J. Biochem. 121, 464-476 (1997)

# Characterization of a Gene Family Encoding Cysteine Proteinases of *Sitophilus zeamais* (Maize Weevil), and Analysis of the Protein Distribution in Various Tissues Including Alimentary Tract and Germ Cells

Ichiro Matsumoto,*[,1] Yasufumi Emori,[†] Keiko Abe,* and Soichi Arai*

*Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, and [†]Department of Biophysics and Biochemistry, Faculty of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113

Received for publication, September 4, 1996

**We have identified and characterized a gene family comprising at least four genes encoding cathepsin L-like cysteine proteinases (SCPs) in *Sitophilus zeamais*, Coleoptera. A cDNA clone corresponding to a major mRNA species of the gene family in adult insects codes for a cysteine proteinase (CP) of 338 amino acid residues, and two genomic clones code for similar CPs of 331 residues whose COOH-terminal structures are seven amino acids shorter. These three SCPs show high sequence similarity to one another as well as to other insect and mammalian cathepsin L-like proteinases. A polyclonal antibody raised against bacterially expressed SCP was used as a probe to examine the molecular forms and distribution of the enzyme. SCP exists in both proenzyme and mature forms in larvae, pupae, and adults, and the proenzyme converts *in vitro* into the mature form at acidic pH. Immunohistochemical analysis showed that SCP is present in several tissues including alimentary organs and germ cells. In alimentary organs, SCP is distributed in the gastric caeca, but not the midgut. SCP is also present in genital organs, especially in oocytes and nurse cells, where it exists at high levels. These results indicate that SCP plays a variety of physiological roles including a role in food digestion.**

**Key words: alimentary tract, cysteine proteinase, gene family, germ cells, maize weevil.**

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

Many cysteine proteinases (CPs) have been characterized and their involvement in diverse intracellular and extracellular events is being disclosed (*1-6*). CPs can be divided into several structural families; these include papain-type proteinases, newly identified ICE/CED3 proteases, and asparagine-specific endoproteases (*5-7*). The papain family is the largest, and its members are widely distributed in animals, plants, and microorganisms (*8-18*). The papain family comprises several subfamilies, such as cathepsins B-like, H-like, and L-like subfamilies. These cathepsins have long been studied as factors involved in physiologically normal proteolysis and also as possible triggers for some metabolic disorders (*1, 2, 19-22*). Recently, cathepsins K, C, and S have been found as lysosomal enzymes belonging to the papain family, suggesting that this family has phylogenically diverged (*7, 11-13*).

The activities of these mammalian CPs are regulated by specific endogenous proteinaceous inhibitors termed cystatins. The inhibitory mechanisms of cystatins and the crystal structures of CP-cystatin complexes have also been studied (*23-25*). The CP-cystatin system in mammals thus appears to be involved in the regulation of physiological protein degradation in intracellular and, in some cases, extracellular spaces. The CP-cystatin system is thought to exist in plants as well, and we have identified a plant cystatin in rice seeds, termed oryzacystatin, and characterized its biochemical and molecular properties (*26-28*). Subsequently, another cystatin in rice seeds was identified and we proposed that a phytocystatin family could be added to the conventional cystatin superfamily that was based exclusively on animal cystatins (*29, 30*). This proposition has been reinforced by the further identification of related cystatins in other plants (*31-33*). Besides phytocystatins, studies on endogenous target proteinases have also been performed. In rice seeds, we have identified three different CPs, termed oryzains $\alpha$, $\beta$, and $\gamma$, which show significant sequence similarities to animal cathepsins H and L (*16*). These findings suggest that the CP-cystatin system in plants might be involved in physiological processes such as protein maturation during seed ripening and proteolysis during germination.

Furthermore, an antiviral activity is observed for phytocystatins (*34, 35*). In addition, a biodefensive function against insect pests has been suggested from the observation that the propagation of cereal insect pests, *Callosobruchus chinensis* and *Riptortus clavatus*, is inhibited by the addition of oryzacystatins to their diet (*36*). Although it is known that CPs, together with serine and aspartic protein-

[1] To whom correspondence should be addressed. Tel: +81-3-3812-2111 (ext. 5129), Fax: +81-3-5802-8897

Abbreviations: CP(s), cysteine proteinase(s); E-64, *N*-(L-3-carboxy-2,3-*trans*-carboxyoxirane-2-carbonyl)-L-leucyl-(4-guanidino)butane; PMSF, phenylmethane sulfonyl fluoride; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcription-PCR; SCP(s), *Sitophilus zeamais* cysteine proteinase(s); *scp*, gene for *Sitophilus zeamais* cysteine proteinase.

ases, are present in extracts from the alimentary tracts of plant insect pests (*37, 38*), there has been little analysis of their molecular and biochemical traits. We thus started to investigate digestive CPs in insects in order to elucidate the biochemical and molecular mechanisms of the above biodefense phenomena. As a first step, we attempted to identify CPs in insect pests that can be inhibited by oryzacystatin. For this, we used *Drosophila melanogaster* as a model insect, since various molecular and genetic methods are applicable, and since anatomical descriptions are available and the physiological significance of various tissues and their developmental processes have been established in *D. melanogaster*. We identified a cathepsin L-like CP, termed DCP1, as an efficient target enzyme for any phytocystatin (*39*). We also showed that DCP1 mRNA is localized in several tissues and organs, especially in the alimentary tract, such as in the midgut and gastric caeca, suggesting that DCP1 may function as a major enzyme necessary for food protein digestion.

In the present study, we have identified a gene family encoding CPs of a grain insect pest belonging to Coleoptera, *Sitophilus zeamais* (maize weevil), that feeds on corn or rice seeds. The encoded CPs, termed SCPs, show significant amino acid sequence similarity to DCP1 and other cathepsin L-like CPs including mammalian cathepsin L and *Bombyx mori* CP (*39-41*). Here we also report the tissue distribution of SCP observed using a specific antibody raised against recombinant SCP expressed in *Escherichia coli*, and suggest that it could play several distinct roles, including roles in digestion and oogenesis.

## MATERIALS AND METHODS

*Isolation and Nucleotide Sequencing of Genomic Clones Encoding Cysteine Proteinase*—A genomic library of *S. zeamais* was constructed in λEMBL3 according to a standard procedure described previously (*39*), and screened with a DNA fragment encoding *D. melanogaster* cysteine proteinase under less stringent conditions, in that the final washing was carried out at 55°C in 2×SSC (20×SSC is 3 M NaCl and 0.3 M trisodium citrate, pH 7.0) containing 0.1% SDS. Positive plaques were subcloned into pUC18 plasmid vector and their nucleotide sequences were determined with a DNA sequencer (model 373A, Applied Biosystems) according to the manufacturer's instructions.

*Genomic Southern Blot Hybridization*—Approximately 10 μg of genomic DNA prepared from adult *S. zeamais* as described previously (*39*) was digested with the appropriate restriction enzymes, fractionated on agarose gels, denatured, and transferred onto nylon membrane filters (Hybond-N, Amersham) according to the method of Southern (*42*). Filters were prehybridized and hybridized at 65°C with a $^{32}$P-labeled *Hin*cII-*Nco*I fragment of pSCPg3 (see "RESULTS" and Fig. 1A). The filters were finally washed at 65°C in 0.1×SSC containing 0.2% SDS.

*Northern Blot Hybridization*—Total RNA was extracted from *S. zeamais* at various developmental stages by the guanidinium-thiocyanate method (*43*) or with ISOGEN RNA purification reagent (Nippongene). Approximately 10 μg of RNA was denatured and electrophoresed in a formaldehyde-containing agarose gel. After electrophoresis, the RNA was transferred onto a nylon membrane filter and hybridized with the $^{32}$P-labeled *Hin*cII-*Nco*I fragment (Fig. 1A) at 42°C in 5×SSPE containing 50% formamide (*43*). The filter was finally washed at 65°C in 0.1×SSC containing 0.2% SDS.

*RT-PCR*—*S. zeamais* mRNA was purified by oligo(dT)-cellulose column chromatography (*43*) and cDNA was synthesized from 0.2 μg of mRNA with a first-strand cDNA synthesis kit (Pharmacia). The following three oligonucleotide primers common to both pSCPg2 and pSCPg3 (see Fig. 1B) were synthesized: S1 (5′-AATGTAGCTCACAAGAA-3′), containing one mismatched base between pSCPg2 and pSCPg3; S2 (5′-GCTAGTTATCCTTAGGTT-3′); and S (5′-CAGTTGAAAAGTAGCT-3′). The syntheses were conducted with a DNA Synthesizer (Oligo 1000, Beckman) for 3′-RACE. PCR amplifications were performed using the synthesized primers and oligo$(dT)_{18}$ primer (20 pmol of each primer) with the annealing temperature set at 48°C (*44*). Amplified cDNA was blunt-ended with Klenow fragment and T4 DNA polymerase, phosphorylated with T4 polynucleotide kinase, digested with appropriate restriction enzymes, and ligated into pBluescript II plasmid vector. The nucleotide sequences of the inserts were determined as described above.

*Antibody Preparation*—An *S. zeamais* genomic DNA fragment of the cysteine proteinase gene (pSCPg3 shown in Fig. 1A) encoding the amino acid sequence from 167 to 331 (Fig. 1) was inserted into the polylinker site downstream of the *lacZ* promoter of pUC18 in the same orientation and in-frame with *lacZ*. Recombinant protein production was induced with isopropyl-β-D-thiogalactopyranoside, and the total protein was subjected to 10% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (*45*). The protein band visualized by staining with Coomassie Brilliant Blue R-250 was excised and crushed to a paste. The paste was sonicated extensively, mixed with an equal volume of Freund's complete adjuvant (Nacalai Tesque), and injected into a male New Zealand white rabbit. Four weeks later, a booster injection was administered in Freund's incomplete adjuvant (Nacalai Tesque). Serum was obtained 10 days after the booster and checked for immunoreactivity (*46*). Affinity purification of the resulting antibodies was performed with recombinant protein fused to glutathione-S-transferase (GST) (*47*). For this, the same DNA fragment described above was inserted downstream of the GST gene of pGEX-3X (Pharmacia Biotech) in the same orientation and in-frame. GST fusion protein was expressed in *E. coli* AD202 kindly provided by Prof. T. Saito (*48*), subjected to 10% SDS-PAGE, and then electroblotted onto a polyvinylidene difluoride (PVDF) membrane (Immobilon-P, Millipore) (*46*). The membrane strips containing the GST fusion protein were excised and incubated with anti-serum against SCP. The antibody bound to the membrane was eluted with 0.2 M glycine/HCl buffer (pH 2.75), immediately neutralized with 1 M Tris/HCl containing NaOH (pH 8.0), and then dialyzed against PBS (140 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$, pH 7.5) for 3 h at 0°C. Affinity-purified antibody was absorbed with a control extract of *E. coli* AD202 not harboring the recombinant plasmid.

*Detection of CPs in Insects*—Total extracts from living *S. zeamais* at larval, pupal, and adult stages, and total extracts of adult *D. melanogaster* were prepared by homogenizing the insects in five volumes of SDS-PAGE loading buffer composed of 0.1 M Tris/HCl (pH 6.8), 0.5 M β-mercapto-



Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

ethanol, 4% SDS, 20% glycerol, and 0.1% BPB. Total proteins were electrophoresed in 10% SDS-polyacrylamide gel (*45*) and electroblotted onto a PVDF membrane (*46*); the membrane was immunoreacted with affinity-purified antibody against SCPs (1:500 dilution) and the immunoreacted proteins were visualized using peroxidase-conjugated goat anti-(rabbit IgG) antibody and an ECL kit (Amersham).

*In Vitro Processing of SCP*—A crude extract of adult *S. zeamais* was prepared as follows. Adult insects were homogenized in 10 mM Tris/HCl buffer (pH 7.5) containing 1 mM EDTA and 2 mM $\beta$-mercaptoethanol, and debris was removed by centrifugation. A saturated ammonium sulfate solution was added to the supernatant to give 50% saturation. The mixture was incubated for 1 h at 0°C, and the precipitate formed was collected by centrifugation, dissolved in a small amount of extraction buffer, and dialyzed against the same buffer. This crude SCP fraction was stored for 2 or 3 days at 4°C and then incubated at 37°C for 30 min in 0.1 M sodium formate buffer (pH 3.5) containing 0.2% hemoglobin or casein to prevent autodegradation. In some cases, cysteine proteinase inhibitors such as E-64 or leupeptin (Peptide Institute) or other proteinase inhibitors were added to the assay solution. The reaction was terminated by adding 2×SDS-PAGE loading buffer, and the samples were subjected to SDS-PAGE and immunoblot analysis as described above.

*Immunohistochemistry*—Tissues and organs were dissected from *S. zeamais* larvae and adults in PBS and fixed with 4% paraformaldehyde in PBS containing 0.1% Triton X-100. Fixed tissues and organs were preincubated with 3% normal goat serum in PBS containing 0.1% Triton X-100, and then incubated with the anti-SCP antibody prepared as above at a dilution of 1:500. The samples were then processed by a standard procedure using an ABC kit (Vector Lab.) with 3,3′-diaminobenzidine as a chromogenic substrate (Pierce) in the presence of $NiCl_2$ and $CoCl_2$.

*In Situ Hybridization to D. melanogaster*—*In situ* hybridization of dissected tissues was carried out essentially as described (*39*). Ovaries were dissected from female *D. melanogaster* Canton-S, fixed in 4% paraformaldehyde in PBS containing 0.1% Triton X-100, dehydrated in methanol, re-fixed, and treated with 50 $\mu$g/ml of proteinase K in PBS containing 0.1% Tween 20. The proteinase-treated tissues were subjected to hybridization using a digoxygenin-labeled antisense RNA probe for DCP1 (*39*). Anti-digoxygenin antibody conjugated to alkaline phosphatase (Boehringer) was allowed to react with the chromogenic substrates, and 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium chloride were used to visualize the signals.

## RESULTS

*Isolation and Characterization of Genes Encoding Sitophilus zeamais Cysteine Proteinase (SCP)*—Screening of an *S. zeamais* genomic library with a DNA probe for *D. melanogaster* cysteine proteinase (DCP1) (*39*) yielded several positive clones. Three that showed strong signals, termed λSCPg2, λSCPg3, and λSCPg12, were subcloned into plasmid vector pUC18 (resultant plasmids; pSCPg2, pSCPg3, and pSCPg12, respectively), and their restriction maps were analyzed. Since two clones, pSCPg2 and pSCPg12, showed the same restriction maps in the overlapping regions spanning approximately 11 kbp, and were thus found to be clones originating from the same genomic region, we further analyzed pSCPg2 as a representative. Two clones, pSCPg2 and pSCPg3, were examined by restriction enzyme mapping and Southern hybridization analysis using DCP1 probe. The restriction map of pSCPg2 was found to be distinct from that of pSCPg3 when several enzymes such as *Bam*HI, *Pst*I, *Mlu*I, *Eco*RV, *Bgl*II, and *Ban*III were used (Fig. 1A). However, the probe-positive regions of the two clones showed several identical restriction sites; they shared *Spe*I, *Pvu*II, *Eco*RI, *Hinc*II, *Nco*I, *Kpn*I, *Sac*I, and *Hind*III sites. Thus, nucleotide sequences of the relevant regions of pSCPg2 and pSCPg3 were determined.

Both clones contained long open reading frames (ORFs) encoding CPs similar to mammalian cathepsin L and *Drosophila* CP (DCP1) as described later (Fig. 5). The proteins encoded by these two genes were termed SCPg2 and SCPg3, corresponding to pSCPg2 and pSCPg3, respectively. The approximately 1.5 kbp sequences of the two clones were very similar to each other, with only eight nucleotide differences in the ORFs and four differences in their deduced amino acid sequences (Fig. 1B). There were also significant similarities in the putative 5′- and 3′-noncoding regions. About 250 nucleotides of the 5′-noncoding regions bordering the ORFs were very similar, but no significant similarity was found in the sequences further upstream. In the 3′-noncoding regions, approximately 200 nucleotides bordering the termination codon were highly similar, and 300 bp sequences downstream of the coding regions also showed lower but significant similarities that decreased gradually.

*Estimation of Copy Number of the Genes Encoding SCPs by Genomic Southern Blot Analysis*—We performed genomic Southern blot analysis to identify DNA fragments of the two clones obtained from the genome of *S. zeamais* and to estimate the copy numbers of related genes. The protein coding regions of the two SCP clones (pSCPg2 and pSCPg3) had the same restriction maps although their 5′-upstream regions apparently contained different sites (Fig. 1A). Thus, the two genes were expected to be distinguishable by genomic Southern blot analysis using two restriction enzymes, one directed to the coding region and the other to the 5′-upstream region. Examination of several restriction sites in the upstream regions of pSCPg2 and pSCPg3 showed three enzymes with distinct sites: *Ban*II sites were located 2.0 and 2.2 kbp from an *Nco*I site in the coding region in pSCPg2 and pSCPg3, respectively, and *Mva*I sites were located 2.0 and 1.8 kbp from the *Nco*I site in pSCPg2 and pSCPg3, respectively. A *Ban*I site was located 1.95 kbp from the *Nco*I site in pSCPg2, but was not found in the upstream region of pSCPg3.

Using these three restriction enzymes together with *Nco*I, a genomic Southern experiment was carried out under stringent hybridization and washing conditions. As shown in Fig. 2A, two bands were observed in each lane; these should correspond to *S. zeamais* genes for these two clones termed *scp*2 and *scp*3, respectively. In addition to these expected bands, one or two other bands were observed. This indicates that in addition to the two genomic clones obtained, at least one unidentified and highly related gene is contained in the *S. zeamais* genome. In order to

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

Fig. 1. **Nucleotide and deduced amino acid sequences of *Sitophilus zeamais* cysteine proteinases.** A: Restriction maps of two genomic DNA fragments encoding cysteine proteinases of *S. zeamais*. Open boxes indicate open reading frames. Restriction enzymes are as follows; Bm, *Bam*HI; Ps, *Pst*I; Ml, *Mlu*I; S, *Spe*I; Pv, *Pvu*II; E, *Eco*RI; Hc, *Hinc*II; N, *Nco*I; K, *Kpn*I; Sc, *Sac*I; Hd, *Hin*dIII; RV, *Eco*RV; Bg, *Bgl*II; BIII, *Ban*III; M, *Mva*I; BI, *Ban*I; BII, *Ban*II. Asterisks denote the partially identified restriction sites. A *Hinc*II-*Nco*I DNA fragment from pSCPg3 used in the genomic Southern and Northern blot analyses as a probe is indicated below the map of pSCPg3. The lengths of probe-positive fragments corresponding to each clone in the genomic Southern blot analysis (see Fig. 2A) are shown above each restriction map. B: Nucleotide and deduced amino acid sequences of pSCPg2 and pSCPg3. Dots indicate identical nucleotides and amino acid residues between pSCPg2 and pSCPg3; hyphens indicate gaps inserted to maximize the match. Nucleotide sequences are numbered from the first letter of the initiation ATG codon. Deduced amino acid sequences are also numbered from the initiation Met. An in-frame termination codon in the 5′-upstream region of each gene is underlined. The shaded sequences in pSCPg3 show the sense primers used in 3′-RACE (see "MATERIALS AND METHODS").

**A**

λSCPg2: 5′ Bm Ps Ml Sp* Hd* RV* Bg 3′; Mv—N; BII—N; BI—N; Hd—Hd; 1kbp; Pv* E* Hc* Sp* K E* E*; N* Pv* Sc; 0.5kbp

λSCPg3: 5′ RV Bg* RV Hc* Sp* Hd* BIII Bg* 3′; Mv—N; BII—N; (no BI)—N; Hd—Hd; 1kbp; Pv* E* Hc* Sp* K E* E*; N Pv* Sc; 0.5kbp

**B**

```
pSCPg2                                                                                     CTATTTTTTATTTTTA  -468
pSCPg3  GTCGATGATAGCTGAATAAAATTCGGAAGTA-TAAACAAGCTAAAACTCACGCACA-GCGGAAACAAG-TACGCCGAAAAAT...CAA.C.A.G..G  -475

pSCPg2  GGTGGGGAAGATGCAGAAATATCGTTTCCTATACATAGAACGGA-AGCCCAGTTTTATAAAAACGTGGCTAGTGATGAGGAAGCAGAAAAACCCAAGAAAGTTGATAAAAAATCTCACAA  -349
pSCPg3  CAA--AA..CTAA..C..G..-AACAAAG.CA.GT.TA....A.C..TGA.C...G.GTTG.TTTC..G.CAA..AC.AA..CA..TC...-T...--..CAAA.C....CG.TCA.T..  -361

pSCPg2  AGGACTTAGTAAA-TTAGCTGACTTAGCAGAGAC-AATCAATAGA-TGGGAAGATGATA--TGAGCCATCCAG-----TAAGTTTTACATAAATCATTTTATATTTTTCTTTTTA-TTTT  -240
pSCPg3  CC.G.AG.CA...CGC.CGA.CGG.....C.AC.TTC..GTAC..C..CT..AC..T..CT.CCAA...AA..TAACA...AG..............................T....  -241

pSCPg2  TTATTTATTTTTTTTATTTATTTACTATTCAATAGGTGATTTCTCAAAGTACTAGTACTACCACACCTAAAAAGACATGGAAGCCACCCGAACCACAGCCTCCTATCACAAATAATCGAG  -120
pSCPg3  ..............................................................................A.........................................  -121

pSCPg2  AAAGAAATCATTTTGCAGACAGTCCGTAAATAAGTAGTTCCATTCCAATTTATGGTCACTCAGTTGAAAAGTAGCTGTACAATCTTATAGACTTGTTATTTATCATATTTCACT-AAAAA  -1
pSCPg3  ..................................................................................................................A....  -1

pSCPg2  ATGAAGCTGTTACTAATTCTGGCAGCTGTAGTCATAAGCTGTCAAGCAGTCTCTTTCTACGATTTAGTACAAGAGCAATGGTCATCATTTAAGATGCAACATTCCAAAAATTACGACAGC  120
         M  K  L  L  L  I  L  A  A  V  V  I  S  C  Q  A  V  S  F  Y  D  L  V  Q  E  Q  W  S  S  F  K  M  Q  H  S  K  N  Y  D  S   40
pSCPg3  ........................................................................................................................  120
         .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .   40

pSCPg2  GAGACTGAGGAAAGATTCCGAATGAAAATATTTATGGAAAACGCCCACAAAGTTGCCAAACATAGCAAACTCTTCTCCCAAGGCTTTGTCAAGTTCAAGCTGGGTTTGAACAAGTACGCC  240
         E  T  E  E  R  F  R  M  K  I  F  M  E  N  A  H  K  V  A  K  H  S  K  L  F  S  Q  G  F  V  K  F  K  L  G  L  N  K  Y  A   80
pSCPg3  ..........................................A.............................................................................  240
         .  .  .  .  .  .  .  .  .  .  .  .  .  .  D  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .   80

pSCPg2  GACATGTTACACCATGAATTCGTCTCTACGCTTAACGGTTTCAACAAAACCAAGAACAATATCCTTAAAGGCAGTGACTTGAACGATGCAGTAAGGTTTATCTCCCCTGCAAATGTCAAA  360
         D  M  L  H  H  E  F  V  S  T  L  N  G  F  N  K  T  K  N  N  I  L  K  G  S  D  L  N  D  A  V  R  F  I  S  P  A  N  V  K  120
pSCPg3  ..T.......................................................................................A.............A.............  360
         .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  120

pSCPg2  TTGCCTGATACTGTTGACTGGAGGGACAAAGGAGCTGTCACTAAAGTCAAAGACCAAGGTCACTGCGGGTCGTGTTGGAGTTTCAGTGGTAGCGGTTCACTCGAAGGCCAGCACTTCCGT  480
         L  P  D  T  V  D  W  R  D  K  G  A  V  T  K  V  K  D  Q  G  H  C  G  S  C  W  S  F  S  G  S  G  S  L  E  G  Q  H  F  R  160
pSCPg3  ........................................................................................................................  480
         .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  160

pSCPg2  AAAACCGGCAAACTAGTGTCCCTTAGCGAACAAAACTTGGTAGATTGTTCAGGAAGATACGGCAACAACGGCTGCAACGGCGGTCTCATGGACAACGCCTTCCGTTACATCAAAGACAAC  600
         K  T  G  K  L  V  S  L  S  E  Q  N  L  V  D  C  S  G  R  Y  G  N  N  G  C  N  G  G  L  M  D  N  A  F  R  Y  I  K  D  N  200
pSCPg3  .....T............................................................C.....................................................  600
         .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  T  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  200

pSCPg2  GGCGGTATCGACACGGAACAGTCCTACCCCTACCTAGCCGAGGACGAGAAATGCCACTACAAAACCCAGAACAGCGGCGCAACCGACAAGGGCTTCGTAGACATCGAAGAAGGCAACGAA  720
         G  G  I  D  T  E  Q  S  Y  P  Y  L  A  E  D  E  K  C  H  Y  K  T  Q  N  S  G  A  T  D  K  G  F  V  D  I  E  E  G  N  E  240
pSCPg3  ........................................................................................................................  720
         .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  240

pSCPg2  GATGACCTTAAGGCAGCAGTGGCTACCGTAGGCCCCATTTCGATAGCCATCGACGCCAGCTACGAAACCTTCCAACTGTACTCGGATGGGGTGTACAGCGATCCTGAATGTATCTCACAA  840
         D  D  L  K  A  A  V  A  T  V  G  P  I  S  I  A  I  D  A  S  Y  E  T  F  Q  L  Y  S  D  G  V  Y  S  D  P  E  C  I  S  Q  280
pSCPg3  .....................................G..................................................................G...............  840
         .  .  .  .  .  .  .  .  .  .  .  .  V  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  S  .  .  280

pSCPg2  GAACTGGACCATGGGGTGTTGGTAGTGGGATACGGTACCAGCGACGATGGTCAGGACTACTGGTTGGTGAAGAATTCGTGGCGACCCAGCTGCGGATTGAACGGATACATCAAGATGGCC  960
         E  L  D  H  G  V  L  V  V  G  Y  G  T  S  D  D  G  Q  D  Y  W  L  V  K  N  S  W  R  P  S  C  G  L  N  G  Y  I  K  M  A  320
pSCPg3  ........................................................................................................................  960
         .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  320

pSCPg2  CGGAATCAAGATAACATGTGCGGAGTTGCATCTTAGGCTAGTTATCCTTAGGTTTAGAGGATATAGGACAGCGTTTGTGATTTTTATTAGGGTTTAGAATTCCTATCTCTAAGAGCTCTC  1080
         R  N  Q  D  N  M  C  G  V  A  S  *                                                                                     331
pSCPg3  ........................................................................................................................  1080
         .  .  .  .  .  .  .  .  .  .  .                                                                                        331

pSCPg2  TATTTTTTTTTGCTATCAATTTAGTTTTGTAAGATATTTTATGTCATCTCCTGAAGAGATTCTTTTTTAAGCTTCCATTAATAATAATAATAATAA-TAATTAATAACAGTT-ACTAACC  1198
pSCPg3  ...........T...........................A.G....-T..T......................................G.CT...TA.......C.T....T  1199

pSCPg2  TGGAG-ACGGACTTTACAATGTACCGTGGCAAATTGTATGCACAATTTATAAAAGTTTTCCCAAATCTTAAACACTTAAGCGAGCAACGCATCGCAGATCAACGCAGGACCATAGTAAGT  1317
pSCPg3  ..T..T.TATTTA...G..GTA.GA..AAA.TG.CT.CA.AT..GCCGC....GAAGAAA.....A.G.GGGT.A.A.T..---T....A.A.AA.T..AGT.TT.TT.GA..G..CT.-  1316

pSCPg2  AGGAATCTTCTTTCTCAAGAGACACTGGCCAGAATCAAGGAACAAATAAGACTAGAAATCAACCCAGAAACAACTATTGACATACCGCCACAACCACCGCAGAGCGAAAGCATA-CCTGA  1436
pSCPg3  .T.--..GAAA..T.--.C..TTCT..T....T..AC.ATTTGTG...CCTA..A..G.TT.TATT.T..A.TG.C..T..C.-.AAAT.TG.GTTTTAA..-.TTTTTTTT..TTA.C.  1430

pSCPg2  AGTAGAGACAGAGACTATACTACCGACA-GAGATTACACTACCGGCCGAACAAAACCCACTTGG                                                         1498
pSCPg3  CT..T.CTT.A.T...T..A.TGGA.T.AA.A..GTTT...TA.TT-A......TT.TCAA..ACAATAATAATAATAATAATAAAATAAGAAAGGATGGAAAGGGAAGAAGAGGCCAAT  1550
```

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

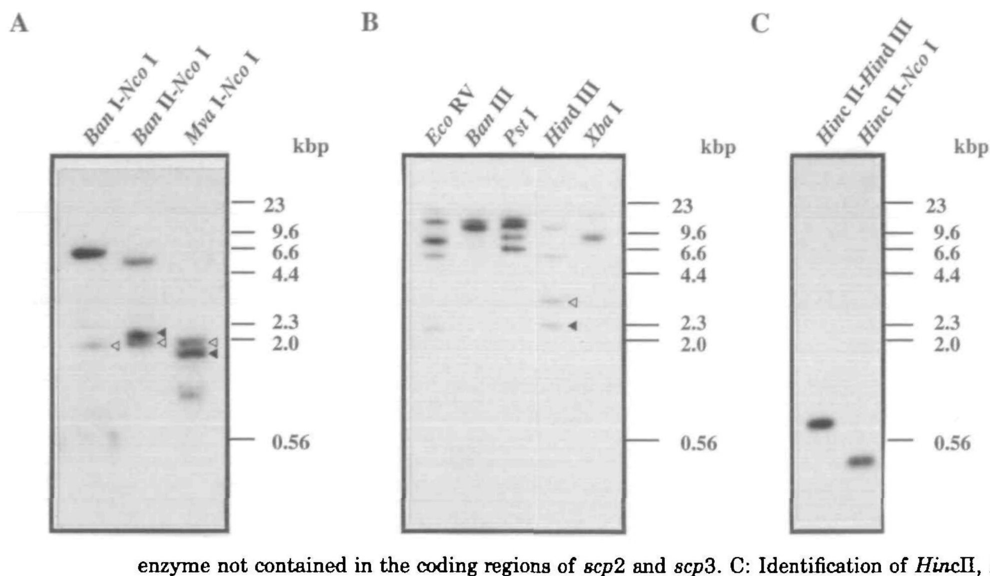


Fig. 2. **Genomic Southern blot analyses.** Chromosomal DNA from *S. zeamais* (approximately 10 $\mu$g) was digested with the restriction enzyme(s) indicated above each lane. The probe used was the $^{32}$P-labeled *Hinc*II-*Nco*I fragment of pSCPg3 (shown in Fig. 1A). Open and filled triangles in A and B indicate the positive signals corresponding to *scp*2 and *scp*3, respectively. A: Identification of two genes, *scp*2 and *scp*3, and other related genes. Several restriction enzymes showing different restriction maps in the 5′-upstream regions were used together with *Nco*I. B: Estimation of copy number for the SCP gene family. Genomic DNA was digested with each restriction enzyme not contained in the coding regions of *scp*2 and *scp*3. C: Identification of *Hinc*II, *Nco*I, and *Hind*III sites in SCP genes.

know the precise copy number, genomic Southern analysis was again performed using different restriction enzymes that did not digest the coding regions. As shown in Fig. 2B, four distinct bands were detected in the *Eco*RV, *Pst*I, and *Hind*III lanes, which indicates that the *S. zeamais* genome contains at least four related genes. Furthermore, the signals of some bands were stronger than the signals of the other bands in the *Eco*RV and *Pst*I lanes, indicating that the copy number of related genes may be five or six.

To examine the structural similarities among the five or six genes concerning the restriction sites observed commonly in *scp*2 and *scp*3, genomic Southern analysis was carried out again. In the lanes corresponding to *Hinc*II-*Hind*III and *Hinc*II-*Nco*I digestion, single bands were detected at 780 and 480 bp, respectively, consistent with the restriction maps of cloned *scp*2 and *scp*3 genes (Fig. 2C). This suggests that these restriction sites are common to all SCP genes and that none of the SCP genes contains an intron in this 780 bp region.

*Identification and Analysis of a Major mRNA Species in the SCP Gene Family*—Since plural CP genes appear to form a gene family in the *S. zeamais* genome as described above, it is important to determine which genes are transcribed. First, Northern blot analysis was performed to detect mRNA(s) for SCP(s). As shown in Fig. 3, a single band of 1.6 kb was detected in larvae, pupae, and adults under stringent hybridization and washing conditions. This suggests that only one copy of an SCP gene is transcribed, or else that plural transcripts of SCP genes have almost the same length. In addition, it was shown that SCP genes are continuously transcribed throughout development, similar to the case of DCP1 (*39*).

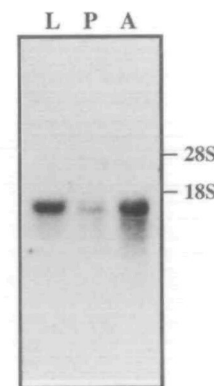


Fig. 3. **Expression of *scp* mRNAs during developmental stages of *Sitophilus*.** Five micrograms each of total RNA from *S. zeamais* at various developmental stages was separated by electrophoresis, blotted onto a nylon filter, and hybridized with the $^{32}$P-labeled *Hinc*II-*Nco*I fragment described in Fig. 1A. The positions of the 18S and 28S ribosomal RNA markers are indicated at the right.

To determine the nucleotide sequence(s) of the expressed gene(s), we performed 3′-RACE using two specific primers corresponding to the COOH-terminal region of the ORF (primer S1) and 3′-noncoding region (primer S2). When cDNA for adult *S. zeamais* mRNA was used as a template, a band of 550 bp for primer S1 and a band of 370 bp for primer S2 were amplified primarily. These amplified cDNA fragments were cloned and their nucleotide sequences were determined. The sequences of many clones revealed that the amplified fragments of 3′-RACE derived essentially from a single mRNA species, a finding unexpectedly different from the finding of two genomic clones, pSCPg2 and pSCPg3. Among more than 100 clones of 3′-RACE analyzed, there were no clones found for *scp*2 and *scp*3. Since the cDNA clones obtained only coded for the COOH-terminal region of CP, long 3′-RACE was performed using another primer corresponding to the 5′-noncoding sequence of pSCPg2 and pSCPg3 (primer S described in Figs. 1B and 4A) to give a longer cDNA fragment for the mRNA of the 3′-RACE clone. Again, only one sequence covering the whole coding sequence was obtained by long 3′-RACE; this sequence was the same as those obtained by primers S1 and S2 in the overlapping regions and coded for a CP (Fig. 4, A and B). The gene and the encoded protein for this major mRNA species were termed *scp*1 and SCPc1, respectively. Both the nucleotide and amino acid sequences of the cDNA in the coding region were very similar to those of pSCPg2 and pSCPg3 as described below. The nucleotide sequence of pSCPc1 was similar to those of pSCPg2 and



Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

**A**

Pv E Hc Sp N K Pv Sc Hd

S S1 S2 oligo (dT)18

pS1dT

pS1dT short

pS2dT

pSCP-c1

**B**

```
                                   CAGTTGAAAAGTAGCTGTACAACCTTATAGACTTAGTTATTTATCATATTTCACTAAGAAA      -1
ATGAAGCTGTTTCTAATTCTGGCAGCTGTAGTCATAAGCTGTCAAGCAGTCTCTTTCTACGATTTAGTACAAGAGCAATGGTCATCATTT   90
M  K  L  F  L  I  L  A  A  V  V  I  S  C  Q  A  V  S  F  Y  D  L  V  Q  E  Q  W  S  S  F      30
AAGATGCAACATTCCAAAAATTACGACAGCGAGACTGAGGAAAGGTTTAGAATGAAGATATTTATGGAAAACGCCCACAAAGTTGCCAAA  180
K  M  Q  H  S  K  N  Y  D  S  E  T  E  E  R  F  R  M  K  I  F  M  E  N  A  H  K  V  A  K      60
CATAACAAACTCTTCTCCCAAGGCTTTGTCAAGTTCAAGCTGGGTTTGAACAAGTACGCCGACATGTTACACCATGAATTCGTCTCTACG  270
H  N  K  L  F  S  Q  G  F  V  K  F  K  L  G  L  N  K  Y  A  D  M  L  H  H  E  F  V  S  T      90
CTTAACGGTTTCAACAAAACCAAGAACAATATTCTAAAAGGTAGTGACTTGAACGATGCAGTAAGATTTATCTCCCCTGCAAATGTCAAA  360
L  N  G  F  N  K  T  K  N  N  I  L  K  G  S  D  L  N  D  A  V  R  F  I  S  P  A  N  V  K     120
TTGCCTGATACTGTTGACTGGAGGGACAAAGGAGCTGTCACTGAAGTCAAAGACCAAGGTCACTGCGGGTCGTGTTGGAGTTTCAGTGCT  450
L  P  D  T  V  D  W  R  D  K  G  A  V  T  E  V  K  D  Q  G  H  C  G  S  C  W  S  F  S  A     150
ACTGGTTCACTCGAAGGTCAGCACTTCCGTAAAACCGGCAAACTAGTGTCCCTTAGCGAACAAAACTTGGTAGATTGTTCAGGAAGATAC  540
T  G  S  L  E  G  Q  H  F  R  K  T  G  K  L  V  S  L  S  E  Q  N  L  V  D  C  S  G  R  Y     180
GGCAACAACGGCTGCAACGGCGGTCTCATGGACAACGCCTTCCGTTACATCAAAGACAACGGCGGTATCGACACGGAAAAGTCCTACCCC  630
G  N  N  G  C  N  G  G  L  M  D  N  A  F  R  Y  I  K  D  N  G  G  I  D  T  E  K  S  Y  P     210
TACCTAGCCGAGGACGAGAAATGCCACTACAAAGCCCAGAACAGCGGCGCAACCGACAAGGGCTTTGTAGACATCGAAGAAGCCAACGAA  720
Y  L  A  E  D  E  K  C  H  Y  K  A  Q  N  S  G  A  T  D  K  G  F  V  D  I  E  E  A  N  E     240
GATGACCTTAAGGCTGCAGTGGCCACCGTAGGCCCCGTTTCTATTGCCATTGATGCCAGCCACGAAACCTTCCAACTGTACTCGGATGGA  810
D  D  L  K  A  A  V  A  T  V  G  P  V  S  I  A  I  D  A  S  H  E  T  F  Q  L  Y  S  D  G     270
GTCTACAGTGATCCTGAATGTAGCTCACAAGAACTAGACCATGGGGTGTTGGTAGTGGGATACGGTACCAGCGACGATGGTCAGGACTAC  900
V  Y  S  D  P  E  C  S  S  Q  E  L  D  H  G  V  L  V  V  G  Y  G  T  S  D  D  G  Q  D  Y     300
TGGTTGGTGAAAAATTCGTGGGGACCCAGCTGGGGATTGAACGGATACATCAAGATGGCCAGGAATCAAGATAACATGTGCGGAGTTGCA  990
W  L  V  K  N  S  W  G  P  S  W  G  L  N  G  Y  I  K  M  A  R  N  Q  D  N  M  C  G  V  A     330
TCTCAGGCTAGTTATCCTTTGGTTTAGAGGAAATAGGACAGCGTTTGTGATTTTTATTAGGGTTTAGAATACCTATCTCTAAGAGCTCTC 1080
S  Q  A  S  Y  P  L  V  *                                                                     338
TATTTTTTTTTGTTATCAATTTAGTTTTGGAAGATATTTTATATGATCTTTGTAGAGATTCTTTTTTAAGCTTCCATTAAAAACGACTAA 1170
AATCATTAAATAAGTAAATAATGTATTACATATTTTTTCCTTGCCATTTTCAAGTAATTCAACATTATATATTAGTTACACGAGCGGTAT 1260
                            ↑
CGACAACACGATAATGTGTTTGTGATTTGGATTTTATATTTATTGTTTTTGTTGTATTAGTTTGTTTTCTATGATTCTCATTCTTCCATC 1350
                            C
AAAAATAATAATATATTTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA                                   1406
```

**C**

```
pSCPc1  CAGGCTAGTTATCCTTTGGTTTAGAGGAAATAGGACAGCGTTTGTGATTTTTATTAGGGTTTAGAATACCTATCTCTAAGAGCTCTCTA-
pSCPg2  ■■■..............A...........T.......................................T.....................T
pSCPg3  ■■■..............A...........T.......................................T.....................T

pSCPc1  TTTTTTTTGTTATCAATTTAGTTTTGGAAGATATTTTATATGATCT-TTGTAGAGATTCTTTTTTAAGCTTCCATTAAAAACGACTAAAA
pSCPg2  .........C.................T............G.C....CC..A.........................T..TA.TA.T..
pSCPg3  ...........................T...................-..............................T..TA.TA.T..

pSCPc1  TCATTAAATAAGTAAATAATGTATTACATATTTTTTCCTTGCCATTTTCAAGTAATTCAACATTATATATTAGTTACACGAGCGGTATCG
pSCPg2  .A..A.TTA.TAAC.G.T.CTA.CCTGGAGACGGACTT.ACAATG.AC.GT.GC.AATTGT..GCACA...TA.A.A.GTTTTCCC.AAT
pSCPg3  .A.G.CTT..TAA.C.GTTCA.TAACTCGTAG.A.ATT.ATTAGAAGTA.GAGT.AAATGTC..CAGATCAGCCGCT.AAGAA.AA.C.A

pSCPc1  ACAACACGATAATGTGTTTGTGATTTGGATTTTATATTTATTGTTTTTGTTGTATTAGTTTGTTTTCTATGATTCTCATTCTTCCATCAA
pSCPg2  CTT.A..AC.T.A.C.AGCAACGCA.C.CAGA.CA.CGC.GGACCA.A..AAGTAGGAA.CT.C..TCTCA.GAGA..C.GGC.AGAATC
pSCPg3  .A.CGTG.G....AATGC.AACGAAACA.A.A..GTC.....AGAA.G..CTG..GTCGAAA...CAC.GTTC.TGTCCAG.ATACAATT

pSCPc1  AAAATAATAATATATTTCACAApoly(A)
pSCPg2  ..GG..C..ATA.GACT.G..ATCAACCCAGAAACAACTATTGACATACCGCCACAACCACCGCAGAGCGAAAGCATACCTGAAGTAGAG
pSCPg3  TGTG.TACC....AAAGTTT.TATTGTAAAATGTCTTTACCTCAAATATGAGTTTTAAAGGTTTTTTTTTATTATCACTTATACTTAAAT

pSCPg2  ACAGAGACTATACTACCGACAGAGATTACA-CTACCGGCCGAACAAAACCCACTTGGNTTATCTNTNCGNANTTCAAACCCTAATTNCGT
pSCPg3  ACTTTAATTGGAATAGGTAATGTTTCTATAGTTAAACAAATTCTCAATGACAATAATAATAATAATAATAAAATAAGAAAGGATGGAAAG

pSCPg2  TTTATGGGTCCAATCCCAACCTCAGACCATCACTTCCGANAAATTCCTTCCACAGCACAAACAGNACCAATCATCG-------------
pSCPg3  GGAAGAAGAGGCCAATCAAAGCATTAATAGAATGGAATATAAACAAATAACCATTGACGAAATTTGCGAAACAGCACAGAAAACACACAA
```

Fig. 4. **Nucleotide and deduced amino acid sequences of a major cysteine proteinase (CP) mRNA in adult *Sitophilus zeamais*.** A: Restriction map of cDNA clones obtained by 3′-RACE. First-strand cDNA was synthesized from poly(A)-rich RNA of *S. zeamais*. 3′-RACE was performed using the sense oligonucleotide primers and oligo(dT)$_{18}$ primers shown by arrowheads. The open box indicates the coding region of pSCPc1. Clones obtained by 3′-RACE are shown below the restriction map. Abbreviations of restriction enzymes are the same as in Fig. 1A. The filled triangle in A and the arrow in B indicate a minor poly(A) addition site. B: Nucleotide and deduced amino acid sequences of pSCPc1. Numbering is the same as described in Fig. 1B. The relevant in-frame termination codon in the 5′-noncoding sequence is underlined. Potential poly(A) addition signals are double-underlined. The letters C and T at nucleotide position 1287 are polymorphic residues among the cDNA clones. C: Comparison of the 3′-noncoding nucleotide sequences of a pSCPc1 and two genomic DNA fragments of pSCPg2 and pSCPg3. Nucleotides are aligned from the termination codon of pSCPg2. The termination codon of each clone is shaded. Unidentified nucleotides are indicated by N. Dots indicate nucleotide identity with pSCPc1. The poly(A) addition signals in pSCPc1 are underlined, and potential poly(A) addition signals in pSCPg2 and pSCPg3 are double-underlined.

pSCPg3 in a region of about 150 nucleotides in the 3′-noncoding region, but different in the downstream 200 nucleotides (Fig. 4C).

*Comparison of the Amino Acid Sequences of SCPs and Other Related CPs*—As described above, two genomic DNA clones and a cDNA clone encoding CPs were obtained, and it was evident that the major mRNA species in adult *S. zeamais* is transcribed from *scp1*. The nucleotide sequences of these three clones were very similar; in the coding region, SCPc1 showed 89.4% similarity to SCPg2 and 88.8% to SCPg3, while the similarity between SCPg2 and SCPg3 was as high as 97.6%.

The deduced amino acid sequences of the three SCPs (331 residues in SCPg2 and SCPg3, 338 residues in SCPc1) were also very similar. Both SCPg2 and SCPg3 differed from SCPc1 at 13 amino acid residues, while only four



Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

differences were found between SCPg2 and SCPg3 (Fig. 5). The amino acid sequences of the three SCPs were also similar to those of other known insect CPs and human cathepsin L. SCPc1, for example, showed significant amino acid sequence similarity to various CPs: 66.6% to *Sarcophaga* cathepsin L-like CP (*Sar*CP) (*40*), 66.3% to *Bombyx* CP (BCP) (*41*), 50.4% to human cathepsin L (*49*) in the entire encoded protein, and 72.8% to DCP1 in the putative mature enzyme region (*39*). Although similarities were observed over the entire CPs, significant differences were observed near the COOH termini of SCPg2 and SCPg3; the COOH-terminal sequences of both SCPg2 and SCPg3 were shorter by seven amino acids. SCPc1 thus had a COOH-terminal sequence similar to those of DCP1, BCP, and mammalian cathepsin L, whereas the sequences in SCPg2 and SCPg3 were shorter (Fig. 5).

To confirm the statistical relationship between SCPs and other known proteins, homology search was carried out using the GenBank CDS data base (release 88.0). This yielded many known vertebrate and invertebrate CPs including BCP (*41*), lobster CPs (48.7 to 51.9%) (*50*), and several CPs of plant origin (34.0 to 45.1%) (*13–15*). SCPc1 also showed significant similarity to mammalian cathepsins: 48.6% to human cathepsin S (*11*), 44.6% to human cathepsin K (*12*), and 39% to rat cathepsin H (*9*) in the whole amino acid sequence. In each case, the COOH-terminal sequences of SCPg2 and SCPg3 were the shortest among known CPs and SCPc1.

*Identification of Two Forms of SCP Protein Using a Specific Antibody and an In Vitro Conversion Experiment at Acidic pH*—Next, to investigate the molecular forms and tissue distribution of SCPs, we prepared anti-SCP antibody raised against an SCP fragment corresponding to the putative mature protein region (residues 165 to 331 of SCPg3) expressed in *E. coli*, and used it to detect SCP(s) in immunoblot and immunohistochemical analyses. Since the amino acid sequences of the three SCPs are highly similar to one another and their molecular masses are within 1 kDa, the antibody was expected to detect all SCP isoforms and it might be difficult to discriminate the three or more SCPs in terms of their molecular masses. However, since the major mRNA transcribed in adult is *scp1*, which occupies more than 90% of the 3′-RACE clones, only SCPc1 would be detected if the translational efficiency and rate of degradation were similar.

After affinity purification and absorption to control *E. coli* proteins, the antibody recognized two bands, one at 49 kDa and the other at 37 kDa, in immunoblot analyses of total extracts from larvae, pupae, and adults of *S. zeamais* (Fig. 6A). The antibody also recognized two bands in DCP1 (*39*), which is similar to SCPs in amino acid sequence (Fig. 6C). It was thus possible to examine the tissue distribution

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

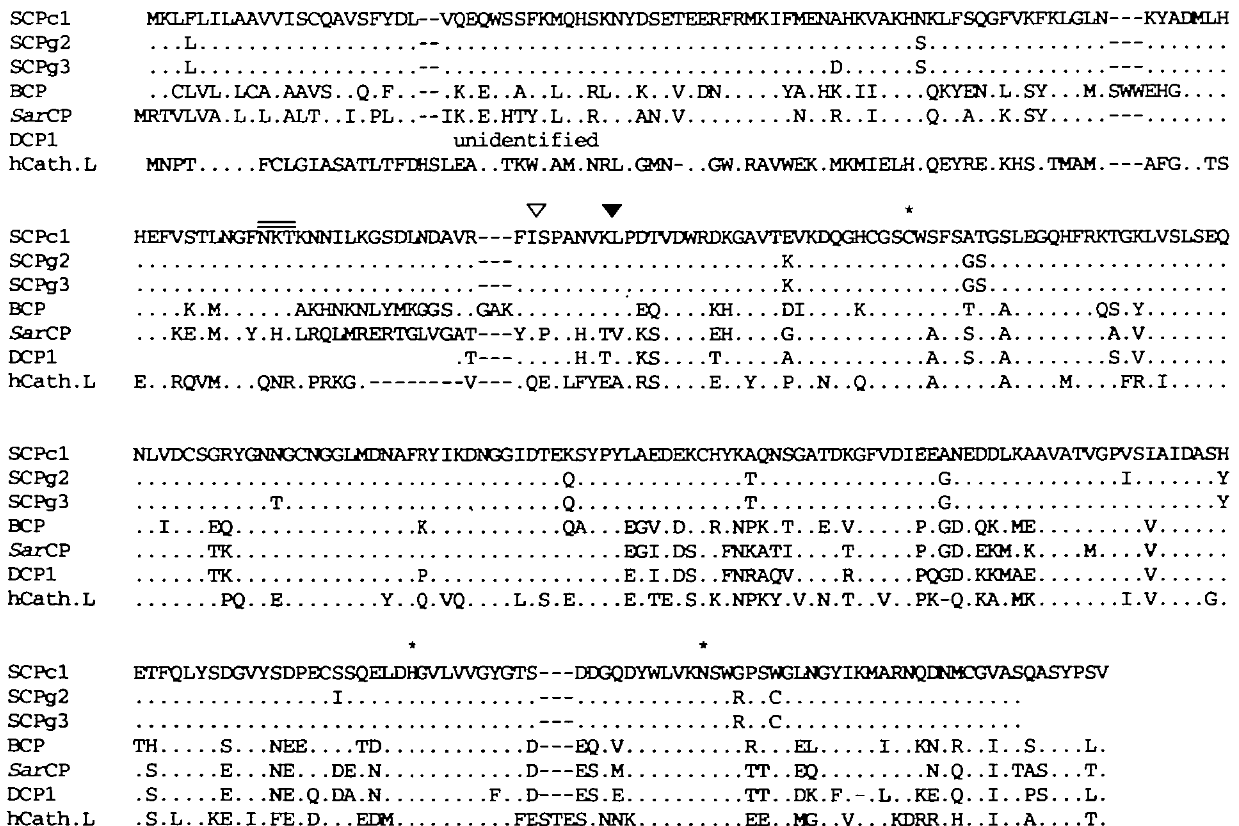


Fig. 5. **Alignment of the amino acid sequences of *S. zeamais* cysteine proteinases (SCPs) and related cysteine proteinases.** Seven sequences are aligned for maximal matching by inserting gaps denoted by hyphens. BCP, *Sar*CP, DCP1, and hCath. L indicate *Bombyx mori* cysteine proteinase (*41*), *Sarcophaga peregrina* cathepsin L-like cysteine proteinase (*40*), *Drosophila melanogaster* cysteine proteinase 1 (*39*), and human cathepsin L (*49*), respectively. Periods denote residues identical to those of SCPc1. The open and closed arrowheads indicate the cleavage sites of the proenzymes of BCP and *Sar*CP, respectively. A potential *N*-glycosylation site is indicated by the double-line. Residues of the catalytic triad are indicated by asterisks.

of DCP1 in *D. melanogaster* as well as SCP in *S. zeamais*. By analogy with BCP (*41*), it was likely that the 49 and 37 kDa bands corresponded to a proenzyme and a mature form of SCP, respectively, and this was confirmed by further experiments as described below. However, the molecular weights estimated from the migration on SDS-PAGE were larger than the calculated sizes of the encoded proteins; the molecular weights of proenzyme and mature enzyme were estimated as 35,867 and 24,261, respectively, if the cleavage sites of the pre- and pro-peptides are the same as those of BCP (*41*). Similar discrepancies have been observed for BCP and *Sar*CP (*40, 41*). This is probably due to the anomalous mobility of SCP on SDS-PAGE; post-translational modifications such as glycosylation might have small effects, because SCP fragments expressed in *E. coli* show the same anomalous mobility on SDS-PAGE (data not shown).

As shown in Fig. 6A, the total larval and adult extracts contained almost equal amounts of 49 and 37 kDa forms. However, the pupal extract contained more 49 kDa form SCP than 37 kDa form.

To investigate precisely the relationship between the two molecular forms of SCP detected here, we prepared a crude SCP fraction and investigated the relationship between the two forms by an *in vitro* conversion experiment. Crude extracts of adult *S. zeamais* obtained as described in "MATERIALS AND METHODS" showed the same SCP forms as total extracts immediately after ammonium sulfate fractionation and dialysis (data not shown). Using this fraction, we tried to separate the 49 kDa form from the 37 kDa form under various conditions. We obtained a crude preparation of the 49 kDa form by storing the extract at 4°C for a few days; this was probably due to the rapid autodegradation of the 37 kDa form, although a small amount of a 47 kDa band was observed (Fig. 6B, lane 1). Using this 49 kDa preparation, we examined the conversion of the 49 kDa form under acidic conditions (pH 3.5). As shown in Fig. 6B, the 49 kDa form was converted to the 37 kDa form in a reaction partially inhibited by cysteine proteinase inhibitors such as E-64 and leupeptin (Fig. 6B, lanes 2–4); the 49 kDa form was converted into a 39 kDa intermediate in the presence of these inhibitors. However, serine proteinase inhibitors, such as PMSF and aprotinin, and an acid proteinase inhibitor, pepstatin, had no effect (data not shown). Pre-treatment of the 49 kDa preparation with E-64 or leupeptin prior to transfer into acidic pH resulted in a complete inhibition of the conversion (Fig. 6B, lanes 5 and 6). These results show that the 49 and 37 kDa forms of SCP are proenzyme and mature enzyme forms, respectively, suggesting that the proteolytic reaction should be catalyzed by a cysteine proteinase present in the preparation, probably by intra- or intermolecular proteolysis by SCP itself. This conversion may occur through multistep reactions producing several intermediates such as 47 and 39 kDa forms, similar to the case of BCP (*51*).

*Tissue Distribution of SCP*—Since we had previouly analyzed the mRNA localization of DCP1 in *D. melanogaster* (*39*) and the anti-SCP antibody recognized DCP1, we first tried to investigate the tissue distribution of the DCP1 protein in *D. melanogaster* to evaluate the specificity of the antibody and to compare the protein distribution with that of the mRNA. As shown in Fig. 7, in a late stage embryo *D. melanogaster* (stage 16), the midgut and garland cells were strongly stained (Fig. 7A); in a 3rd instar larva, the gastric caeca were stained together with garland cells and midgut (Fig. 7B). These observations are essentially consistent with the mRNA distribution reported previously (*39*). However, weak signals were also observed in neighboring tissues such as the proventriculus in the larva and the hindgut in the embryo, which might indicate that the DCP1 protein is secreted into neighboring spaces since those tissues did not show any expression of the DCP1 mRNA (*39*).

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

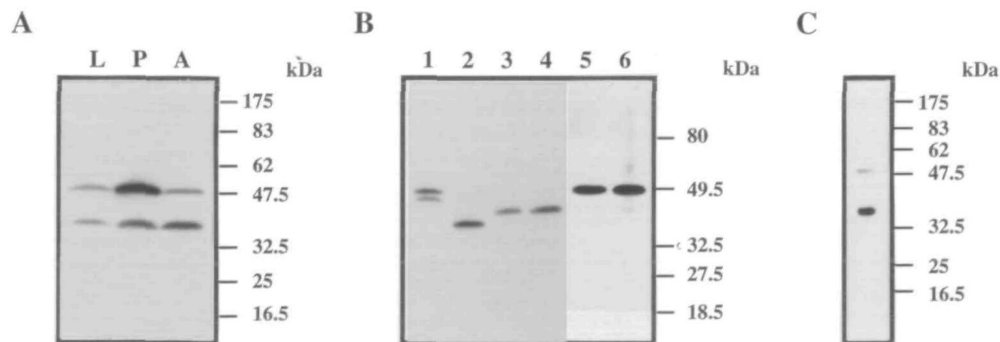


Fig. 6. **Molecular forms of *Sitophilus* cysteine proteinase (SCP) and *in vitro* conversion.** Total proteins of *S. zeamais* at various developmental stages and crude SCP fractions from adult *S. zeamais* were prepared and subjected to SDS-polyacrylamide gel electrophoresis, electroblotted onto polyvinylidene difluoride membranes, and immunoreacted with affinity-purified antiserum against *Sitophilus* cysteine proteinase(s). Signals were detected using peroxidase-conjugated anti-(rabbit IgG) antibody and an ECL kit (Amersham). Prestained molecular weight markers (New England Biolabs or Bio-Rad) are indicated at the right. A: Immunoblot analysis of total *S. zeamais* proteins by anti-SCP antibody. The letters above the lanes indicate the developmental stages as follows: L, larvae; P, pupae; A, adults. B: *In vitro* conversion of the SCP proenzyme into the mature form after incubation at acidic pH. Crude 49 kDa proenzyme fractions were incubated at 37°C for 30 min in 0.1 M sodium formate buffer (pH 3.5) and subjected to immunoblot analysis using affinity-purified anti-SCP antibody. Lane 1, crude proenzyme fraction of adult *S. zeamais*; lane 2, crude proenzyme fraction incubated in 0.1 M sodium formate buffer (pH 3.5) containing 0.2% casein for 30 min at 37°C in the absence of any inhibitors; lane 3, crude proenzyme fraction incubated in the presence of 0.1 mM E-64; lane 4, crude proenzyme fraction incubated in the presence of 0.1 mM leupeptin; lane 5, crude proenzyme fraction pre-incubated with 0.1 mM E-64 for 10 min at 4°C and incubated under the same conditions as lane 2; lane 6, crude proenzyme fraction pre-incubated with 0.1 mM leupeptin for 10 min at 4°C and incubated under the same conditions as lane 2. C: Immunoblot analysis of total proteins of adult *D. melanogaster* by anti-SCP antibody.

Next, the tissue distribution of SCP in *S. zeamais* was surveyed immunohistochemically. Since there is more structural information about adult alimentary organs than those of other growth stages in *S. zeamais*, dissected tissues from adults were used. The presence of SCP was observed in the gastric caeca but not in the midgut (Fig. 7, C and D) (*52*). Furthermore, amorphous cell clusters, probably fat body assemblies in which nutrients are stored and processed, were also strongly stained (Fig. 7E). Thus, it was evident that SCP is contained in the alimentary tissues of *S. zeamais* as in the case of DCP1, although the midgut of *S. zeamais* lacked CP, which was different from the case of DCP1 in *D. melanogaster*.

Since Yamamoto and Takahashi (*53*) reported that BCP is abundant in the ovary of *Bombyx mori*, we also examined these organs for the occurrence of CPs in *D. melanogaster*

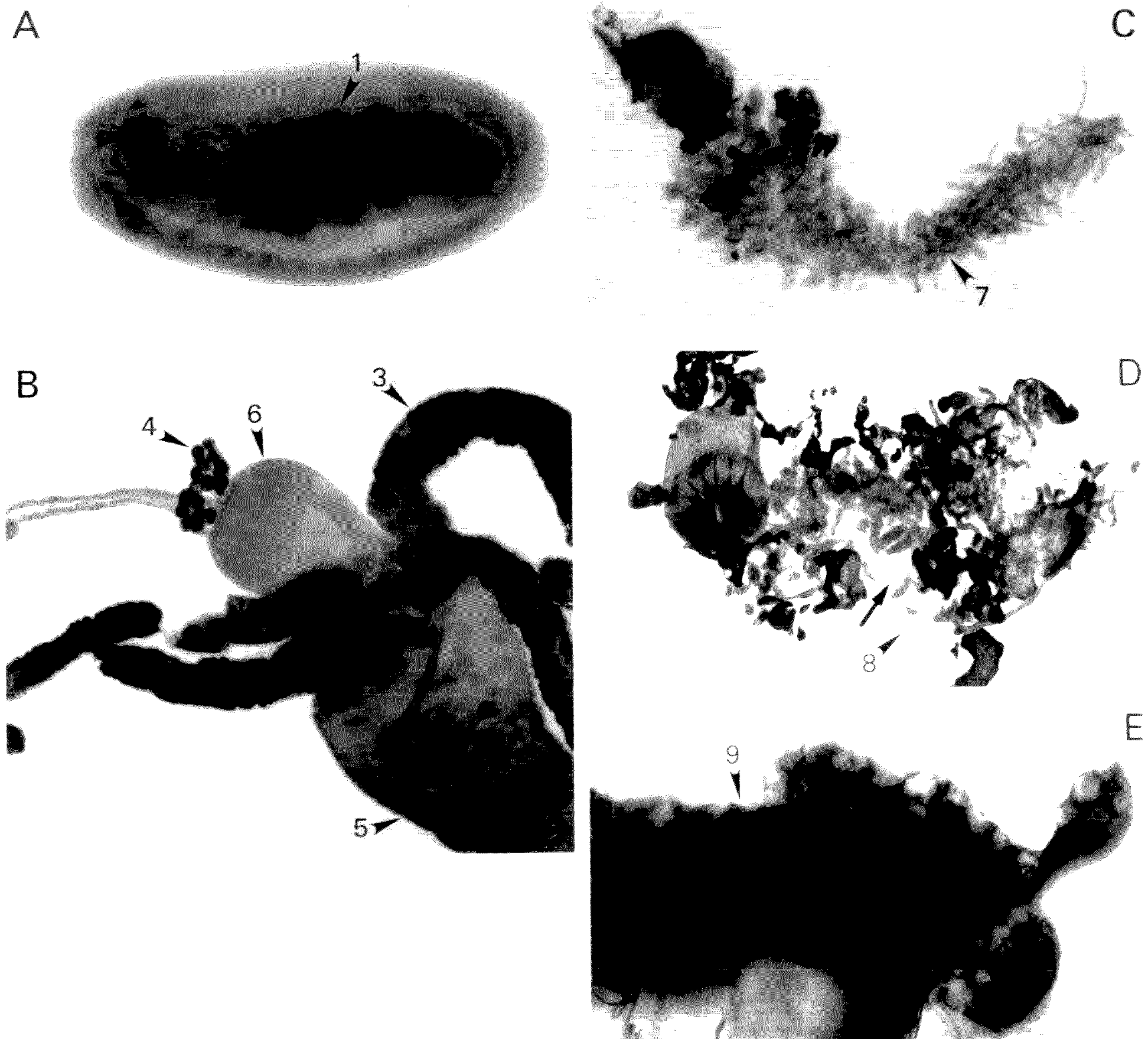


Fig. 7. **Localization of insect cysteine proteinases in alimentary organs.** Dechorionated embryo and dissected tissues from *D. melanogaster* (A, B), and dissected tissues from *S. zeamais* (C-E) were fixed with paraformaldehyde and immunoreacted with purified antiserum against *S. zeamais* cysteine proteinase(s). Signals were visualized using biotin-conjugated anti-(rabbit IgG) antibody and an ABC kit (Vector Lab.). Arrowheads indicate positive signals. A: Whole mount staining of *D. melanogaster* embryo at stage 16. Dark brownish signals are observed in the midgut (arrowhead 1), a winding tubular organ halfway to the posterior, and in garland cells (arrowhead 2) clustered in front of the midgut. B: Larval alimentary organs of *D. melanogaster*. Strong signals are observed in gastric caeca (four diverticula rooted at the junction between proventriculus and midgut, indicated by arrowhead 3), garland cells (clustered cells circling the esophagus in front of the proventriculus, indicated by arrowhead 4), and midgut (arrowhead 5). Weak signals are also observed in an anterior half of the proventriculus (arrowhead 6). C and D: Alimentary tracts of *S. zeamais* (*52*). Adult gastric caeca of *S. zeamais* (arrowheads 7 and 8), multiple short diverticula around the midgut, are stained. No significant signals are detected in the midgut indicated by the arrow. E: Immunostaining of amorphous cell clusters attached to the larval alimentary tract of *S. zeamais*. Probable fat bodies are strongly stained (arrowhead 9). Micrographs were taken at these magnifications: A and E, ×200; B, C, and D, ×100.

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

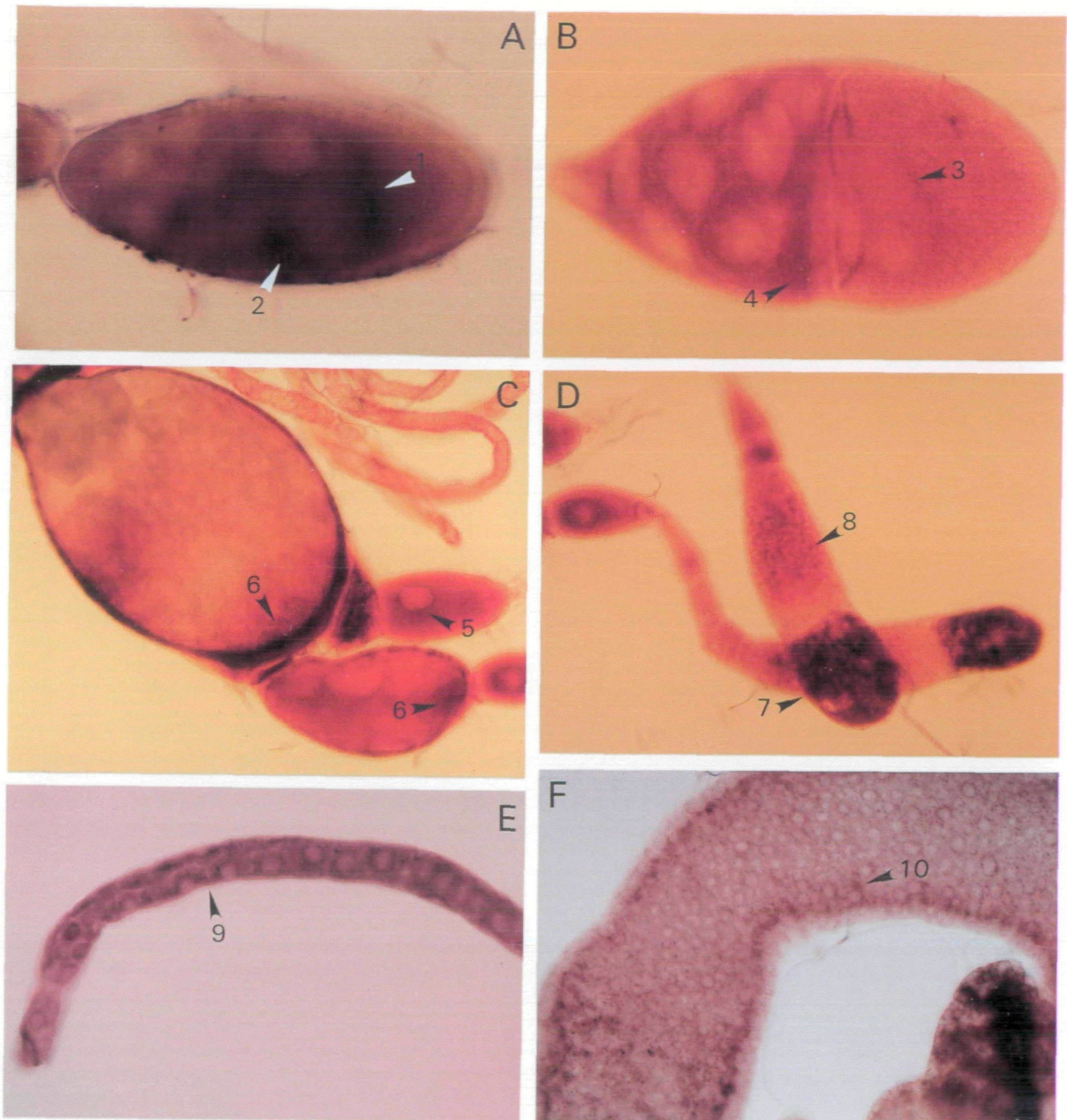


Fig. 8. **Occurrence of cysteine proteinases in ovary as detected by immunohistochemical analyses in *D. melanogaster* and *S. zeamais*.** *In situ* hybridization (A) and immunostaining (B-G) were carried out using insect ovaries as described in "MATERIALS AND METHODS." Arrowheads indicate positive signals. A: *In situ* hybridization of *Drosophila* ovary. Both oocyte (arrowhead 1, whitened) and nurse cells (arrowhead 2, whitened) are stained, but follicle cells surrounding the oocyte show no significant signals. B: Immunostaining of *Drosophila* ovary. Signals are observed in both oocyte (arrowhead 3) and nurse cells (arrowhead 4), but not in follicle cells as in A. C: Immunostaining of later stages of *S. zeamais* ovary (*54*). A region surrounding the karyosome of an immature oocyte (arrowhead 5) and the surface adjacent to a premature oocyte of a mature oocyte (arrowheads 6), especially the borders with the neighboring immature oocyte, are strongly stained. D: Immunostaining of early stages of *S. zeamais* ovary. The tropharium region is stained, and two distinct regions in terms of signal intensity are observed: strong signals halfway to the terminal filament (arrowhead 7) and weaker signals in other regions (arrowhead 8). E and F: Immunostaining of other tissues of *S. zeamais*. Cytoplasm in the cells of a long slender tubular organ of adult (arrowhead 9), probably a malpighian tubule, is stained (E). A large tubular organ in larva (arrowhead 10), probably the alimentary tube, is stained (F). Micrographs were taken at these magnifications: A, B, E, and F, ×200; C and D, ×100.

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

and *S. zeamais*. In the case of *D. melanogaster*, DCP1 was detected in the germline cells of the ovary, nurse cells and oocyte, but not in somatic follicle cells, as also found for the distribution of the mRNA (Fig. 8, A and B). In *S. zeamais*, strong signals were observed in both immature and mature oocytes; significant signals were observed around the karyosome of immature oocytes, and in mature oocytes and the tropharium region (Fig. 8, C and D) (*54*). With regard to the intracellular localization of SCP in oocytes, strong signals were found on the surface of the oocytes and at the boundary with neighboring oocytes. In the tropharium region, positive signals appeared to originate from nurse cells, and not from interstitial cells belonging to somatic cells (Fig. 8C). The end of the tropharium, which borders the terminal filament and contains germ cells in the initial developmental stages (*54*), was strongly stained (Fig. 8D).

The occurrence of SCP was also observed in some other tissues. A long slender tubular organ, probably the malpighian tubule, and larval large tubular organs were significantly stained (Fig. 8, E and F), although the anatomical information about these cells and tissues is so poor that definite identification remains impossible.

## DISCUSSION

In this study, we have identified and characterized a CP gene family in *Sitophilus*, an insect pest that inflicts damage on corn and rice seeds during storage. The three CP sequences we determined showed high similarity to one another and to other known CPs of insect and mammalian origin.

The gene family appears to consist of four to six related genes, among which a single gene, termed *scp*1, is transcribed into a major mRNA species in adults. The other genes might be transcribed under different transcriptional control, although we were not able to detect their mRNAs in adults by 3′-RACE experiments. Of course, it is possible to recognize SCP genes other than *scp*1 as pseudogenes. However, at least two genes, termed *scp*2 and *scp*3, which we analyzed at the nucleotide sequence level, contain intact coding sequences that are not interrupted by termination codons and do not contain frame-shifts. In addition, the 5′- and 3′-noncoding sequences of *scp*2 and *scp*3, probably important for basal transcriptional and/or translational regulation, are similar to those of *scp*1. From these observations, we think it likely that *scp*2 and *scp*3 are transcribed into minor mRNAs existing transiently during development and/or in limited cell populations. Our speculation is also consistent with the structural feature that the far 5′-upstream sequences of *scp*2 and *scp*3 are totally different beyond about 250 nucleotides from the initiation codons and might contain enhancer sequences and be responsible for differential expression. This possibility can be analyzed in the future by establishing a method for obtaining large numbers of embryos, larvae, and pupae and by using specific probes for each gene, or by enhancer assays using genomic upstream sequences.

Concerning the evolution of this gene family, the most intriguing feature is the very high similarity of the three *scp* genes sequenced as shown in Figs. 1 and 4. Additionally, Southern blot analyses indicate that the other one to three *scp* genes would also have very similar sequences, because nearly equal signals were observed under stringent conditions (Fig. 2). Although the physiological significance of this multiplicity has not yet been clarified, it should have originated by gene amplification and subsequent conservation of at least the coding regions and near borders during the evolution of Coleoptera or *Sitophilus*. In this respect, it is very interesting that genomic Southern blot analysis using *D. melanogaster* DNA showed the DCP1 gene to be essentially a single copy gene (unpublished data). This suggests that the amplification of the CP gene occurred after the divergence of Coleoptera and Diptera. Further phylogenic studies on the copy number of cathepsin L-like CP genes using various insects, together with comparative structural, biochemical, and physiological studies of insect CPs, may confirm this speculation.

Immunohistochemical analyses demonstrated two noteworthy features common to insect CPs. One is that both SCP and DCP1 occur in alimentary organs. As described in "RESULTS," there are several minor differences in the tissue distributions of SCP and DCP1, but both CPs are present in abundance in alimentary organs, especially in the tissues involved in digestive functions such as gastric caeca. Considering that biological observations have shown that CPs as well as serine and aspartic proteinases play roles in the digestive processes of insects including Coleoptera and Diptera (*55, 56*), it appears likely that SCP is a digestive proteinase. This is consistent with our previous report that gut extracts of *S. zeamais* contain a proteolytic activity similar in biochemical characteristics to cathepsin L (*57*). The other significant feature of insect CPs is that SCP and DCP1 occur in germline cells of the ovary, nurse cells and oocytes. A similar observation has been made for BCP (*53*), although BCP also occurs in somatic cells of the ovary (follicle cells) where no significant immunostaining was observed for SCP or DCP1. In any case, mature oocytes of these insects contain significant amounts of CPs. This suggests that insect CPs are commonly involved in proteolytic reactions during oocyte maturation and/or early development after fertilization.

On the other hand, immunoblot analysis of *S. zeamais* extracts using anti-SCP antibody also revealed another informative feature of SCP. Molecular forms of SCP observed in *S. zeamais* extracts showed that SCP exists as both a proenzyme and a mature enzyme, and our *in vitro* study showed that the proenzyme is converted into the mature enzyme under acidic conditions. The *in vivo* existence of two molecular forms and their processing at acidic pH may reflect the secretion of the SCP protein and the regulation of SCP activity. For example, it is speculated that SCP is synthesized and stored as a proenzyme at neutral pH in some organs such as gastric caeca, and the proenzyme is converted into the active form after secretion into neighboring places that require SCP activity, such as the digestive fluid of the foregut where the pH is possibly acidic. This hypothesis is consistent with the observation that in Coleoptera, especially in curculionidae, the pH of the digestive fluid of the anterior midgut containing the gastric caeca is neutral and that of the foregut region is somewhat acidic (*56*). In this regard, it would be interesting to prepare an antibody specific for the proenzyme form and to use it as an immunohistochemical probe.

Other functions are also expected from the tissue distribution of CPs. The probable malpighian tubule of *Sitophilus* as well as garland cells and macrophages of *Droso-*

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

*phila* contain significant amounts of CPs, suggesting that insect CPs may be involved in cleansing processes as previously indicated (*39*).

Taken together, it is likely that *Sitophilus* CP acts as a secreted or lysozomal proteinase and has various functions depending on the type of cell and tissue. These include a genital function in the ovary, digestion in alimentary organs, and cleansing in the malpighian tubule. Additional physiological and biochemical analyses will be needed to confirm the functions and enzymatic profiles of SCPs.

We thank Dr. Hiroshi Fujii, National Food Research Institute, for providing the insect *Sitophilus zeamais*.

REFERENCES

1. Katunuma, N. and Kominami, E. (1983) Structures and functions of lysosomal thiol proteinases and their endogenous inhibitor. *Curr. Top. Cell. Regul.* **22**, 71-101
2. Bond, J.S. and Butler, P.E. (1987) Intracellular proteases. *Annu. Rev. Biochem.* **56**, 333-364
3. Yuan, J., Shaham, S., Ledoux, S., Ellis, H.M., and Horvitz, H.R. (1993) The C. elegans cell death gene *ced*-3 encodes a protein similar to mammalian interleukin-1$\beta$-converting enzyme. *Cell* **75**, 641-652
4. Miura, M., Zhu, H., Rotello, R., Hartwieg, E.A., and Yuan, J. (1993) Induction of apoptosis in fibroblasts by IL-1$\beta$-converting enzyme, a mammalian homolog of the C. elegans cell death gene *ced*-3. *Cell* **75**, 653-660
5. Kumar, S. (1995) ICE-like proteases in apoptosis. *Trends Biochem. Sci.* **20**, 198-202
6. Hara-Nishimura, I., Shimada, T., Hiraiwa, N., and Nishimura, M. (1995) Vacuolar processing enzyme responsible for maturation of seed proteins. *J. Plant Physiol.* **145**, 632-640
7. Berti, P.J. and Storer, A.C. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases. *J. Mol. Biol.* **246**, 273-283
8. Segundo, B.S., Chan, S.J., and Steiner, D.F. (1985) Identification of cDNA clones encoding a precursor of rat liver cathepsin B. *Proc. Natl. Acad. Sci. USA* **82**, 2320-2324
9. Ishidoh, K., Imajoh, S., Emori, Y., Ohno, S., Kawasaki, H., Minami, Y., Kominami, E., Katunuma, N., and Suzuki, K. (1987) Molecular cloning and sequencing of cDNA for rat cathepsin H. *FEBS Lett.* **226**, 33-37
10. Ishidoh, K., Towatari, T., Imajoh, S., Kawasaki, H., Kominami, E., Katunuma, N., and Suzuki, K. (1987) Molecular cloning and sequencing of cDNA for rat cathepsin L. *FEBS Lett.* **223**, 69-73
11. Petanceska, S. and Devi, L. (1992) Sequence analysis, tissue distribution, and expression of rat cathepsin S. *J. Biol. Chem.* **267**, 26038-26043
12. Tezuka, K., Tezuka, Y., Maejima, A., Sato, T., Nemoto, K., Kamioka, H., Hakeda, Y., and Kumegawa, M. (1994) Molecular cloning of a possible cysteine proteinase predominantly expressed in osteoclasts. *J. Biol. Chem.* **269**, 1106-1109
13. Ishidoh, K., Muno, D., Sato, N., and Kominami, E. (1991) Molecular cloning of cDNA for rat cathepsin C. *J. Biol. Chem.* **266**, 16312-16317
14. Cohen, L.W., Coghlan, V.M., and Dihel, L.C. (1986) Cloning and sequencing of papain-encoding cDNA. *Gene* **48**, 219-227
15. Rogers, J.C., Dean, D., and Heck, G.R. (1985) Aleurain: a barley thiol protease closely related to mammalian cathepsin H. *Proc. Natl. Acad. Sci. USA* **82**, 6512-6516
16. Watanabe, H., Abe, K., Emori, Y., Hosoyama, H., and Arai, S. (1991) Molecular cloning and gebberellin-induced expression of multiple cysteine proteinases of rice seeds (Oryzains). *J. Biol. Chem.* **266**, 16897-16902
17. Williams, J.G., North, M.J., and Mahbubani, H. (1985) A developmentally regulated cysteine proteinase in *Dictyostelium discoideum*. *EMBO J.* **4**, 999-1006
18. Campetella, O., Henriksson, J., Aslund, L., Frasch, A.C.C., Pettersson, U., and Cazzulo, J.J. (1992) The major cysteine proteinase (cruzipain) from *Trypanosoma cruzi* is encoded by multiple polymorphic tandemly organized gene located on different chromosomes. *Mol. Biochem. Parasitol.* **50**, 225-234
19. Mason, R.W., Gal, S., and Gottesman, M.M. (1987) The identification of the major excreted protein (MEP) from a transformed mouse fibroblast cell line as a catalytically active precursor form of cathepsin L. *Biochem. J.* **248**, 449-454
20. Chauhan, S.S., Goldstein, L.J., and Gottesman, M.M. (1991) Expression of cathepsin L in human tumors. *Cancer Res.* **51**, 1478-1481
21. Tagami, K., Kakegawa, H., Kamioka, H., Sumitani, K., Kawata, T., Lenercic, B., Turk, V., and Katunuma, N. (1994) The mechanism and regulation of procathepsin L secretion from osteoclasts in bone resorption. *FEBS Lett.* **342**, 308-312
22. Golde, T.E., Estus, S., Younkin, L.H., Selkoe, D.J., and Younkin, S.G. (1992) Processing of the amyloid protein precursor to potentially amyloidogenic derivatives. *Science* **255**, 728-730
23. Turk, V. and Bode, W. (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett.* **285**, 213-219
24. Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., and Turk, V. (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* **7**, 2593-2599
25. Stubbs, M.T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., and Turk, V. (1990) The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J.* **9**, 1939-1947
26. Abe, K., Emori, Y., Kondo, H., Suzuki, K., and Arai, S. (1987) Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). *J. Biol. Chem.* **262**, 16793-16797
27. Abe, K., Emori, Y., Kondo, H., Arai, S., and Suzuki, K. (1988) The $NH_2$-terminal 21 amino acid residues are not essential for the papain-inhibitory activity of oryzacystatin, a member of the cystatin superfamily. *J. Biol. Chem.* **263**, 7655-7659
28. Kondo, H., Emori, Y., Abe, K., Suzuki, K., and Arai, S. (1989) Cloning and sequence analysis of the genomic DNA fragment encoding oryzacystatin. *Gene* **81**, 259-265
29. Kondo, H., Abe, K., Nishimura, I., Watanabe, H., Emori, Y., and Arai, S. (1990) Two distinct cystatin species in rice seeds with different specificities against cysteine proteinases. *J. Biol. Chem.* **265**, 15832-15837
30. Abe, K., Kondo, H., Watanabe, H., Emori, Y., and Arai, S. (1991) Oryzacystatins as the first well-defined cystatins of plant origin and their target proteinases in rice seeds. *Biomed. Biochim. Acta* **50**, 637-641
31. Abe, M., Abe, K., Kuroda, M., and Arai, S. (1992) Corn kernel cysteine proteinase inhibitor as a novel cystatin superfamily member of plant origin. *Eur. J. Biochem.* **209**, 933-937
32. Fernandes, K.V.S., Sabelli, P.A., Barratt, D.H.P., Richardson, M., Xavier-Filho, J., and Shewry, P.R. (1993) The resistance of cowpea seeds to bruchid beetles is not related to level of cysteine proteinase inhibitors. *Plant Mol. Biol.* **23**, 215-219
33. Waldron, C., Wegrich, L.M., Merlo, P.A.O., and Walsh, T.A. (1993) Characterization of a genomic sequence coding for potato multicystatin, an eight-domain cysteine proteinase inhibitor. *Plant Mol. Biol.* **23**, 801-812
34. Kondo, H., Ijiri, S., Abe, K., Maeda, H., and Arai, S. (1992) Inhibitory effect of oryzacystatins and a truncation mutant on the replication of poliovirus in infected Vero cells. *FEBS Lett.* **299**, 48-50
35. Aoki, H., Akaike, T., Abe, K., Kuroda, M., Arai, S., Okumura, R., Negi, A., and Maeda, H. (1995) Antiviral effect of oryzacystatin, a proteinase inhibitor in rice, against herpes simplex virus type 1 in vitro and in vivo. *Antimicrob. Agents Chemother.* **39**, 846-849
36. Kuroda, M., Ishimoto, M., Suzuki, K., Kondo, H., Abe, K., Kitamura, K., and Arai, S. (1996) Oryzacystatins exhibit growth-inhibitory and lethal effects on different species of bean insect pests, *Callosobruchus chinensis* (Coleoptera) and *Riptortus*

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012

*clavatus* (Hemiptera). *Biosci. Biotech. Biochem.* **60**, 209-212
37. Murdock, L.L., Brookhart, G., Dunn, P.E., Foard, D.E., Kelley, S., Kitch, L., Shade, R.E., Shukle, R.H., and Wolfson, J.L. (1987) Cysteine digestive proteinases in coleoptera. *Comp. Biochem. Physiol.* **87B**, 783-787
38. Oppert, B., Morgan, T.D., Culbertson, C., and Kramer, K.J. (1993) Dietary mixtures of cysteine and serine proteinase inhibitors exhibit synergistic toxicity toward the red flour beetle, *Tribolium castaneum*. *Comp. Biochem. Physiol.* **105C**, 379-385
39. Matsumoto, I., Watanabe, H., Abe, K., Arai, S., and Emori, Y. (1995) A putative digestive cysteine proteinase from *Drosophila melanogaster* is predominantly expressed in the embryonic and larval midgut. *Eur. J. Biochem.* **227**, 582-587
40. Homma, K., Kurata, S., and Natori, S. (1994) Purification, characterization, and cDNA cloning of procathepsin L from the culture medium of NIH-Sape-4, an embryonic cell line of *Sarcophaga peregrina* (flesh fly), and its involvement in the differentiation of imaginal discs. *J. Biol. Chem.* **269**, 15258-15264
41. Yamamoto, Y., Takimoto, K., Izumi, S., Toriyama-Sakurai, M., Kageyama, T., and Takahashi, S.Y. (1994) Molecular cloning and sequencing of cDNA that encodes cysteine proteinase in the eggs of the silkmoth, *Bombyx mori*. *J. Biochem.* **116**, 1330-1335
42. Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503-517
43. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
44. Frohman, M.A., Dush, M.K., and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002
45. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
46. Harlow, E. and Lane, D. (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York
47. Emori, Y. and Saigo, K. (1994) Calpain localization changes in coordination with actin-related cytoskeletal changes during early embryonic development of *Drosophila*. *J. Biol. Chem.* **269**, 25137-25142
48. Nakano, H., Yamazaki, T., Ikeda, M., Masai, H., Miyatake, S., and Saito, T. (1994) Purification of glutathione S-transferase fusion proteins as a non-degraded form by using a protease-negative *E. coli* strain, AD202. *Nucleic Acids Res.* **22**, 543-544
49. Gal, S. and Gottesman, M.M. (1988) Isolation and sequence of a cDNA for human pro-(cathepsin L). *Biochem. J.* **253**, 303-306
50. Laycock, M.V., MacKay, R.M., Fruscio, M.D., and Gallant, J.W. (1991) Molecular cloning of three cDNAs that encode cysteine proteinases in the digestive gland of the American lobster (*Homarus americanus*). *FEBS Lett.* **292**, 115-120
51. Takahashi, S.Y., Yamamoto, Y., Shionoya, Y., and Kageyama, T. (1993) Cysteine proteinase from the eggs of the silkmoth, *Bombyx mori*: identification of a latent enzyme and characterization of activation and proteolytic processing *in vivo* and *in vitro*. *J. Biochem.* **114**, 267-272
52. Baker, J. E., Woo, S.M., and Byrd, R.V. (1984) Ultrastructural features of the gut of *Sitophilus granarius* (L.) (Coleoptera: Curculionidae) with notes on distribution of proteinases and amylases in crop and midgut. *Can. J. Zool.* **62**, 1251-1259
53. Yamamoto, Y. and Takahashi, S.Y. (1993) Cysteine proteinase from *Bombyx* eggs: role in programmed degradation of yolk proteins during embryogenesis. *Comp. Biochem. Physiol.* **106B**, 35-45
54. King, R.C. and Büning, J. (1985) The origin and functioning of insect oocytes and nurse cells in *Comprehensive Insect Physiology, Biochemistry, and Pharmacology, Vol. 1. Embryogenesis and Reproduction* (Kerkut, G.A. and Gilbert, L.I., eds.) pp. 37-82, Pergamon Press, Headington Hill Hall, Oxford
55. Liang, C., Brookhart, G., Feng, G.H., Reeck, G.R., and Kramer, K.J. (1991) Inhibition of digestive proteinases of stored grain coleoptera by oryzacystatin, a cysteine proteinase inhibitor from rice seed. *FEBS Lett.* **278**, 139-142
56. Terra, W.R. and Ferreira, C. (1994) Insect digestive enzymes: properties, compartmentalization and function. *Comp. Biochem. Physiol.* **109B**, 1-62
57. Irie, K., Hosoyama, H., Takeuchi, T., Iwabuchi, K., Watanabe, H., Abe, M., Abe, K., and Arai, S. (1996) Transgenic rice established to express corn cystatin exhibits strong inhibitory activity against insect gut proteinases. *Plant Mol. Biol.* **30**, 149-157

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012