V. RESULTS

5.1. Isolation and characterization of the *hsp70* gene from *P. marneffei*

In a previous study by Pongpom (2004), a cDNA library was constructed in ZipLox vector from the yeast phase of *P. marneffei*. Three positive clones encoding Hsp70 (H1, H2 and H3) were obtained by immunoscreening of an expression library (20,000 pfu), using anti-*Histoplasma capsulatum* Hsp70 monoclonal antibody. An approximately 1.9 kb cDNA (H3) clone was sequenced and compared with the other *hsp70* gene. A 323 nucleotide sequence from the 5' region showed high similarity (>90%) to fungal *hsp70* gene including *Aspergillus nidulans, Trichophyton rubrum, Paracoccidioides brasiliensis* and *Histoplasma capsulatum*. Comparison with other *hsp70* genes revealed that the cDNA clone had an incomplete 5' end.

To isolate the full length *hsp70* cDNA clone, a 1.9 kb insert (clone H3), released by *Not*I-*Sal*I digestion, was used as a probe to screen the other clone containing *hsp70* gene from the cDNA library (10,000 pfu) by using ECL Direct Nucleic Acid Labeling and Detection Kit (Amersham Pharmacia Biotech)**.** Twenty-one positive clones from primary screening were screened for the longest insert size by PCR using T7 and H70-2 primers (Fig. 6). Secondary screening was performed on a putative full length *hsp70* cDNA**,** ph20, to isolate a pure phage containing *hsp70* cDNA, which was then sequened and characterized. The cloned ph20 was 2,204 nucleotides in length and contained 96 bases at the 5' untranslated region (UTR). The cDNA had a single open reading frame (ORF). An open reading frame of *hsp70* composed of 1,911 basepairs was delimited by the well-known start (ATG) and stop (TAA) codons. The ATG codon at base 93 encoded the presumed initiating methionine. This amino acid was in the appropriate position of a consensus translation start (Kozak, 1986). The polypeptide of 636 amino acids had a calculated molecular mass of 69.5 kDa and a pI of 5.03. The cDNA included 201 bp in the 3'UTR, exclusive of the poly-A tail. The stop codon TAA was located at position 2256.

In order to find out the existence of any other intron in the *P. marneffei hsp70* gene, a genomic fragment corresponding to the entire region of ORF presented in the cDNA sequence was isolated by polymerase chain reaction (PCR) using PM-F and PM-R primers and sequenced (Fig. 7 and Table 3). Comparison of the nucleotide sequences from cDNA library and genomic DNA showed compatibility of both nucleotide sequences, except for two non-homologous regions from genomic DNA between nucleotides 838-894 and 898-1,095 (Fig. 7). Analysis of these regions strongly suggested that they are introns of 57 bp and 198 bp, respectively. Their 5' and 3' ends, shown in Figure 7, conformed to the basic consensus GT/AG for eukaryotic splice donor and acceptor sites (Mount, 1982; Breathnach and Chambon, 1981). Both regions contain an internal putative splice box which matches the filamentous fungus consensus sequence (NNCTPuPy) upstream from the 3'end of introns (Gurr and Unkles, 1987). Surprisingly, a microexon between two introns contains only 3 bases.

From Table 4, it can be seen that in fungal *hsp70* genes encoding Hsp70s from different cellular compartments, different numbers of introns are present, ranging from 1 in *Blastocladiella emersonii* (L22497) and *Rhizopus nigricans* (AY147870) *hsp70* genes to up to 8 in *Emericella nidulans* (X98931) *hsp70* gene. Characterization of the intron size range for intron-containing fungal *hsp70* gene is shown in Table 5. The frequency of the size distribution of intron-containing fungal *hsp70* gene was high; between 40-79, with a dominant frequency between 50-59 nucleotides. Moreover, almost all intron-containing fungal *hsp70* genes show at least 1 intron present in the amino acid terminal region in Hsp70 conserved-GIDLGTTYSCV (residues 6-16)-sequence, adenosine triphosphate-binding domain, such as V7 (*R. nigricans*, AY147871 and AY147869), S14 (*Puccinia graminis*, U26597), D13 (*B. emersonii*, L22497), G6 (*H. capsulatum*, U46464; *P. brasiliensis*, U91560), I7 (*Pneumocystis carinii f.sp carinii*, U80967), G15 (*E. nidulans*, X98931) and L9 (*C. neoformans* var.*neoformans*, AB126638) amino acid. It should be noted that this is a moderate conserved-characteristic of intron-containing fungal *hsp70* gene.

5.2 Characteristics of the deduced amino acid sequence

Deduced amino acid sequence of a clone containing the full-length *hsp70* transcript displayed strong homology to Hsp70s of both prokaryotic and eukaryotic origins. Alignment of the predicted *P. marneffei* protein sequence (*Pm*Hsp70) with reported sequences of Hsp70 by ClustalW analysis revealed high level of homology. Figure 8 showed the comparison of the *Pm*Hsp70 with known Hsp70 sequences. Positions in which a gap was introduced to optimize homology were showed as dashes. High identity was found when comparing PmHsp70 with other fungal Hsp70s. The sequence showed homology to Hsp70 sequences from other pathogenic fungi (87% to *P. brasiliensis* and 86% to *H. capsulatum* and *T. rubrum*), with the highest score identical to Hsp70 from *A. nidulans* (88 % identity). The presence of conserved sequence motifs of the Hsp70 family was observed in the amino acid sequence (Fig. 8): (1) GIDLGTTYSCV (residues 6-16); (2) NEPTAA (residues 172-177); and (3) DLGGGTFD (residues 197-204) (Rensing and Maier, 1994); (4) a KRKYKKDLTTNARALRR motif (residues 244-260), described as a putative nuclear localization signal (NLS) of eukaryotic Hsp70 homologues (Dingwall and Laskey, 1991); (5) a RARFEE motif (residues 297-302), which serves as a signature to eukaryotic non-organellar stress-70 proteins. The EELD amino acid present at the C-terminal of the protein was the same as those seen in *A. nidulans* (AY960135) and *Ustilago maydis* (EAK84826), whereas the other organism present the EEVD motif (highly conserved among members of the Hsp70 and Hsp90 protein families). Moreover, we also found the presence of five conserved sequence motifs in amino acid sequence of all 32 fungal cytoplasmic Hsp70s as followed; (I) TVPAYFNDSQRQ (residues 143-154); (II) GGSTRIP (residues 336-442); (III) VAYGAAVQA (residues 366-374); (IV) GIETAGG (residues 400-406); (V) DNQPGVLIQV (residues 431-440). The motifs I, II and III could be found in other family members of fungal Hsp70s, except that the motifs IV and V are found only fungal cytosolic Hsp70s.

5.3. Phylogenetic analysis

Since Hsp70 homologues are among the most highly conserved proteins, they are useful tools for phylogenetic analysis. The complete deduced protein sequences of PmHsp70 and 48 other fungal Hsp70 sequences obtained from GenBank were used for the phylogenetic tree construction. Figure 9 shows the deduced phylogeny using neighbour-joining analysis of amino acid sequences in MEGA3 program (Kumar *et al.*, 2004). The resulting tree shows 4 distinct branches. Three branches include the Hsp70 family members of different intracellular compartments (cytoplasmic, endoplasmic reticulum and mitochondria) and fourth branch represents the ribosome-associated Hsp70s. The *P. marneffei* Hsp70 was clustered with the fungal cytosolic Hsp70s and it was closely related with *A. nidulans* Hsp70.

5.4. Southern blot analysis of *P. marneffei* genomic DNA

To obtain the information about the genomic organization of the gene corresponding to *hsp70*, Southern blot analysis was carried out using *P. marneffei* genomic DNA (F4 strain). Total DNA was digested with the restriction endonucleases including *Bam*HI, *Sal*I, *Hin*dIII, *Bgl*II, *Eco*RI, *Nde*I, *Xho*I, *Bss*SI and *Kpn*I . The digested DNA was fractionated on a 1% agarose gel and hybridized to the probes encoding PmHsp70, as shown in Figure 10B. High-stringency washes were used in the hybridization procedure. The uncut genomic DNA of *P. marneffei* (F4 strain) gave a strong-sharp band that indicated a complete genomic DNA. The enzymes that have no restriction site in the *Pm hsp70* gene (*Sal*I, *Bgl*II and *Hin*dIII) resulted more than one band, whereas only *Bam*HI displayed one band matched with the expected band. The other enzymes that have one, two and three restriction site(s) (*Eco*RI, *Nde*I, *Xho*I, *Bss*SI and *Kpn*I) resulted the number of positive band more than the expected for single copy from the restriction maps of *Pmhsp70* (Fig. 10A). The results indicated that *P. marneffei* genome contained multiple copies of the *hsp70* gene.

5.5. Morphological changes during phase transition *in vitro*

To investigate saprobic and parasitic phase transition *in vitro*, conidial suspension of *P. marneffei* was cultured in liquid medium (BHI) at different temperatures. At 25°C (Fig. 11), spores of *P. marneffei* rapidly grew by swelling within the first 6 h. Most germ tubes were produced between 6-12 h and then they grew apically to hyphae. Hyphae growth was fully established in 3-4-day culture. The growth of *P. marneffei* in these conditions indicated saprobic phase transition. At 37°C (Fig. 11), the cellular differentiation stage was slower than at 25°C. Spore growth occurred within the first 12 h. Germ tubes were formed by polarization between 12 to 24 h. Then, some of germ tubes rapidly produced highly branched hyphae cells. Cellular differentiation to yeast-like cells began between 24 to 48 h. At 48 h, most germ tubes and branched hyphae cells went through arthroconidiation and segmentation. Most completed yeast cells were observed at 96 h after germination. This condition revealed the parasitic phase transition of *P. marneffei*. In addition, cellular differentiation from mycelial cells to yeast-like cells was observed. The dynamics of morphological changes paralleled with parasitic phase transition. However, the occurrence of yeast-like cell differentiation was earlier than those seen in parasitic phase transition. Moreover, complete yeast cells that habored with heat shock (39°C) and severe heat shock (42°C) within 3 h did not affect the morphology of this fungus and it could grow in SDA plate at 25°C.

5.6. Northern blot analysis of *hsp70* mRNA level during saprobic and parasitic phase transition *in vitro* and under heat shock conditions

To investigate the differential expression of the *P. marneffei hsp70* gene at various stages and under heat shock condition, a Northern blot analysis was performed in which total RNA was probed with a cDNA fragment encoding the whole ORF of the Hsp70. The results showed a strong hybridization band at conidia sample, indicating an accumulation of *hsp70* transcripts in this stage (the dormant form of the organism) (Fig. 12). The faint bands were observed during the development to saprobic mycelial phase at 25°C (Fig. 12). It revealed that the *hsp70* transcript was downregulated during growth in natural conditions, with the exception of an upregulation at 48 h. A shift in the incubation temperature for the mycelium or spore from 25°C to 37°C, which required the fungus to

undergo the transition to yeast form, led to an immediate increase in the *hsp70* mRNA level (Fig. 12 and Fig. 13). However, the *hsp70* transcripts rapidly decreased paralleling with the differentiation of hyphae to yeast cells at 48 h, and until most cells were yeast cells at 72 hours.

The expression pattern of the *hsp70* gene was observed at normal homeostatic conditions (37^{o}C) and during simulated human fever (39^{o}C). When the incubation temperature of completed yeast cells (37^{o}C, 96h) upshifted from 37^{o}C to 39^{o}C (Fig. 14), the *hsp70* mRNA rapidly increased within 30 min to the maximum level, and then moderately decreased at 3 h. However, a severe heat shock condition of 42^{o}C resulted in lowering the *hsp70* transcript. It revealed the over-limitation condition to produce *hsp70* RNA of *P. marneffei*. In contrast to mycelial to yeast phase transition, decreasing the incubation temperature for the yeast cells from 37^{o}C to 25^{o}C did not affect the *hsp70* mRNA level.

5.7. Analysis of *hsp70* mRNA processing during saprobic and parasitic phase transition *in vitro*

The presence of putative introns in the *hsp70* gene led to investigate their occurrence in *hsp70* mRNA and splicing during saprobic and parasitic phase transition. Reverse transcription-PCR with primers flanking two introns revealed the presence of expected bands of spliced *hsp70* mRNA and intron(s) contained *hsp7*0 mRNA. A strong 617 bp band indicated the mature *hsp70* mRNA. The other weak higher band was likely corresponding to unspliced mRNAs. In order to confirm both intron-free and intron-containing fragments, size identity in a PCR using *hsp70*-containing plasmid (ph20) and genomic DNA as templates were performed (Fig. 15). The intron-free and intron-containing fragments were also confirmed by the DNA sequencing. Surprisingly, the DNA sequence of intron-containing fragment showed that it contained only intron II of *P. marneffei hsp70* gene (Fig. 16). A small population of intronII-unspliced *hsp70* mRNA was found during mycelial phase transition and during yeast phase transition. However, the unspliced *hsp70* mRNA level was low at yeast phase.

```
   1  cctgtacactatcttctatcctcatttccttactttcccgtgataccccctcttcccatac
  61  acacccacatacacacaacatacacattcaccATGGCCCCCGCTATCGGTATCGATTTGG
                                      M  A  P  A  I  G  I  D  L  G  10
 121  GAACCACCTACTCCTGCGTGGGTGTCTTCCGTGATGACCGTATCGAAATCATTGCCAACG
        T  T  Y  S  C  V  G  V  F  R  D  D  R  I  E  I  I  A  N  D  30
 181  ATCAGGGTAACCGAACCACCCCCTCGTTCGTTGCCTTCACCGACTCCGAGCGTCTCATTG
        Q  G  N  R  T  T  P  S  F  V  A  F  T  D  S  E  R  L  I  G  50
 241  GTGATGCTGCCAAGAATCAGGTCGCCATGAACCCTCACAACACAGTCTTCGATGCTAAGC
        D  A  A  K  N  Q  V  A  M  N  P  H  N  T  V  F  D  A  K  R  70
 301  GTTTGATCGGCCGCAAATTCTCCGATCCTGAGGTCCAGGCTGATGCCAAGCACTTCCCTT
        L  I  G  R  K  F  S  D  P  E  V  Q  A  D  A  K  H  F  P  F  90
 361  TCAAGATCATCGAGAAGGCCACCAAGCCCGTTATCGAGGTCGAGTTCAAGGGTGAGGTCA
        K  I  I  E  K  A  T  K  P  V  I  E  V  E  F  K  G  E  V  K  110
 421  AGCAGTTCACACCTGAGGAAATCTCTTCCATGGTCTTGATCAAGATGCGTGAGACTGCTG
        Q  F  T  P  E  E  I  S  S  M  V  L  I  K  M  R  E  T  A  E  130
 481  AGGCCTACCTCGGTGGTACCGTTAACAACGCTGTCATCACTGTCCCCGCCTACTTCAACG
        A  Y  L  G  G  T  V  N  N  A  V  I  T  V  P  A  Y  F  N  D  150
 541  ACTCCCAGCGTCAGGCCACCAAGGATGCTGGTCTCATTGCCGGTTTGAACGTCCTCCGTA
        S  Q  R  Q  A  T  K  D  A  G  L  I  A  G  L  N  V  L  R  I  170
 601  TCATCAACGAACCTACTGCTGCCGCCATTGCCTACGGTCTCGACAAGAAGGTTGAGGGTG
        I  N  E  P  T  A  A  A  I  A  Y  G  L  D  K  K  V  E  G  E  190
 661  AGCGCAACGTTCTCATCTTCGATCTTGGTGGTGGTACCTTCGATGTCTCTCTCCTCACCA
        R  N  V  L  I  F  D  L  G  G  G  T  F  D  V  S  L  L  T  I  210
 721  TCGAGGACGGTATCTTCGAGGTCAAGGCCACCGCCGGTGACACTCACTTGGGTGGTGAGG
        E  D  G  I  F  E  V  K  A  T  A  G  D  T  H  L  G  G  E  D  230
 781  ACTTCGACTCTCGCCTTGTCAACCACTTTGCCTCCGAGTTCAAGAGGAAATATAAGA***gt***a
        F  D  S  R  L  V  N  H  F  A  S  E  F  K  R  K  Y  K  K     249
 841  tgtgatatacccccatctgttttacaagaattgataagctaacaactttcac***ag***AGG***gt***a
                                                              D     250
 901  cgtcaaaaatccgaaatgaattctcattccctcccgtctcttgcattcttatttgatatt
 961  tggagcttggctctggagacttcctcagaaccatctgaagaatgcatatgaagactccca
1021  atatgaatccatctccttccctttcatccgcacttgactcttgaagaactcctcagatgc
1081  taacatgtgaaat***ag***ATTTGACCACCAATGCTCGTGCTCTCCGCCGTCTCCGCACTGCCT
                          L  T  T  N  A  R  A  L  R  R  L  R  T  A  C  265
1141  GTGAGCGTGCTAAGCGTACCCTCTCTTCTTCCGCCCAGACCTCCATTGAGATCGACTCTC
        E  R  A  K  R  T  L  S  S  S  A  Q  T  S  I  E  I  D  S  L  285
1201  TCTTCGAGGGTATTGACTTCTACACCTCCATCACCCGTGCTCGTTTCGAGGAGCTCTGCC
        F  E  G  I  D  F  Y  T  S  I  T  R  A  R  F  E  E  L  C  Q  305
1261  AGGATCTCTTCCGTTCCACCATGGAGCCCGTCGAGCGTGTCCTCCGTGATGCCAAGACCG
        D  L  F  R  S  T  M  E  P  V  E  R  V  L  R  D  A  K  T  D  325
1321  ACAAGTCTTCTGTCCACGAAATCGTCTTGGTCGGTGGTTCCACCCGTATCCCCAAGATCC
        K  S  S  V  H  E  I  V  L  V  G  G  S  T  R  I  P  K  I  Q  345
1381  AGAAGCTCGTCACCGACTTCTTCAACAAGGAGCCCAACAAGTCCATCAACCCCGATGAGG
        K  L  V  T  D  F  F  N  K  E  P  N  K  S  I  N  P  D  E  A  365
1441  CTGTTGCCTACGGTGCTGCCGTCCAGGCTGCTATCCTTTCTGGTGACACTTCTTCCAAGT
        V  A  Y  G  A  A  V  Q  A  A  I  L  S  G  D  T  S  S  K  S  385
1501  CCACCAACGAAATCTTGCTTCTCGACGTTGCTCCTCTCTCCGTCGGTATTGAGACTGCTG
        T  N  E  I  L  L  L  D  V  A  P  L  S  V  G  I  E  T  A  G  405
1561  GAGGTGTCATGACTCCTCTCATCAAGCGCAACACCACCATCCCCACCAAGAAGTCCGAGA
        G  V  M  T  P  L  I  K  R  N  T  T  I  P  T  K  K  S  E  T  425
1621  CCTTCTCCACCTACTCCGACAACCAGCCCGGTGTGTTGATTCAGGTCTACGAGGGTGAGC
        F  S  T  Y  S  D  N  Q  P  G  V  L  I  Q  V  Y  E  G  E  R  445
1681  AGTGCTCGTACCAAGGCAACAACTTGCTCGGCAAGTTCGAGCTCACTGGCATCCCCCCTG
        A  R  T  K  D  N  N  L  L  G  K  F  E  L  T  G  I  P  P  A  465
1741  CTCCTCGTGGTGTTCCTCAGATCGAGGTCACCTTCGACATGGACGCCAACGGTATCATGA
        P  R  G  V  P  Q  I  E  V  T  F  D  M  D  A  N  G  I  M  N  485
1801  AACGTCTCTGCCGTTGGAAGGGTACCGGTAAGAGCAACAAGATTGTCATCACCAACGACA
        V  S  A  V  E  K  G  T  G  K  S  N  K  I  V  I  T  N  D  K  505
```

```
1861  AGGGCCGTCTCTCCAAGGAGGAAATCGAGCGCATGTTGGCTGAGGCTGAGAAGTACAAGG
        G  R  L  S  K  E  E  I  E  R  M  L  A  E  A  E  K  Y  K  A 525
1921  CTGAGGATGAGGCCGAGGCTACCCGTATCCAGGCCAAGAACGGTCTTGAGTCTTATGCTT
        E  D  E  A  E  A  T  R  I  Q  A  K  N  G  L  E  S  Y  A  Y 545
1981  ACTCCCTCAAGAACACCCTCAGCGAGGGCAAGCTCACCATTGACGCTTCCGAGAAGGAGA
        S  L  K  N  T  L  S  E  G  K  L  T  I  D  A  S  E  K  E  K 565
2041  AGCTCGAGGCTGAGATCGAAAAGACCATCAGCTGGCTCGACAGCAACCAGACCGCCACCA
        L  E  A  E  I  E  K  T  I  S  W  L  D  S  N  Q  T  A  T  K 585
2101  AGGACGAGTACGAGGCTCAGCAGAAGGAGCTCGAGAGCGTTGCTAACCCCATCATCTCCG
        D  E  Y  E  A  Q  Q  K  E  L  E  S  V  A  N  P  I  I  S  A 605
2161  CAGCCTATGGCGGAGCAGGAGCTGCTCCCGGAGCTACCGGTGCTTCTGCTACCCGCGAGG
        A  Y  G  G  A  G  A  A  P  G  A  T  G  A  S  A  T  R  E  A 625
2221  CTGATGAGGTTGAGGAGCGCCCTGAGGAGCTCGACTAAacgcctaatgagtttgaagttt
        D  E  V  E  E  R  P  E  E  L  D  *                       636
2281  gatgagctttggagtttttaaaattcttttcatttgtcttatcatttccttttgttggcc
2341  aaaaattccttgctatgcgagattcggcgctaatgtcatgacggttatgttttatgagca
2401  tagatgcgcaatgcgctaatcaatcaattccttcttattctttgaaaaaaaaaaaaaaa
```

Figure 7. The nucleotide sequence and deduced amino acid sequence of the *hsp70* gene from *P. marneffei*. The exon is indicated by upper case letters, while the intron, 5' and 3' non-translated nucleotides are indicated by lower case letters. Base numbers are on the left, and amino acid numbers are on the right. The putative splice boxes in the intron sequences are marked by rectangles. Nucleotides in bold italics represent the conserved 5' and 3' consensus of the introns.

Table 4. GenBank accession numbers of fungal Hsp70 sequences used in this study and numbers of introns in fungal *hsp70* genes

Organism	*hsp70* gene sequence accession number	No. of introns	Reference
Ascomycota			
Penicillium marneffei	AY960136	2	Kummasook *et al.*, 2005* this study
Paracoccidioides brasiliensis	U91560	2	da Silva *et al.*, 1999
Ajellomyces capsulatus (*Histoplasma capsulatum*)	U46464	3	Allendoerfer *et al.*, 1996
Trichophyton rubrum	AF952391[@]	2	Rezaie *et al.*, 2000
Aspergillus awamori	Y12504	3	Hijarrubia *et al.*, 1997
Aspergillus oryzae	AB030231	3	Kasuya *et al.*, 1999
Aspergillus niger	Y08868	3	van Gemeren *et al.*, 1997
Emericella nidulans (*Aspergillus nidulans*)	X98931	8	Kramer, 1996*
Neurospora crassa	U10443 Y09011	4 5	Kapoor *et al.*, 1995 Techel *et al.*, 1998
Pneumocystis carinii f.sp carinii	U80967 L46790	2 4	Stedman *et al.*,1998 Stedman and Buck, 1996
Pneumocystis carinii f.sp ratti	U40994	3	Stedman and Buck, 1996
Zygomycota			
Rhizopus nigricans	AY147871	≥2	Cernila *et al.*, 1999
(*Rhizopus stolonifer*)	AY147870	1	Cernila *et al.*, 1999
	AY147869	3	Cernila *et al.*, 2003
Basidiomycota			
Puccinia graminis	U26597	4	Staples, 1995*
Cryptococcus neoformans var.*neoformans*	AB126638	5	Kakeya *et al.*, 2003*
Chytridiomycota			
Blastocladiella emersonii	L22497	1	Stefani and Gomes, 1995

@ Acession number of cDNA

* Indicates direct submission to GenBank. These entries are not included in "References"

Table 5. Frequency of intron lengths from fungal intron-containing *hsp70* genes

Length (base) of introns	**Frequency of *hsp70* introns**
40 – 49	5
50 - 59	29
60 - 69	10
70 - 79	5
80 - 89	1
90 - 99	-
100 - 109	3
110 - 119	-
120 - 129	-
130 - 139	-
140 - 149	-
150 - 159	-
160 - 169	3
170 - 179	1
180 - 189	1
190 - 199	1

```
Pb-AAB53051      ---MA--PAIGIDLGTTYSCVGIFRDDRIEIIANEQGNRNTPSFVAFTDTERLIGDAAKN 55
Pb-AAK66771      ---MA--PAIGIDLGTTYSCVGIFRDDRIEIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Ac-AAC05418      ---MA--PAVGIDLGTTYSCVGIFRDDRIEIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Tr-AAD08909      ---MA--PAVGIDLGTTYSCVGIFRDDRIEIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Tv-AAQ83701      ---MA--PAVGIDLGTTYSCVGIFRDDRIEIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Dt-CAA57452      ---MA--PAIGIDLGTTYSCVGIYRDDRIEIIANDQGNRTTPSFVAFTDTERLIGDSAKN 55
Gz-EAA70431      ---MG--PAVGIDLGTTYSCVGIFREDRCDIIANDQGNRTTPSFVGFTDTERLIGDAAKN 55
Hj-AAP40020      ---MG--PAVGIDLGTTYSCVGIFREDRCDIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Mg-EAA55301      ---MA--PAVGIDLGTTYSCVGIFRDDRIEIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Nc-AAA82183      ---MA--PAVGIDLGTTYSCVGVFREDRCEIIANDQGNRTTPSFVAFTDTERLVGDAAKN 55
Pm-AY960135      ---MA--PAIGIDLGTTYSCVGVFRDDRIEIIANDQGNRTTPSFVAFTDSERLIGDAAKN 55
An-EAA62310      ---MA--PAVGIDLGTTYSCVGVFRDDRIDIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Pc-AAD00455      -MTFE--KAIGIDLGTTFSCVGVWQNDRVEIIANDQGNRTTPSYVAFSDTERLIGDAAKN 57
Cn-AAW42238      ---MT--KAIGIDLGTTYSCVAVWQNDRVEIIANDQGNRTTPSYVAFNDSERLIGDAAKN 55
Cn-EAL21768      ---MT--KAIGIDLGTTYSCVAVWQNDRVEIIANDQGNRTTPSYVAFNDSERLIGDAAKN 55
Cn-AAW42202      ---MV--KAVGIDLGTTYSCVAVWQNDRVEIIANDQGNRTTPSYVAFNDSERLIGDAAKN 55
Cc-CAA72797      ---MS--KAIGIDLGTTYSCVAVWQQDRVEIIANDQGNRTTPSYVAFTDSERLIGDAAKN 55
Um-EAK84826      ---MT--KAIGIDLGTTYSCVAVWQNDRVEVIANDQGNRTTPSYVAFTDSERLIGDAAKN 55
Rn-AAF13877      ---MSEGKSVGIDLGTTYSCVGVWQNDRVEIIANDQGNRTTPSYVGFTDSERLIGDAAKN 57
Rn-AAN52148      ---MSEGKSVGIDLGTTYSCVGVWQNDRVEIIANDQGNRTTPSYVGFTDSERLIGDAAKN 57
Rn-AAN52149      ---MS--KAVGIDLGTTYSCVGVWQNDRVEIIANDQGNRTTPSYVAFTDTERLIGDAAKN 55
Pg-AAB93665      ---MT--KAIGIDLGTTYSCVGVWQNDRVEIIANDQGNRTTPSYVAFTDTERLIGDSAKN 55
Sp-CAA20787      MS-----KSIGIDLGTTYSCVGHFSNNRVEIIANDQGNRTTPSYVAFTDTERLIGDAAKN 55
Sp-CAA93590      MS-----KSIGIDLGTTYSCVGHFSNNRVEIIANDQGNRTTPSYVAFTDTERLIGDAAKN 55
Ca-CAA82929      ---MS--KAVGIDLGTTYSCVAHFANDRVEIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Ca-EAK94611      ---MS--KAVGIDLGTTYSCVAHFANDRVEIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Sc-AAA63574      ---MS--KAVGIDLGTTYSCVAHFANDRVEIIANDQGNRTTPSYVAFTDTERLIGDAAKN 55
Sc-AAC37398      ---MS--RAVGIDLGTTYSCVAHFSNDRVEIIANDQGNRTTPSYVAFTDTERLIGDAAKN 55
Pa-AAB63968      ---MS--KAVGIDGGTTYSCVAHFANDRVEIIANDQGNRTTPSFVAFTDTERLSGDAAKN 55
Pa-CAA82570      ---MS--KAVGIDLGTTYSCVAHFVNDRVEIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Sc-CAA31393      ---MS--KAVGIDLGTTYSCVAHFANDRVDIIANDQGNRTTPSFVAFTDTERLIGDAAKN 55
Sc-CAA31394      ---MS--KAVGIDLGTTYSCVAHFSNDRVDIIANDQGNRTTPSFVGFTDTERLIGDAAKN 55
                        ::*** ***:***. : ::* ::***:****.***:*.*.*:*** **:***

Pb-AAB53051      QVAMNP-QTVFEAKRLIGRKFADPEVQADMKHFPFKITDKGGNPKIQVEFR-GEGKHFTP 113
Pb-AAK66771      QVAMNPSNTVFDAKRLIGRKFADPEVQSDMKHFPFKVIDKAGKPVISVEFK-GEEKQFTP 114
Ac-AAC05418      QVAMNPANTVFDAKRLIGRKFADPEVQADMKHFPFKITDKGGKPVIQVEFK-GETKEFTP 114
Tr-AAD08909      QVAMNPINTVFDAKRLIGRKFNDAEVQADMKHFPFKVVEKSGKPIVQVEFK-GEEKQFTP 114
Tv-AAQ83701      QVAMNPINTVFDAKRLIGRKFNDAEVQADMKHFPFKVVEKSGKPIVQVEFK-GEEKQFTP 114
Dt-CAA57452      QVAINPHNTVFDAKRLIGRKFQDAEVQADMKHFPFKVIEKAGKPVTQVEFK-GETKDFTP 114
Gz-EAA70431      QVAMNPQNTVFDAKRLIGRKFADAEVQADMKHFPFTIIDKAGKPAIEVEFK-GEKKTFTP 114
Hj-AAP40020      QVAMNPQNTVFDAKRLIGRKFADPEVQADMKHFPFKVIDKNSKPVIEVEFK-GEKKTFTP 114
Mg-EAA55301      QVAMNPHNTVFDAKRLIGRKFSDAEVQADMKHFPFKVVERSGKPVIEVEFK-GESKQFTP 114
Nc-AAA82183      QVAMNPANTVFDAKRLIGRKFSDPEVQADMKHFPFKVIDRGGKPVIQVEFK-GETKVFTP 114
Pm-AY960135      QVAMNPHNTVFDAKRLIGRKFSDPEVQADAKHFPFKIIEKATKPVIEVEFK-GEVKQFTP 114
An-EAA62310      QVAMNPHNTVFDAKRLIGRRFGDAEVQADMKHWPFKVVDKSGKPIIEVEFK-GETKQFTP 114
Pc-AAD00455      QVAMNPSNTVFDAKRLIGRKFGEPEVQADMKHWPFKVVNKDGKPYIEVEFK-GESKTFTP 116
Cn-AAW42238      QVAMNPYNTVFDAKRLIGRKFNDAEVQADMKHWPFKVVEKGGKPVIKVEYK-GEEKTFTP 114
Cn-EAL21768      QVAMNPYNTVFDAKRLIGRKFNDAEVQADMKHWPFKVVEKGGKPVIKVEYK-GEEKTFTP 114
Cn-AAW42202      QVAMNPYNTVFDAKRLIGRKFEDAEVQADMKHWPFKVIDRAGKPAIQVEYR-GEEKVFTP 114
Cc-CAA72797      QVAMNPYNTVFDAKRLIGRKFEDAEVQSDMKHWPFKVISRLAS-LHQVEYARGEE--FSP 112
Um-EAK84826      QVAMNPHNTVFDAKRLIGRKFDDAEVQSDMKHWPFEVTSVGGKPQIRIEYK-GEKKTFTP 114
Rn-AAF13877      QVAMNPYNTVFDAKRLIGRRFNDAEVQADMKHWSFKVVEKESKPIIQVEYK-GETKTFTP 116
Rn-AAN52148      QVAMNPYNTVFDAKRLIGRRFNDAEVQADMKHWSFKVVEKESKPIIQVEYK-GETKTFTP 116
Rn-AAN52149      QVAMNPHNTVFDAKRMIGRRYNEQEIQEDMKHWPFKIVEKDGKPNIQVDHK-GETKIFTP 114
Pg-AAB93665      QVAMNPHNTVFDAKRLIGRKFADAEVQSDIKHFPFTVFDKGGKPYIRVQYR-GEDKEFSP 114
Sp-CAA20787      QVAMNPHNTIFDAKRLIGRKFDDPEVQSDMKHWPFKVISKDGKPVLQVEYK-GETKTFTP 114
Sp-CAA93590      QVAMNPHNTIFDAKRLIGRRFNDPEVQSDMKHWPFKVIEKDGKPLIQVEFK-GETKTFTP 114
Ca-CAA82929      QAAMNPANTVFDAKRLIGRKFDDPEVINDAKHFPFKVIDKAGKPVIQVEYK-GETKTFSP 114
Ca-EAK94611      QAAMNPANTVFDAKRLIGRKFDDHEVQGDIKHFPFKVVDKASKPMIQVEYK-GETKTFSP 114
Sc-AAA63574      QAAMNPHNTVFDAKRLIGRKFDDPEVTNDAKHYPFKVIDKGGKPVVQVEYK-GETKTFTP 114
Sc-AAC37398      QAAINPHNTVFDAKRLIGRKFDDPEVTTDAKHFPFKVISRDGKPVVQVEYK-GETKTFTP 114
Pa-AAB63968      QAAMNPANTVFDAKRLIGRKFDDPEVQGDIKHFPFKVVDKSGKPQIQVEFK-GETKVFTP 114
Pa-CAA82570      QAAMNPANTVFDAKRLIGRKFDDPEVQNDIKHFPFKVVEKGGKPHIQVEFK-GETKVFTP 114
Sc-CAA31393      QAAMNPSNTVFDAKRLIGRNFNDPEVQADMKHFPFKLIDVDGKPQIQVEFK-GETKNFTP 114
Sc-CAA31394      QAAMNPANTVFDAKRLIGRNFNDPEVQGDMKHFPFKLIDVDGKPQIQVEFK-GETKNFTP 114
                 *.*:** :*:*:***:***.: : *:  * **:.* : .   .    ::.  **   *:*
```

```
Pb-AAB53051      DNRLVNHFVNEFKRKHK----------------------KDLSSN-ARALRRLRTACERA 267
Pb-AAK66771      DNRLVNHFVNEFKRKHK----------------------KDLSSN-ARALRRLRTACERA 268
Ac-AAC05418      DNRLVNHFVSEFKRKFK----------------------KISPAERARALRRSPTACERA 269
Tr-AAD08909      DNRLVNHFVNEFKRKNK----------------------KDLSTN-ARALRRLRTACERA 268
Tv-AAQ83701      DNRLVNHFVNEFKRKNK----------------------KDLSTN-ARALRRLRTACERA 268
Dt-CAA57452      DNRLVNHFSNEFKRKHK----------------------KDLSDN-ARALRRLRTACERA 268
Gz-EAA70431      DNRLVNHFVNEFKRKHK----------------------KDLSTN-VRALRRLRTACERA 268
Hj-AAP40020      DNRLVNHFVNEFKRKHKKVSSSTHVAPHSLDARHMLTPDRDLSTN-ARALRRLRTACERA 290
Mg-EAA55301      DNRLVTHFANEFKRKHK----------------------KDLTTN-ARALRRLRTACERA 268
Nc-AAA82183      DNRLVNHFVQEFKRKDK----------------------KDLSTN-ARALRRLRTACERA 268
Pm-AY960135      DSRLVNHFASEFKRKYK----------------------KDLTTN-ARALRRLRTACERA 268
An-EAA62310      DNRLVNHFVTEFKRKHK----------------------KDLSTN-ARALRRLRTACERA 268
Pc-AAD00455      DNRLVQHFVQEFKRKHK----------------------KDISGN-PRALRRLRTACERA 270
Cn-AAW42238      DNRLVNHFVQEFKRKNK----------------------KDLSSN-ARALRRLRTACERA 268
Cn-EAL21768      DNRLVNHFVQEFKRKNK----------------------KDLSSN-ARALRRLRTACERA 268
Cn-AAW42202      DNRLVNHFVQEFKRKNK----------------------KDLSSN-ARALRRLRTACERA 268
Cc-CAA72797      DNRLVNHFVQEFKRKNK----------------------KDITSN-ARALRRLRTACERA 266
Um-EAK84826      DNRLVNHFVQEFKRKNK----------------------KDLTTN-ARALRRLRTACERA 268
Rn-AAF13877      DNRLVAHFMQEFKRKFK----------------------KDITGN-ARAIRRLRTACERA 270
Rn-AAN52148      DNRLVSHFMQEFKRKFK----------------------KDITGN-ARAIRRLRTACERA 270
Rn-AAN52149      DNRLVDHFIQEFKRKFK----------------------KDITGN-ARAVRRLRTACERA 268
Pg-AAB93665      DNRLVNHFVQEFKRKHK----------------------KDLSSN-PRALRRLRTACERA 268
Sp-CAA20787      DSRLVNHFIQEFKRKNK----------------------KDITGN-ARAVRRLRTACERA 268
Sp-CAA93590      DSRLVNHFAQEFKRKNK----------------------KDITGN-ARAVRRLRTACERA 268
Ca-CAA82929      DNRLVNFFIQEFKRKNK----------------------KDISTN-QRALRRLRTACERA 269
Ca-EAK94611      DNRLVNFFIQEFKRKNK----------------------KDISTN-QRALRRLRTACERA 268
Sc-AAA63574      DSRLVNFLAEEFKRKNK----------------------KDLTTN-QRSLRRLRTAAERA 268
Sc-AAC37398      DNRLVNHLATEFKRKTK----------------------KDISNN-QRSLRRLRTAAERA 268
Pa-AAB63968      DNRLVNHFINEFKRKNK----------------------KDICGN-QRALRRLRTACERA 268
Pa-CAA82570      DNRLVNHFANEFKRKYK----------------------KDLTTN-QRALRRLRTACERA 270
Sc-CAA31393      DNRLVNHFIQEFKRKNK----------------------KDLSTN-QRALRRLRTACERA 267
Sc-CAA31394      DNRLVNHFIQEFKRKNK----------------------KDLSTN-QRALRRLRTACERA 267
                 *.*** .:  ***** *                      :   :  *::**  **.***

Pb-AAB53051      KRTLSSAAQTSIEIDSLYEGIDFYTSITRARFEELCQDLFRSTMDPVERVLRDAKIDKSS 327
Pb-AAK66771      KRTLSSAAQTSIEIDSLYEGIDFYTSITRARFEELCQDLFRSTMDPVERVLRDAKIDKSS 328
Ac-AAC05418      KRTLSSAAQTSIEIDSLYEGIDFYTSITRARFEELCQDLFRSTMEPVERVLRDAKIDKSS 329
Tr-AAD08909      KRTLSSAAQTSIEIDSLYEGVDFYTSITRARFEELCQDLFRSTMEPVERVLRDAKIDKSS 328
Tv-AAQ83701      KRTLSSAAQTSIEIDSLYEGVDFYTSITRARFEELCQDLFRSTMEPVERVLRDAKIDKSS 328
Dt-CAA57452      KRTLSSSAQTSIEIDSLFEGIDFFTSNTRARFEEVGQDLFRGNMEPGERTLRDDKIDKSS 328
Gz-EAA70431      KRTLSSSAQTSIEIDSLFEGIDFYTSITRARFEELCQDLFRSTIQPVDRVLTDAKIDKSL 328
Hj-AAP40020      KRTLSSSAQTSIEIDSLYEGIDYYTSITRARFEELCQDLFRSTIQPVDRVLADAKIDKSQ 350
Mg-EAA55301      KRTLSSSAQTSIEIDSLYEGIDFYTSITRARFEELCQDLFRSTLQPVDRVLTDAKIDKAQ 328
Nc-AAA82183      KRTLSSSAQTSIEIDSLFEGIDFYTSITRARFEELCQDLFRSTLQPVDRVLTDAKIDKSQ 328
Pm-AY960135      KRTLSSSAQTSIEIDSLFEGIDFYTSITRARFEELCQDLFRSTMEPVERVLRDAKTDKSS 328
An-EAA62310      KRTLSSAAQTSIEIDSLFEGIDFYTSITRARFEELCQDLFRGTMEPVERVLRDAKIDKSS 328
Pc-AAD00455      KRSLSSSTQTSIEIDSLYEGIDLYTSITRARFEELCQDLFRGTMEPVEKVLRDAKIDKSS 330
Cn-AAW42238      KRTLSSAAQTSIEIDSLFDGIDFYTSITRARFEELCQDLFRSTMDPVEKVLRDSKIDKSS 328
Cn-EAL21768      KRTLSSAAQTTIEIDSLFDGIDFYTSITRARFEELCQDLFRSTMDPVEKVLRDSKIDKSS 328
Cn-AAW42202      KRTLSSAAQTSIEIDSLFDGIDFYTSITRARFEELCQDLFRSTMDPVEKVLRDSKIDKSS 328
Cc-CAA72797      KRTLSSAAQTSIEIDSLYDGIDFYTSITRARFEELCQDLFRSTMEPVEKVLRDSKIDKSS 326
Um-EAK84826      KRTLSSAAQTTIEIDSLFEGIDFYTSITRARFEELCGDLFSHTIEPVEKVLRDSKIDKGS 328
Rn-AAF13877      KRTLSSAAQTTIEIDSLFEGVDFYTSLTRARFEELNQDLFRNTMEPVEKVLRDSKIDKSQ 330
Rn-AAN52148      KRTLSSAAQTTIEIDSLFEGVDFYTSLTRARFEELNQDLFRNTMEPVEKVLRDSKIDKSQ 330
Rn-AAN52149      KRTLSSAAQTSIEIDSLFEGVDFYTSLTRARFEELNQDLFRNTMEPVEKVLRDSKLDKGA 328
Pg-AAB93665      KRTLSSAAQTTIEIDSLFEGVDFYTSLTRARFEELCQDLFRSTLEPVEKVLRDAKIDKAA 328
Sp-CAA20787      KRTLSSSAQASIEIDSLFEGIDFYTSITRARFEELCADLFRKTMEPVERVLRDSKVDKAS 328
Sp-CAA93590      KRTLSSSAQASIEIDSLYEGIDFYTSITRARFEELCADLFRNTMEPVEKVLRDSKIDKSS 328
Ca-CAA82929      KRTLSSSAQTSIEIDSLYEGIDFYTSITRARFEELCADLFRSTLDPVGKVLADAKIDKSQ 329
Ca-EAK94611      KRTLSSSAQTSIEIDSLYEGIDFYTSITRARFEELCADLFRSTLEPVDKVLSDAKIDKSK 328
Sc-AAA63574      KRTLSSSAQTSIEIDSLFEGIDFYTSITRARFEELCADLFRSTLEPVEKVLADSKLDKSQ 328
Sc-AAC37398      KRALSSSSQTSIEIDSLFEGMDFYTSLTRARFEELCADLFRSTLEPVEKVLKDSKLDKSQ 328
Pa-AAB63968      KTTLSSSAQTSLEIDSLYEGIDFYTSITRARFEELCQDLFRSTLDPVEKVLKDAKLDKSQ 328
Pa-CAA82570      KRTLSSSAQTSVEIDSLYEGIDFYTSITRARFEELCQDLFRSTLDPVEKVMRDGKLDKSQ 330
Sc-CAA31393      KRTLSSSAQTSVEIDSLFEGIDFYTSITRARFEELCADLFRSTLDPVEKVLRDAKLDKSQ 327
Sc-CAA31394      KRTLSSSAQTSVEIDSLFEGIDFYTSITRARFEELCADLFRSTLDPVEKVLRDAKLDKSQ 327
                 * :***::*:::*****::*:* :** *******:  ***  .::*  :.: * * **.
```

```
                                II                                           III
Pb-AAB53051      VHEIVLVGGSTRIPRIQKLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILSGDSTSKSTN 387
Pb-AAK66771      VHEIVLVGGSTRIPRIQKLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILSGDTTSKSTN 388
Ac-AAC05418      VHEIVLVGGSTRIPRIQKLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILSGDTSSKSTN 389
Tr-AAD08909      VHEIVLVGGSTRIPKIQKMVSDFFNGKEPNKSINPDEGVAYGAAVQAAILSGDTSSKSTN 388
Tv-AAQ83701      VHEIVLVGGSTRIPKIQKMVSDFFNGKEPNKSINPDEAVAYGAAVQAAILSGDTSSKSTN 388
Dt-CAA57452      VHEIVLGGGSTRIPKVQKLVSDFFNGKEPCKSINPDEAVAYGAAVQAAILSGDTSSKSTK 388
Gz-EAA70431      VHEIVLVGGSTRIPKVQKLITDYFNGKEPNKSINPDEAVAYGAAVQAAILSGDTSSKATN 388
Hj-AAP40020      VHEIVLVGGSTRIPRIQKPITDYFNGKEPNKSINPDEAVAYGAAVQAAILSGDTTSKSTN 410
Mg-EAA55301      VHEIVLVGGSTRIPRIQKLITDYFNGKEPNKSINPDEAVAYGAAVQAAILSGDTSSKSTN 388
Nc-AAA82183      VHEIVLVGGSTRIPRIQKLISDYFDGKEPNKRINPDEAVAYGAAVQAAILSGDTSSKSTS 388
Pm-AY960135      VHEIVLVGGSTRIPKIQKLVTDFFN-KEPNKSINPDEAVAYGAAVQAAILSGDTSSKSTN 387
An-EAA62310      VHEIVLVGGSTRIPKIQRLVSDYFN-KEANKSINPDEAVAYGAAVQAAILSGDTSSKSTN 387
Pc-AAD00455      VHEIVLVGGSTRIPRVQKLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILSGDTSEK-TQ 389
Cn-AAW42238      VNEIVLVGGSTRIPKIQKLVSDMFSGREPNRSINPDEAVAYGAAVQAAILTGDTSEA-TQ 387
Cn-EAL21768      VNEIVLVGGSTRIPKIQKLVSDMFSGREPNRSINPDEAVAYGAAVQAAILTGDTSEA-TQ 387
Cn-AAW42202      VNEIVLVGGSTRIPKIQKLVSDMFSGREPNRSINPDEAVAYGAAVQAAILTGDTSEA-TQ 387
Cc-CAA72797      VNEIVLVGGSTRIPKIQKLVSDMFSGREPNRSINPDEAVAYGAAVQASILNGETSEA-TQ 385
Um-EAK84826      VHEIVLVGGSTRIPKVQKLLTDFFNGRELNKSINPDEAVAYGAAVQAAILSGDTSEK-TQ 387
Rn-AAF13877      VHDIVLVGGSTRIPKVQKLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILSGDTSEK-TQ 389
Rn-AAN52148      VHDIVLVGGSTRIPKVQKLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILSGDTSEK-TQ 389
Rn-AAN52149      VHDIVLVGGSTRIPKVQKLVSDYFNGKEPNKSINPDEAVAYGAAVQAAILTGNTSEK-TD 387
Pg-AAB93665      VHEIVLVGGSTRIPRIVKLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILTGDTSEK-TQ 387
Sp-CAA20787      VNEIVLVGGSTRIPRVQKLVSDFFNGKEPCKSINPDEAVAYGAAVQAAVLTGDTSEK-TQ 387
Sp-CAA93590      VNEIVLVGGSTRIPRIQKLVSDFFNGKEPCKSINPDEAVAYGAAVQAAILVGDTSEK-TQ 387
Ca-CAA82929      VEEIVLVGGSTRIPKIQKLVSDFFNGKELNKSINPDEAVAYGAAVQAAILTGDTSSK-TQ 388
Ca-EAK94611      VDEIVLVGGSTRIPKVQKLVSDYFNGKEPNRSINPDEAVAYGAAVQAAILSGDTSSK-TQ 387
Sc-AAA63574      IDEIVLVGGSTRIPKVQKLVSDFFNGKEPNRSINPDEAVAYGAAVQAAILTGDQSST-TQ 387
Sc-AAC37398      IDEIVLVGGSTRIPKIQKLVSDFFNGKEPNRSINPDEAVAYGAAVQAAILTGDQSTK-TQ 387
Pa-AAB63968      VNEIVLVGGSTRIPKVQKLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILTGDTSSK-TQ 387
Pa-CAA82570      VAEIVLVGGSTRIPKIQKLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILTGDTSSK-TQ 389
Sc-CAA31393      VDEIVLVGGSTRIPKVQKLVTDYFNGKEPNRSINPDEAVAYGAAVQAAILTGDESSK-TQ 386
Sc-CAA31394      VDEIVLVGGSTRIPKVQKLVTDYFNGKEPNRSINPDEAVAYGAAVQAAILTGDESSK-TQ 386
                 : :*** *******:: : ::* *. :*  : *****.*********::* *: :   *.

                                       IV                                V
Pb-AAB53051      EILLLDVAPLSVGIETAGGVMTPLIKRNTTIPTKKSETFSTFADNQPGVLIQVFEGERAR 447
Pb-AAK66771      EILLLDVAPLSVGIETAGGVMTPLIKRNTTIPTKKSETFSTFADNQPGVLIQVFEGERAR 448
Ac-AAC05418      EILLLDVAPLSLGIETAGGVMTPLIKRNTTIPTKKSETFSTFSDNQPGVLIQVFEGERAR 449
Tr-AAD08909      EILLLDVAPLSLGIETAGGVMTPLIKRNTTIPTKKSETFSTFSDNQPGVLIQVFEGERAR 448
Tv-AAQ83701      EILLLDVAPLSLGIETAGGVMTPLIKRNTTIPTKKSETFSTFSDNQPGVLIQVFEGERAR 448
Dt-CAA57452      EILLLDVAPLSLGIETAGGVMTALIKRNTTIPTKKSETFSTFSDNQPGVLIQVFEGERAR 448
Gz-EAA70431      EILLLDVAPLSLGIETAGGMMTKLIPRNTTIPTKKSEVFSTFSDNQPGVLIQVYEGERQR 448
Hj-AAP40020      EILLLDVAPLSLGIETAGGMMTKLIPRNTTIPTKKSEVFSTFSDNQPGVLIQVYEGERQR 470
Mg-EAA55301      EILLLDVAPLSLGIETAGGMMTKLIPRNTTIPTKKSEVFSTFSDNQPGVLIQVYEGERQR 448
Nc-AAA82183      EILLLDVAPLSLGIETAGGMMTKLIPRNTTIPTKKSEVFSTFSDNQPGVLIQVYEGERQR 448
Pm-AY960135      EILLLDVAPLSVGIETAGGVMTPLIKRNTTIPTKKSETFSTYSDNQPGVLIQVYEGERAR 447
An-EAA62310      EILLLDVAPLSVGIETAGGVMTPLVKRNTTIPTKKSETFSTYSDNQPGVLIQVYEGERAR 447
Pc-AAD00455      DLLLVDVAPLSMGIETAGGVMTPLIKRNTTVPTRKSEVFSTYSDNQPGVLIQVFEGERPR 449
Cn-AAW42238      DLLLLDVAPLSMGIETAGGIMTPLIKRNTTVPTKKSEVFSTYSDNQPGVLIQVFEGERAK 447
Cn-EAL21768      DLLLLDVAPLSMGIETAGGIMTPLIKRNTTVPTKKSEVFSTYSDNQPGVLIQVFEGERAK 447
Cn-AAW42202      DLLLLDVAPLSMGIETAGGIMTPLIKRNTTVPTKKSEVFSTYSDNQPGVLIQVFEGERAK 447
Cc-CAA72797      DLLLLDVAPLSLGIETAGGIMTPLIKRNTTVPTKKSEIFSTYADNQPGVLIQVYEGERAK 445
Um-EAK84826      DLLLLDVAPLSMGIETAGGVFTPLIKRNTTVPTKKSEIFSTYADNQPGVLIQVFEGERAR 447
Rn-AAF13877      DLLLLDVAPLSLGIETAGGVMTALIKRNTTVPTKKSEVFSTYADNQPGVLIQVFEGERAR 449
Rn-AAN52148      DLLLLDVAPLSLGIETAGGVMTALIKRNTTVPTKKSEVFSTYADNQPGVLIQVFEGERAR 449
Rn-AAN52149      SLLLLDVAPLSLGIETAGGVMTPLIKRNTTVPTKKSEVFSTYADNQPGVLIQVYEGERAR 447
Pg-AAB93665      DLLLLDVSPLSLGIETAGGVMTPLIKRNTTVPTKKSEIFSTYSDNQPGVLIQVYEGERAR 447
Sp-CAA20787      DLLLLDVAPLSMGIETAGGVMTPLIKRNTTIPTKKSEVFSTYSDNQPGVLIQVFEGERAR 447
Sp-CAA93590      DLLLLDVAPLSLGIETAGGVMTPLIKRNTTIPTKKSEVFSTYADNQPGVLIQVFEGERAR 447
Ca-CAA82929      DILLLDVAPLSLGIETAGGIMTKLIPRNSTIPTKKSETFSTYADNQPGVLIQVFEGERAK 448
Ca-EAK94611      DLLLLDVAPLSLGIETAGGIMTKLIPRNSTIPTKKSETFSTYADNQPGVLIQVFEGERAQ 447
Sc-AAA63574      DLLLLDVAPLSLGIETAGGIMTKLIPRNSTIPTKKSEVFSTYADNQPGVLIQVFEGERTR 447
Sc-AAC37398      DLLLLDVAPLSLGIETAGGIMTKLIPRNSTIPTKKSETFSTYADNQPGVLIQVFEGERTR 447
Pa-AAB63968      DLLLLDVAPLSLGIETAGGVMTKLIPRNTTIPTKKSEIFSTYSDNQPGVLIQVYEGERAR 447
Pa-CAA82570      DLLLLDVAPLSLGIETAGGVMTKLIPRNTTIPTKKSEIFSTYSDNQPGVLIQVYEGERAK 449
Sc-CAA31393      DLLLLDVAPLSLGIETAGGVMTKLIPRNSTIPTKKSEIFSTYADNQPGVLIQVFEGERAK 446
Sc-CAA31394      DLLLLDVAPLSLGIETAGGVMTKLIPRNSTIPTKKSEVFSTYADNQPGVLIQVFEGERAK 446
                 .:**:**:***:*******::* *: **:*:**:*** ***::**********:**** :
```

```
Pb-AAB53051      TKENNLLGKFELTGIPPAP-GVPQIEVTFDVDANGIMNVSALEKGIGKTNKIVITNDKGR 506
Pb-AAK66771      TKDNNLLGKFELTGIPPAPRGVPQIEVTFDVDANGIMNVSALEKGTGKTNKIVITNDKGR 508
Ac-AAC05418      TKDNNLLGKFELTGIPRAR-GVPQIEVTFDVDANGIMNVSALEKGTRKTNKIVITNDKGR 508
Tr-AAD08909      TKDNNLLGKFELTGIPPAPRGVPQIEVTFDVDANGIMNVSALEKGTGKTNKIVITNDKGR 508
Tv-AAQ83701      TKDNNLLGKFELTGIPPAPRGVPQIEVTFDVDANGIMNVSALEKGTGKTNKIVITNDKGR 508
Dt-CAA57452      TKDINLMGKFELSGIRPAPRGVPQIEVTFDLDANGIMNVSALEKGTGKTNKIVITNDKGR 508
Gz-EAA70431      TKDNNLMGKFELTGIPPAPRGVPQIEVTFDLDANGIMNVSAVEKGTGKSNKIVITNDKGR 508
Hj-AAP40020      TKDNNLLGKFELTGIPPAPRGVPQIEVTFDVDANGIMNVSAVEKGTGKSNKIVITNDKGR 530
Mg-EAA55301      TKDNNLLGKFELTGIPPAPRGVPQIEVTFDLDANGIMNVSAVEKGTGKTNQITITNDKGR 508
Nc-AAA82183      TKDNNLLGKFELTGIPPAPRGVPQIEVTFDVDANGIMNVSALEKGTGKTNQITITNDKGR 508
Pm-AY960135      TKDNNLLGKFELTGIPPAPRGVPQIEVTFDMDANGIMNVSAVEKGTGKSNKIVITNDKGR 507
An-EAA62310      TKDNNLLGKFELTGIPPAPRGVPQIEVTFDLDANGIMNVSAVEKGTGKTNKITITNDKGR 507
Pc-AAD00455      TRDCHLLGCFDLTGIPPAPRGVPQIEVTFDVDANSILNVSAVEKGTGKTNKIEIKNEKGR 509
Cn-AAW42238      TKDCNLLGKFELSGIPPAPRGVPQIEVSFDIDANGILNVNAADKSTGKTSKITITNDKGR 507
Cn-EAL21768      TKDCNLLGKFELSGIPPAPRGVPQIEVSFDIDANGILNVNAADKSTGKTSKITITNDKGR 507
Cn-AAW42202      TKDCNLLGKFDLSGIPPAPRGVPQIEVSFDVDANGILNVNAADKSTGKSSKITITNDKGR 507
Cc-CAA72797      TKDCNLLGKFELSGIPPAPRGVPQIEVSFDVDANGILNVNASDKTTGKSNRITITNDKGR 505
Um-EAK84826      TKDNNLLGKFELSGIPPAPRGVPQIEVTFDVDANAILNVSAAEKGTGKSEKITIRNDKGR 507
Rn-AAF13877      TKDNNILGKFELSGIPPAPRGVPQVEVTFDVDANGILNVSALDKTTGKSNKITITNDKGR 509
Rn-AAN52148      TKDNNILGKFELSGIPPAPRGVPQIEVTFDVDANGILNVSALDKTTGKSNKITITNDKGR 509
Rn-AAN52149      TKDNNLLGKFELSGIPPAPRGIPQVEVTFDVDANGILNVNALDKTTGKSNKITITNDKGR 507
Pg-AAB93665      TKDNNLLGKFELAGIPPAPRGVPQIEVTFDIDANGILNVSASDKTTGKSNRITITNDKGR 507
Sp-CAA20787      TKDCNLLGKFELSGIPPAPRGVPQIEVTFDVDANGILNVSALEKGTGKTQKITITNDKGR 507
Sp-CAA93590      TKDCNLLGKFELSGIPPAPRGVPQIEVTFDVDANGILNVSALEKGTGKTQKITITNDKGR 507
Ca-CAA82929      TKDNNLLGKFELSGIPPAPRGVPQIEVTFDIDANGILNVSALEKGTGKTQKITITNDKGR 508
Ca-EAK94611      TKDNNLLGKFELSGIPPAPRGVPQIEVTFDIDANGILNVSALEKGTGKTQKITITNDKGR 507
Sc-AAA63574      TKDNNLLGKFELSGIPPAPRGVPQIEVTFDIDANGILNVSAVEKGTGKSNKITITNDKGR 507
Sc-AAC37398      TKDNNLLGKFELSGIPPAPRGVPQIDVTFDIDANGILNVSALEKGTGKSNKITITNDKGR 507
Pa-AAB63968      TKDNNLLGKFELSGIPPAPRGVPQIEVTFDIDANGILNVSAVEKGTGKSEKITITNDKGR 507
Pa-CAA82570      TKDNNLLGKFELSGIPPAPRGVPQIEVTFDIDANGILNVSAVEKGTGKSQKITITNDKGR 509
Sc-CAA31393      TKDNNLLGKFELSGIPPAPRGVPQIEVTFDVDSNGILNVSAVEKGTGKSNKITITNDKGR 506
Sc-CAA31394      TKDNNLLGKFELSGIPPAPRGVPQIEVTFDVDSNGILNVSAVEKGTGKSNKITITNDKGR 506
                 *:: :::* *:*:**  *  *:**::*:**:*:*.*:**.* :*   *:.:* * *:***

Pb-AAB53051      LSKEEIERMLAEAEKYKAEDEAEASRISAKNGLESYAYSLRNTISDSKVDEKLDASDKEK 566
Pb-AAK66771      LSKEEIERMLAEAEKYKAEDEAEASRISAKNGLESYAYSLRNTISDSKVDEKLDASDKEK 568
Ac-AAC05418      LSKEEIERMLAEAEKYKAEDEAEASRIRPKNGLESYAYSLRNSLRHSKVDEKLEAGDKEK 568
Tr-AAD08909      LSKEEIERMLAEAEKYKAEDEAETARIGAKNGLESYAYSLKNTLSDSKVDEKLDAADKEK 568
Tv-AAQ83701      LSKEEIERMLAEAEKYKAEDEAETARIGAKNGLESYAYSLKNTLSDSKVDEKLDAADKEK 568
Dt-CAA57452      LSKEEIERMLADAEKYKEEDEAEAGRIQAKNGLESYAYSLKNTVSDPKVEEKLSAEDKET 568
Gz-EAA70431      LSKEEIERMLNDAEKYKEEDEAEGKRVAAKNGLESYAYSLRNTLSDPKVEEKIEASDKET 568
Hj-AAP40020      LSKEEIERMLSDAEKYKEEDEAEAKRVSAKNGLESYAYSLRNTLNDPKVDEKIEAADKET 590
Mg-EAA55301      LSKEDIERMLAEAEKFKEEDEAEAARVSAKNGLESYAYSLRNTLQDSKVDEKLDAADKEK 568
Nc-AAA82183      LSKEEIERMLAEAEKFKEEDEAEAKRVAAKNGLESYAYSLRNTLSDSKVDEKLDAADKEK 568
Pm-AY960135      LSKEEIERMLAEAEKYKAEDEAEATRIQAKNGLESYAYSLKNTLSEG--KLTIDASEKEK 565
An-EAA62310      LSKEDIERMLADAEKYKAEDEAEAARIQAKNGLESYAYSLKNTISEG--QLQISEDDKKK 565
Pc-AAD00455      LSKEEIERMIQEAEKYKAEDEEEAQRVSAKNALESYAYSLKNTVSDSKVSDKISETEKSK 569
Cn-AAW42238      LSKEEIERMLAEAERFKAEDEAAAATVQAKNGLESYAYSLKTTLSDN--KDKFDSADHET 565
Cn-EAL21768      LSKEEIERMLAEAERFKAEDEAAAATVQAKNGLESYAYSLKTTLSDN--KDKFDSADHET 565
Cn-AAW42202      LSKEEIERMLAEAEKFKAEDEAAAATVQAKNGLESYSYSLKTTLSDN--QDKFDAADHET 565
Cc-CAA72797      LSKEEIERMLAEAEKYAAEDKEVAERVQAKNALESYTYSLKSTLSEN--ADKFSAEDKET 563
Um-EAK84826      LSSEQIEEMLKQAEQFAEEDKQALERTQAKNGLESYIYNVRNTTNEPQLKDKLEAADKEA 567
Rn-AAF13877      LSKEDIERMVNDAEKYKAEDEAAASRIQAKNGLESYAYNLRNTLQDERSVGALPEEDKTK 569
Rn-AAN52148      LSKEDIERMVNDAEKYKAEDEEAASRIQAKNGLESYAYNLRNTLQDEKVGGALPEEDKTK 569
Rn-AAN52149      LSKEDIERMVNDAEKFKAEDEAAAARIQARNGLESYAYNLRNSLQDEKLATKIEAGDKEK 567
Pg-AAB93665      LSKEEIERMVNEAEKYKAEDEEATARITSRNALESYAYNLRNSLTDEKLADKFDAADKKK 567
Sp-CAA20787      LSKEEIDRMVAEAEKYKAEDEAESGRIQAKNHLESYAYSLRNSLDDPNLKDKVDASDKET 567
Sp-CAA93590      LSKEEIDRMVSEAEKYKAEDEAETSRIQAKNHLESYAYSLRNSLDDPNLKDKVDASDKEA 567
Ca-CAA82929      LSKEEIDKMVSEAEKFKEEDEKEAARVQAKNQLESYAYSLKNTINDGEMKDKIGADDKEK 568
Ca-EAK94611      LSKEEIEKMVSEAEKFKEEDEKEASRVQAKNQLESYAYSLKNTLGEEQFKSKLDASEIEE 567
Sc-AAA63574      LSKEDIDKMVAEAEKFKAEDEQEAQRVQAKNQLESYAFTLKNSVSENNFKEKVGEEDARK 567
Sc-AAC37398      LSKDDIDRMVSEAEKYRADDEREAERVQAKNQLESYAFTLKNTINEASFKEKVGEDDAKR 567
Pa-AAB63968      LSQEEIERMVSEAEKYKEEDEKEAKRIAAKNGLESYAYSLKQTTSEKQFEEKVEASKREA 567
Pa-CAA82570      LSKEEIDRMVAEAEKYKEEDEKEAARIAAKNGLESYAYSLKQTASEKQFEEKVDASKRES 569
Sc-CAA31393      LSKEDIEKMVAEAEKFKEEDEKESQRIASKNQLESIAYSLKNTISEAG--DKLEQADKDT 564
Sc-CAA31394      LSKEDIEKMVAEAEKFKEEDEKESQRIASKNQLESIAYSLKNTISEAG--DKLEQADKDA 564
                 **.::*:.*: :**::  :*:       .:* ***  :.:: :  .      .   .
```

```
Pb-AAB53051      LKTEIDKTVSWLDENQTATKEEFEAQQKELESVANPIMMKFYGAGG---APGAGFPGAG- 622
Pb-AAK66771      LKTEIDKTVSWLDENQTATKEEFEAQQKELESVANPIMMKFYGAGGEGGAPGAGFPGAG- 627
Ac-AAC05418      LKSEIDKTVQWLDENQTATKEEYESQQKELEAVANPIMMKFYAGGE--GAPGG-FPGAG- 624
Tr-AAD08909      LKSEIDKVVAWLDDNQTATKEEYESQQKELEGVANPIMMKFYGAGGEGGAPGG-FPGAGA 627
Tv-AAQ83701      LKSEIDKVVAWLDDNQTATKEEYESQQKELEGVANPIMMKFYGAGGEGGAPGG-FPGAGA 627
Dt-CAA57452      LTGAIDKTVAWIDENQTATKEEYEAEQKQLESVANPVMMKIYGAEG--GAPGG-MPGQG- 624
Gz-EAA70431      LTAEIDKVVQWLDDNQQATREEYEEHQKELEGKANPIMMKFYGAGG--EGAPGGMPG-MP 625
Hj-AAP40020      LKSEIDKIVQWLDDNQQASTEEYESHQKELEGVANPIMMKFYGAGG--E---GGMPGGMP 645
Mg-EAA55301      LNAEINKIVSWLDESQQATKEEYEEHQKELEAVANPIMMKFYGAGG--AP--GGMPGAPG 624
Nc-AAA82183      LKSEIDKIVAWLDENQQATREEYEERQKELEAIANPIMMKFYGAGG--AP--GGMP---G 621
Pm-AY960135      LEAEIEKTISWLDSNQTATKDEYEAQQKELESVANPIISAAYGGAG-------AAPG--- 615
An-EAA62310      VSDKIDEVISWLDNNQTAEKDEYESQQKELEGVANPIISAAYAAAG-------GAPGGAA 618
Pc-AAD00455      LEKAISDVTSWLETNQTATKEEYTSKQKDLETIAGPIMMKLYQSGE-------GVPGMG- 621
Cn-AAW42238      LSKKVEEVINALDTMQSASKEEFESLQKELEGIANPIMTKFYGASG-------GMPGGA- 617
Cn-EAL21768      LSKKVEEVINALDTMQSASKEEFESLQKELEGIANPIMTKFYGASG-------GMPGGA- 617
Cn-AAW42202      LSKKVDEVISSLDTMQSASKEEFESLQKELEAVANPIMTKFYGAQG-------GAPG--- 615
Cc-CAA72797      LQKAVDDCIKFLDTADSASKDEIESHQKELEALSGPIMQRFYCSTG-------GAPGGAP 616
Um-EAK84826      LEKIVKEGIEWLDSNTTASTDELKDKQKEIEEQVNPIMTKIYSAAG-------GAPGGMP 620
Rn-AAF13877      LNAAVDESIKWLDESQEASKEEYESKQKELEEIANPIMMKFYQQAG-------GAPGG-- 620
Rn-AAN52148      LNAAVDESIKWLDESQEASKEEYESKQKELEEVANPIMMKFYQQAG-------GAPGG-- 620
Rn-AAN52149      LESAVKEAIDWMDNSQEASKEEYESRQKELEEVANPIMMKLYQGEG-------GMPGGG- 619
Pg-AAB93665      LEDAVNSTISWLDNSQEASKEEYEEHQKELEAVANPIMQKLYAGAG-------GAPGGAP 620
Sp-CAA20787      VDKAVKETIEWLDSNTTAAKDEFEAKQKELESVANPIMAKIYQAGG--------APGGMP 619
Sp-CAA93590      IDKAVKETIEWLDHNTTAAKDEYEDKQKELEGVANPIMAKIYQAGG--------APGGAP 619
Ca-CAA82929      LTKAIDETISWLDASQAASTEEYEDKRKELESVANPIISGAYGAAGGAPGGAGGFPGAG- 627
Ca-EAK94611      VTKAADETIAWLDSNQTATQEEFADQQKELESKANPIMTKAYQAGATPSGAAGAAP--G- 624
Sc-AAA63574      LEAAAQDAINWLDASQAASTEEYKERQKELEGVANPIMSKFYGAAGGA-PGAG------- 619
Sc-AAC37398      LETASQETIDWLDASQAASTDEYKDRQKELEGIANPIMTKFYGAGAGAGPGAGESG---- 623
Pa-AAB63968      FTKACDDTIAWLDENQTATAEEYDDKRKELEQAGNEVLKDLYAEGGVPG---G-APG--- 620
Pa-CAA82570      LNKAIEETISWLDNNQSATTDEYEDKRKELEGIANDALKDLYAAGGVPG---GAAPG--- 623
Sc-CAA31393      VTKKAEETISWLDSNTTASKEEFDDKLKELQDIANPIMSKLYQAGGAPGGAAGGAPG--- 621
Sc-CAA31394      VTKKAEETIAWLDSNTTATKEEFDDQLKELQEVANPIMSKLYQAGGAP---EGAAPG--- 618
                 .    ..    ::    *  :*    *:::   .  :   *

Pb-AAB53051      ------GPGGFPG-----AGVGGAHSGGDDGPTVEEVD---------------------- 649
Pb-AAK66771      ------GPGGFPG-----AGAGGAHSGGDDGPTVEEVD---------------------- 654
Ac-AAC05418      ------GPGGFPGG----PGAGHASGGGDDGPTVEEVDLKFPMLPLPWQLSVRKMHRPFF 674
Tr-AAD08909      -----GGPGGFPG-----AGAGGAAAH-DDGPTVEEVD---------------------- 654
Tv-AAQ83701      -----GGPGGFPG-----AGAGGAAAH-DDGPTVEEVD---------------------- 654
Dt-CAA57452      --------AGAPP-----PGAG------DDGPTVEEVD---------------------- 643
Gz-EAA70431      -----GGPGGFPG-----AGGPAPGAGGDDGPTVEEVD---------------------- 653
Hj-AAP40020      -----GGAGGFPG-----AGG-APHQGGDDGPTVEEVD---------------------- 672
Mg-EAA55301      ----GAPGGFPG-----AGG-APGAGGDDGPTVEEVD---------------------- 651
Nc-AAA82183      -----AAPGGFP-------GG-APGSNDNEGPTVEEVD---------------------- 646
Pm-AY960135      ----------ATG-------ASATREADEVEERP<b>EELD</b>---------------------- 636
An-EAA62310      -----PGAGAAPG-------GGAGFRNDGVVEEN<b>EELD</b>---------------------- 644
Pc-AAD00455      ------GACSQPG--------GFPGTTDDHGPTIEEVD---------------------- 645
Cn-AAW42238      -----GAPGGFPGG------GAGGGAAQEEGPSVEEVD---------------------- 644
Cn-EAL21768      -----GAPGGFPGG------GAGGGAAQEEGPSVEEVD---------------------- 644
Cn-AAW42202      -----GAPGGFPG--------AGGAPAQEEGPSVEEVD---------------------- 640
Cc-CAA72797      ----GGAPGGFPG-------AGGPGASHEDGPSVEEVD---------------------- 643
Um-EAK84826      ----GGAPGAAPG---------GAAPGGDDGPTV<b>EELD</b>---------------------- 645
Rn-AAF13877      --APGAAPGGFPGA----G-AAPGSTDETDGPSIEEVD---------------------- 651
Rn-AAN52148      --APGAAPGGFPGA----G-AAPGSTDETDGPSIEEVD---------------------- 651
Rn-AAN52149      GMPGGGAPGGFPGG----------DTGGEPTVEEVD---------------------- 645
Pg-AAB93665      GGFPGGAPGGFPGG----------APAGEDGPSVEEVD---------------------- 648
Sp-CAA20787      G----AAPGAAPG------AAPGAAPGGDNGPEVEEVD---------------------- 647
Sp-CAA93590      G----GMPGGAPG------GAPG---GADNGPEVEEVD---------------------- 644
Ca-CAA82929      ---------GFPGGAPGAGGPGGATGGESSGPTVEEVD---------------------- 656
Ca-EAK94611      ---------GFPG--------GAAPEPSNDGPTVEEVD---------------------- 645
Sc-AAA63574      ---------PVPG----AGAGPTG-APDNGPTVEEVD---------------------- 642
Sc-AAC37398      ---------GFPGS---MPNSGATGGGEDTGPTVEEVD---------------------- 649
Pa-AAB63968      ---------GFPGA------GGAPSTEETQGPTVEEVD---------------------- 643
Pa-CAA82570      ---------GFPGA------GGAAPGAD-QGPSVEEVD---------------------- 645
Sc-CAA31393      ---------GFP--------GGAPPAPEAEGPTVEEVD---------------------- 642
Sc-CAA31394      ---------GFP--------GGAPPAPEAEGPTVEEVD---------------------- 639
                                     .                **:*
```

```
Pb-AAB53051      ------------------------------
Pb-AAK66771      ------------------------------
Ac-AAC05418      LFLLFLIFLIFLILFLFYFFLPVRFNESCFS 705
Tr-AAD08909      ------------------------------
Tv-AAQ83701      ------------------------------
Dt-CAA57452      ------------------------------
Gz-EAA70431      ------------------------------
Hj-AAP40020      ------------------------------
Mg-EAA55301      ------------------------------
Nc-AAA82183      ------------------------------
Pm-AY960135      ------------------------------
An-EAA62310      ------------------------------
Pc-AAD00455      ------------------------------
Cn-AAW42238      ------------------------------
Cn-EAL21768      ------------------------------
Cn-AAW42202      ------------------------------
Cc-CAA72797      ------------------------------
Um-EAK84826      ------------------------------
Rn-AAF13877      ------------------------------
Rn-AAN52148      ------------------------------
Rn-AAN52149      ------------------------------
Pg-AAB93665      ------------------------------
Sp-CAA20787      ------------------------------
Sp-CAA93590      ------------------------------
Ca-CAA82929      ------------------------------
Ca-EAK94611      ------------------------------
Sc-AAA63574      ------------------------------
Sc-AAC37398      ------------------------------
Pa-AAB63968      ------------------------------
Pa-CAA82570      ------------------------------
Sc-CAA31393      ------------------------------
Sc-CAA31394      ------------------------------
```

Figure 8. Alignment of the amino acid sequences. Amino acid sequences were obtained from *Penicillium marneffei* (Pm), *Saccharomyces cerivisiae* (Sc), *Paracoccidioides brasiliensis* (Pb), *Ajellomyces capsulatus* (Ac), *Trichophyton rubrum* (Tr), *Trichophyton verrucosum* (Tv), *Davidiella tassiana* (Dt), *Gibberella zeae* (Gz), *Hypocrea jecorina* (Hj), *Magnaporthe grisea* (Mg), *Neurospora crassa* (Nc), *Aspergillus nidulans* (An), *Pneumocystis carinii* (Pc), *Cryptococcus neoformans* (Cn), *Cryptococcus curvatus* (Cc), *Ustilago maydis* (Um), *Rhizopus nigricans* (Rn), *Puccinia graminis* (Pg), *Schizosaccharomyces pombe* (Sp), *Candida albicans* (Ca) and *Pichia angusta* (Pa). The alignment was performed using ClustalW program. The vital conserved sequence motifs are underlined. The putative conserved sequence motifs for fungal cytoplasmic Hsp70s are boxed.

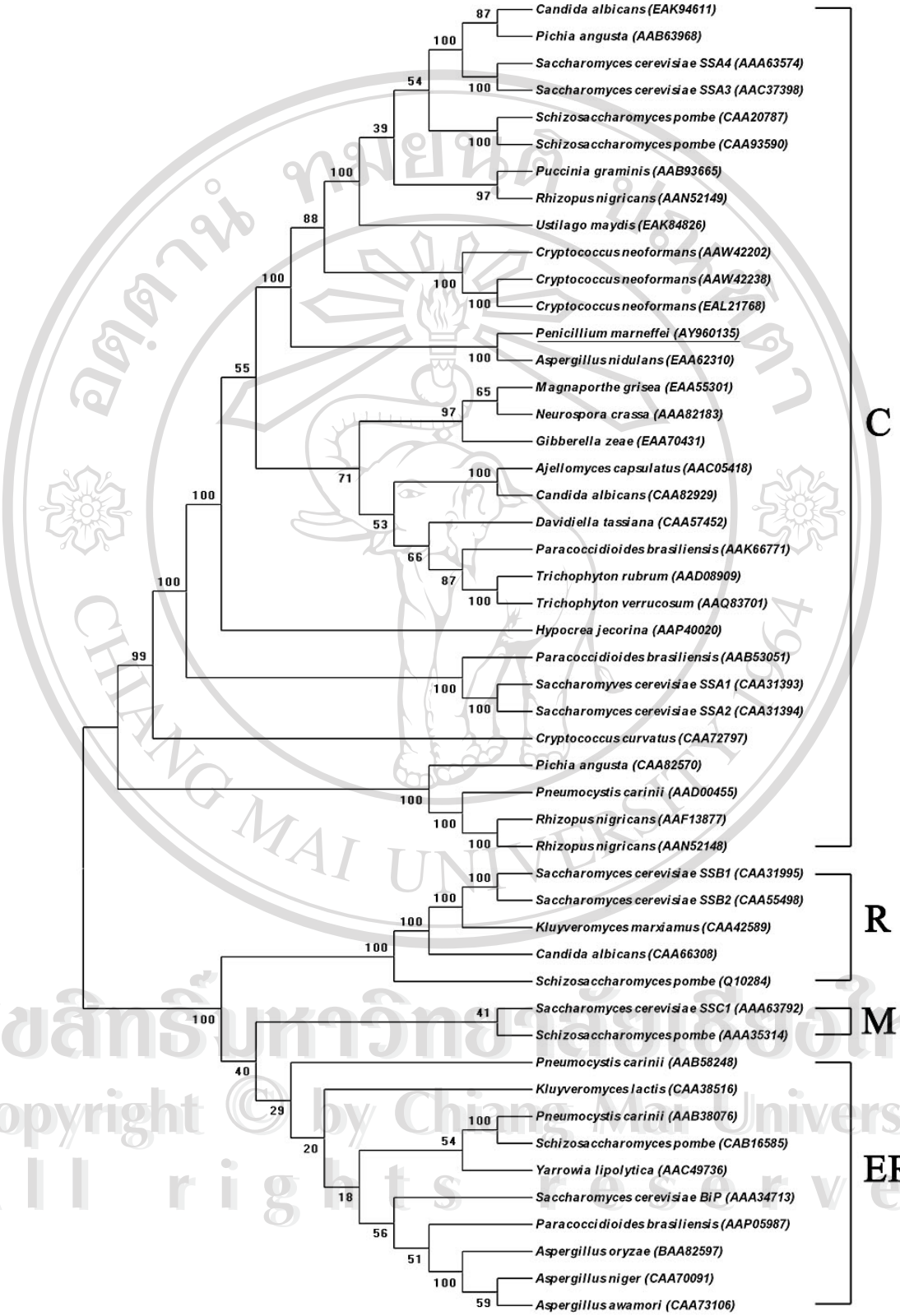

Figure 9. Neighbor-joining phylogenetic tree of fungal heat shock protein 70 (Hsp70s). Alignment of complete deduced amino acid sequences and construction of phylogenetic tree was done with MEGA3 program. Intracellular localization of Hsp70s is indicated with letters: M, mitochondria; R, associated with ribosomes; ER, endoplasmic reticulum; C, cytosol. GenBank accession numbers of protein sequences are shown in brackets.

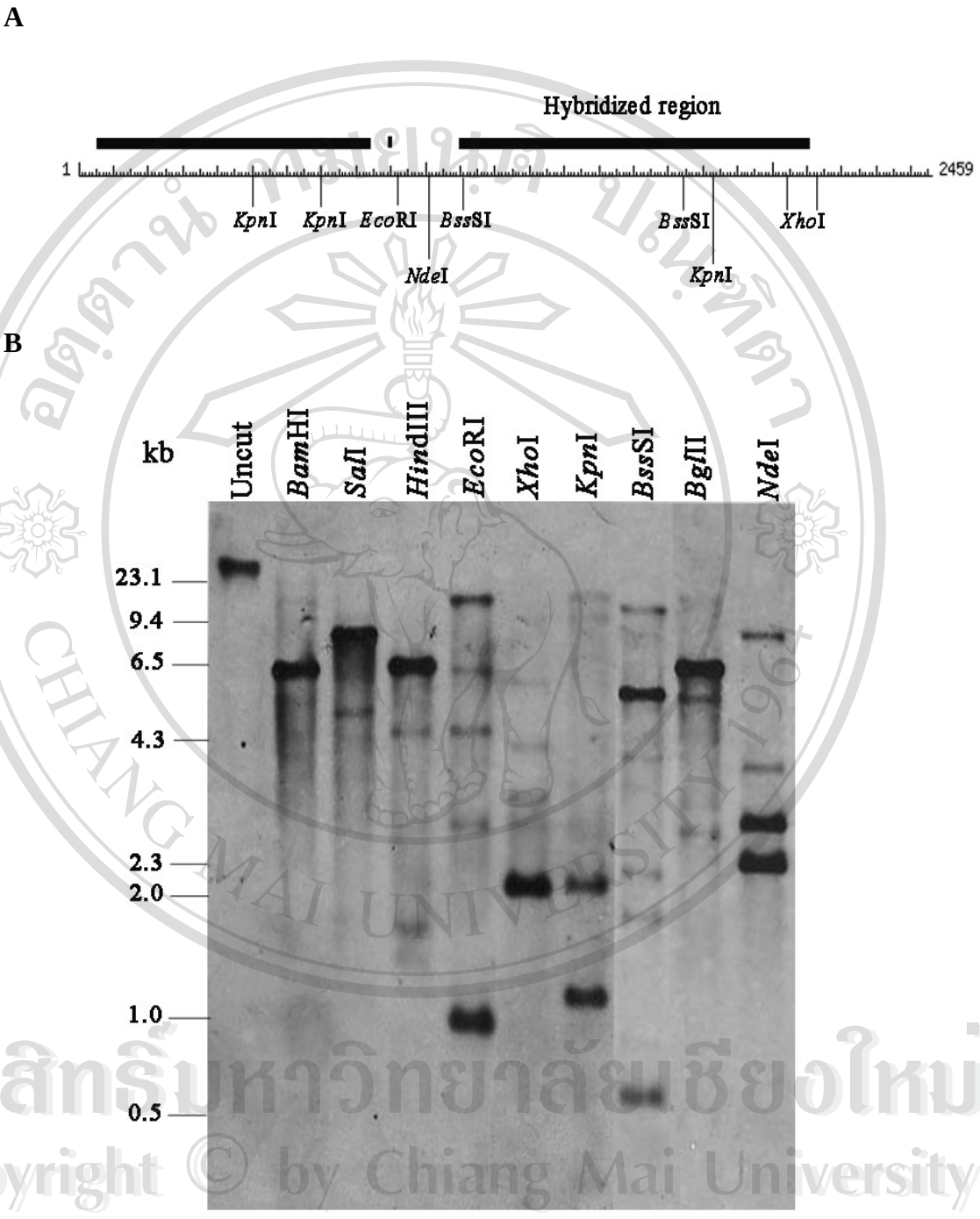

Figure 10. Southern blot analysis of *hsp70*. Restriction map (A) shows the cut sites of restriction enzymes used in Southern blot analysis (B). Immobilized digested genomic DNA of *P. marneffei* F4 strain was probed with *hsp70* gene fragment under high stringency condition. Hybridized region is shown in the black bar. Sizes of Lamda *Hin*dIII marker are shown in the left panel.

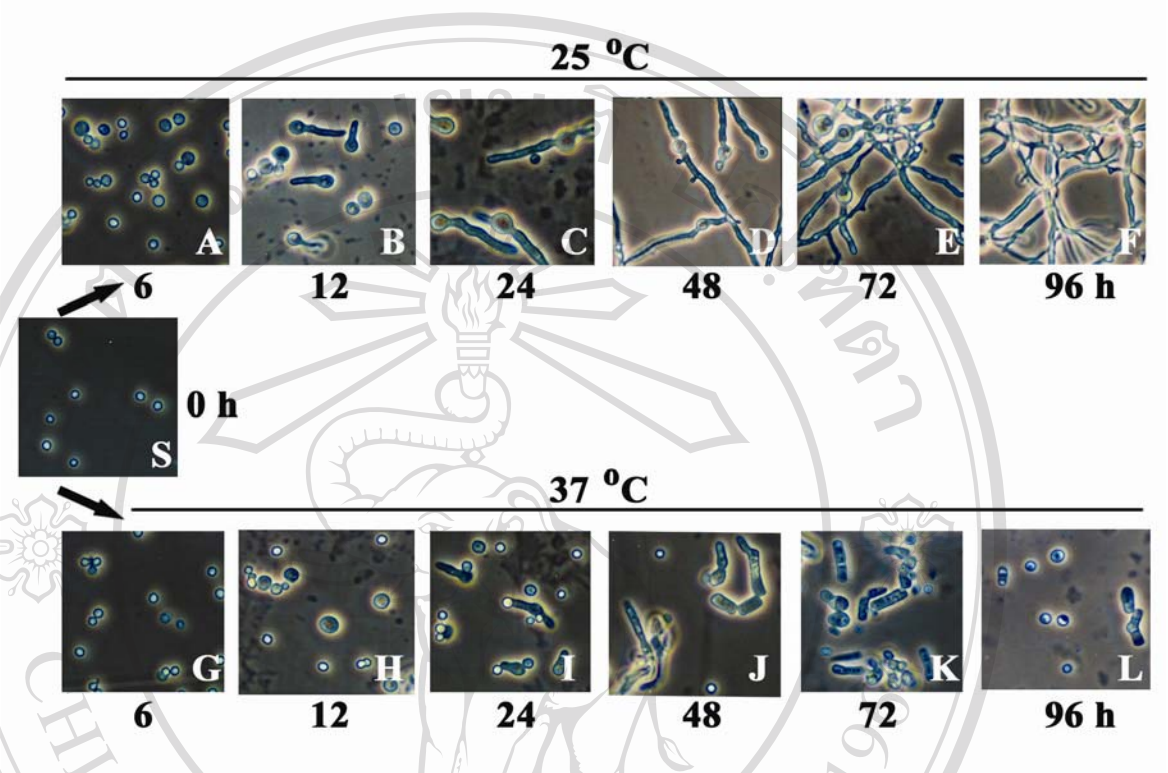

Figure 11. Saprobic-parasitic-phase transition of *P. marneffei*. Conidia (S) were incubated at 25°C or 37°C in brain heart infusion broth. At 25°C (A to F), conidia were swelled within 6 h. Germ tubes were produced between 6-12 h. Hyphae growth was fully established in 3-4-day culture. At 37°C (G to L), conidia growth occurred within the first 12 h. Germ tubes were formed by polarization between 12 to 24 h. Cellular differentiation to yeast-like cells began 48 h. Most completed yeast cells were observed at 96 h after germination.

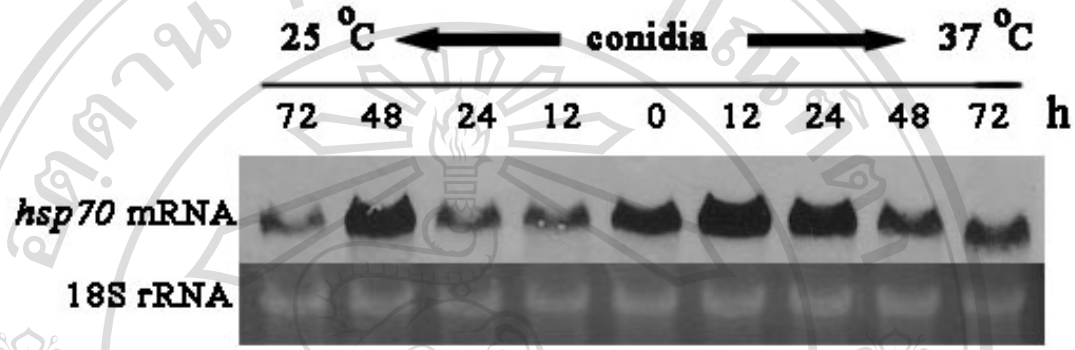

Figure 12. Expression of *hsp70* during mold and yeast phase transition by Northern hybridization. Total RNA was extracted from spore, mycelial and yeast cells at different time intervals during the temperature-induced morphological transition of the fungus. *P. marneffei hsp70* cDNA was used as a probe (upper panel) in Northern hybridization. The 18S rRNA loading control is shown in the lower panel.

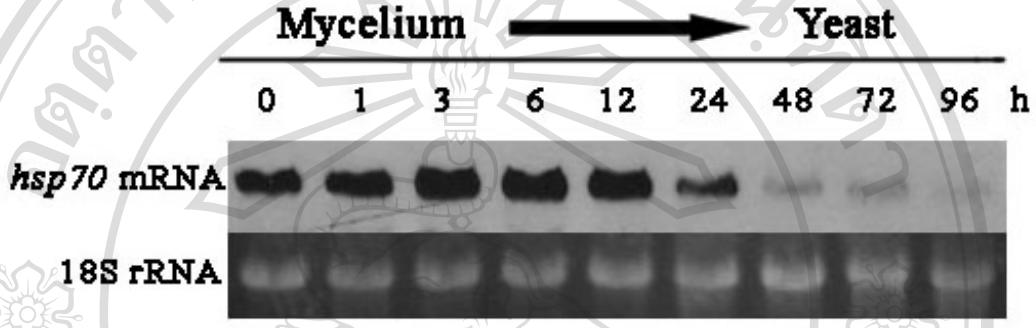

Figure 13. Expression of *hsp70* during mycelial transition to yeast. Total RNA was extracted from complete hyphae (25°C, 72 h) and cells encountered to the yeast inducible temperature (37°C) for 0.5, 1 and 3 h. The *P. marneffei hsp70* cDNA was used as a probe in Northern hybridization (upper panel). The 18S rRNA loading control is shown in the lower panel.

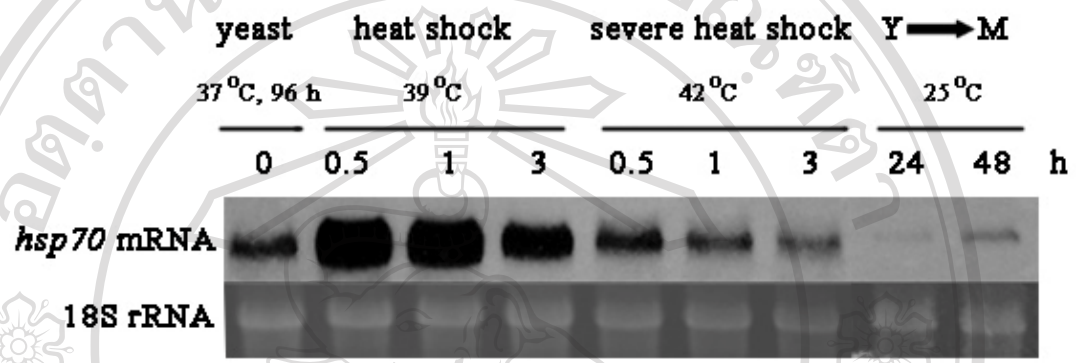

Figure 14. Expression of *hsp70* during heat shock and the reverse to room temperature conditions. Total RNA was extracted from complete yeast cells (37 ^{o}C, 96 h); cells encountered the heat shock (39^{o}C) and severe heat shock (42^{o}C) conditions for 0.5, 1 and 3 h; and after reverse to room temperature for 24 and 48 h. The *P. marneffei hsp70* cDNA was used as a probe in Northern hybridization (upper panel). The 18S rRNA loading control is shown in the lower panel.

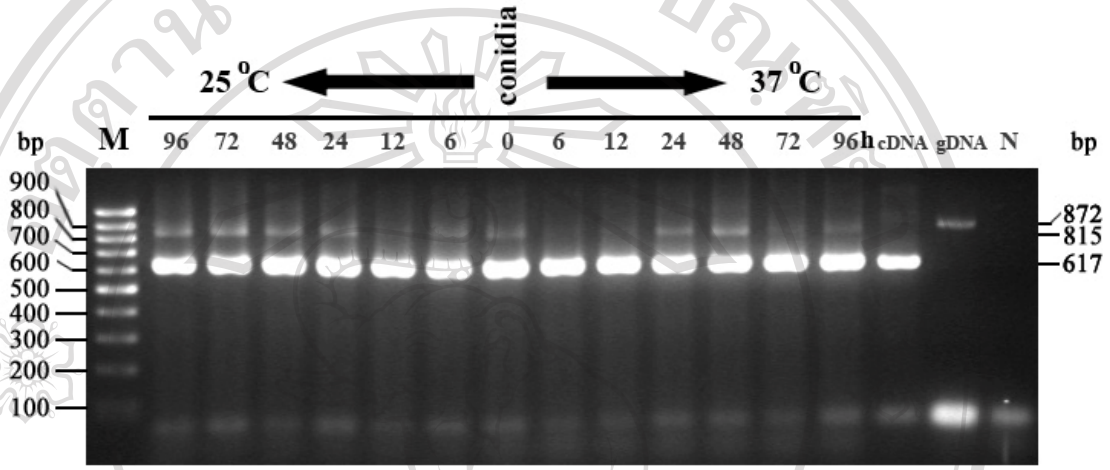

Figure 15. Analysis of *hsp70* mRNA splicing during saprobic (to 25°C) and parasitic (to 37°C) phase transition by RT-PCR. The molecular marker indicated on the left and the expected sizes of mature *hsp70* mRNA (617 bp), intronII-unspliced *hsp70* mRNA (815 bp), and a fragment from PCR which contained 2 introns (872 bp) are indicated on the right. Control templates; cDNA (complementary DNA), gDNA (genomic DNA) and N (negative control).

```
RNAM13R+F       CAGTCTTCGATGCTAAGCGTTTGATCGGCCGCAAATTCTCCGATCCTGAGGTCCAGGCTG 60
G20M13F+R       CAGTCTTCGATGCTAAGCGTTTGATCGGCCGCAAATTCTCCGATCCTGAGGTCCAGGCTG 60
17M13F+R        CAGTCTTCGATGCTAAGCGTTTGATCGGCCGCAAATTCTCCGATCCTGAGGTCCAGGCTG 60
                ************************************************************
                ---------->
                PMH70-F2

RNAM13R+F       ATGCCAAGCACTTCCCTTTCAAGATCATCGAGAAGGCCACCAAGCCCGTTATCGAGGTCG 120
G20M13F+R       ATGCCAAGCACTTCCCTTTCAAGATCATCGAGAAGGCCACCAAGCCCGTTATCGAGGTCG 120
17M13F+R        ATGCCAAGCACTTCCCTTTCAAGATCATCGAGAAGGCCACCAAGCCCGTTATCGAGGTCG 120
                ************************************************************

RNAM13R+F       AGTTCAAGGGTGAGGTCAAGCAGTTCACACCTGAGGAAATCTCTTCCATGGTCTTGATCA 180
G20M13F+R       AGTTCAAGGGTGAGGTCAAGCAGTTCACACCTGAGGAAATCTCTTCCATGGTCTTGATCA 180
17M13F+R        AGTTCAAGGGTGAGGTCAAGCAGTTCACACCTGAGGAAATCTCTTCCATGGTCTTGATCA 180
                ************************************************************

RNAM13R+F       AGATGCGTGAGACTGCTGAGGCCTACCTCGGTGGTACCGTTAACAACGCTGTCATCACTG 240
G20M13F+R       AGATGCGTGAGACTGCTGAGGCCTACCTCGGTGGTACCGTTAACAACGCTGTCATCACTG 240
17M13F+R        AGATGCGTGAGACTGCTGAGGCCTACCTCGGTGGTACCGTTAACAACGCTGTCATCACTG 240
                ************************************************************

RNAM13R+F       TCCCCGCCTACTTCAACGACTCCCAGCGTCAGGCCACCAAGGATGCTGGTCTCATTGCCG 300
G20M13F+R       TCCCCGCCTACTTCAACGACTCCCAGCGTCAGGCCACCAAGGATGCTGGTCTCATTGCCG 300
17M13F+R        TCCCCGCCTACTTCAACGACTCCCAGCGTCAGGCCACCAAGGATGCTGGTCTCATTGCCG 300
                ************************************************************

RNAM13R+F       GTTTGAACGTCCTCCGTATCATCAACGAACCTACTGCTGCCGCCATTGCCTACGGTCTCG 360
G20M13F+R       GTTTGAACGTCCTCCGTATCATCAACGAACCTACTGCTGCCGCCATTGCCTACGGTCTCG 360
17M13F+R        GTTTGAACGTCCTCCGTATCATCAACGAACCTACTGCTGCCGCCATTGCCTACGGTCTCG 360
                ************************************************************

RNAM13R+F       ACAAGAAGGTTGAGGGTGAGCGCAACGTTCTCATCTTCGATCTTGGTGGTGGTACCTTCG 420
G20M13F+R       ACAAGAAGGTTGAGGGTGAGCGCAACGTTCTCATCTTCGATCTTGGTGGTGGTACCTTCG 420
17M13F+R        ACAAGAAGGTTGAGGGTGAGCGCAACGTTCTCATCTTCGATCTTGGTGGTGGTACCTTCG 420
                ************************************************************

RNAM13R+F       ATGTCTCTCTCCTCACCATCGAGGACGGTATCTTCGAGGTCAAGGCCACCGCCGGTGACA 480
G20M13F+R       ATGTCTCTCTCCTCACCATCGAGGACGGTATCTTCGAGGTCAAGGCCACCGCCGGTGACA 480
17M13F+R        ATGTCTCTCTCCTCACCATCGAGGACGGTATCTTCGAGGTCAAGGCCACCGCCGGTGACA 480
                ************************************************************

RNAM13R+F       CTCACTTGGGTGGTGAGGACTTCGACTCTCGCCTTGTCAACCACTTTGCCTCCGAGTTCA 540
G20M13F+R       CTCACTTGGGTGGTGAGGACTTCGACTCTCGCCTTGTCAACCACTTTGCCTCCGAGTTCA 540
17M13F+R        CTCACTTGGGTGGTGAGGACTTCGACTCTCGCCTTGTCAACCACTTTGCCTCCGAGTTCA 540
                ************************************************************
                                                    Intron I
RNAM13R+F       AGAGGAAATATAAGA-------------------------------------------- 555
G20M13F+R       AGAGGAAATATAAGAGTATGTGATATACCCCCATCTGTTTTACAAGAATTGATAAGCTAA 600
17M13F+R        AGAGGAAATATAAGA-------------------------------------------- 555
                ***************

RNAM13R+F       ------------AGG------------------------------------------- 558
G20M13F+R       CAACTTTCACAGAGGGTACGTCAAAAATCCGAAATGAATTCTCATTCCCTCCCGTCTCTT 660
17M13F+R        ------------AGGGTACGTCAAAAATCCGAAATGAATTCTCATTCCCTCCCGTCTCTT 603
                            ***
                                                    Intron II
RNAM13R+F       ------------------------------------------------------------
G20M13F+R       GCATTCTTATTTGATATTTGGAGCTTGGCTCTGGAGACTTCCTCAGAACCATCTGAAGAA 720
17M13F+R        GCATTCTTATTTGATATTTGGAGCTTGGCTCTGGAGACTTCCTCAGAACCATCTGAAGAA 663

RNAM13R+F       ------------------------------------------------------------
G20M13F+R       TGCATATGAAGACTCCCAATATGAATCCATCTCCTTCCCTTTCATCCGCACTTGACTCTT 780
17M13F+R        TGCATATGAAGACTCCCAATATGAATCCATCTCCTTCCCTTTCATCCGCACTTGACTCTT 723

RNAM13R+F       --------------------------------ATTTGACCACCAATGCTCGTGCTCTCC 585
G20M13F+R       GAAGAACTCCTCAGATGCTAACATGTGAAATAGATTTGACCACCAATGCTCGTGCTCTCC 840
17M13F+R        GAAGAACTCCTCAGATGCTAACATGTGAAATAGATTTGACCACCAATGCTCGTGCTCTCC 783
                                                ************************
```

```
RNAM13R+F       GCCGTCTCCGCACTGCCTGTGAGCGTGCTAAG 617
G20M13F+R       GCCGTCTCCGCACTGCCTGTGAGCGTGCTAAG 872
17M13F+R        GCCGTCTCCGCACTGCCTGTGAGCGTGCTAAG 815
                ********************************
```

PMH70-REV2

Figure 16. Comparison of the nucleotide sequences for confirming the sequences of RT-PCR products. Partially nucleotide sequences of suspected mature *hsp70* mRNA (RNAM13R+F), intronII-unspliced *hsp70* mRNA (17M13F+R), and intronI+II containing *hsp7*0 gene (G20M13F+R) were used. The alignment was performed using ClustalW program. Nucleotide sequences of intron I are indicated by a line and intronII are in bold. Primer binding sites are indicated by arrows.