 Heat Treatment

238 For semi-quantitative RT-PCR, rotifers were batch-cultured in 400 mL at 25˚C. In the
239 exponential growth phase, approximately 6×10^4 individuals were randomly divided into two
240 groups and concentrated to about 10 mL by filtration through 20 µm mesh filters. One group (heat-
241 treated group) was transferred to culture medium at 30˚C for 10 min, collected with a 20 µm mesh
242 filter, and stored in ISOGEN (Nippon Gene, Tokyo, Japan) at -80˚C. The control group was
243 prepared by the same procedure at 25˚C.

244 For quantitative real-time PCR, 8 rotifers bearing 2 or 3 eggs were selected from a batch
245 culture population and inoculated into a well of a 12-well plate (Nunc, Rochester, NY) containing 1
246 mL of culture medium at 25˚C. The 12-well plate was put on a water bath at 40˚C for 10 min, and
247 then transferred to a water bath at 25˚C. Rotifers were sampled before heat treatment (control), just
248 after heat treatment (time 0), and 2, 3, 4, 6, and 8 h after heat treatment (n = 3, each group contained
249 8 individuals). Rotifer samples were stored in 500 µL ISOGEN (Nippon Gene) and stored at -80˚C
250 until use. Rotifers used for *in situ* hybridization were prepared by the same method except the heat
251 shock temperature was 30˚C.

252

253 cDNA Cloning of HSP70

254 5' RACE for the partial sequence of *B. plicatilis* HSP70 gene [AB076052; (Kaneko et al.
255 2002)] was carried out using a GeneRacer kit (Invitrogen, Carlsbad, CA) with gene-specific primers
256 rHSP70_5RACE1, rHSP70_5RACE2, and rHSP70_5RACE3 (supplementary table S1). Total RNA
257 was extracted from a rotifer population in an exponential growth phase and used for 1st strand
258 cDNA synthesis using the rHSP70_5RACE1 primer and the GeneRacer kit (Invitrogen). First round
259 PCR was performed at 94˚C for 3 min followed by 30 cycles of 94˚C for 30 s, 55˚C for 30 s and 72
260 ˚C for 1 min. The final extension step was performed at 72˚C for 5 min. The 20 µL-reaction mixture
261 contained about 1 µg of the 1st strand cDNA, 2 µL of Ex *Taq* buffer (Takara, Shiga, Japan), 1 U of
262 Ex *Taq* DNA polymerase, 4 nmol of dNTP mixture, 10 pmol of rHSP70_5RACE2 primer and
263 GeneRacer 5' primer (Table 1). The amplified products were diluted 10-fold with sterile distilled
264 water and used as the template for the nested PCR with rHSP70_5RACE3 and GeneRacer 5' nested
265 primers (supplementary table S1). The PCR conditions were same as the first round PCR except for
266 the template and primers. The amplified products were subcloned into pGEM-T vecter (Promega,
267 Madison, WI) and sequenced as reported previously (Yoon et al. 2008).

268 3' RACE for rotifer HSP70 gene was carried out as follows. An aliquot of first strand cDNA
269 was synthesized from rotifers in an exponential phase using an oligo-dT adapter primer
270 (supplementary table S1). PCR was performed with the first strand cDNA as a template and with
271 rHSP70_3RACE1 and AUAP primers. PCR was carried out with 30 cycles of 94˚C for 30 s, 60˚C
272 for 30 s, and 72˚C for 2 min. The final extension step was performed at 72˚C for 5 min. The 20 µl of
273 reaction mixture contained 5 pmol of forward and reverse primers, approximately 1 µg of first
274 strand cDNA, 4 nmol of dNTP mixture, 2 µL of 10 × PCR buffer [100 mM Tris-HCl, pH8.3, 500
275 mM KCl, 15 mM $MgCl_2$, 0.01% (w/v) gelatin] and 0.2 U of Ex *Taq* DNA polymerase (Takara).
276 Nested PCR was performed under the same conditions using the PCR product diluted 100-fold and
277 the rHSP70_3RACE2 primer. The amplified products were subcloned into pGEM-T vecter
278 (Promega) and sequenced as described above. The nucleotide sequence of the open reading frame

279 was confirmed by a single PCR using rHSP70_full_f1 and rHSP70_full_r1 primers (supplementary
280 table S1). Conditions of the single PCR were same as described in 5' RACE.
281
282 Semi-Quantitative RT-PCR
283 RNA extraction and 1st strand cDNA synthesis were performed as described above. Primers
284 rHSP70_gapF and rHSP70_gapR were designed to amplify the region encoding the GGMP repeat
285 of *B. plicatilis* HSP70-1 and HSP70-2 genes (Table 1). The β-actin gene was used as the internal
286 control with specific primers rActinF and rActinR reported previously (Kaneko et al. 2005). PCR
287 was carried out at 94˚C for 3 min followed by 20 – 40 cycles of 94˚C for 30 s, 58˚C for 30 s and
288 72˚C for 30 s. A 20-μL reaction mixture contained approximately 1 μg of 1st strand cDNA, 10 pmol
289 of rHSP70_gapF and rHSP70_gapR, 2 μL of Ampli *Taq* buffer (Applied Biosystems, Foster City,
290 CA), 4 nmol of dNTP mixture, and 0.2 U of Ampli *Taq* DNA polymerase (Applied Biosystems).
291 The PCR products were separated with an 8% acrylamide gel and visualized by ethidium bromide
292 staining.
293
294 Quantitative Real-Time PCR
295 Rotifers were collected 0, 2, 4, 6, and 8 h after the heat treatment, and stored in ISOGEN
296 (Takara) at -80˚C until use (three replicates for each time point, one replicate contained eight
297 individuals). Rotifers collected before the heat treatment were used as the control. Quantitative real-
298 time PCR was performed as reported previously (Kaneko et al. 2011). Primers for the β-actin gene
299 were described previously (Kaneko et al. 2005); primers for HSP70 genes were designed using the
300 Primer Express software ver. 2 (Applied Biosystems) (supplementary table S1). R version 3.5.1. and
301 the multcomp package (Hothorn et al. 2008) was used for one-way analysis of variance (ANOVA)
302 followed by the Dunnett's test on the Macintosh platform.
303
304 *In situ* Hybridization

305 Rotifers were fixed in PBS containing 4% paraformaldehyde 4 h after the heat treatment. *In*
306 *situ* hybridization was performed as reported previously (Kaneko et al. 2011), using DIG-labeled
307 RNA probes synthesized as follows. Primers HSP70_insituF and HSP70_insituR were designed to
308 amplify the DNA fragment of 2026 - 2103 nt of *B. plicatilis* HSP70-1 gene (supplementary table
309 S1). PCR was performed in a 20 µL reaction mixture containing approximately 1 µg of the genomic
310 DNA of *B. plicatilis,* 10 pmol each of the HSP70_insituF and HSP70_insituR primers, 2 µL of 10 ×
311 PCR buffer, 4 nmol dNTP mixture, and 0.2 U of Ex *Taq* DNA polymerase (Takara). The
312 amplification was carried out with the initial denaturation at 94˚C for 3 min and 40 cycles of 94˚C
313 for 30 s, 60˚C for 30 s, and 72˚C for 1.5 min, followed by the final extension step at 72˚C for 5 min.
314 The amplified products were subcloned into pGEM-T easy vector (Promega). DIG-labeled RNA
315 probes were synthesized with a DIG RNA labeling kit (Roche) according to the manufacturer's
316 instructions.
317
318 Phylogenetic Analysis
319 A multiple sequence alignment of HSP70 and HSC70 amino acid sequences was produced
320 using Clustal Omega, Expresso, and M-Coffee with default parameters Bayesian phylogenetic
321 analysis was performed using MrBayes (v3.2.6) on the CIPRES scientific gateway v3.3 (Miller et
322 al. 2010) with four chains each for 10^7 generations. Every 1000th trees were sampled, and the first
323 25% of samples were discarded as a burn-in. Bayesian consensus trees were visualized with the
324 TreeView software v1.4.3. The ML tree was constructed on the MEGA7 software (Kumar et al.
325 2016) with the LG + G model, which was selected as the best-fit model according to Bayesian
326 information criterion (BIC) by Prottest v3.4.2 (Darriba et al. 2011). Bootstrap resampling analysis
327 from 1000 replicates was used to evaluate internal branches.
328
329 Synonymous and Nonsynonymous Substitution Rates
330 R version 3.6.2. was used for the calculation of the synonymous and nonsynonymous

substitution rates (the kaks function), Kruskal-Wallis test (kruskal.test function), and post-hoc test (pairwise.wilcox.test function) on the Macintosh platform. Heat maps were created in Microsoft Excel 2016 (Redmond, MA).

Author Contributions

Conceived and designed the experiments: TY, GK. Performed experiments and data analyses: EMY, FLJ, DMW, GK. Contributed reagents/materials/analysis tools: EH, DMW, GK. Wrote the paper: DMW, GK. Read and approved the final manuscript: EMY, TY, FLJ, EH, DMW, GK.

Acknowledgements

We express our sincere thanks to Dr. Kristin Gribble, Marine Biological Laboratory, for helpful discussions and comments on the manuscript. The authors are grateful to Cynthia Rodriguez and Jamie Wilson, University of Houston-Victoria, for their help in gene expression analyses. GK was supported by funding from M.G. and Lillie A. Johnson Foundation, Victoria, Texas.

Figure Legends

Fig. 1. Semi-quantitative reverse transcription-PCR *B. plicatilis* HSP70 genes. A representative result is shown out of three independent experiments. (A) Location of primers rHSP70_gapF and rHSP70_gapR that amplify cDNA fragment of 112 and 124 bp from HSP70-1 and HSP70-2 cDNAs, respectively. (B) Polyacrylamide gel electrophoresis patterns of the RT-PCR product. Cycle numbers within a linear range of PCR amplification were determined to be 24 to 28 cycles for both cDNAs by preliminary experiments on the basis of signal intensities of amplified products by RT-PCR. (C) Signal intensities of HSP70-1 and HSP70-2 genes standardized to those of β-actin.

Fig. 2. Expression analyses of *B. plicatilis* HSP70 genes. (A) Heat stress-induced expression of HSP70 quantified by real-time PCR. Beta-actin gene was used as the internal control. Bars

357 represent standard errors from three replications. One-way ANOVA detected significant effects of
358 heat treatment (F = 3.281, df = 5, P = 0.0426). Dunnett's multiple comparison was used to detect
359 significant differences between the control and other groups. Difference between control and two
360 hours group was statistically significant (t = 3.246, P = 0.028). (B) *In situ* hybridization for *B.*
361 *plicatilis* HSP70 genes. Rotifers were fixed 4 h after the heat treatment.
362
363 **Fig. 3.** Bayesian tree of HSP70 family members. Nodes discussed in the text indicated. Numbers on
364 the branches indicate the posterior probability support for each node. Stress-inducible and non-
365 inducible genes are shown in red and blue, respectively. No data on the stress inducibility is
366 available for genes shown in black. Approved HGNC names of human HSP70 family members are
367 shown in yellow boxes for reference. The DDBJ/EMBL/GenBank accession numbers and other
368 information are summarized in supplementary table S2. Clades 1 to 6 include only cytosolic
369 members of the HSP70 family.
370
371 **Fig. 4.** Logo visualization of mitochondrial-, endoplasmic reticulum-, and cytosolic-specific
372 sequence near the N terminus of the HSP70 family members.
373
374 **Fig. 5.** Comparison of the amino acid sequences of the HSP70 family members from invertebrates.
375 Nodes in the phylogenetic tree (fig. 3) are indicated on the right margin, and serine residues specific
376 to node 3 are highlighted. Stress-inducible and non-inducible genes are shown in red and blue,
377 respectively. No data on the stress inducibility is available for genes shown in black. The accession
378 numbers and other information are summarized in supplementary table S2.
379
380 **Fig. 6.** Selection pressure on cytosolic HSP70 family member genes. (A) Nonsynonymous
381 substitution rates (Ka values) between 34 HSP70 family member genes. Gene number 1 is the fruit
382 fly HSP70 Bb (AF295957) gene. Accession numbers for other genes are indicated in the figure.

383 Clades and nodes correspond to those in fig. 3. (B) Boxplots for the Ka/Ks values averaged for each
384 clade. Only Ka/Ks values calculated from inter-cluster pairs were averaged. Statistical differences
385 were calculated by Kruskal-Wallis test (chi-squared = 88.704, df = 5, $p < 2.2e-16$) followed by the
386 non-parametric post-hoc tests (pairwise Wilcox test with P value adjustment by the Holm method).
387 Clades sharing same letters are not significantly different at the 5% level of significance. One
388 outlier in clade 3 is not included in the box plot (see Supplementary Material 1).
389

390 **References**

391 Baird NA, Turnbull DW, Johnson EA. 2006. Induction of the heat shock pathway during hypoxia
392 requires regulation of heat shock factor by hypoxia-inducible factor-1. *J Biol Chem*
393 281:38675-38681.
394 Boorstein WR, Ziegelhoffer T, Craig EA. 1994. Molecular evolution of the HSP70 multigene
395 family. *J Mol Evol* 38:1-17.
396 Cellura C, Toubiana M, Parrinello N, Roch P. 2006. HSP70 gene expression in *Mytilus*
397 *galloprovincialis* hemocytes is triggered by moderate heat shock and *Vibrio anguillarum*, but
398 not by *V. splendidus* or *Micrococcus lysodeikticus*. *Dev Comp Immunol* 30:984-997.
399 Ceyhun SB, Sentürk M, Ekinci D, Erdoğan O, Ciltaş A, Kocaman EM. 2010. Deltamethrin
400 attenuates antioxidant defense system and induces the expression of heat shock protein 70 in
401 rainbow trout. *Comp Biochem Physiol, C Toxicol Pharmacol* 152:215-223.
402 Cottin D, Ravaux J, Leger N, Halary S, Toullec JY, Sarradin PM, Gaill F, Shillito B. 2008. Thermal
403 biology of the deep-sea vent annelied *Paralvinella grasslei*: in vivo studies. *J Exp Biol*
404 211:2196-2204.
405 Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of
406 protein evolution. *Bioinformatics* 27:1164-1165.
407 Daugaard M, Rohde M, Jäättelä M. 2007. The heat shock protein 70 family: Highly homologous
408 proteins with overlapping and distinct functions. *FEBS Lett* 581:3702-3710.
409 Deane EE, Woo NYS. 2009. Modulation of fish growth hormone levels by salinity, temperature,
410 pollutants and aquaculture related stress: a review. *Rev Fish Biol Fish* 19:97-120.
411 Demand J, Lüders J, Höhfeld J. 1998. The carboxy-terminal domain of Hsc70 provides binding

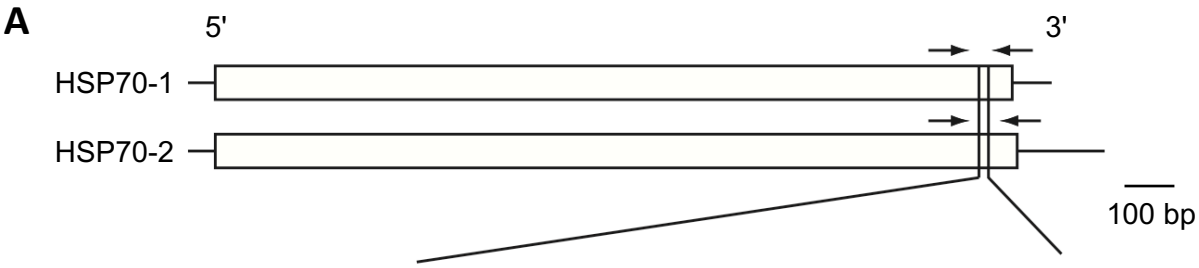
412 sites for a distinct set of chaperone cofactors. *Mol Cell Biol* 18:2023-2028.
413 Florin L, Becker KA, Sapp C, Lambert C, Sirma H, Müller M, Streeck RE, Sapp M. 2004. Nuclear
414 translocation of papillomavirus minor capsid protein L2 requires Hsc70. *J Virol* 78:5546-5553.
415 Fröbius AC, Funch P. 2017. Rotiferan Hox genes give new insights into the evolution of metazoan
416 bodyplans. *Nat Commun* 8:9.
417 Ganguly A, Han X, Das U, Wang L, Loi J, Sun J, Gitler D, Caillol G, Leterrier C, Yates JR. 2017.
418 Hsc70 chaperone activity is required for the cytosolic slow axonal transport of synapsin. *J Cell*
419 *Biol* 216:2059-2074.
420 Garbuz DG, Yushenova IA, Zatsepina OG, Przhiboro AA, Bettencourt BR, Evgen'ev MB.
421 2011. Organization and evolution of hsp70 clusters strikingly differ in two species of
422 Stratiomyidae (Diptera) inhabiting thermally contrasting environments. *BMC Evol Biol* 11:74.
423 Gutierrez JA, Guerriero V. 1995. Chemical modifications of a recombinant bovine stress-inducible
424 70 kDa heat-shock protein (Hsp70) mimics Hsp70 isoforms from tissues. *Biochem J* 305:197-
425 203.
426 Hartl FU, Bracher A, Hayer-Hartl M. 2011. Molecular chaperones in protein folding and
427 proteostasis. *Nature* 475:324-332.
428 Heschl MFP, Baillie DL. 1989. Characterization of the hsp70 multigene family of *Caenorhabditis*
429 *elegans*. *DNA* 8:233-243.
430 Hothorn T, Bretz F, Westfall P. 2008. Simultaneous inference in general parametric models. *Biom J*
431 50:346-363.
432 Jayasena S, Chandrasekharan N, Karunanayake EH. 1999. Molecular characterisation of a hsp70
433 gene from the filarial parasite *Setaria digitata*. *Int J Parasitol* 29:581-591.
434 Jedlicka P, Mortin MA, Wu C. 1997. Multiple functions of *Drosophila* heat shock transcription
435 factor *in vivo*. *EMBO J* 16:2452-2462.
436 Judge ML, Botton ML, Hamilton MG. 2011. Physiological consequences of the supralittoral fringe:
437 microhabitat temperature profiles and stress protein levels in the tropical periwinkle *Cenchritis*
438 *muricatus* (Linneaus, 1758). *Hydrobiologia* 675:143-156.
439 Kaneko G, Kinoshita S, Yoshinaga T, Tsukamoto K, Watabe S. 2002. Changes in expression
440 patterns of stress protein genes during population growth of the rotifer *Brachionus plicatilis*.
441 *Fish Sci* 68:1317-1323.

442 Kaneko G, Yoshinaga T, Gribble KE, Mark Welch D, Ushio H. 2016. Measurement of survival time
443 in *Brachionus* rotifers: synchronization of maternal conditions. *J Vis Exp* 113:e54126.
444 Kaneko G, Yoshinaga T, Yanagawa Y, Kinoshita S, Tsukamoto K, Watabe S. 2005. Molecular
445 characterization of Mn-superoxide dismutase and gene expression studies in dietary restricted
446 *Brachionus plicatilis* rotifers. *Hydrobiologia* 546:117-123.
447 Kaneko G, Yoshinaga T, Yanagawa Y, Ozaki Y, Tsukamoto K, Watabe S. 2011. Calorie restriction-
448 induced maternal longevity is transmitted to their daughters in a rotifer. *Funct Ecol* 25:209-
449 216.
450 Kaushik S, Cuervo AM. 2012. Chaperone-mediated autophagy: a unique way to enter the lysosome
451 world. *Trends Cell Biol* 22:407-417.
452 Kim R-O, Rhee J-S, Won E-J, Lee K-W, Kang C-M, Lee Y-M, Lee J-S. 2011. Ultraviolet B retards
453 growth, induces oxidative stress, and modulates DNA repair-related gene and heat shock
454 protein gene expression in the monogonont rotifer, *Brachionus* sp. *Aquat Toxicol* 101:529-539.
455 Kourtidis A, Drosopoulou E, Nikolaidis N, Hatzi VI, Chintiroglou CC, Scouras ZG. 2006.
456 Identification of several cytoplasmic HSP70 genes from the Mediterranean mussel (*Mytilus*
457 *galloprovincialis*) and their long-term evolution in Mollusca and Metazoa. *J Mol Evol* 62:446-
458 459.
459 Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version
460 7.0 for bigger datasets. *Mol Biol Evol* 33:1870-1874.
461 Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. *Mol Biol Evol*
462 25:1307-1320.
463 Lee JM, Lee KG, Choi HS, Kim ES, Keum B, Seo YS, Jeen YT, Chun HJ, Lee HS, Um SH. 2018.
464 Increased heat shock protein 70 expression attenuates pancreatic fibrosis induced by dibutyltin
465 dichloride. *Scand J Gastroenterol* 53:1404-1410.
466 Liu L, Cheng T-y, Yang Y. 2017. Cloning and expression pattern of a heat shock cognate protein 70
467 gene in ticks (*Haemaphysalis flava*). *Parasitol Res* 116:1695-1703.
468 Luan W, Li F, Zhang J, Wen R, Li Y, Xiang J. 2010. Identification of a novel inducible cytosolic
469 Hsp70 gene in Chinese shrimp *Fenneropenaeus chinensis* and comparison of its expression
470 with the cognate Hsc70 under different stresses. *Cell Stress Chaperones* 15:83-93.
471 Ma K-X, Chen G-W, Lou H, Fei L-N. 2009. Cloning and expression analysis of hsp70 gene from

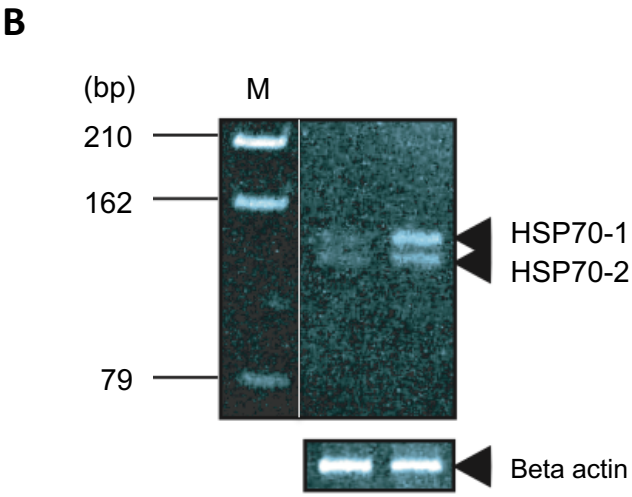
472 freshwater planarian *Dugesia japonica*. *Biologia* 64:1018-1024.
473 Marchler G, Wu C. 2001. Modulation of *Drosophila* heat shock transcription factor activity by the
474 molecular chaperone DROJ1. *EMBO J* 20:499-509.
475 Miernyk JA. 1997. The 70 kDa stress-related proteins as molecular chaperones. *Trends Plant Sci*
476 2:180-187.
477 Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of
478 large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop
479 (GCE); New Orleans, LA. p. 1-8.
480 Milner CM, Campbell RD. 1990. Structure and expression of the three MHC-linked HSP70 genes.
481 *Immunogenetics* 32:242-251.
482 Moribe Y, Oka K, Niimi T, Yamashita O, Yaginuma T. 2010. Expression of heat shock protein 70a
483 mRNA in *Bombyx mori* diapause eggs. *J Insect Physiol* 56:1246-1252.
484 Neumann S, Ziv E, Lantner F, Schechter I. 1993. Regulation of HSP70 gene expression during the
485 life cycle of the parasitic helminth *Schistosoma mansoni*. *Eur J Biochem* 212:589-596.
486 Nikolaidis N, Nei M. 2004. Concerted and nonconcerted evolution of the Hsp70 gene superfamily
487 in two sibling species of nematodes. *Mol Biol Evol* 21:498-505.
488 Piano A, Asirelli C, Caselli F, Fabbri E. 2002. Hsp70 expression in thermally stressed *Ostrea edulis*,
489 a commercially important oyster in Europe. *Cell Stress Chaperones* 7:250.
490 Ravi V, Kubofcik J, Bandopathyaya S, Geetha M, Narayanan R, Nutman T, Kaliraj P. 2004.
491 *Wuchereria bancrofti*: cloning and characterization of heat shock protein 70 from the human
492 lymphatic filarial parasite. *Exp Parasitol* 106:1-10.
493 Sanders BM. 1993. Stress proteins in aquatic organisms - an environmental perspective. *Crit Rev*
494 *Toxicol* 23:49-75.
495 Ścieglińska D, Pigłowski W, Mazurek A, Małusecka E, Żebracka J, Filipczak P, Krawczyk Z. 2008.
496 The HspA2 protein localizes in nucleoli and centrosomes of heat shocked cancer cells. *J Cell*
497 *Biochem* 104:2193-2206.
498 Simoncelli F, Morosi L, Di Rosa I, Pascolini R, Fagotti A. 2010. Molecular characterization and
499 expression of a heat-shock cognate 70 (Hsc70) and a heat-shock protein 70 (Hsp70) cDNAs in
500 Rana (Pelophylax) lessonae embryos. *Comp Biochem Physiol, A Mol Integr Physiol* 156:552-
501 560.

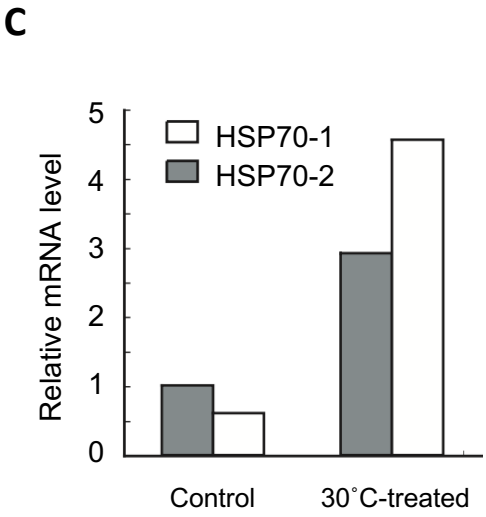
502 Snutch TP, Heschl MFP, Baillie DL. 1988. The *Caenorhabditis elegans* hsp70 gene family - a
503 molecular genetic-characterization. *Gene* 64:241-255.
504 Sørensen JG, Kristensen TN, Loeschcke V. 2003. The evolutionary and ecological role of heat
505 shock proteins. *Ecol Lett* 6:1025-1037.
506 Sorger PK, Pelham HR. 1988. Yeast heat shock factor is an essential DNA-binding protein that
507 exhibits temperature-dependent phosphorylation. *Cell* 54:855-864.
508 Struck TH, Wey-Fabrizius AR, Golombek A, Hering L, Weigert A, Bleidorn C, Klebow S,
509 Iakovenko N, Hausdorf B, Petersen M. 2014. Platyzoan paraphyly based on phylogenomic
510 data supports a noncoelomate ancestry of Spiralia. *Mol Biol Evol* 31:1833-1849.
511 Xiu Y, Feng J, Lu W, Liu D, Wu T, Zhu H, Liu P, Li W, Ren Q, Gu W. 2014. Identification of a
512 novel cognate cytosolic Hsp70 gene (MnHsc70-2) from oriental river prawn *Macrobrachium*
513 *nipponense* and comparison of its expressions with the first cognate Hsc70 (MnHsc70-1)
514 under different stresses. *Cell Stress Chaperones* 19:949-961.
515 Xu H-H, Chen X-L, Yang Y-H, Zhou X, Yang J-X. 2017. Heat shock protein 70 from the rotifer
516 *Asplanchna brightwelli*: sequencing, phylogenetic analysis and expression in response to
517 oxidation stressors. *Isr J Ecol Evol*:1-7.
518 Xu M, Jin T, Chen L, Zhang X, Zhu G, Wang Q, Lin Z. 2019. Mortalin is a distinct bio-marker and
519 prognostic factor in serous ovarian carcinoma. *Gene* 696:63-71.
520 Xu Q, Zou Q, Zheng H, Zhang F, Tang B, Wang S. 2011. Three heat shock proteins from
521 *Spodoptera exigua*: gene cloning, characterization and comparative stress response during heat
522 and cold shocks. *Comp Biochem Physiol B Biochem Mol Biol* 159:92-102.
523 Yabu T, Imamura S, Mohammed MS, Touhata K, Minami T, Terayama M, Yamashita M. 2011.
524 Differential gene expression of HSC70/HSP70 in yellowtail cells in response to chaperone-
525 mediated autophagy. *FEBS J* 278:673-685.
526 Yamashita M, Hirayoshi K, Nagata K. 2004. Characterization of multiple members of the HSP70
527 family in platyfish culture cells: molecular evolution of stress protein HSP70 in vertebrates.
528 *Gene* 336:207-218.
529 Yokoyama Y, Hashimoto H, Kubota S, Kuriyama A, Ogura Y, Mizuta S, Yoshinaka R, Toyohara H.
530 2006. cDNA cloning of Japanese oyster stress protein homologous to the mammalian 78-kDa
531 glucose regulated protein and its induction by heatshock. *Fish Sci* 72:402-409.

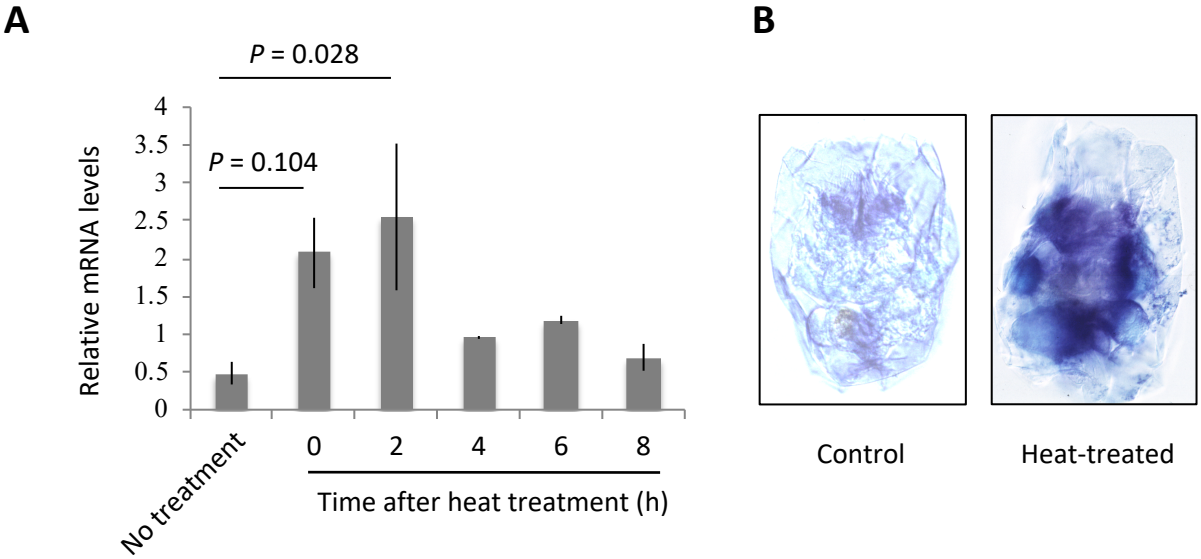
532 Yoon SH, Itoh Y, Kaneko G, Nakaniwa M, Ohta M, Watabe S. 2008. Molecular characterization of
533 Japanese sillago vitellogenin and changes in its expression levels on exposure to 17 beta-
534 estradiol and 4-tert-octylphenol. *Mar Biotechnol* 10:19-30.
535 Yoshinaga T, Minegishi Y, Rumengan IFM, Kaneko G, Furukawa S, Yanagawa Y, Tsukamoto K,
536 Watabe S. 2004. Molecular phylogeny of the rotifers with two Indonesian *Brachionus*
537 lineages. *Coast Mar Sci* 29:45-56.
538 Zafarullah M, Wisniewski J, Shworak NW, Schieman S, Misra S, Gedamu L. 1992. Molecular
539 cloning and characterization of a constitutively expressed heat-shock-cognate hsc71 gene from
540 rainbow trout. *Eur J Biochem* 204:893-900.
541 Zhang Q, Jiang L, Mu W, Li S, Wei S. 2016. Gene cloning and expression analysis of an inducible
542 heat shock protein 70 gene from the polychaete *Perinereis aibuhitensis* under Cu^{2+} stress. *J*
543 *Appl Ecol* 27:1628-1638.
544 Zheng G, Dong S, Hou Y, Yang K, Yu X. 2012. Molecular characteristics of HSC70 gene and its
545 expression in the golden apple snails, *Pomacea canaliculata* (Mollusca: Gastropoda).
546 *Aquaculture* 358-359:41-49.
547 Zuiderweg ER, Hightower LE, Gestwicki JE. 2017. The remarkable multivalency of the Hsp70
548 chaperones. *Cell Stress Chaperones* 22:173-189.
549

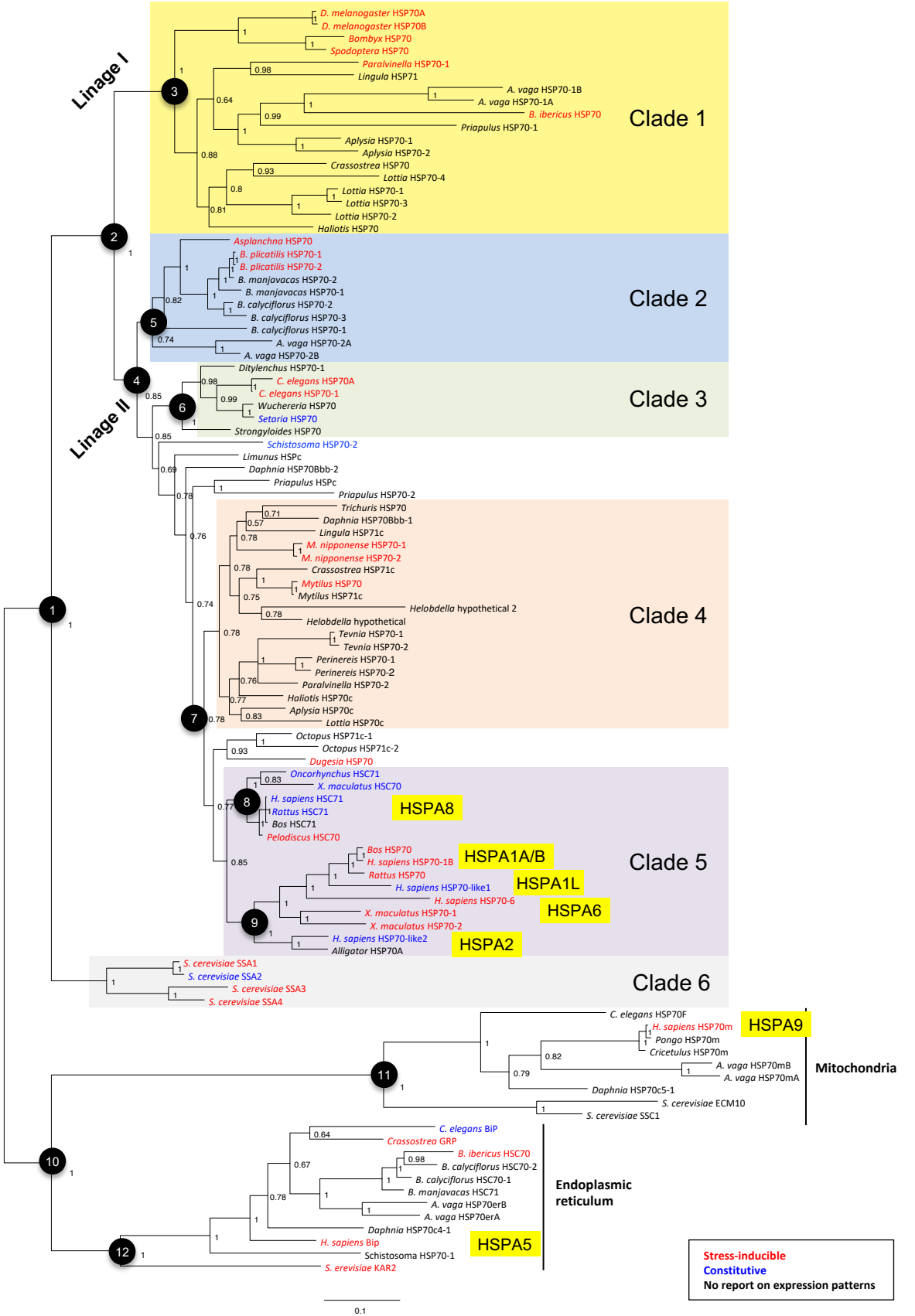



```
GTGGCATGCCTGGTGGTATGCCTGG------------CGGTATGCCTGGTGGAATGCCTG 1968
GTGGTATGCCTGGTGGTATGCCTGGTGGTATGCCGGGCGGTATGCCTGGTGGAATGCCTG 1979
```

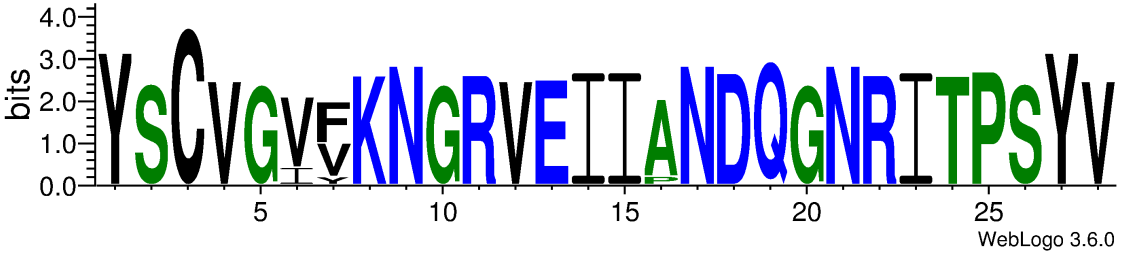




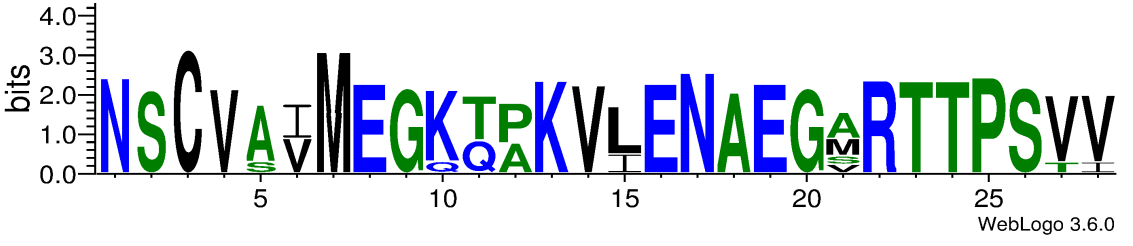

A
P = 0.028
P = 0.104
Relative mRNA levels
4
3.5
3
2.5
2
1.5
1
0.5
0
No treatment
0
2
4
6
8
Time after heat treatment (h)
B
Control
Heat-treated


Linage I
Linage II
Clade 1
D. melanogaster HSP70A
D. melanogaster HSP70B
Bombyx HSP70
Spodoptera HSP70
Paralvinella HSP70-1
Lingula HSP71
A. vaga HSP70-1B
A. vaga HSP70-1A
B. ibericus HSP70
Priapulus HSP70-1
Aplysia HSP70-1
Aplysia HSP70-2
Crassostrea HSP70
Lottia HSP70-4
Lottia HSP70-1
Lottia HSP70-3
Lottia HSP70-2
Haliotis HSP70
Clade 2
Asplanchna HSP70
B. plicatilis HSP70-1
B. plicatilis HSP70-2
B. manjavacas HSP70-2
B. manjavacas HSP70-1
B. calyciflorus HSP70-2
B. calyciflorus HSP70-3
B. calyciflorus HSP70-1
A. vaga HSP70-2A
A. vaga HSP70-2B
Clade 3
Ditylenchus HSP70-1
C. elegans HSP70A
C. elegans HSP70-1
Wuchereria HSP70
Setaria HSP70
Strongyloides HSP70
Schistosoma HSP70-2
Limunus HSPc
Daphnia HSP70Bbb-2
Priapulus HSPc
Priapulus HSP70-2
Clade 4
Trichuris HSP70
Daphnia HSP70Bbb-1
Lingula HSP71c
M. nipponense HSP70-1
M. nipponense HSP70-2
Crassostrea HSP71c
Mytilus HSP70
Mytilus HSP71c
Helobdella hypothetical 2
Helobdella hypothetical
Tevnia HSP70-1
Tevnia HSP70-2
Perinereis HSP70-1
Perinereis HSP70-2
Paralvinella HSP70-2
Haliotis HSP70c
Aplysia HSP70c
Lottia HSP70c
Octopus HSP71c-1
Octopus HSP71c-2
Dugesia HSP70
Clade 5
Oncorhynchus HSC71
X. maculatus HSC70
H. sapiens HSC71
Rattus HSC71
Bos HSC71
Pelodiscus HSC70
HSPA8
Bos HSP70
H. sapiens HSP70-1B
HSPA1A/B
Rattus HSP70
H. sapiens HSP70-like1
HSPA1L
H. sapiens HSP70-6
HSPA6
X. maculatus HSP70-1
X. maculatus HSP70-2
H. sapiens HSP70-like2
Alligator HSP70A
HSPA2
Clade 6
S. cerevisiae SSA1
S. cerevisiae SSA2
S. cerevisiae SSA3
S. cerevisiae SSA4
C. elegans HSP70F
H. sapiens HSP70m
HSPA9
Pongo HSP70m
Cricetulus HSP70m
A. vaga HSP70mB
A. vaga HSP70mA
Daphnia HSP70c5-1
S. cerevisiae ECM10
S. cerevisiae SSC1
Mitochondria
C. elegans BiP
Crassostrea GRP
B. ibericus HSC70
B. calyciflorus HSC70-2
B. calyciflorus HSC70-1
B. manjavacas HSC71
A. vaga HSP70erB
A. vaga HSP70erA
Daphnia HSP70c4-1
H. sapiens Bip
Schistosoma HSP70-1
S. erevisiae KAR2
HSPA5
Endoplasmic reticulum
Stress-inducible
Constitutive
No report on expression patterns
0.1

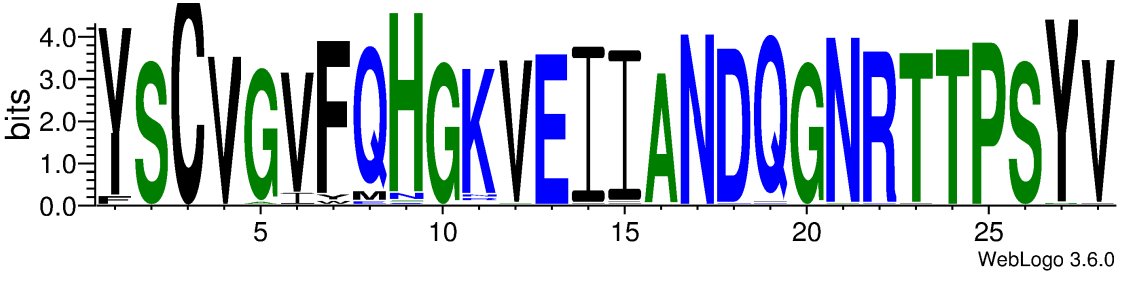
Mitochondrial



Endoplasmic reticulum



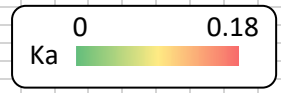
Cytosol



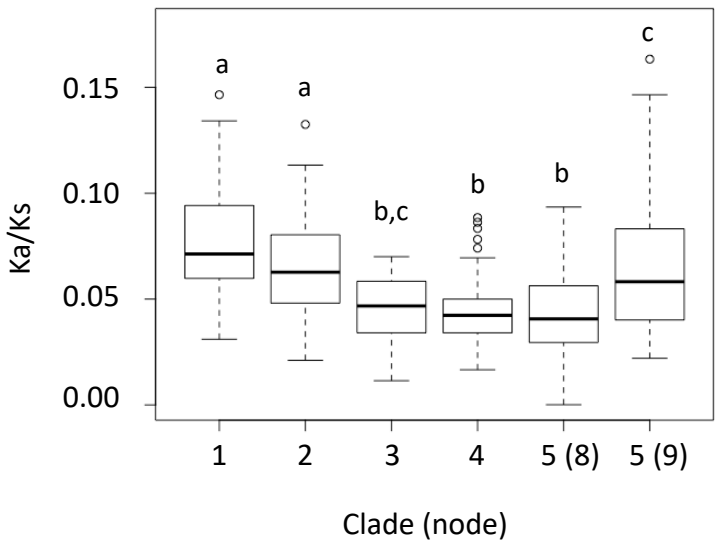
			Node
D. melanogaster HSP70A	KGERNVLIFDLGGGTFDVSILTIDE----G**S**LFEVRSTAGDT	224	
D. melanogaster HSP70B	KGERNVLILDLGGGTFDVSILTIDE----G**S**LFEVRSTAGDT	224	
Bombyx HSP70	KGERNVLIFDLGGGTFDVSILTIDE----G**S**LFEVKSTAGDT	224	
Spodoptera HSP70	KGERNVLIFDLGGGTFDVSILTIDE----G**S**LFEVRATAGDT	224	
Paralvinella HSP70-1	SGEKNVLIFDLGGGTFDVSILSIAD----G**S**LFEVKSTAGDT	229	
Lingula HSP71	QGEKNVLIFDLGGGTFDVSILSIAD----G**S**LFEVRSTAGDT	228	
A. vaga HSP70-1B	HRDQNILIYDLGGGTFDVSVLTISGGDVSG**S**VFEVKATAGDT	238	
A. vaga HSP70-1A	QRDQNILIYDLGGGTFDVSILTISGGDAGG**S**VFEVRSTAGDT	238	
B. ibericus HSP70	KRRKNVLIYDLGGGTFDVSVLSIND----G-LFEVLSTNGNT	224	
Priapulus HSP70-1	AGERNVLVYDLGGGTFDVSVLTIDE----GAMFEVRSTAGDT	226	**3**
Aplysia HSP70-1	KGEKHVLIFDLGGGTFDVSVLAIDE----G**S**IFEVKATAGDT	227	
Aplysia HSP70-2	SGEKNVLIFDLGGGTFDVSVLTIDE----G**S**MFEVKATAGDT	230	
Crassostrea HSP70	SGEKNVLIFDLGGGTFDVSILTIDE----G**S**IFEVRSTAGDT	228	
Lottia HSP70-4	DTEQHILIFDLGGGTFDVSILAIED----G-VFEVLATAGDT	227	
Lottia HSP70-1	SGEKNVLIFDLGGGTFDVSILTIDE----G**S**LFEVKSTAGDT	225	
Lottia HSP70-3	SGEKNVLIFDLGGGTFDVSILTIDE----G**S**LFEVKSTAGDT	226	
Lottia HSP70-2	SGEKNVLIFDLGGGTFDVSILAIDE----G**S**LFEVKSTAGDT	226	
Haliotis HSP70	KGEKNVLIFDLGGGTFDVPVLTIDE----G**S**MFEVRSTAGDT	231	
B. plicatilis HSP70-1	GGEKNILIFDLGGGTFDVSILTIEE----G-IFEVKSTAGDT	226	
B. manjavacas HSP70-1	AGEKNILIFDLGGGTFDVSILTIED----G-IFEVKATAGDT	226	**5**
A. vaga HSP70-2A	SGERNILIFDLGGGTFDVSVLKIEE----G-IFEVKSTAGDT	184	
C. elegans HSP70A	HGERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	227	
Wuchereria HSP70	HGERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	226	**6**
Setaria HSP70	HGERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	226	
Strongyloides HSP70	ATERNVLIFDLGGGTFDVSVLTIED----G-IFEVKSTAGDT	227	
Schistosoma HSP70-2	GGERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	224	
M. nipponense HSP70-1	GGERNVLIFDLGGGTFDVSILTIED----D-IFEVKSTAGDT	226	
Mytilus HSP70	GGERNVLIFDLGGGTFDVSILTIKD----G-IFEVKSTSGDT	227	
Tevnia1 HSP70-1	GGEQNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	226	
Perinereis1 HSP70-1	GGERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	225	
Lottia HSP70c	GGERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	226	
Dugesia HSP70	GIERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	226	
Oncorhynchus HSC71	GAERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	226	
H. sapiens HSC71	GAERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	226	**8**
Pelodiscus HSC70	GAERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	226	
H. sapiens HSP70-1B	KGERNVLIFDLGGGTFDVSILTIDD----G-IFEVKATAGDT	226	
H. sapiens HSP70-like1	QGERHVLIFDLGGGTFDVSILTIDD----G-IFEVKATAGDT	228	**9**
X. maculatus HSP70-1	RGERNVLIFDLGGGTFDVSILTIED----G-IFEVKSTAGDT	228	
Alligator HSP70A	AGEKNVLIFDLGGGTLDVSILTIED----G-IFEVKSTAGDT	229	

A

Ka			Clade 1					Clade 2			Clade 3				Clade 4						Clade 5 (node 8)					Clade 5 (node 9)						
	Gene No.		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Clade 1	2	*D. melanogaster* HSP70Aa (AF295933)	0.00																													
	3	*Spodoptera* HSP70 (FJ754276)	0.09	0.09																												
	4	*Bombyx* HSP70 (AB035326)	0.10	0.10	0.03																											
	5	*Paralvinella* HSP70-1 (EF580992)	0.16	0.16	0.14	0.14																										
Clade 2	6	*B. plicatilis* _HSP70-1 (AB775784)	0.15	0.15	0.14	0.14	0.15																									
	7	*B. plicatilis* _HSP70-2 (AB775785)	0.15	0.15	0.14	0.14	0.15	0.00																								
	8	*A. brightwelli* _HSP70 (KX119429)	0.16	0.15	0.15	0.16	0.15	0.06	0.06																							
Clade 3	9	*Wuchereria* (AF167352)	0.16	0.15	0.15	0.16	0.16	0.08	0.09	0.09																						
	10	*Setaria* (AF079360)	0.15	0.15	0.15	0.15	0.16	0.08	0.08	0.09	0.01																					
	11	*C. elegans* HSP70A (M18540)	0.16	0.16	0.16	0.16	0.18	0.09	0.09	0.11	0.05	0.05																				
	12	*C. elegans* HSP-1 (NM_070667)	0.15	0.15	0.15	0.15	0.17	0.08	0.08	0.09	0.04	0.03	0.01																			
Clade 4	13	*M. nipponense* HSP70-1 (KC460343)	0.16	0.15	0.15	0.15	0.15	0.07	0.07	0.08	0.08	0.08	0.09	0.08																		
	14	*M. nipponense* HSP70-2 (DQ660140)	0.15	0.15	0.15	0.15	0.15	0.07	0.07	0.07	0.08	0.08	0.09	0.08	0.01																	
	15	*Mytilus* HSP70 (AY861684)	0.16	0.16	0.16	0.16	0.15	0.09	0.09	0.07	0.09	0.09	0.11	0.10	0.07	0.07																
	16	*Perinereis* HSP70-1 (KU255783)	0.15	0.15	0.17	0.16	0.15	0.09	0.09	0.08	0.10	0.10	0.11	0.10	0.07	0.07	0.06															
	17	*Paralvinella* HSP70-2 (EF580993)	0.17	0.17	0.17	0.17	0.16	0.10	0.10	0.09	0.11	0.11	0.12	0.11	0.08	0.08	0.07	0.06														
	18	*Haliotis* HSP70c (FJ812176)	0.15	0.15	0.16	0.16	0.16	0.10	0.10	0.09	0.10	0.10	0.12	0.11	0.08	0.08	0.06	0.06	0.07													
Clade 5 (node 8)	19	*Oncorhynchus* HSC71 (AAB21658)	0.16	0.16	0.16	0.16	0.16	0.08	0.08	0.08	0.09	0.09	0.11	0.09	0.06	0.06	0.07	0.06	0.07	0.08												
	20	*X. maculatus* HSC70 (AB062115)	0.18	0.17	0.16	0.16	0.15	0.09	0.09	0.08	0.11	0.11	0.12	0.11	0.07	0.07	0.07	0.08	0.09	0.08	0.04											
	21	Human_HSPA8 (NP_006588)	0.15	0.15	0.14	0.15	0.15	0.07	0.07	0.07	0.09	0.09	0.10	0.09	0.07	0.06	0.06	0.07	0.07	0.07	0.03	0.05										
	22	*Rattus* HSC71 (NM_024351)	0.15	0.15	0.14	0.15	0.15	0.07	0.07	0.07	0.09	0.09	0.10	0.09	0.06	0.06	0.06	0.07	0.07	0.07	0.03	0.05	0.00									
	23	*Pelodiscus* HSC70 (HQ219723)	0.15	0.15	0.14	0.15	0.14	0.07	0.07	0.07	0.09	0.08	0.10	0.09	0.06	0.06	0.06	0.07	0.07	0.06	0.03	0.04	0.00	0.00								
Clade 5 (node 9)	24	*Bos* HSP70 (U09861)	0.15	0.15	0.15	0.15	0.16	0.10	0.10	0.09	0.10	0.10	0.12	0.11	0.09	0.09	0.09	0.10	0.11	0.10	0.08	0.09	0.07	0.07	0.07							
	25	Human_HSPA1A (NM_005345)	0.15	0.15	0.15	0.15	0.15	0.10	0.09	0.09	0.10	0.09	0.11	0.10	0.08	0.08	0.09	0.09	0.11	0.10	0.07	0.09	0.06	0.06	0.06	0.00						
	26	*Rattus* HSP70_1A (NM_031971)	0.15	0.15	0.15	0.15	0.16	0.10	0.10	0.09	0.10	0.10	0.12	0.11	0.09	0.09	0.09	0.09	0.11	0.10	0.08	0.09	0.07	0.07	0.07	0.01	0.01					
	27	*X. maculatus* HSP70-1 (AB062113)	0.17	0.17	0.17	0.16	0.16	0.10	0.10	0.09	0.11	0.10	0.12	0.11	0.09	0.09	0.09	0.10	0.09	0.10	0.07	0.08	0.06	0.06	0.06	0.07	0.07	0.07				
	28	*X. maculatus* HSP70-2 (AB062114)	0.16	0.16	0.16	0.17	0.16	0.10	0.10	0.09	0.11	0.10	0.12	0.11	0.09	0.09	0.09	0.10	0.11	0.10	0.08	0.09	0.07	0.07	0.07	0.07	0.07	0.07	0.06			
	29	Human_HSPA1L (NM_005527)	0.15	0.15	0.15	0.16	0.16	0.10	0.10	0.10	0.10	0.10	0.12	0.11	0.09	0.09	0.10	0.10	0.11	0.11	0.09	0.10	0.08	0.08	0.08	0.04	0.04	0.04	0.07	0.07		
	30	Human_HSPA6 (NP_002146)	0.16	0.16	0.16	0.17	0.17	0.12	0.12	0.11	0.12	0.12	0.14	0.13	0.12	0.11	0.11	0.11	0.13	0.12	0.12	0.12	0.10	0.10	0.10	0.07	0.07	0.07	0.12	0.11	0.09	
	31	Alligator_HSP70 (AB306279)	0.16	0.16	0.16	0.16	0.17	0.09	0.09	0.08	0.10	0.10	0.12	0.10	0.09	0.09	0.09	0.09	0.10	0.10	0.08	0.09	0.07	0.07	0.07	0.09	0.09	0.09	0.08	0.09	0.10	0.11



B



1: Invertebrate HSP70s
2: Rotifera HSP70s
3: Nematode HSP70s
4: Invertebrate HSP/HSC70s
5 (8): Vertebrate HSC70s
5 (9): Vertebrate HSP70s

Table 1. Support for major nodes from each approach

		ClustalOmega		Expresso		M-Coffee	
		MrBayes	RAxML	MrBayes	RAxML	MrBayes	RAxML
1	Cytosolic	100	99	100	99	100	100
2	Metazoan	100	95	100	99	100	94
3	Lineage 1	(a)	99	100	99	100	100
4	Lineage 2	(b)	41	85	67	(b)	(c)
5	Rotifer	100	48	74	48	100	65
6	Nematode	94	93 (d,e)	100 (e)	96 (e)	100 (e)	86 (e)
7	Chordate		42	85	35		
8	HSC	100	91	100	86	100	31
9	HSP	100	88	100	83	100	87
10	Outgroup	100	99	100	99	100	100
11	Mitochondrial	100	100	100	97	100	100
12	ER	100	99	100	100	100	100

(a) Rotifer genes (node 5) is included in node 3 with 83% support
(b) Rotifer genes (node 5) sister group to node 3 with 80% support, other paraphyly
(c) Nematode genes (node 6) basal; rotifer genes sister group to node 3 with 46% support, other paraphyly
(d) 36% with Trichuris (immediate outgroup to node 6)
(e) Trichuris not in node 6