

# Cloning and Expression Analysis of Four Heat Shock Protein Genes in Ericerus pela (Homoptera: Coccidae)

Authors: Liu, Wei-Wei, Yang, Pu, Chen, Xiao-Ming, Xu, Dong-Li, and Hu, Yan-Hong

Source: Journal of Insect Science, 14(142): 1-9

Published By: Entomological Society of America

URL: https://doi.org/10.1093/jisesa/ieu032

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at <u>www.bioone.org/terms-of-use</u>.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

# RESEARCH

# Cloning and Expression Analysis of Four Heat Shock Protein Genes in *Ericerus pela* (Homoptera: Coccidae)

# Wei-Wei Liu,\* Pu Yang,\* Xiao-Ming Chen,<sup>1</sup> Dong-Li Xu, and Yan-Hong Hu

Research Institute of Resources Insects, Chinese Academy of Forestry, Key laboratory of Cultivating and Utilization of Resources Insects of State Forestry Administration, Kunming 650224, China

<sup>1</sup>Corresponding author, e-mail: cafcxm@139.com \*These authors contributed equally to this work

Subject Editor: Qili Feng

J. Insect Sci. 14(142): 2014; DOI: 10.1093/jisesa/ieu032

**ABSTRACT.** To explore the function of small heat shock protein genes (*shsps*) and *hsp70* in *Ericerus pela*, we cloned the full-length cDNA sequences of *hsp21.5*, *hsp21.7*, *hsp70*, and *hsc70* and the genomic sequence of *hsc70*. Open reading frames of the four *hsps* were 570, 564, 1,908, and 1,962 base pairs (bp), respectively, which encode proteins with calculated molecular mass of 21.5, 21.7, 69.8, and 71.6 kDa. Amino acid sequence analysis revealed the presence of the conserved Hsp motifs in all four proteins. The genomic DNA of *hsc70* had four introns. *ep-hsp21.5* was orthologous and *ep-hsp21.7* was species specific. Expression of all four transcripts during heat or cold stress and development was examined by quantitative real-time polymerase chain reaction. All four *hsps* were upregulated during heat or cold stress in female adults, indicating a correlation between the four *hsps* and heat or cold-stress tolerance in female adults. *ep-hsp21.7* and *ep-hsp21.7* were upregulated during heat stress in male larvae, implying a correlation between the two *hsps* and heat-stress tolerance in males, and *ep-hsp21.5*, *ep-hsp20*, and *ep-hsc70* were upregulated in females, which indicates their possible role in the developmental regulation of *E. pela*.

Key Words: Ericerus pela, heat shock protein, heat stress, cold stress, development

The Chinese white wax scale insect, Ericerus pela Chavannes (Homoptera: Coccidae), is a coccid insect famous for its role in wax production. The species is highly sexually dimorphic with holometamorphic males and hemimetabolic females. Females develop through the egg, first- and second-stage larva, and adult stages during their life. The female adults only have rudimentary legs and cannot move during the adult period. However, males undergo egg, first- and second-stage larva, propupa, pupa, and adult stages, emerging as winged adults. These white wax scale insects produce only one generation annually, and the females produce eggs in May (spring), which then develop through two larval stages. The insects enter the adult stage in August (summer). Female adults endure the winter and produce offsprings in the subsequent May before they die. The female adult life spans over >180 d, with a duration that lasts three seasons: summer, autumn, and winter. During its life, the female adults of *E. pela* face summer heat stress and winter cold stress. Therefore, tolerance of female adults to high and low temperatures is critical to the survival and reproduction of *E. pela*. In contrast, adult males have a life span of <1 wk, and they die just after copulation in August. In male larvae, the second instar is the longest developmental stage of 70-80 d, during which they continuously secrete large amounts of wax. During this period, heat stress and high humidity can enhance male larvae to secrete wax, forming a thick wax layer on the surface of the body (Zhang and Liu 1997, Chen et al. 2007a, Chen and Feng 2009, Chen 2011). The larvae and the pupae stay in the wax layer until eclosion. To increase wax production, E. pela are usually transported to areas of high temperature and humidity to facilitate wax secretion by the male larvae. Therefore, tolerance of second-stage male larvae to high temperature stress is critically important for the wax production and male survival.

*E. pela* has been bred for over 1,000 yr, and this insect species is widely distributed in most parts of China and elsewhere, Japan, and the Korean peninsula. For wax production, high-quality females are produced in regions with an annual average temperature of  $11-16^{\circ}$ C, with the lowest temperature between -5 and  $-2^{\circ}$ C, an annual rain fall of

800-1,200 mm/a, and an annual light period of 1,900-2,500 h/a. In attempts to introduce E. pela to various geographical regions, we found that E. pela was more adaptable and plastic with respect to the environment than most of other insects. It is widely distributed in the subtropical, warm temperate and temperate areas. In Changchun City of China, where the lowest temperature is below  $-30^{\circ}$ C and the lowest average monthly temperature is about -13°C, E. pela can survive and produce offsprings. In Chongqing City, where the highest temperature is 44°C and the highest average monthly temperature is about 26°C, E. pela can also survive and carry on reproduction (Zhang and Liu 1997, Chen and Feng 2009, Chen 2011). This remarkable tolerance of E. pela to extreme hot and cold temperatures is important for wax production. E. pela also provides a model to study the mechanism of extreme temperature stress adaptation in insects (Lalouette et al. 2011, Dunning et al. 2012, Reiskind and Zarrabi 2012). Most studies on E. pela have focused on white wax production and related ecology (Jiao and Zhao 1999; Chen et al. 2008, 2011); however, the molecular mechanisms underlying the adaptability of E. pela to extreme temperatures have not been reported.

Heat shock proteins (Hsps) are induced by a diverse array of stresses, such as desiccation (Lopez-Martinez et al. 2009), hypoxia (Michaud et al. 2011), chemicals (Planelló et al. 2011), metals (Zhang and Zhang 2012), ultraviolet radiation (Sang et al. 2012), parasitic infections (Shim et al. 2008), and high population densities (Wang et al. 2007). Hsps act as molecular chaperones and play critical roles in facilitating correct protein folding or refolding, preventing aggregation of denatured proteins, degrading misfolded or aggregated proteins, and transporting and assembling or disassembling proteins (Gething and Sambrook 1992, Sanders 1993). Hsps are divided into several families according to sequence homology and molecular mass, including Hsp90, Hsp70, Hsp60, Hsp40, and small Hsp (sHsp), and each family comprises one or several members. Among the Hsps, Hsp70 is well studied and is the major Hsp family that comprises both inducible and constitutive protein members. The constitutive proteins are referred to

<sup>©</sup> The Author 2014. Published by Oxford University Press on behalf of the Entomological Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

as heat shock cognate proteins (Hscs), which do not show significant change in response to stress. Proteins of the shsp family range in size from 12 to 42 kDa, most of which are smaller than 30 kDa and more diverse than other inducible Hsps in the sequence, except a conserved  $\alpha$ -crystallin domain of ~90 amino acids (Sørensen et al. 2003, Stahl et al. 2007, Lü and Wan 2011, Hanazono et al. 2012). In insects, Hsps are known to confer temperature tolerance. In addition to this function, Hsps are also integral to normal cell growth and development, and the expression patterns of Hsps are in general consistent with enhanced stress such as cold winter temperatures encountered by insects (Feder and Hofmann 1999, Morrow and Tanguay 2012). Considering the critical role of Hsps in response to irreversible stress conditions in many insects, it is presumed that Hsps can also be involved in the temperature adaptation of *E. pela*. To explore this possibility, the full-length cDNA sequences of four hsps (hsp21.5, hsp21.7, hsp70, and hsc70) were selected from the transcriptome of E. pela and cloned. In addition, the hsc70 genomic sequence was also characterized. Transcriptional expression patterns of the four E. pela hsps under temperature stress and during various developmental stages were examined by quantitative real-time polymerase chain reaction (qRT-PCR). Our results of this study will promote greater understanding of the basis for the adaptation and population variations within E. pela at a molecular level.

# **Materials and Methods**

**Insects.** *E. pela* samples were collected from the Chinese private *Ligustrum lucidum* in the Research Institute of Resources Insects.

Samples of *E. pela* were collected at the following developmental stages: a mixture of male and female eggs, first- and second-stage larvae of male and female, male pupae, male adults and female adults at early- (FA1), middle- (FA2), and late stages (FA3).

To induce heat stress, female adults and male second-stage larvae, which have the longest developmental stages for each sex, were exposed to the following temperatures for 1 h: 45°C, 42°C, 39°C, 36°C, and 33°C. To simulate cold stress, temperatures were set at  $-20^{\circ}$ C,  $-15^{\circ}$ C,  $-10^{\circ}$ C,  $-5^{\circ}$ C,  $0^{\circ}$ C, and 5°C. Female adults and male second-stage larvae maintained at 25°C were used as controls. Immediately after treatment, all samples were separately homogenized in TRIzol reagent and then stored at  $-70^{\circ}$ C until RNA extraction. Eggs were used as a control group. Three replicates were prepared for each experimental group.

**cDNA Synthesis and RACE.** Total RNA from adult females was isolated using TRIzol reagent (Invitrogen, California, USA) according to the manufacturer's instructions and used as template to generate cDNA using SMARTer RACE (Rapid Amplification of cDNA Ends) cDNA Amplification kit (Clontech, CA). Specific primers used for RACE were designed based on *hsp* sequences identified in the *E. pela* transcriptome (Yang et al. 2012). Primer sequences for *hsp21.5*, *hsp21.7*, *hsp70*, and *hsc70* used in 5'- and 3'-RACE are listed in Table 1. PCR products from 5'- and 3'-RACE reactions were cloned into the pGEM T-Easy vector (Promega, Shanghai, China) and sequenced. Assembly of all sequences of cDNA products was done by using DNAMAN software (version 6.0, Lynnon Corporation; http://www.lynnon.com/pc/pcmain\_new.html).

**Amplification of Genomic DNA.** Genomic DNA was isolated from adult female *E. pela* using DNeasy Blood & Tissue kit (Qiagen, Dusseldorf, Germany). Primers (Table 1) located in the 5'- and 3'-untranslated regions (UTRs) of *hsc70* cDNA were used to amplify *hsc70* genomic DNA. Thermal cycling conditions were as follows: 94°C for 4 min followed by 31 cycles of 94°C for 1 min, 56°C for 1 min, and 72°C for 5 min, with a final extension at 72°C for 10 min.

**Sequence Analysis.** Amino acid sequences of insect Hsps were downloaded from NCBI after BLAST search (http://www.ncbi.nlm.cih.gov/ blast/Blast.cgi). Sequence alignments were conducted by using DNAMAN. Open reading frames (ORFs) were identified using the ORF Finder (http://www.ncbi.nlm.cih.gov/gorf/gorf.html). Molecular mass of the predicted proteins was calculated using the Compute pI/Mw program available online (http://www.expasy.ch/tools/pi\_tool.html), and *E. pela* Hsps were named according to their predicted molecular mass.

#### Table 1. Primers used in the study

		,
Gene		5'- to 3'-Primer sequences
Primers use	d in RACE	
hsp20.5	5′-1	GGGCCTCAACTGTCAATACACCATCT
	5′-2	TTGTCCACCGTCTTGACCACGATTTCCT
	3'-1	CGACTAGCACCTCAAATACCGT
	3'-2	TGAAGCTGCGATTCGACGTGAG
hsp20.7	5'-1	TAGCTGAGGTATTCGTGCGAGTGAT
	5′-2	CCTCGATGACCAAATAATCGTCCACAAT
	3'-1	GTTCAGCAATTCAAACCCGAAGAAAT
	3′-2	CGCCAATTCACACGTCGTTACAGATT
hsp70	5'-1	CGATTTCCTCCTTTGACAGTCTTCCT
	5′-2	GTCGATCTAAACAAGTCCGAGCACAAT
	3'-1	GATTCAAGGATGAGGACGAGAAACAG
	3′-2	GAACACGTTGGCCGAGAAAGATGAGTAT
hsc70	5′-1	TCCAACCGGCGAATCGTTTCGTTACAT
	5′-2	GTCTTTGCTAAGTCTACCCTTGTCGTT
	3'-1	GCCAACGGTATTTTGAATGTGAGT
	3′-2	CGCCGGTTGGACTCGAATCAATTAGCT
Primers use	d in genomic DNA cl	one
hsc70	5′	ACGGTGAAACTGAAACTGAAG
	3′	CTCCAGCGACATTGTATTATTC
Primers use	d in PT DCP	
hsp20.5	F	GATAACGAGTTCAGCAGCAT
risp20.5	R	TGGATACGAGTATTTGAGGTG
hsp20.7	F	GCCAATTCACACGTCGTTAC
11sp20.7	R	CGGAGCAACAATCGATAGGA
hsp70	F	CTCGTGGTATCCCTCAAATC
lisp70	R	CCTCCTTTGACAGTCTTCCT
hsc70	F	ACGGTATTTTGACAGTCTTCCT
115070	F	TCTTCAACGGTGCTCTTCAT
P actio	F	CCACGAGACGACCTACAAT
$\beta$ -actin	F	CGATCCATACGGAGTACT
	n	COALCCATACOOAOTACT

A neighbor-joining tree was constructed using the Poissoncorrection distance method (Nei and Kumar 2000) based on the fulllength amino acid sequences. Alignment of the selected amino acid sequences was first performed using the CLUSTALX1.83 software (Thompson et al. 1997) and then imported to MEGA5.0. Phylogenetic trees were generated using the neighbor-joining method and bootstrapped with 1,000 iterations to evaluate the branch strength of the tree. Bootstrap percentages of 1,000 replicates that were >50% were shown at the nodes (Tamura et al. 2011). sHsp from *Arabidopsis thaliana* and *Glycine max* were used as outgroups.

**qRT-PCR Analysis.** Specific primer pairs were designed based on the sequences of *hsps* and  $\beta$ -actin from *E. pela* (Table 1). qRT-PCR was performed on a CFX96 Real-Time PCR Detection System in 20 µl reactions that included 10 µl SsoFast EvaGreen Supermix (Bio-Rad, CA), 1 µl of each of the forward and reverse primers (20 µM), and 2 µl of cDNA template. PCR cycling conditions were as follows: initial denaturation at 95°C for 3 min, followed by 30 cycles at 95°C for 10 s and 55°C for 20 s. Melting curves were generated to ensure specificity of the PCR products with temperatures from 65 to 95°C, increased in 0.5°C, 5 s increments. Relative gene expression levels of each *hsp* were calculated using the  $2^{-\Delta\Delta Ct}$  method (Livak and Schmittgen 2001) with  $\beta$ -actin as an internal standard for normalization. The least significant difference test at *P* < 0.01 level was performed with the DPS statistical program to analyze gene expression variations (Tang and Feng 2010).

# Results

**Sequence Analysis of** *ep-hsp21.5* and *ep-hsp21.7*. The full-length sequences of *ep-hsp21.5* and *ep-hsp21.7* (GenBank accession numbers: KC161298 and KC161299, respectively) were 997 and 873 bp in length and contained 570 and 564 bp ORFs, respectively. The ORF of *ep-hsp21.5* encoded a putative protein of 189 amino acids with a calculated molecular mass of 21.5 kDa, and the ORF of *ep-hsp21.7* encoded a deduced protein of 187 amino acids with a molecular mass of 21.7 kDa. *ep-hsp21.5* contained a 287-bp 5'-UTR and 140-bp 3'-UTR,

whereas *ep-hsp21.7* had a 150-bp 5'-UTR and 159-bp 3'-UTR. The termination codon (TGA) was located at 855 and 712 bp, followed by the polyadenylation signal AATAA (Huang et al. 2008) in *ep-hsp21.5* and ATTAA (Beaudoing et al. 2000) in *ep-hsp21.7* (Fig. 1). Blastp comparisons of the deduced ep-Hsp21.5 amino acid sequence shared a 77% identity with sHsp of *Spodoptera litura* (ADK55519), and ep-Hsp21.7 shared a 55% identity with the putative sHsp from *Maconellicoccus hirsutus* (ABM55532). Multiple sequence alignment showed that all sequences contained the conserved  $\alpha$ -crystallin domain (Fig. 1), indicating that the isolated *ep-hsp21.5* and *ep-hsp21.7* sequences were *shsps*.

The phylogenetic tree analysis showed that ep-Hsp21.7 was clustered into the species-specific sHsp group, whereas ep-Hsp21.5 was clustered into the orthologous sHsp group (Fig. 2).

**Sequence Analysis of** *ep-hsp70* and *ep-hsc70*. Full-length cDNA sequences of *ep-hsp70* (GenBank accession number: KC161300) and *ep-hsc70* (GenBank accession number: KC161301) were 2,355 and 2,417 bp in length with 1,908 and 1,962 bp, respectively. *ep-hsp70* encoded a deduced protein of 635 amino acids with a calculated molecular mass of 69.8 kDa. *ep-hsc70* encoded a deduced protein of 653 amino acids with a calculated molecular mass of 71.6 kDa. *ep-hsp70* contained a 182-bp 5'-UTR and 265-bp 3'-UTR, whereas *ep-hsc70* contained a 181-bp 5'-UTR and 274-bp 3'-UTR. The termination codon (TAA) was present at 2,088 (*hsp70*) and 2,141 (*hsc70*) bp, followed by the polyadenylation signal (AATAAA) and a poly (A) tail. Three characteristic signature sequences of the Hsp70 family were observed in the amino acids sequences of ep-Hsp70 and ep-Hsc70, including IDLGTTYS, IFDLGGGTFDVSIL, and IVLVGGSTRIPKIQK (Figs. 3

and 4). The conserved motif EEVD usually present in the *C*-terminus of cytoplasmic Hsp70 was altered to DEVD in ep-Hsp70. A Blastp search revealed that the deduced amino acid sequence of ep-Hsp70 shared an 81% identity with *Bemicia tabaci* Hsp70 (ADO14473) and ep-Hsc70 shared a 92% identity with *Nilaparvata lugens* Hsc70 (ADE34170).

The phylogenetic tree analysis of Hsp70 showed that ep-Hsp70 was clustered together with inducible Hsp70, whereas ep-Hsc70 was clustered with cognate Hsc70 of other species. However, both ep-Hsp70 and ep-Hsc70 were clustered with cytoplasmic Hsp70 (Fig. 5).

The genomic DNA sequence of hsc70 (GenBank accession number: KC161301) had four introns, which were 1,139, 255, 514, and 94 bp in length. The first intron was the longest from 173 to 1,311 bp and located in the 5'-UTR, whereas the remaining three introns were located in the coding regions, from 1,526 to 1,780 bp, 1,930 to 2,443 bp, and 2,652 to 2,745 bp, respectively (Fig. 4). The intron–exon junctions of the first three introns followed the GT-AG rule, whereas the fourth intron–exon junction was AG-TC.

**Expression of the** *ep-hsps* in Female and Male *E. pela* Under Temperature Stress. qRT-PCR revealed that in females, the mRNA expression levels of all four *hsps* were correlated with heat stress. In adult females at 33°C, the expression of *hsp21.7* and *hsp70* were drastically upregulated by 65.02 and 145.31 folds, respectively, over the control expression at 25°C (Fig. 6B and C). mRNA expression of *hsp21.5* and *hsc70* was also significantly upregulated, when compared with the control (Fig. 6A and D).

In male second-stage larvae, mRNA levels of *hsp21.7* increased when the temperatures were between 33 and 39°C and peaked at 39°C,

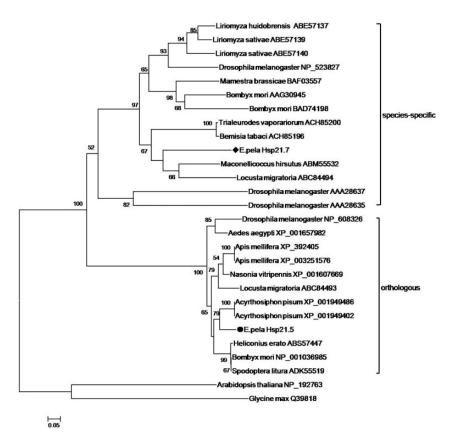
1	AC	ATG	GGG	GAGT	TTA	TCA	TTG	ATA	ACG	GTO	GTGC	TAG	CAC	GACO	GTTC	TTC	TTC	CAGO	GTT	TAT	TT	CAC	GAA	TTI	TTI	TCG	CCC	сто	GTT	TTTA
91	CG	TTT	CAT	TAG	TTG	CGA	TTI	ATT	CAT	TTO	TTT	ATT	TCI	TT	TGG	TGG	TG/	ATTG	AAT	TTC	GCC	TTC	TTT	TAT	TGI	AAT	TGO	AGC	GCG	ATTI
181	TC	AGC	AGT	TTT	AAC	TGA	TT	TTC	TAC	GAA	TTT	TTI	TCI	CTT	TGG	TTT	TTO	CGC	TTC	TAC	TAC	тст	TTT	TTI	CCA	TCG	TT	GTT	TTT	TGCI
271	TA	GTG	CT/	тα	ACA	CAA	ATG	TCT	GAC	AGO	GGA	CGT	'AA/	CAC	CAAT	ATC	CC/	ATC	AAA	TTO	GGG	GAC	TTC	AAC	GTO	ATT	GAT	AAC	GAG	TTC/
1							M	S	D	S	G	R	K	Н	N	I	Р	I	K	L	G	D	F	N	V	Ι	D	N	Е	F
361	GC	AGC	ATI	CGC	GAA	CGG	TTC	GAC	GCT	GAA	ATG	CGT	AA/	ATO	GAA	GAC	GAC	GATG	ACC	AAA	TTO	CGT	TCT	GAA	TTO	ATG	AAT	CGC	GAG	TCGA
25	S	S	I	R	Е	R	F	D	A	E	M	R	K	M	E	D	E	M	Т	K	F	R	S	E	L	M	N	R	Е	S
451	AC	TTT	TTC	CAAA	ACA	TCG	ACT	AGC	ACC	TCA	AAT	ACC	GT/	TCO	CAGC	AGT	TC/	CTA	GAC	CAAT	CAC	GGA	GCG	CAG	CAC	CAAG	AAC	ACC	TGG	TTGO
55	N	F	F	K	Т	S	Т	S	Т	S	N	Т	V	S	S	S	S	L	D	N	Q	G	А	Q	H	K	N	Т	W	L
541	AC	AAC	ATT	TAAT	TCG	coo	сто	ATC	CAG	GA/	GAT	GGC	GAG	GGG	CAAA	AGT	CTO	GAAG	CTG	CGA	TTO	GAC	GTG	AGC	CAA	TAC	CAC	CCC	GAG	GAA/
85	D	N	I	N	S	Р	L	I	Q	E	D	G	D	G	K	S	L	K	L	R	F	D	V	S	Q	Y	Q	Р	E	E
631	TC	GTG	GTO	CAAG	ACG	GTG	GAC	AAC	AAA	TTA	TTG	GTI	CAT	GCT	TAAA	CAC	GA/	GAA	AAA	TCC	GAC	ACA	AAA	TCC	GTO	TAC	CGA	GAA	TAC	AACO
115	I	V	V	K	Т	V	D	Ν	K	L	L	V	H	A	K	H	E	E	K	S	D	Т	K	S	V	Y	R	E	Y	N
721	GT	GAA	TTO	TTG	TTG	coo	AAA	GGT	ACG	AAT	CCC	GAG	ACC	CATO	CAAG	TCG	TCO	STTG	AGC	AAA	GAT	GGT	GTA	TTG	AC/	GTI	GAG	GCC	CCCG	TTAC
145	R	Е	F	L	L	Р	K	G	Τ	N	Р	E	T	I	K	S	S	L	S	K	D	G	V	L	T	V	E	A	Р	L
811	CG	GTT	СТС	GCA	GGT	CCA	GAA	AAG	CTG	ATA	CCA	ATI	GAG		AAA	TGA	AC/	CTC	TGI	ATT	GA/	TCG	CCC	AGA	AAG	TAT	TTI	TTA	ATA	ATTI
175	Р	V	L	Α	G	Р	E	К	L	I	Р	I	D	K	K	*														
901	AA	AAT	CGI	AAT	TTA	AAA	AAC	AAA	AAA	GTA	TAT	TAT	GT/	TT	TAA	AGA	AA/	AAG	AAA	CTA	AA/	GGA	GAA	CCA	AAA	AAA	AAA	AAA	AAA	AAAA
991	AA	AAA	AA																											

#### ep-hsp21.5

ACATGGGGGTTTTATACTGATTCAGTATTCGTAGTATTTTTGCCGTTAGTCGTTATTGCATGTAGTTTGAATCGTTTAAAGTGAAGTGT 1 91 CATTTCGTACTCGATAACTGCTAGTTAGTGTTCAAAAGCGTCTTCAATTTCTTGGCAAAGATGTCTGTACTACCGTATGCGTTGAACGAC MSVLPYALND 181 TGGTGGTACGATGCCCGTGATCCTTTTAGTCGTCTGTACGATCAACACCTTCGGTATGGGTGTTTGATGATGATGATATCATGCTGAATAGA 11 W W Y D A R D P F S R L Y D Q H F G M G V F D D D I M L N R 271 CCTTCGTGGGCACCTCTTACCGTTCCTTTGCGGTCTGGATACCTGCGGACGCTGAGACCGTTTGTGTTGGAAGACAGTGGTGTCTCGTCA 41 P S W A P L T V P L R S G Y L R T L R P F V L E D S G V S S 361 GTCGAAAATGAGGAAGATAAAGTGAAGATCAATCTAGACGTTCAGCAATTCAAACCCGAAGAAATTAGCGTTAAAATTGTGGACGATTAT 71 VENEEDKVKINLDVQQFKPEEISVKIVDDY 451 TTGGTCATCGAGGGAAACCACGAAGAAAAACAAGATCAGCACGGTTACGTATCGCGCCAATTCACACGTCGTTACAGATTGCCGGAAAAT 101 I E G N H E E K Q D Q H G Y V S R Q F T R R Y R L P E N 541 ATCATAAAAGATGATATTTCCAGTTCGATTTCTTCAGATGGAGTCCTATCGATTGTTGCTCCGAAAAAAACCGGAAGCCATAAGAAATCAA 131 IIKDDISSSISSDGVLSIVAPKKPEAIRNQ 631 ACTGAGAGGGAAAATTCCGATCACTCGCACGAATACCTCAGCTATAAAAACATAATAAGGGGGAGTGAAAAATGAGAGAAAAAAATAGAATCGT 161 TERQIPITRTNTSAIKHNKGSENERKK\* 721 CGAATTCGGCATCGTTATTCGTAACAGCTTATTATTTTCCCTTTGAAATTCTATGTGGACTGATCAGGCGTGTAGAGCTGTTTTTATGGT 811 

#### ep-hsp21.7

Fig. 1. Complete nucleotide and deduced amino acid sequences of *ep-hsp21.5* and *ep-hsp21.7* cloned from *E. pela*. Translation termination codon is marked with an asterisk, putative polyadenylation signal is boxed, and the alpha-crystallin domain is underlined.



**Fig. 2.** Phylogenetic analysis of Hsp20 sequences. Some orthologous sequences used in this analysis are from Li et al. (2009). Sequence labels are denoted by the species names and GenBank accession numbers. Black diamond indicates the sequence of ep-Hsp21.7, and black round indicates the sequence of ep-Hsp21.5.

181     0       271     0       230     4       231     0       30     4       641     0       90     5       91     12       541     0       721     6       721     1       901     1       240     1       1210     1       1301     2       300     1       1371     2       1381     1       1381     1       1381     1       1381     1       1381     1       1381     1       141     1       1431     1       1431     1       1431     1       1445     1	CAATG K CCCAACA A N CCTTTG A L GGCCTT F N CTTCTC A L TCCAT E N CTCTCCA C N CTCCCA T H GAGCCC R A GAGCCT R A CTCACS D N CTCCAC C N CCCACCA C N CCCACACA C N C CCCACACA C N C CCCACACA C N C C C C C N C N C N C N C N	CCCAN P CATCO D AATCO N TTTTA F CATCO CCCAN CCCAN R AATCA L CCCAN R AATCA L CCCAN R AATCA L CCCAN R AATCO L	CAAAAA AGATC XAAGAA Q G COCAGC Q COCAGC Q COCAGC Q COCAGC Q COCAGC Q COCAGC Q COCAGC Q COCAGC Q COCAA S E AAGAA X ACGAA X C COCAC Q COCAA X C COCAC Q COCAA X C C C C C C C C C C C C C C C C C C C	AATTA N AATTA I AACAA T CCCCC R TTTCG G ACTAT G GAGT E CTAT L GAATTA G GAGT E CTAT L GAATTA G GAGT C C C C C C C C C C C C C C C C C C C	A V CAACO R T I ATGA T I AATGA N D CAACT T I AATGA AAAATT S L TCAAC F K K CCCCT S S TCCAAC F S S S TCCAAC S S S S TCCAAC S S TCCAAC S S TCCAAC S S TCCAAC S S S S S S S S S S S S S S S S S S S	GGGTT GGACA T TTTTC F GAAA K TACC T GAAA K C GAACA R GAACA R GAACA C F GAACA S GGTCG	ATTC T COCC/ P COCC/ P COCC/ P COCC/ COCC/ P COCC/ COCC/ COCC/ P COCC/ P COCC/ P COCC/ P COCC/ P COCC/ P COCC/ P COCC/ P COCC/ P COCC/ P COCC/ P COCC/ P COCC/ COCC/ P COCC/ P COCC/ COCC/ P COCC/ COCC/ P COCC/ COCC/ P COCC/ COCC/ P COCC/ COCO	AAAGCA BAAAGCA BAAAGCA CR BAAAGCA CR BAACCA CR CAAAAGCA CAAAAAGCA CAAAAGCA CAAAAAGCA CAAAAAGCA CAAAAAGCA CAAAAAGCA CAAAAAGCA CAAAAAAGCA CAAAAAGCA CAAAAAGCA CAAAAAGCA CAAAAAAGCA CAAAAAAGCA CAAAAAAGCA CAAAAAAGCA CAAAAAAAA	TTGGGL G ACGT Y V AACG K R CGAA P R CTGAA A E CCGGA G D TTAG G D TTAG G D TTAG K R K CTAA A N TTAG F R	CACC TCCCC A TCCL L AACT L AACT TCCC A AACC AACC	TTTT F AATT I CCAG Q CTAC Y CATC I AATT I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI C CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I CCAA CATTI I C C C C C C C C C C C C C C C C C	TAT Y ACO T G G G C C C C C C C C C C C C C C C C	TCGT S GAC/ D CCGCC R GAC/ R GAC/ R GAC/ C GGCT G GAC/ C C GGCT G GAC/ N C GAC/ N C GAC/ N C GAC/ N C GAC/ N C GAC/ N C GAC/ N C GAC/ N C GAC/ N C GAC/ C R C GAC/ C C GAC/ C C GAC/ C C GAC/ C C GAC/ C C GAC/ C C GAC/ C C C C C C C C C C C C C	C C T C C T T R T A T A C C C T T A A C C A C C A C C T T A A C C T T A C C T T R T A T A C C T T C C T T R T A C C C T T C C T T C C T T C C T T C C T T C C T T R T A C C C T T C C T T R T A T A C C T T R T A T A T A T A T A T A T A T A T	V SAGA E TATC Y AAAC K AAAC N TTCC F CATT H CATC H CATC L CTTCC	GTGT G V AGGCT R L GAGGT G V GTTAC V I GTTAC V I GTTAC I G GTCAC I G G GTCAC I G G GTCAC I G G G G G G G G G G G G G G G G G G	ATGC GATT GATT GATT GAA K TGC A C C C C C C C C C C C C C C C C C C	Q A G G A A A A A A A A A A A A A A A A	CAAG Q GATG D ATCA I ACTI T GTCA V ATTA GGAAT D GATI D GATI D GATI D AAGG	GGG/ GGCTC A T T T T T T T T T T T T T T T T T	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	STCGA V E AAGAA K N GATAT D N CCTGA P E STTCC V F CCAAC P I GATGT D V AACCA T A ACCAA T F AAATT	AMATC I ITCAG I Q COCCT I	I SGTAG V CATT H AATCA I TACT Y CCAAC A AATCC I CCAAC V CCAAC K
1       2271     0       361     0       361     0       361     0       541     0       120     5       120     5       120     5       120     5       120     5       120     5       120     5       120     5       120     5       120     5       120     5       120     5       120     5       120     5       120     5       120     5       1310     1       1351     7       3900     6       14410     6       4200     1       4501     1       4501     1	W CCCAACA A N CCCTTG C CTTG C S S S S S S S S S S S S S S S C	P GAATC D AAATCA N F TTTA F ATCA M TTTA F ATCA A CGCGT A A AAAAC K CCGGT R A AAAAC K CCGGT R A AAAAC L L	K M AAGGA Q G CCAAGG P R CCGTT T V TGCTC V L CTCAG S Q ACGGA Y G GCGAA Y G GCGAA Y G GCGAA CGAACG R T TTGAA F E ACGAA R T TTGAA F C ACGAA C R T TTGAA C R T C TTGAA C C AAGGA Y C C CCGTT T V L CTCAC C C C C C C C C C C C C C C C C C	P AAATC N AAATA N AATAA T AACCA T TCCGCC R TTCCG C R TTCCG C R TTCCG C R TTCCG C R TTCCG C R TTCCG C R TTCCG C C R TTCCG C C R T C C C C C C C R T A ATAA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N AATTA N N N AATTA N N N AATTA N N N AATTA N N N AATTA N N N N	A V GAAG R T LCAAT T I ATGAT N D AAAT R M CAAGC Q A GAAGC Q A GAAGC S L TCGCT S L TCGCT S S TCTCG C L C C TCTTCT V L	G GACA T TTTC F TAAC N GAAA K TACC T GAAA K GAAA R GAGA R GACA T CTCCG S GGCG	T CCCC/ P CCCC/ D CCCC/ S CCCC/ S CCCC/ C C CCCC/ C C CCCC/ C C CCC/ C C C CCC/ C	D AGCT S GCCA A AAAGC K T S GACG C T AAAAGC T AAAAGC T T S AACCC T T S AACCC C T S AACCC C T S AACCC A T S S ACCCA A A AAGCC T S S C C A A AAGCC T S S C C A A AAGCC T S S C C A A AAGCC T S S C C A A AAGCC T S S C C A A A C C C C A A A C C C C A A A C C C C A A A C C C C A A A C C C C A A A C C C S T S S ACCC S S A C C S S C A A A C C C S T S S ACCC S S ACCC S S ACCC S S C S ACCC S S C S ACCC S S C S ACCC S S C S S C C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S	L G ACGT Y V AACG K R CCGAA P R CCGAA P R CCGAA CCGG A G GTGA A G GTGA A G GTGA CCGG A G GTGA CCGG A G GTGA CCGG A C CCGA A C CCGA A C CCGA A C CCGA A C CCGA A C CCGA A C CCGA A C CCGA A C CCGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCCGGA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C CCGGAA C C CCGAA C C CCGAA C C CCGAA C C CCGAA C C CCGGAA C C CCGGAA C C CCGGAA C C CCGGAA C C C C	T CCCC A TTCL L AACTO T TCCC A ACTO T TCCC A CAAA T T ACAA T T ACAT T ACAT T A CAAT T T CCCC A A TCL A C T C CCCC A A TCL A C T C C C A A C T C C A A C C A A C C A A C C A A C C A A C C A A C C A A C C A A C C A A C C A C A C A C A C A C C A A C C C A C A C C A C A C C A C C C A C C C C A C C C C A C	T TTTT F MATT I CCAG Q TACC Y CATC I MAAT N GCC A T T T GAA E GAA T T	Y ACO T CGTC G CTCO L CCCO A CGTC V V CCTO C C CCO A T C C C C C C C C C C C C C C C C C C	S GAC/ D CCCCC R GACT E GGCTC G GCTC G CTCG/ L [ GAC/ C CTCG/ D AAC/ N GATC D CCCCC C CTCG/ C CTCG/ C CCCCC C R C CCCCC C R C CCCCC C R C C CCCCC C R C C CCCCC C R C C CCCCC C R C C CCCCC C R C C CCCCC C R C	C ACCC T T CGTT R ATA G GTA A G GTA A C CAG	V GAGA E TATC Y AAAC R AAAC R AAAC R C TTCC F T C ATC H C ATC H C TTCC L GTTCC	G V AGOCT R L GATGA D D GGAGG G V GTTAC V L GATTT V L GATTT V L GATTT D L CTAGG L G CGTGC R A GAOGC D A GAAAA	W GATT I TTCCC P CAAAA K TTGAT D CAACGC R GGGGT G G G G CTTCCA L CAACGC C L CAACGC	Q GGAAA K K AAAAA K GGGGA G G G G G G G G	Q GATC D ATCA I GTCA V ATTA I GGAA G GATT D CCGAC R GATT D D AAAGG	G A A T T T T T T T T T T T T T T T T T	K GCA/A Q Q TCCC S ACAC TCCC S ACAC TCCC TCCC GAC/ D ACAC TCCC TCCC TCCC TCCC TCCC TCCC TC	V E AAGAA E N GATAT D N CCTGA P E GTTCC V E CCAAC P I GATGI D V AACCAA T A AACCAA T E	E I ATCAGE ATCAGE ATCAGE A ACCAGE C A ACCAGE C A A CAGECCA A A CAGECCA A A CAGECCA A A CAGECCA A CAGECCA A A A A A A A A A A A A A	I GTAC V CATTI H MATC/ I TACTI Y CATCA A A A A A CATCA A CATCA A CATTA CATCA A CATTA
2271     2       30     J       30     J       30     J       30     J       41     G       90     J       91     J       92     J       93     J       941     G       95     J       96     J       9721     G       186     J       9721     G       186     J       991     G       991     G <td>CCCAAC A N CTTTG GGCCT W P GGTTCT S S S S CTCTCC A L CTCAC F N CTCTCCAC F N CTCTCCA A L C N CCCCAT F N CCCCAT F N CCCCAT F N CCCCAT F N CCCCAT C N C CCCCCA C N C C CCCCCA C N C C C C C C C C C C C C C C C C C C</td> <td>CATC D AAATCA F ATCA GCATT D CCCA A ATCA L CCCGA K CCCGCI R A ATCA L CCCGCI R A ATCA L L</td> <td>AAGGA Q G CCAATI</td> <td>AAATC N AAATA N AATTA I AACCA T CCCCC R TACCA C CCCCC R TACCA C CCCCC R TACCA C CCCCC C R TACCA C CCCCC C R TACCA C R C CCCCC C R TACCA T C CCCCC R C C CCCCC R C C C C C C C C</td> <td>CAACO R T CAAT T I ATGAT N D AAAT R M CAACA Q A ACAAC Q A ACAAC D R CCCT S L TCAAC F R F CCCTC S S TCTCC CCTCC S S TCTCC CCTCC TCTCC CCCTCC S S</td> <td>GACA T TTTC F GAAA GAAA TACC T GAAA T GAAA C GAAA C C C C C C C C C C C C C C</td> <td>CCCC/ P CCATC D S CCAC/ S CCAC/ E CCACC C C CCC/ P CCACC C C C CCC/ P CCACC/ C C CCC/ P CCACC/ C C C C C C C C C C C C C C C C</td> <td>AGCT S GCCA A AAAGC R T GAACG T GAACG T GAACG T GAACG V V T GAACG V V C C AAAGC V V C C AAAGCC R T GAACG C T C C C A A AAGCC R T C C A A AAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A C C C C A C C C C A C C C C C C</td> <td>ACGT Y V AAACG K R CTGA A E CTGA A E CTGA A E CTGA A G GTGA C TTAGG K R CTAA A N TTAG F R</td> <td>CGCC A TCT. L ACTC L AACTC T TCCC A CAAA CAAA T ACAA T ACAA T AACC A CAAA T T ACAA T T ACAA T T A CAAA T T T C C A T T T C C A T T C L A ACTC A T T C L A ACTC ACTC ACTC ACTC A ACTC A</td> <td>FTTT F MATT I CCAG Q TAC Y CATC I MAAT N N GCC A T T T GAA E SACG T</td> <td>ACCO T G G G G G G C C C C C C C C C C C C C</td> <td>GACA D CCCCC R GAGT G GGT G GGT G GGT C CCCA D D CCCCA D CCCCA D CCCCA C GGT C CCCA C G GGT C CCCCC C C C C C C C C C C C C</td> <td>ACCCC T CCTT R FATA SGTA G GTAA G TTAA C T AACCC T AACCC T AACCCC A CCAG</td> <td>EAGA E TATC Y AAAC R R AAAC R R AAAC R R AAAC F TATC H CATC H CATC L GTTC</td> <td>R L GATGA D C GATGA G V GTTAC V I GATTI V L GATTI D L CTAGG C GTGC R A GAGGC D A GAAAA</td> <td>GATT TCCCC P GAAA K TTGAT D C C C C C C C C C C C C C C C C C C</td> <td>GAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAAAA</td> <td>GATCA D ATCA I ACTI T GGCA V ATTA I GGCA I GGATI D CGAC R GATI D AAAGG</td> <td>A ACTO T T TCT F ATC/ I ACA T T TCC/ I ACA T T TCC/ F TCC/ F C TC/ F SACO</td> <td>GCA/ Q Q TCGX S ACAC T GAGA T TC GAGA/ R TAT/ Y GCG/</td> <td>A GAAATI</td> <td>TCAG GCT Q GCCT A ACCC A A CTCCT A A CTCCT A A CTCCT A C A C</td> <td>CATA V CATT H AATC/ I TTACT Y CAACA CAACA CAACA S CAACA K</td>	CCCAAC A N CTTTG GGCCT W P GGTTCT S S S S CTCTCC A L CTCAC F N CTCTCCAC F N CTCTCCA A L C N CCCCAT F N CCCCAT F N CCCCAT F N CCCCAT F N CCCCAT C N C CCCCCA C N C C CCCCCA C N C C C C C C C C C C C C C C C C C C	CATC D AAATCA F ATCA GCATT D CCCA A ATCA L CCCGA K CCCGCI R A ATCA L CCCGCI R A ATCA L L	AAGGA Q G CCAATI	AAATC N AAATA N AATTA I AACCA T CCCCC R TACCA C CCCCC R TACCA C CCCCC R TACCA C CCCCC C R TACCA C CCCCC C R TACCA C R C CCCCC C R TACCA T C CCCCC R C C CCCCC R C C C C C C C C	CAACO R T CAAT T I ATGAT N D AAAT R M CAACA Q A ACAAC Q A ACAAC D R CCCT S L TCAAC F R F CCCTC S S TCTCC CCTCC S S TCTCC CCTCC TCTCC CCCTCC S S	GACA T TTTC F GAAA GAAA TACC T GAAA T GAAA C GAAA C C C C C C C C C C C C C C	CCCC/ P CCATC D S CCAC/ S CCAC/ E CCACC C C CCC/ P CCACC C C C CCC/ P CCACC/ C C CCC/ P CCACC/ C C C C C C C C C C C C C C C C	AGCT S GCCA A AAAGC R T GAACG T GAACG T GAACG T GAACG V V T GAACG V V C C AAAGC V V C C AAAGCC R T GAACG C T C C C A A AAGCC R T C C A A AAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A AAAGCC R T C C A C C C C A C C C C A C C C C C C	ACGT Y V AAACG K R CTGA A E CTGA A E CTGA A E CTGA A G GTGA C TTAGG K R CTAA A N TTAG F R	CGCC A TCT. L ACTC L AACTC T TCCC A CAAA CAAA T ACAA T ACAA T AACC A CAAA T T ACAA T T ACAA T T A CAAA T T T C C A T T T C C A T T C L A ACTC A T T C L A ACTC ACTC ACTC ACTC A ACTC A	FTTT F MATT I CCAG Q TAC Y CATC I MAAT N N GCC A T T T GAA E SACG T	ACCO T G G G G G G C C C C C C C C C C C C C	GACA D CCCCC R GAGT G GGT G GGT G GGT C CCCA D D CCCCA D CCCCA D CCCCA C GGT C CCCA C G GGT C CCCCC C C C C C C C C C C C C	ACCCC T CCTT R FATA SGTA G GTAA G TTAA C T AACCC T AACCC T AACCCC A CCAG	EAGA E TATC Y AAAC R R AAAC R R AAAC R R AAAC F TATC H CATC H CATC L GTTC	R L GATGA D C GATGA G V GTTAC V I GATTI V L GATTI D L CTAGG C GTGC R A GAGGC D A GAAAA	GATT TCCCC P GAAA K TTGAT D C C C C C C C C C C C C C C C C C C	GAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAA KAAAAAA	GATCA D ATCA I ACTI T GGCA V ATTA I GGCA I GGATI D CGAC R GATI D AAAGG	A ACTO T T TCT F ATC/ I ACA T T TCC/ I ACA T T TCC/ F TCC/ F C TC/ F SACO	GCA/ Q Q TCGX S ACAC T GAGA T TC GAGA/ R TAT/ Y GCG/	A GAAATI	TCAG GCT Q GCCT A ACCC A A CTCCT A A CTCCT A A CTCCT A C A C	CATA V CATT H AATC/ I TTACT Y CAACA CAACA CAACA S CAACA K
30     301     301       361     60     4       451     60     4       120     5     6       120     5     6       120     5     6       120     5     6       1300     4     1       901     721     6       901     721     6       901     721     6       901     721     7       901     7     7       901     7     7       901     7     7       901     7     7       901     7     7       901     7     7       901     7     7       300     6     1       1351     7     390       4200     8     1       14410     1     1       450     8     1       1631     1     1	A N CTTTG GGCCT W P GTTCT S S S TCAAC F N CTCTCC A L TCAAT L N GGCCC R A GGACCT R A GTGAC S D TGAAC L N	D MAATC N TTTA F AATGC M CAATCA M CAATCA K CAATCA CAATCA K CAATCA CAATCA K CAATCA K CAATCA K CAATCA K CAATCA K CAATCA K CAATCA CAATCA K CAATCA K CAATCA CAATCAATCA CAATCAAT	Q G COGAGO P R COCOTI T V TGCTCAO S Q CACCGA Y G CACCGA Y G CACCGA X G CACCGA R T TTGAA F E AACGAA CR T TTGAA F C CAATI	N AATA N ATTA I ACCGCC R TTCGC C C C C C C C C C C C C C C C C C	R T CAAT T I AAGAT N D AAAT R M CAAGC Q A AAAGC Q A AAAGC Q A AAAGC S L CCCTO S S CCCTO S S TCCAG S S TCCCTO S S CCCTO S S S CCCTO S S S CCCTO S S S S S S CCCTO S S S S S S S S S S S S S S S S S S S	T TTTCC F TAACC N GAAA K TACCC T GAAA K GAACC F GAGA R GACCC T CTCCG S GCGTG	P CGATC D TTCG/ S GGAGA E CAAAAC E CAAAAC E CAAAGT K CAACC T C CAACC T D CGGTC	S GCCA A AAAGC K T SACG D D AAAGC K D D GAAG K V T T CACA V CAAG E T T GGAT	Y V AACCG K R CCGAA P R CTGA A E CCGG A G GTGA GTGA GTGA GTGA GTGA GTGA CTAA A N TTAG F R	A TTCT L A A CT T T C C A A C A A C A A C A A C A A C A A C A C A C T T C C A C C C A C T C C C C	F AATT I CAGG Q CTACC Y CATCC I AAATT N NGCCC A T I CGAA E GACG T	T GGTG GGTG V CCTO L CCCO A GGTT G TCA S ATO I TTG L	D CCCCC R GAGT G GGTT G GGTT G GACA D GATC D GTCC	T R R G G G G G T T A C C C C C C C C C C C C C	E TATC Y XAAC K MAAC K MAAC F CATC H CATC H CATC L STTC	R L SATGA D D SGAGT G V STTAC V I STTTI V L SATTI D L CTAGG C CTGC R A SACCC D A SAAAA	I TCCC P GAA/ K TGAAT D ACGC R G G G G G G G G C TTCC A G G C TTCC C A C G G C T C G A C C C C C C C C A C C C A C C C A C C C A C C C A C C C A C	G G G G G G G G G G G G G G G G G G G	D ATCA I I GTCA V ATTA GGAA G GATT D CCGAC R GATT D AAACC	A T T T T T T T T T T T T T T T T T T T	A Q PCCX S ACAC T GACA F TCX F GACA D A GACA R T A C A C A C A C A C A C A C A C A C A	K N GATAT D M CCTGA P E GTTCC V F CCAAC P T GATGT D V AACCGC T A ACCGCC T A ACCGCC T A ACCGAA T F	A A A A A A A A A A A A A A A A A A A	V CATT H AATC/ I TACT Y X COXA A AATCX I TCG/A/ V X CAAG/ S CAAG/ K
361 6   60 7   451 6   90 7   541 6   120 5   721 6   721 7   180 7   210 1   991 6   2240 1   1081 6   3000 1   1171 6   3300 1   1351 7   390 6   1441 6   420 1   1531 1   450 1	CTTTIG A L GGCCT W P GTTCTS S S TCAAC F N CTCAC F N CTCAC F N CTCAC F N CTCAC F N CCCAT F A GAGCCT R A GAGCCT R A GAGCCT R A GTGAC S D TGAAC L N	AATCA N TTTA F ATGG GGGT D D GGGGT A ATCA I CTGA R CCGGT R AATTA R I CCGGT L	CCAACGA P R CCGTT T V TGCTCAC S Q CACGGA S E CACGGA S E CACGGA S E CACGACGA R T TTGCAC F E CACGAC R T TTGCAC F E CACGAC R T TTGCAC F R CCCAT I CCCCAC R T TTGCAC F R CCCCAC S Q CACGAC R T T TCCCAC R T T T CCCCAC S Q C CACGAC R T T T C CCCAC S Q C C C C C C C C C C C C C C C C C C C	AATA N ATTA I AACGA T COCCC R CACTA C C GAGT E C GAGT E C GAGT E C GAGT E C GAGT C C G A T C G C G C C C C C C C C C C C C C C C	CAAT T I ATGAT N D AAAT R M AAAGC Q A AAAGC Q A AAAGC Q A ACAAG D K CCCTO S L CCCTO S S TCCAG CCCTO S S TCCCTO S S TCCTO S S TCC TCCTO S S TCC S S S S S S S S S S S S S S S S S S S	TTTC F TAAC N GAAA K TACC T GAAT N ATTC GAGA R GACC T CTCG S GGTG	CAACC D TTCG/ S GGAGA E CAAAAC E CAAAAC E CAACC C C C CAACC C T C CAACC C T C CAACC C C C	GCCA A AAAGC K GACCG D AAAAG D AAAAG D AAAAG V FACA Y GAAAG E FTGT L L GGAT	AACG K R CGAA P R CTGA A E CTGA G D TTAG G D TTAG K R CTAA A N TTAG F R	TCL L AACT L AACC T TGCO A CAAA T ACAAA T AGAA T AGAA T AGAA T AGAA S	I I I I I I I I I I I I I I I I I I I	GGTU GGTG V CTOO L GGCO A GGTU G CGTU G CTCA S ATOO I L	CCCCC R GAGI E GGTG GGTG GGTG CTCA L GACA D GATC D GTCC	R FATA Y G G TAA G TAA C T A A C C C A C C A C C A C	FATC Y K MAAC K K MAAC F CATC H CATC H CATC L STTC	CATGA D D CAGTA G V CTTAC V I CTTAC V I CTTTI D L CTACG CCTCC R A CCTCC D A CCAAAA	GAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAA R GCGG A A A A A A A A A A A A A A A A A	ATCA I ACTI T GTCA V ATTA GATI D GATI D GATI D AAGG	T T T T T T T T T T T T T T T T T T T	Q Q TCOX S ACAC T GACAC F GACAC F GACAC F GACAC P AGACAC F C ACAC	CATAT D M CCTGA P E GTTCC V F CCAAC P T GATGT D V AACCG T AA ACCGA T F AAATT	CGCCT A AGAAA E E CTGCT A CAGCC A A CTGCT A C A CTGCT A C A C A C A C A C A C A C A C A C A	I TAC I TA TAC I TAC I TAC I TAC I TAC I TAC I T
60     451     6       4451     6     6       541     6     1     6       120     5     6     1       120     5     6     1     1       150     6     7     1     6       111     1     1     1     1     1       120     1     901     7     2     1     1       180     1	A L GGCCT W P GTTCT S S TCAAC F N CTCAC F N CTCAC A L TCAAT L N CCCAT F H GAGCC R A GAGCT R A GAGCT R A GTGAC S D TGAAC L N	N TTTA F ATGG GGGT D GGGT A ATGA I CTGA I CTGA R CGGT R R ATTA I CTGT L	P R CCCTT T V TCCCCAC S Q CACCCA Y C CACCCA S E AACGAT K D CCAAT TTGAA F E AACGAC N D CAAT	N ATTA I ACGA T CCCC R TTGG L GGAT G GGAT C CATT L GAAT E CTGG V	T I ATGAT N D AAAT K M CAAGC Q A ACAAG C Q A C C C T C ACAA S L T C ACAA S L T C ACAA S L T C ACAA S L T C ACAA C T C C C C T C T C C C C C C	F TAAC N GAAA K TACC T GAAT S GACA R GACC T CTCG S GGTG	D TCGA S GAGA E AAAAC K TTGA L CGAGG T CGACT D CGGTC	A AAGC F GACG D AAAG B GTCC V FACA Y GAAG E FTGT L GGAT	K R CCGAA P R CTGA A E CCGG A G GTGA G D TTAG G D TTAG K R CTAA A N TTAG F R	L AACTO L AAACO TTGCO A CAAA K TACAAA T AAGA L D CAAT I AATCO S	I CCAG Q TAC Y CATC I AAAT N MGCC A T TGAA E GACG T	G G TO L CCO A G TCA S ATO I C TTG L	R GAGI G G G G G G G G G G G G G G G G G	R FATA G G FTAA C T A A C C C A C C A	Y K K K K K K K K K K K K K K K K K K K	D E GAGT G V GTTAC V T GTTTT V L GATTT D L TTAGG L G CGTGC R A GAGGC D A GAGAAA	GAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	R R R R R R R R R R R R R R R R R R R	I ACTI T GTCA V ATTA I GGAA G GATI D CGAC R GATI D AAGG	T F T T T T T T T T T T T T T T T T T T	Q FCGA S ACAC F GACA F GACA D AGAA R F AGAA R F AGAA R F AGAA R F GACA C F GACA C ACAC	D W CCTGA P E GTTCC V F CCAAC P T GATGT D V AACCGC T A ACCGC T A ACCGA T B AACT	A A A A A A A A A A A A A A A A A A A	H IATC/ I Y XGOO A A I I XGOO A I X XGOO A I X XGOO A I X XGOO A I X X X X X X X X X X X X X X X X X X
4451   0     90   0     90   0     901   0     150   1     150   1     150   1     150   1     150   1     150   1     1100   1     210   1     2240   1     991   0     2240   1     991   0     3300   1     11261   1     390   0     1441   0     420   1     1531   1     450   1	GGCCT GGCCT GTTCTS S S TCAAC F N CCCAT CCCAT H GAGCCT R A GAGCT R A GAGCT R A GTGAC S D N CCCAT L N CCCCAT L N CCCCAT L N C CCCAT L N C C C N C C L N C C C L N C C C L N C C C L N C C C C L N C C C L N C C C C C L N C C C C L N C C C C C L N	TTTA F ATGG M GGATI D GGCGI A ATCA I CTGA I CTGA R AAAAC K R AATTA I CTGI I L	CCCCTA T V TGCTCA S Q CCCCAC S Q CACCGAA Y G CCCCAA S E CACCGAA S E S E CACCGAA S E S E S E S E S E S E S E S E S E S E	ATTA I ACGA T CCCC R TTGG L GGAT G G G G G CTAT L G AATT E C CATTA C G C C C C C C C C C C C C C C C C C	ATGAT N D AAATG R M AAAATG R M AAAAATG R M A AAAAATG R M A AAAAA F R R CCCTO S L TCCAAG F R TCCTO S S S TGTG C L C CCTO S S S TGTG C L C CCTO S S S TGTG C L C C C C C C C C C C C C C C C C C C	TAAC N GAAA K TACC T GAAT R GACA F GAGA R GACC T CTCG S GGTG	TCGA S GAGA E AAAG K CAAG E AAAG K CAACC T CAACC T CAACC T CAACC	AAGC F ACCG T GACG D AAAG K GAAG F C C C C C C C C C C C C C C C C C C	CGAA P R CTGA A E CCGG A G GTGA G D TTAG G D TTAG K R CTAA A N TTAG F R	ACTO AACO TGCO A CAAA TACAAA TACAAA TACAAA TACAAA TACAAA TACAAA TACAAA TACAAA TACAAA TACAAA TACAAA TACAAAAA TACAAAAA TACAAAAA TACAAAAA TACAAAAA TACAAAAAAAA	CAG Q TAC Y CATC I MAAT N GCC A T T GAA E SACG T	CTO L CCO A CTT V CCT C C C C C C C C C C C C C C C	GAGI GGTI GGTI GGTI GGTI GACA D GACA D GATC D GTGC	FATA G G TTAA L ATCT T AACTC T AACTC A C A C A C A	K K K K K K K K K K K K K K K K K K K	GAGAGA G V GTTAC V I GTTTT V L GATTT D L TTAGG L G CGTGC R A GAGGC D A GAAAA	GAAAA R TGAT D ACGC R GGGT G G TGGA G CTTC AGGT G AGCC	AAAA R GCCG A CATA I GCCG G G G G G G G G G G G G G G G G	ACTI T GTCA V ATTA I GGAA G GATI D GATI D AAGG	F F ACA T T T T T T T T T T T T T	F GAGA F GAGA F GAGA R TAT/ Y GOG/	P E GTTCC V F CCAAC P T GATGT D V AACCGC T A ACCGC T A ACCGAA T B AAAATT	AGAA E E TGCT A CAGCC A CTCA TGTCA C CAGCC A A CAGCC A A CAGCC A C A CAGCC A C C C C	I TACT Y COOL A A TOCA Y COAL E TOCA S CAAC
90     90       541     6       541     20       631     1       150     6       631     1       150     6       811     1       210     1       901     6       270     6       1081     6       300     6       1171     6       300     1       3151     4       420     1       1351     4       450     4	F P GTTCT S S TCAAC F N CTCTCC A L TCAAT C N CCCAT T H GAGCC R A GAGCCT R A GTGAC S D TGAAC L N	F ATGG M GATI D GCGI A ATCA I CTGA R ATTA I CTGI L	T V TGCTCAG S Q ACCGAA S E AAGAAT K D CGAACGA R T TTGGAA F E AACGAC N D CAATT	I ACGA T CCCC R TTGG L GGAT E CTAT L GAAT E CTAT E CTAT C C C C C C C C C C C C C C C C C	N D AAATO K M AAAGC Q A ACAAO D K CGCTU S L TCAAO F K CCGTO S S TGTO S S TGTO C TGTTO V L	N GAAA TACC T GAAT N ATTC F GAGA R GACC T CTCG S GGTG	S GAGA E CAAAC K TTGA L CAAGI K CACC T C GACI D CGGTC	K ACCG T SACG D AAAG K STCC V FACA Y SAAG E TTGT L GGAT	P R CTGA A E CCGG GTGA G D TTAG CTAA A N TTAG F R	L AACC T T CAAL T ACAAL T ACAAL T ACAAL T ACAAL T ACAAL T ACAAL T ACAAL T ACAAL T ACAAL T T C AACC ACAAL T T C C AACC ACAAL T T C C AACAAL T T C C AACAAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T T C C AAL T C C AAL T C C AAL T C C C C C C C C C C C C C C C C C C	Q TAC Y CATC I MAAT N GCC A A T T GAA E GACG T	V CTO L CCO A CTT G TCA S CTTG I L	E GGTG G GGTT G GCTGA L CTGA L C GACA D GACA D GATG	Y G FTAA L ATCI A ACTC T A ACTC N SCTC A CCAG	R AATC N TTCC F TTCC H CATC H CATC L STTCC	G V GTTAC V I GTTTI V L GATTI D L TTAGG L G CGTGC R A GAGGC D A GAAAA	R GGGT G G G G G G G G G G G G G G G G G	R A A A A A A A A A A A A A A A A A A A	T GTCA V ATTA I GGAA G GATI D CGAC R GATI D AAGG	F I ACA N CAT T T T C F C T C/ L T C T C F C C T C/ C T C/ T C/ T C/ T C	S ACAC T GAC E TTO F GAC D AGA/ R TAT/ Y GCG/	P E GTTCC V F CCAAC P I GATGI D V AACCGC N F ACCGCC T A ACCCAA T F AAAATI	E E TGCT A CAGCO A CGTCA V S CGTCA V CAGCO A A AAATT CAGCI S	I Y X X X X X X X X X X X X X X X X X X
541 6   120 5   631 1   150 1   721 6   811 1   210 1   210 1   210 1   210 1   210 1   1081 6   300 1   1171 6   3300 1   1171 6   3300 1   1171 6   3300 1   1151 4   420 1   1531 1   450 1	GTTCT S S TCAAC F N CTCTCC A L TCAAT L N GAGCCAT T H GAGCCCAT T H GAGCCCAT R A GTGACC S D TGAAC L N	ATGG A GATI D GCGI A A TCA A CTGA L CCGGI R A TTA I CCGGI L	V L CTCAG S Q ACCGA Y G GCCGAA S E AAGAT K D CGAACG R T TTGAA F E ACCGAC N D CCAATT	ACGA T CCCCC R TTCGG L GGAT C GAGT E CTAT L GAAT E CTAT C C C C C C C C C C C C C C C C C	AAATC AAACC Q A AACAA D R CGCTU S L TCAAC F R CCGTO S S TGTO CG TO TGTO V L	GAAA TACC T GAAT N ATTC F GAGA R GACC T CTCG S GGTG	GAGA E AAAA K TTGA L GAGG E AAGT K C AAGT K C AAGT K C ACC C D C GGTC	T GACG D AAAG K GTCC V FACA Y GAAG E TTGT L GGAT	CTGA A E CCGG A G GTGA G D TTAG L S AGAG K R CTAA A N TTAG F R	AACC TGCC A CAAC TAC TAC TAC TAC T AGA T CAT T AGA T S	TAC Y AATC I AAAT N GCC A CATT I GAA E GACG T	CTO L CCO A CTTO G TCA S ATO I L	GGTG GGTTI G CTGA L [ GACA D AACA N GATC D GTGC	G G TTAA L ATCT I ACTC T AACCC N SCTC A CCAG	K ATC N TTCC F CATT H CATC H CTCC L STTC	GTTAC V I GTTTI V L GATTI D L TTAGG L G CGTGC R A GACGC D A GAAAA	GGGT GGGT GGGT GGGT GGGT GGGT GGGT GGG	A A A A A A A A A A A A A A A A A A A	GTCA V ATTA GGAA G GATT D CGAC R GATT D AAGG	I ACA N CAT T T T C F C T G/ L T C T C F S ACC	ACAC T GAG E TTC F GAC/ D AGA/ R TAT/ Y GCG/	GTTCC V F CCAAC P I GATGI D V AACCG T A ACCGC T A ACCGA T R AAATI	A CAGCO A CAGCO A CGTCA C CGTCA C CGTCA A C CAGCO A A CAGCO A C CAGCO C C C C C C C C C C C C C C C C C C	TACT Y A A TOTA Y CTA Y CTA Y CTA Y CTA Y CTA Y CTA Y CTA Y S CA A S CA A S CA S CA S CA S CA S
120   \$     631   1     150   []     721   (]     721   (]     721   (]     721   (]     91   (]     901   (]     270   []     991   (]     270   []     1081   (]     300   []     1171   (]     300   []     1351   4]     4390   (]     1441   (]     450   []     450   []	S S TCAAC F N CTCTCC A L TCAAT L N CCCAT T H GAGCC R A GAGCT R A GAGCT R A GTGAC S D TGAAC L N	K GGGTI A GGGGT A CTGA L CGGGT R CGGGT R A TTA L	V L CTCAC S Q ACCGA Y G GCCGAA S E AAGAT K D CCAAT	T CGCC R TTGG L GGAT C GAGT E CTAT L GAAT E GTGG V	K M AAGC Q A AACAAG D K CCCCT S L TCAAG F K CCCTC S S TCTCG L C TCTTC V L	K TACC T GAAT N ATTC F GAGA R GACC T CTCG S GGTG	E AAAC K TTG/ L CGAGC E AAGT K CACC T D CGGTC	T GACG D AAAG K GTCC V FACA Y GAAG E TTGT L GGAT	A E CCCGG GTGA G D TTAG L S AGAG K R CTAA A N TTAG F R	T TGCC A CAA K TAC T AGA D CAT I ATCC S	Y CATC I MAAT N MGCC A CATT I GAAA E GACG T	L GCO A GTT G TCA S ATO I L	G G CTGA L G G A CTGA D A A A C A D G A TGC	G TTAA L ATCI I ACTO T AACO N GCTO A CCAG	K N TCC F CATT H CATC H CATC L CTCC	V I STTTI V L SATTI D L TTAGG L G CGTGC R A SACGC D A SAAAA	D ACGC R GGGT G G G CTTC A G CTTC A G C CTTC A G C C C C C C C C C C C C C C C C C	A CATA I GGCO G G G G G G G G G G G G G C G A C G A C G G C C C G C C C C	V ATTA I GGAA G GATI D CGAC R GATI D AAGG	I N CAT T T T C T C T C T C T C T C T C T C	T GAGA E TTCA F GAC/ D AGA/ R TAT/ Y GCG/	V F CCAAC P T GATGT D V AACCO N F ACCGC T A ACCCAA T F AACCAA	A CAGCO A CTCA CTCA CAGCO A CAGCO A A A A A A T CAGCO A C A C C C C A C C C C A C C C C A C C C C A C C C C A C C C C A C C C C C A C C C C A C C C C A C C C C A C C C C C C A C C C C C A C	Y A A TOTA/ Y CTA/ V CTA/ V CTA/ V CTA/ K
631     1       150     F       721     C       780     C       811     1       210     I       991     C       240     1       991     C       300     F       1081     C       330     S       1261     1       350     C       1351     J       420     F       1531     1       450     J       1621     J	TCAAC F N CTCTCC A L TCAAT CCCAT T H GAGCC R A GAGCT R A GAGCT R A GTGAC S D TGAAC L N	GATI D GCGI A A A CTGA L CCGI R CCGI R A TA A A CCGI L	CTCAC S Q ACCGA Y G GCCAA S E AAGAT K D GCAACG R T TTGAA F E ACCAAC N D CCAATT	R R GGAT G GAGT E CTAT L GAAT E GAGT C GAGT C C C C C C C C C C C C C C C C C C C	AAGC Q A ACAA D K CGCT S L TCAA F K CGCT S S TCTO L C TGTTO V L	TACC T GAAT N ATTC F GAGA R GACC T CTCG S GGTG	AAAA K TTGA L CGAGG E AAGT K CACCC T CGACT D CGGTC	GACG D AAAG K GTCC V FACA Y GAAG E FTGT L GGAT	CCCGG A G GTGA G D TTAG L S AGAG K R CTAA A N TTAG F R	TGCA A CAAA TACA T AGA CAT I ATCA S	ATC I AAT N AGCC A A TATT I GAA E GACG T	A GTT V CGTT G TCA S ATO I CTTG L	GGTT G CTGA L [ GACA D AACA N GATC D GTGC	ITAA L ATCI I ACTO T AACTO N SCTO A CCAG	N TTCC F CATT H CATC H CATC L STTC	V L GATTI D L TTAGG L G CGTGC R A GACGC D A GAAAA	ACGC R GGGT G TGGA G CTTC A G CTTC G AGCT	I G G G G G G G G G G G G G G G G G G G	ATTA I GGAA GATI D CGAC R GATI D AAGG	T T T T T T T T T T T T T T T T T T T	GAGA F GACA D AGAA R TATA Y GCGA	CCAAC P 1 GATGI D V AACCO N F ACCGC T A ACCCAA T F AACCAA	CAGOC A GTCA CGTCA CGTCA CAGOC A A A A A A T GAGC C A C A C C A GTCA C C C A C C C C A C C C A C C C C A C C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C A C C C C A C C C C A C C C C A C	A A TOTA TOTA CAAC E TOC/ S CAAC K
150   F     721   G     721   G     811   1     210   I     210   I     991   G     2270   I     1081   G     300   I     1171   G     330   S     1261   I     330   G     11351   J     420   I     1531   I     4420   I     450   J     1621   J	F N CTCTCC A L TCAAT L N CCCAT T H GAGCCC R A GAGCCT R A GTGAC S D TGAAC L N	D GCGI A A CTGA L CCGGI R CCGGI R A TTA I CTGI L	S Q ACCGAA Y G ACCGAA S E AAGAAT K D CAACGAC R T TTGAA F E AACGAC N D CAATT	R TTGG L GGAT G GAGT E CTAT L GAAT E GTGG V	Q A ACAAA D K CGCT S L TCAAA F K CCCTO S S TCTO L C TCTTO V L	T GAAT N ATTC F GAGA R GACC T CTCG S GGTG	K TTGA L GAGG E AAG1 K AAG1 K C ACCC T GAC1 D GGGTC	D AAAG K STCC V FACA Y GAAG E FTGT L GGAT	A G GTGA G D TTAG L S AGAG K R CTAA A N TTAG F R	A CAAA K TACA T AGA D CAT I ATCC	I N AGCC A CATT I GAA E GACG T	A CGTT V CGT G TCA S ATO I S L	G CTGA L [ GACA D AACA N GATC D GTGC	L ATCI I ACTO T AACTO T AACTO A SCTO A	N F CATT H CATC H CATC L STTC	V L GATTI D L TTAGG L G CGTGC R A GACGC D A GAAAA	R GGGT G TGG/ G CTTC AGGT G AGGC	I G G G G G G G G G G G G G G G G G G G	I GGAA G GATT D CGAC R GATT D AAGG	N T T T T T T T T T T T T T T T T T T T	E F GAC/ D AGA/ R TAT/ Y GCG/	P 1 GATGI D V AACCO N R ACCGO T A ACCAA T R AACT	A GTCA CAGCCA A CAGCCA A AAATTA CAGCTA CAGCTA CAGCTA CAGCTA CAGCCA CAGCCA CAGCCA CAGCCA CAGCCA CAG	A ATCX I CTA/ V CTA/ V CAAC E TCC/ S CAAC/ K
721 (   180 //   811 1   210 [   901 (   240 1   991 (   270 1   1081 (   3300 1   1171 (   3300 1   1351 //   390 (   1441 (   420 1   1531 1   4450 (   1621 //	CTCTCC A L TCAAT L N CCCAT T H GAGCCC R A GAGCCT R A GTGAC S D TGAAC L N	CCGGI R CCGGI R CCGGI R CCGGI R CCGGI R L	ACGGAA Y G GCGAAA S E AAGGAT K D CGAACC R T TTGAA F E ACGAC N D CCAATT	TTGG L GGAT G GAGT E CTAT L GAAT E CTGG V	ACAAA D K CCCT S L TCAAA F K CCCTC S S TCTCC L C TCTTC V L	GAAT N ATTC F GAGA R GACC T CTCG S GGTG	TTGA L CGAGO E AAGI K CACCO T CGACI D CGGTO	AAAG K STCC V FACA Y GAAG E TTGT L SGAT	GTGA G D TTAG L S AGAG K R CTAA A N TTAG F R	CAAA TACA T AGA D CAT I ATCO	AAAT NAGCO A CATT I GAAA E GACG T	GTT V GGT G TCA S ATO I C TTG L	CTGA L [ GACA D AACA N GATC D GTGC	ATCT I ACTC T AACCC N SCTC A CCAG	F CATT H CATC H CTCC L STTC	D L TTAGG L G GTGC R A GACGC D A GAAAA	GGGT G TGG/ G CTTC C L C AGGT G AGCC	GGGGG GGAGA E CGAGA R CGTCA V X X X X X X	GGAA GATT D CGAC R GATT D AAGG	T T T T T T T T T T T T T T T T T T T	F GAC/ D AGA/ R TAT/ Y GCG/	GATGI D V AACCO N R ACCGC T A ACCAA T R AACTAA	CAGCCA CAGCCA A AAATTA CAGCA A AAATTA CAGCA A AAATTA CAGCA CAGCA A AAATTA CAGC	I CGTA/ V CGAAC E TOC/ S CAAC/ K
180   /     811   1     210   1     210   1     220   1     991   (     2240   1     991   (     2270   1     1081   (     3300   1     11261   1     3900   (     11351   2     3900   (     1441   (     4200   1     1531   1     450   )     1621   4	A L TCAAT L N CCCAT T H GAGCCC R A GAGCT R A GTGAC S D TGAAC L N	A I CTGA L CTGA K CGGI R CGGI R I CTGI L	Y G GCGAA S E AAGAT K D CGAACC R T TTGAA F E ACGAC N D CCAATT	L GGAT G GAGT E CTAT L GAAT E GAAT E CTGG	D K CGCT/ S L TCAA F K CCTO S S TGTO L C TGTTO V L	N ATTC F GAGA R GACC T CTCG S GGTG	L CGAGG E AAG1 K CACCO T CGAC1 D CGGTC	K GTCC V FACA Y GAAG E TTGT L GGAT	G D TTAG L S AGAG K R CTAA A N TTAG F R	K TACA T AGA D CAT I ATCO S	N AGCC A TATT I GAAA E GACG T	V G TCA S ATO I C TTG L	L [ GACA D AACA N GATO D GTGO	I ACTC T AACC N GCTC A CCAG	F ATT H CATC H TCC L STTC	D L TTAGG L G CGTGC R A GACGC D A GAAAA	G G G CTTC C C C C C C C C C C C C C C C	G E CGAG R GTC V X	G GATT D CGAC R GATT D AAGG	T F TG/ L TCT F	F GAC/ D AGA/ R TAT/ Y GCG/	D V AACCO N R ACCCC T A ACCCAA T R AACCAA	STGTT CAGOC A A AAATT CAGT CAGT	I GTA/ V CGAAC E TOG/ S CAAG/ K
8111   1     210   1     901   0     901   0     991   0     270   1     991   0     270   1     1081   0     3300   1     1351   4     390   0     1441   0     420   1     1531   1     4450   4     1621   4	TCAAT CCCAT T H GAGCCC R A GAGCTC R A GTGACC S D TGAAC L N	ATCA I CTGA L CAAAC K CCGGI R ATTA I CTGI L	GCGAA S E AAGAT K D CGAACC R T TTGAA F E ACGAC N D CAATT	GGAT G GAGT E CTAT L GAAT E CGAG V	CGCT S L TCAAC F K CGTCC S S TGTCC L C TGTCC V L	ATTC F GAGA R GACC T CTCG S GGTG	CAGG E AAG1 K CACCC T CGAC1 D CGGTC	STCC V FACA Y SAAG E FTGT L SGAT	TTAG L S AGAG K R CTAA A N TTAG F R	TACA T AGA D CAT I ATCO	AGCC A TATT I TGAA E GACG T	G TCA S ATO I L	GACA D AACA N GATC D GTGC	ACTO T AACC N GCTC A CCAG	ATT H CATC H CTCC L STTC	L G CGTGC R A GACGC D A GAAAA	G G CTTC L CAGGT G AGCC	E CGAG R CGAG R CGTC V X	GATI D CGAC R GATI D AAGG	F TG/ L TCI F	GAC/ D AGA/ R TAT/ Y GCG/	AACCG N R ACGGC T A ACCAA T R AAATT	CAGOO A A AAATT CAGOO A A CAGOO A CAGOO A CAGOO A CAGOO A CAGOO A CAGOO A CA	GAAC E TOG/ S AAG/ K
210     I       901     (2       240     1       991     (2       270     F       1081     (2       300     F       1171     (2       330     S       1261     1       360     I       1351     /       420     F       1531     1       4450     M       1621     /	L N CCCAT T H GAGCC R A GAGCT R A GTGAC S D TGAAC L N	I CTGA L CAAAC K CCGGI R ATTA I CTGI L	S E AAGAT K D GAACG R T TTGAA F E ACGAC N D CAATT	G GAGT E CTAT L GAAT E XTGG V	S L TCAAC F K CGTO S S TGTO L C TGTTO V L	F GAGA R GACC T CTCG S GGTG	E AAGT K CACCO T CGACT D CGGTC	V TACA Y GAAG E TTGT L GGAT	L S AGAG K R CTAA A N TTAG F R	T AGA D CAT I ATC	A TATT I TGAA E GACG T	G TCA S ATO I TTG L	D AACA N GATO D GTGO	T N GCTC A CCAG	H ATC H CTCC L STTC	L G GGTGC R A GACGC D A GAAAA	G CTTC L AGGI AGCC	E CGA R TGTO V X	D CGAC R GATI D AAGG	F L TCT F	D AGA/ R TAT/ Y GCG/	N R ACGGC T A ACCAA T R AAATT	AGCCA A A A A A A A A T A A T A A T A A T A A T A A T A A C A C	V E TOG/ S AAG/ K
901 ( 240 1 991 ( 270 F 1081 ( 3300 F 1171 ( 330 S 1261 1 3360 1 1351 4 390 ( 1441 ( 420 F 1531 1 1531 1 1531 2 1531 1 1531 2	CCCAT T H GAGCC R A GAGCT R A GTGAC S D TGAAC L N	CTGA L CGGI R CGGI R ATTA I CTGI L	AAGA1 K D GAACG R T TTGAA F E ACGAC N D CAATI	GAGT E CTAT L GAAT E CTGG	TCAAC F K CCGTCA S S TGTCA L C TGTTCA V L	GAGA R GACC T CTCG S GGTG	AAGT K ACCO T KGACT D KGGTO	FACA Y GAAG E FTGT L GGAT	AGAG K R CTAA A N TTAG F R	AGA D CAT I ATCO S	I I GAA E GACG T	TCA S ATO I TTG L	AACA N GATO D GTGO	N GCTC A CCAG	H TCC L	R A GACGC D A GAAAA	CTTC L AGG1 G AGCC	R R GTC V X	CGAC R GATT D AAGG	L TCT F	R R TAT/ Y GCG/	T A ACCAA T R AAAATT	AGCCA A AAATT CAACTA CGAGT . S	E TCG/ S AAG/ K
240     1       991     2       270     F       1081     2       300     F       11711     C       330     S       1261     1       350     F       1351     A       390     C       1441     C       420     F       1531     A       1621     A	T H GAGCC R A GAGCT R A GTGAC S D TGAAC L N	L AAAAC K CCGGI R ATTA I CTGI L	K D GAACG R T TTGAA F E ACGAC N D CAATT	E CTAT L GAAT E CTGG	F K CGTO S S TGTO L C TGTTO V L	R GACC T CTCG S GGTG	K ACCO T GACT D GGTC	Y GAAG E ITGT L GGAT	K R CTAA A N TTAG F R	D CAT I ATC	I GAA E GACG T	S ATO I TTG L	N GATO D GTGO	N GCTC A CCAG	H TCC L	R A GACGC D A GAAAA	L AGGT G AGCC	R GTC V CTT	R GATI D AAGG	L F GACC	R TAT/ Y GCG/	T A ACCAA T R AAATT	A AATT GAGT S	E TOG/ S AAG/ K
991   0     270   1     1081   0     300   1     1171   0     330   2     1261   1     360   1     1351   4     390   0     1441   0     420   1     1531   4     450   1     1621   4	GAGOC R A GAGCT R A GTGAC S D TGAAC L N	R CGGT R ATTA I CTGI L	GAACG R T TTGAA F E ACGAC N D CAATT	E E CAAT E CTGG	CGTO S S TGTO L C TGTTO V L	GACC T CTCG S GGTG	T GACT D GGTC	GAAG E ITGT L GGAT	CTAA A N TTAG F R	CAT I ATCO	E GACG	I I TTG L	GATO D GTGC	A	L	D A GAAAA	G	V X	GATT D AAGG	F	Y GCG/	T R	AATT I GAGT S	S AAG K
270     F       1081     G       300     F       1171     G       330     S       1261     F       360     F       1351     A       390     G       1441     G       420     F       1531     A       450     M       1621     A	R A GAGCT R A GTGAC S D TGAAC L N	R R ATTA I CTGI	R T TTGAA F E ACGAC N D CAATI	L GAAT E GTGG V	S S TGTG L C TGTT V L	T CTCG S GGTG	T GACT D GGTC	E FTGT L GGAT	A N TTAG F R	I ATCO S	E SACG T	I TTG L	D GTGC	A	L	D A	G	V CTT	D	F	¥ GCG/	T R	GAGT	S AAG/ K
1081   0     300   F     1171   0     330   S     1261   1     360   F     1351   A     390   0     420   F     1531   T     450   L     1621   A	GAGCT R A GTGAC S D TGAAC L N	R ATTA I CTGI L	TTGAA F E ACGAC N D CAATI	GAAT E GTGG	TGTG L C TGTT V L	CTCG S GGTG	GACT D GGTC	FTGT L GGAT	TTAG F R	ATC	ACG	TTG L	GTGC	CCAG	TTO	GAAAA	AGC	TT	AAGG	ACC	GCG/	AAATT	GAGT	AAG/
300   F     1171   G     330   S     1261   T     360   I     1351   A     390   G     1441   G     420   F     1531   T     450   M     1621   A	R A GTGAC S D TGAAC L N	R ATTA I CTGI L	F E ACGAC N D CAATI	E GTGG V	L C TGTTC V L	S CCTC	D	L GGAT	FR	S	т	L											. s	к
11711 ( 330 § 1261 1 360 1 1351 4 390 ( 1441 ( 420 1 1531 1 450 1 1621 4	GTGAC S D TGAAC L N	I CTGI L	ACGAC N D CAATI	CTCC V	TGTTO V L	GGTG	GGTC	GAT					v	Р	v	ER	A	L	к	D	A	K I		
330   330   31261     1261   1     360   1     1351   A     390   C     1441   C     420   F     1531   T     450   L     1621   A	S D TGAAC L N	I CTGI L	N D	v	V L				CCAC	TOG	TATA												-	0.00
1261 1 360 1 1351 4 390 0 1441 0 420 1 1531 1 450 1 1621 4	TGAAC L N	CTG1 L	CAATT			v	0					CCC	AAAA	ATTC	CAAC	CATCT	GCTO	CAG	AATT	TCT	TTC/	ACCG	TAAG	
360 1   1351 4   390 6   1441 6   420 1   1531 1   450 1   1621 4	LN	L		AACC				G	S T	R	I	Р	к	I	Q	HL	L	Q	N	F	F	NG	K	Q
1351 / 390 ( 1441 ( 420 F 1531 7 450 k 1621 /			S T		CGGA	CGAA	GCTG	STCG	CGTA	CGG	GCT	GCO	GTTC	CAAG	CCC	CCAT	TCTC	AGC	GGCG	SATA	ACTA	AGGTC	AGCA	ATTO
390 ( 1441 ( 420 F 1531 7 450 M 1621 /	AGGAT			N	P D	E	A	v	A Y	G	A	A	v	Q	A	AI	L	S	G	D	т	RS	A	I
1441 ( 420 E 1531 1 450 M 1621 4		GTAI	TACTI	GTTG	ACGTO	GGCT	COCT	TGT	CTCT	CGG	ATC	GAA	ACGO	GCCG	GTO	GTGT	CATO	ACC	AAAT	TTA/	ATCO	GAGCG	GAAT	AGCO
420 F 1531 7 450 N 1621 A	Q D	v	LL	v	D V	A	Р	L	S L	G	I	E	т	Α	G	GV	м	т	к	L	I	ER	I N	S
1531 1 450 k 1621 /	GTATT	CCAT	GTAAA	CAGT	CGAA	AATA	TTTA	ACAA	CATA	CGC	GAT	AAT	CAGO	CCCG	CTO	TCAC	CATT	CAA	GTGI	TTO	GAAG	GGTGA	ACGA	GCA
450 M	RI	Р	CK	Q	SK	I	F	т	TY	A	D	N	Q	Р	A	V I	I	Q	v	F	E	GE	R	А
1621 A	TGACC	AAAG	ACAAT	AATT	TGTTG	GGGG	AAAT	FTCG	ATTT	GAC	CGT	ATT	CCGG	CCGG	CTC	CTCG	TGGT	ATC	CCTC	CAA	ATCO	GAGGI	ATCO	TTT/
	м т	к	D N	N	LL	G	к	F	DL	T	G	Ι	Р	Р	А	PR	G	Ι	Р	Q	I	EV	5	F
480 *	ATATG	GATG	CGAAC	GGTA	TTTT	GGAC	GTTA	ACGG	CGAA	GGA	AGC	AGC	ACCO	GGTA	ATC	TGAA	GAAT	ATC	GTTA	ATC	AG	AACGA	CAAA	GGA
- r	NM	D	A N	G	IL	D	v	T	AK	D	s	s	т	G	N	VB	N	I	v	I	к	NE	K	G
1711 (	GACTG	TCAA	AGGAG	GAAA	TCGA	CAGA	ATGO	TTA	ACGA	TGC	GAA	CGA	TTCA	AAGG	ATC	GAGGA	CGAG	AAA	CAGO	CA/	AATO	CGTGI	CCAG	TCG
510 F	RL	S	KE	E	I D	R	м	v	N D	A	E	R	F	к	D	ED	E	к	Q	R	N	RV	Q	S
1801 (	GGAAT	AGTI	TGGAG	GOGT	ACAT	TTTC	CAGO	TGA	AGCA	AGC	CTG	GAT	GAAG	GCCC	GAG	GATAA	GTTC	CCG	GAGA	GTO	GAG/	AAAAG	TTO	TGTO
540 E	RN	5	LE	A	YI	F	Q	v	KQ	A	L	D	E	A	G	DR	L	Р	E	S	E	K S	S	C
1891 (	GCCTC	GAAT	GTGAG	AATA	TGCT	TAAA	TGGT	TTGG	ATTC	GAA	ACG	TTG	GCCC	GAGA	AAG	GATGA	GTAT	GAT	GATA	AGO	CTA	AAGGA	GTTA	CAG
570 E	RL	E	CE	N	ML	к	w	L	DS	N	т	L	A	E	K	DE	Y	D	D	K	L	K E	L	Q
1981 (	GAATT	TGCG	CTCCA	GTCA	TGAC	CAAA	ATTO	CATC	AAGG	TGG	AGC	GGG	GGCG	GGTT	TCCA	ACTGG	GGCT	AAC	AGTO	сто	GTO	GATTO	GGGT	CCA
600 F	RI	С	A P	v	мт	к	Ι	н	Q G	G	s	G	G	G	S	TG	A	N	s	A	R	DS	G	P
2071 0	CGATC	GACG	AAGTA	GATT	AATT	ATAT	ATTT	TTGC	ATTG	GTT	TAG	AGT	AAGT	TTT	CAAG	TTCC	TTTT	GTT	ITTI	CAT	TTC	CTTGA	ACCI	GAC
630 1				1.1.1.1.1	*																			
2161 7	гі	D	EV	D							ATT	TA												TAT
			E V			CTTT	CTTT	TCTT	ттаа				ATTA	ATTT	TTT	TTTT	AGC	GAA	GGAA	ATC	TTO	CGTAI	1111	
2341 /	TGATA	GTCC		TTGT	ATAA																			

Fig. 3. The complete nucleotide and deduced amino acid sequences of *ep-hsp70* cloned from *E. pela*. The three characteristic signature sequences of the Hsp70 family are boxed. Translation termination codon is marked with an asterisk, and the putative polyadenylation signal is underlined.

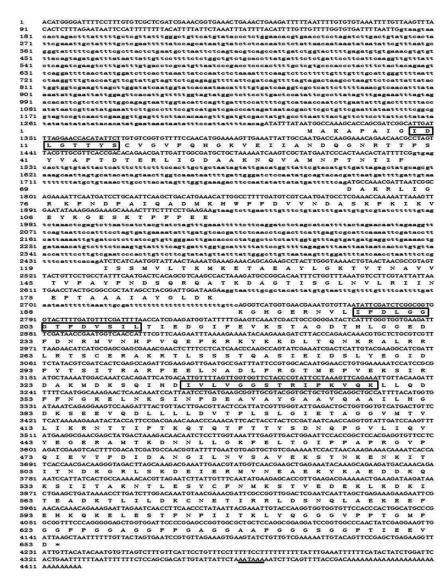


Fig. 4. The nucleotide and deduced amino acid sequences of *ep-hsc70* cloned from *E. pela*. The three characteristic signature sequences of the Hsp70 family are boxed. Translation termination codon is marked with an asterisk, and the putative polyadenylation signal is underlined. Nucleotides in lowercase letters represent the intron region.

(Fig. 7H).

8.45-fold over the control at  $25^{\circ}$ C (Fig. 6F). *hsp70* was also upregulated in males under heat stress (Fig. 6G), whereas expression levels of *hsp21.5* and *hsc70* mRNA were downregulated by the heat stress (Fig. 6E and H).

All four *ep-hsps* were significantly induced by cold stress in female adults. *hsp21.5* was upregulated by the cold treatments with a peak at  $-10^{\circ}$ C, which was a 4.48-fold increase when compared with the control at 25°C, whereas *hsp70* reached a peak at 5°C, 11.88-fold over the control (Fig. 6I and K).

**Expression of** *ep-hsps* in Female and Male *E. pela* at Different Developmental Stages. In females, the mRNA levels of *hsp21.5* at each of larval, pupal, and adult stages were higher than eggs (Fig. 7A). The expression of *hsp21.5* reached a maximum in second-stage larvae with a 23.22-fold increase when compared with eggs. The expression of *hsp21.7* was very low through all stages of the female development (Fig. 7B). *hsp70* was upregulated significantly in the late female adults (Fig. 7C). *hsc70* was upregulated significantly in the first-stage larvae (Fig. 7D).

In males, the mRNA levels of *hsp21.5* in all developmental stages were also remarkably higher than in eggs (Fig. 7E). Specifically, the mRNA level of *hsp21.5* was maximal in adults, showing a 50.21-fold

**Discussion Discussion**In insects, a number of *shsps* are reported to be involved in temperature adaptation and development. For example, six *shsps*, which are classified into three groups with diverse functions, are reported in the common cutworm, *S. litura* (Shen et al. 2011), and three *shsps* are found to be responsible for cold stress and development in the leafminer, *Liriomyza sativa* (Huang et al. 2009). In this study, two *shsps*

nound to be responsible for cold stress and development in the featminer, *Liriomyza sativa* (Huang et al. 2009). In this study, two *shsps* (*hsp21.5* and *hsp21.7*) were identified and cloned. Phylogenetic analysis showed that *hsp21.5* was classified into the orthologous cluster, and *hsp21.7* was classified into the species-specific cluster (Li et al. 2009). Orthologous *shsps*, which include *ep-hsp21.5*, have highly conserved sequences, gene structures, and functions, and they have at least one intron in their genes. Expressions of these *shsps* are usually constitutive

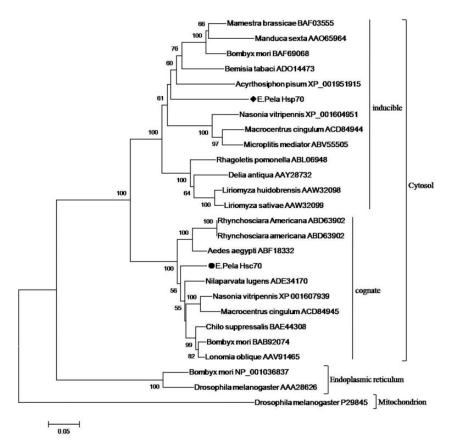
increase over the eggs. The mRNA levels of other hsps peaked in sec-

ond-stage larvae. The mRNA levels of hsp21.7 and hsp70 in second-

stage larvae were remarkably higher than in other developmental stages, increasing 19.31- and 22.08-fold, respectively, when compared

with expression in the eggs (Fig. 7F and G). The hsc70 mRNA peaked

at the second larval stage and then decreased in pupae and adult stages



**Fig. 5.** Phylogenetic analysis of Hsp70 sequences. Some sequences used in this analysis are previously reported by other groups (Huang et al. 2008, Xu et al. 2010, Aruda et al. 2011). Sequence labels are denoted by the species names and GenBank accession numbers. Black diamond indicates the sequence of ep-Hsp70, and black round indicates the sequence of ep-Hsc70.

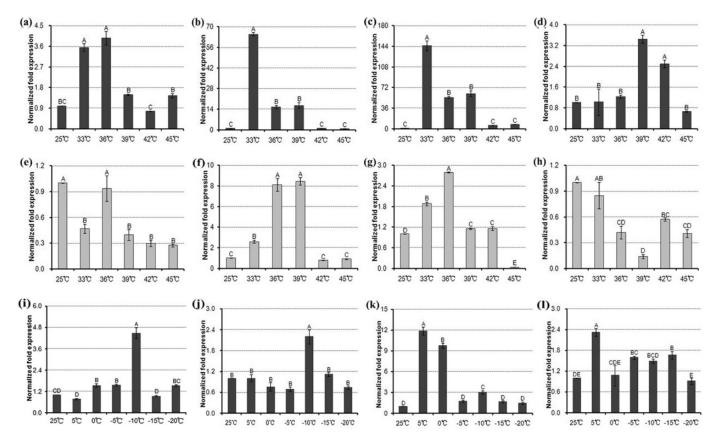
but are not inducible in response to environmental stresses. They are involved in basic metabolic processes. On the other hand, *shsps*, including *ep-hsp21.7*, in the species-specific group are clustered in specific insect order. These intron-less *shsps* are important in responses to temperature stimuli and display remarkable mRNA expression differences in mRNA expression levels during insect development. Although most insect sHsps identified so far are species specific, many orthologous sHsps have also been indentified in insects, nematodes (Lillibridge et al. 1996), amphibians (Franck et al. 2004), and mammals (Kampinga et al. 2009). However, orthologous and species-specific sHsps may be involved in various functions in different insects (Shen et al. 2011).

In *E. pela* females, *ep-hsp21.5* was upregulated in larval stages, and expression in adults was also significantly higher than expression measured in eggs. Similar upregulation was also observed during male insect development, suggesting a correlation between *ep-hsp21.5* and the developmental regulation of female and male *E. pela*. In male larvae, *ep-hsp21.5* was not sensitive to heat stimuli, whereas, in female adults, *ep-hsp21.5* was upregulated significantly by the heat- and cold-stress treatments, which may explain why *E. pela* female adults can persist in  $-30^{\circ}$ C temperatures in Changchun, China. The expression pattern of *ep-hsp21.5* in larval stage was consistent with the expression of orthologous *shsps* in most insects.

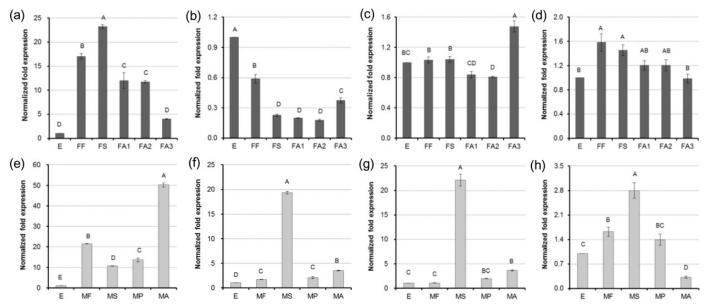
In contrast, *ep-hsp21.7* was more sensitive to the heat or cold stimulus than *hsp21.5*, indicating that this gene may play more important role in temperature-stress tolerance than *hsp21.5*. Our data were also consistent with studies in *Bombyx mori*, in which heat shock triggered a massive increase of six species-specific *shsps* but not the orthologous *hsp21.4* (Sakano et al. 2006). However, *ep-hsp21.7* was upregulated drastically at the second larval stage, when the wax secretion in males is at the maximum. Studies to investigate the ecological adaptability of *E. pela* and analyses of climatic factors that affect wax secretion have suggested that wax secretion is an ecological strategy and a protective response to severe environmental conditions (Chen et al. 2007b). Thus, *ep-hsp21.7* was presumed to be correlated with the protective response to environment stress by secreting wax.

The members of the Hsp70 family are found to be present in the cytoplasm, mitochondria, endoplasmic reticulum, and chloroplast depending on the sequences specific to the each cellular organelle. For example, cytoplasmic Hsp70 has the specific motif EEVD. Phylogenetic analysis revealed the two ep-Hsp70 of *E. pela* were clustered to the group of cytoplasm Hsp70 by their cellular localization rather than insect orders. The expression of *ep-hsp70* was inducible and the expression of *ep-hsp70* was constitutive. Usually, *hsc70* genes have introns, whereas most *hsp70* genes have no introns, which may facilitate their rapid transcription during stress because no mRNA splicing is needed. In this study, *ep-hsp70* was typically induced by temperature shock, especially in adults. It was also drastically upregulated in larval stage of male second-stage larvae. However, significantly high expression level was detected in both female adults (FA3) and male adults, indicating that this gene may be correlated with the development of adults.

Generally, when compared with *hsp70*, *hsc70* is constitutively expressed in many insects and does not significantly change under conditions of stress (Shim et al. 2006, Colinet et al. 2009, Koštál and Tollarová-Borovanská 2009). In *E. pela*, *ep-hsc70* was upregulated during the development of females and males, indicating that it may be associated with development. In addition, *ep-hsc70* did not respond to heat shock at the larval stage (male second-stage larvae), which was consistent with many reports about *hsc70*. However, *ep-hsc70* was sensitive to heat or cold stimulus in adult stages (female adults). The results of this study showed that *ep-hsc70* was responsible to the heat or



**Fig. 6.** Relative mRNA expression levels of each *hsp* under temperature stresses. *hsp21.5* (a), *hsp21.7* (b), *hsp70* (c), *hsc70* (d) in female *E. pela* under heat shock, *hsp21.5* (e), *hsp21.7* (f), *hsp70* (g), *hsc70* (h) in male *E. pela* under heat shock, *hsp21.5* (i), *hsp21.7* (j), *hsp70* (k), *hsc70* (l) in female *E. pela* under cold shock. The values are mean  $\pm$  SE of three replicates for each of the treatments; 25°C is the control. Letters above columns indicate levels of difference significance at *P* < 0.01. The same letters are not significantly different, *P* > 0.01.



**Fig. 7.** Relative mRNA expression levels of *hsps* at different developmental stages in both female and male *E. pela. hsp21.5* (a), *hsp21.7* (b), *hsp70* (c), and *hsc70* (d) in female *E. pela; hsp21.5* (e), *hsp21.7* (f), *hsp70* (g), and *hsc70* (h) in male *E. pela.* E, mixture of female and male eggs; FF, female first-stage larvae; FS, female second-stage larvae; FA1, early female adults; FA2, mid female adults; FA3, late female adults; MF, male first-stage larvae; MS, male second-stage larvae; MP, male pupae; MA, male adults. Eggs are the control. Letters above columns indicate levels of difference significance at P < 0.01. The same letters are not significantly different, P > 0.01.

cold-stress treatments in female adults, suggesting that this gene may be correlated with temperature–stress tolerance of the female insect in extreme temperatures. The special expression pattern of *hsc70* was also reported in *Pteromalus puparum* (Wang et al. 2008) and *Sarcophaga crassipalpis* (Rinehart et al. 2000).

As a chaperone molecule, *hsps* have been thought to be involved in insect developmental processes. Specifically, they play vital roles in cell differentiation, cell cycle regulation, and embryonic development (Mahroof et al. 2005, Tachibana et al. 2005, de Andrade et al. 2009, Zhang and Denlinger 2010). *E. pela* has intriguing biological characteristics, but not much is known about the expression patterns of *hsps* at extreme temperatures. This study revealed that *ep-hsp21.5*, *ep-hsp21.7*, *ep-hsp70*, and *ep-hsc70* may be correlated with heat or cold-stress tolerance in female adults, whereas *ep-hsp21.7* and *ep-hsp70* could be correlated with heat stress in male larvae. The four *ep-hsp70* could be correlated with heat stress in the developmental regulation of males, and *ep-hsp21.5*, *ep-hsp70*, and *ep-hsc70* may be critical for developmental regulation of females. However, the physiological functions of these *ep-hsps* during development of *E. pela* need further investigation.

## Acknowledgments

This work was supported by the Special Fund for Forestry Research in the Public Interest (201204602, 201304808), the National Natural Science Foundation of China (Youth Program; 31000983), the Applied Basic Research Foundation of Yunnan Province (2010ZC235), and a National Nonprofit Institute Research Grant of Research Institute of Resources Insects, Chinese Academy of Forestry (riricaf200904M-3, riricaf2011006M).

### **References Cited**

- Aruda, A. M., M. F. Baumgartner, A. M. Reitzel, and A. M. Tarrant. 2011. Heat shock protein expression during stress and diapause in the marine copepod *Calanus finmarchicus*. J. Insect Physiol. 57: 665–675.
- Beaudoing, E., S. Freier, J. R. Wyatt, J. M. Claverie, and D. Gautheret. 2000. Patterns of variant polyadenylation signal usage in human genes. Genome Res. 10: 1001–1010.
- Chen, X. M. 2011. Regularity of wax secretion, pp.65–79. In X. M. Chen (ed.), Natural population ecology of *Ericerus pela*. Science Press, Beijing, China.
- Chen, X. M., and Y. Feng. 2009. The Chinese white wax scale *Ericerus pela* Chavannes, pp. 29–30. *In* X. M. Chen and Y. Feng (eds.), An introduction to resource entomology. Science Press, Beijing, China.
- Chen, X. M., Z. Wang, Y. Chen, S. Wang, and Y. Feng. 2007a. The main climate factors affecting wax excretion of *Ericerus pela* Chavannes (Homopetera: Coccidae) and an analysis of its ecological adaptability. Acta Entomol. Sin. 50: 136–143.
- Chen, X. M., Z. Wang, Y. Chen, S. Ye, S. Wang, and Y. Feng. 2007b. The impact of environmental factors on the wax excretion by Chinese white wax scale (*Ericerus pela* Chavannes). Acta Ecol. Sin. 27: 103–112.
- Chen, X. M., Z. L. Wang, Y. Chen, J. J. Zhao, S. D. Ye, and S. Y. Wang. 2008. Age-specific life table of Chinese white wax scale (*Ericerus pela*) natural population and analysis of death key factors. Forest Res. 44: 87–94.
- Chen, X. M., Z. L. Wang, Y. Chen, J. J. Zhao, S. D. Ye, and S. Y. Wang. 2011. Design and analysis of a system for breeding white wax scale insects and producing insect wax in the same region. Forest Res. 46: 106–113.
- Colinet, H., S. F. Lee, and A. Hoffmann. 2009. Temporal expression of heat shock genes during cold stress and recovery from chill coma in adult *Drosophila melanogaster*. FEBS J. 277: 174–185.
- de Andrade, A., F. Siviero, P. Rezende-Teixeira, R. V. Santelli, and G. M. Machado-Santelli. 2009. Molecular characterization of a putative heat shock protein cognate gene in *Rhynchosciara americana*. Chromosome Res. 17: 935–945.
- Dunning, L. T., A. B. Dennis, D. C. Park, B. J. Sinclair, R. D. Newcomb, and T. R. Buckley. 2012. Identification of cold-responsive genes in a New Zealand alpine stick insect using RNA-Seq. Comp. Biochem. Physiol. Part D Genomics Proteomics 8: 24–31.
- Feder, M. E. and G. E. Hofmann. 1999. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Annu. Rev. Physiol. 61: 243–282.
- Franck, E., O. Madsen, T. van Rheede, G. Ricard, M. A. Huynen, and W. W. de Jong. 2004. Evolutionary diversity of vertebrate small heat shock proteins. J. Mol. Evol. 59: 792–805.

- Gething, M. J., and J. Sambrook. 1992. Protein folding in the cell. Nature 355: 33–45.
- Hanazono, Y., K. Takeda, M. Yohda, and K. Miki. 2012. Structural studies on the oligomeric transition of a small heat shock protein, StHsp14. 0. J. Mol. Biol. 422: 100–108.
- Huang, L. H., C. Z. Wang, and L. Kang. 2009. Cloning and expression of five heat shock protein genes in relation to cold hardening and development in the leafminer, *Liriomyza sativa*. J. Insect Physiol. 55: 279–285.
- Huang, P. Y., S. T. Kang, W. Y. Chen, T. C. Hsu, C. F. Lo, K. F. Liu, and L. L. Chen. 2008. Identification of the small heat shock protein, HSP21, of shrimp *Penaeus monodon* and the gene expression of HSP21 is inactivated after white spot syndrome virus (WSSV) infection. Fish Shellfish Immunol. 25: 250–257.
- Jiao, Y., and P. Zhao. 1999. The species and community structure of white wax insect natural enemy in different production area in the oviposition period. Acta Entomol. Sin. 21: 1148–1152.
- Kampinga, H.H., J. Hageman, M. J. Vos, H. Kubota, R. M. Tanguay, E. A. Bruford, M. E. Cheetham, B. Chen, and L. E. Hightower. 2009. Guidelines for the nomenclature of the human heat shock proteins. Cell Stress Chaperones 14: 105–111.
- Koštál, V., and M. Tollarová-Borovanská. 2009. The 70 kDa heat shock protein assists during the repair of chilling injury in the insect, *Pyrrhocoris apterus*. PLoS One 4: e4546.
- Lalouette, L., C. Williams, F. Hervant, B. J. Sinclair, and D. Renault. 2011. Metabolic rate and oxidative stress in insects exposed to low temperature thermal fluctuations. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 158: 229–234.
- Li, Z. W., X. Li, Q. Y. Yu, Z. H. Xiang, H. Kishino, and Z. Zhang. 2009. The small heat shock protein (sHSP) genes in the silkworm, *Bombyx mori*, and comparative analysis with other insect sHSP genes. BMC Evol. Biol. 9: 215.
- Lillibridge, C. D., W. Rudin, and M. T. Philipp. 1996. Dirofilaria immitis: ultrastructural localization, molecular characterization, and analysis of the expression of p27, a small heat shock protein homolog of nematodes. Exp. Parasitol. 83: 30–45.
- Livak, K. J., and T. D. Schmittgen. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. Methods 25: 402–408.
- Lopez-Martinez, G., J. B. Benoit, J. P. Rinehart, M. A. Elnitsky, R. E. Lee, and D. L. Denlinger. 2009. Dehydration, rehydration, and overhydration alter patterns of gene expression in the Antarctic midge, *Belgica antarctica*. J. Comp. Physiol. B Biochem. Syst. Environ. Physiol. 179: 481–491.
- Lü, Z. C., and F. H. Wan. 2011. Using double-stranded RNA to explore the role of heat shock protein genes in heat tolerance in *Bemisia tabaci* (Gennadius). J. Exp. Biol. 214: 764–769.
- Mahroof, R., K. Y. Zhu, L. Neven, B. Subramanyam, and J. Bai. 2005. Expression patterns of three heat shock protein 70 genes among developmental stages of the red flour beetle, *Tribolium castaneum* (Coleoptera: Tenebrionidae). Comp. Biochem. Physiol. A Mol. Integr. Physiol. 141: 247–256.
- Michaud, M. R., N. M. Teets, J. T. Peyton, B. M. Blobner, and D. L. Denlinger. 2011. Heat shock response to hypoxia and its attenuation during recovery in the flesh fly, *Sarcophaga crassipalpis*. J. Insect Physiol. 57: 203–210.
- Morrow, G., and R. M. Tanguay. 2012. Small heat shock protein expression and functions during development. Int. J. Biochem. Cell Biol. 44: 1613–1621.
- Nei, M., and S. Kumar. 2000. Molecular evolution and phylogenetics. Oxford University Press, Oxford, United Kingdom.
- Planelló, R., O. Herrero, J. L. Martínez-Guitarte, and G. Morcillo. 2011. Comparative effects of butyl benzyl phthalate (BBP) and di (2-ethylhexyl) phthalate (DEHP) on the aquatic larvae of *Chironomus riparius* based on gene expression assays related to the endocrine system, the stress response and ribosomes. Aquat. Toxicol. 105: 62–70.
- Reiskind, M. H., and A. A. Zarrabi. 2012. Is bigger really bigger? Differential responses to temperature in measures of body size of the mosquito, *Aedes albopictus*. J. Insect Physiol. 58: 911–917.
- Rinehart, J. P., G. D. Yocum, and D. L. Denlinger. 2000. Developmental upregulation of inducible hsp70 transcripts, but not the cognate form, during pupal diapause in the flesh fly, *Sarcophaga crassipalpis*. Insect Biochem. Mol. Biol. 30: 515–521.
- Sakano, D., B. Li, Q. Xia, K. Yamamoto, H. Fujii, and Y. Aso. 2006. Genes encoding small heat shock proteins of the silkworm, *Bombyx mori*. Biosci. Biotechnol. Biochem. 70: 2443–2450.
- Sanders, B. M. 1993. Stress proteins in aquatic organisms: an environmental perspective. Crit. Rev. Toxicol. 23: 49–75.

- Sang, W., W. H. Ma, L. Qiu, Z. H. Zhu, and C. L. Lei. 2012. The involvement of heat shock protein and cytochrome P450 genes in response to UV-A exposure in the beetle *Tribolium castaneum*. J. Insect Physiol. 58: 830–836.
- Shen, Y., J. Gu, L. H. Huang, S. C. Zheng, L. Liu, W. H. Xu, Q. L. Feng, and L. Kang. 2011. Cloning and expression analysis of six small heat shock protein genes in the common cutworm, *Spodoptera litura*. J. Insect Physiol. 57: 908–914.
- Shim, J. K., D. M. Ha, S. K. Nho, K. S. Song, and K. Y. Lee. 2008. Upregulation of heat shock protein genes by envenomation of ectoparasitoid *Bracon hebetor* in larval host of Indian meal moth *Plodia interpunctella*. J. Invertebr. Pathol. 97: 306–309.
- Shim, J. K., D. O. Jung, J. W. Park, D. W. Kim, D. M. Ha, and K. Y. Lee. 2006. Molecular cloning of the heat-shock cognate 70 (Hsc70) gene from the two-spotted spider mite, *Tetranychus urticae*, and its expression in response to heat shock and starvation. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 145: 288–295.
- Sørensen, J. G., T. N. Kristensen, and V. Loeschcke. 2003. The evolutionary and ecological role of heat shock proteins. Ecol. Lett. 6: 1025–1037.
- Stahl, M., M. Retzlaff, M. Nassal, and J. Beck. 2007. Chaperone activation of the hepadnaviral reverse transcriptase for template RNA binding is established by the Hsp70 and stimulated by the Hsp90 system. Nucleic Acids Res. 35: 6124–6136.
- Tachibana, S. I., H. Numata, and S. G. Goto. 2005. Gene expression of heatshock proteins (*Hsp23*, *Hsp70* and *Hsp90*) during and after larval diapause in the blow fly *Lucilia sericata*. J. Insect Physiol. 51: 641–647.
- Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731–2739.
- Tang, Q. Y., and M. G. Feng. 2010. DPS-data processing system: experimental design, statistical analysis and data mining (in Chinese), pp. 82–84. Science Press, Beijing, China.

- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876–4882.
- Wang, H., S. Dong, K. Li, C. Hu, and G. Ye. 2008. A heat shock cognate 70 gene in the endoparasitoid, *Pteromalus puparum*, and its expression in relation to thermal stress. BMB Rep. 41: 388–393.
- Wang, H. S., X. H. Wang, C. S. Zhou, L. H. Huang, S. F. Zhang, W. Guo, and L. Kang. 2007. cDNA cloning of heat shock proteins and their expression in the two phases of the migratory locust. Insect Mol. Biol. 16: 207–219.
- Xu, P., J. Xiao, L. Liu, T. Li, and D. Huang. 2010. Molecular cloning and characterization of four heat shock protein genes from *Macrocentrus cingulum* (Hymenoptera: Braconidae). Mol. Biol. Rep. 37: 2265–2272.
- Yang, P., J. Y. Zhu, Z. J. Gong, D. L. Xu, and X. M. Chen. 2012. Transcriptome analysis of the Chinese white wax scale *Ericerus pela* with focus on genes involved in wax biosynthesis. PLoS One 7: e35719.
- Zhang, C., and H. Liu. 1997. Biological characteristics and ecological environment of *Ericerus pela*, pp. 6–60. *In* C. J. Zhang, and H. Q. Liu (eds.), Chinese white wax scale and wax production technology. China Forestry Publishing House, Beijing, China.
- Zhang, Q., and D. L. Denlinger. 2010. Molecular characterization of heat shock protein 90, 70 and 70 cognate cDNAs and their expression patterns during thermal stress and pupal diapause in the corn ear worm. J. Insect Physiol. 56: 138–150.
- Zhang, Z., and Q. Zhang. 2012. Molecular cloning, characterization and expression of heat shock protein 70 gene from the oyster *Crassostrea hongkongensis* responding to thermal stress and exposure of Cu<sup>2+</sup> and malachite green. Gene 497: 172–180.

Received 21 January 2013; accepted 17 May 2013.