รายงานวิจัยฉบับสมบูรณ์

โครงการ การสร้างรีคอมบิแนนท์คาเทปซินแอล จาก *Paragonimus pseudoheterotremus* เพื่อพัฒนาการตรวจวินิจฉัยโรคพยาธิใบไม้ปอด

โดย นางทิพยรัตน์ อยู่นวล

ธันวาคม 2558

สัญญาเลขที่ TRG5480011

รายงานวิจัยฉบับสมบูรณ์

โครงการ การสร้าง รีคอมบิแนนท์คาเทปซินแอล จาก *Paragonimus pseudoheterotremus* เพื่อพัฒนาการตรวจวินิจฉัยโรคพยาธิใบไม้ปอด

นางทิพยรัตน์ อยู่นวล
คณะเวชศาสตร์เขตร้อน มหาวิทยาลัยมหิดล

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย
และมหาวิทยาลัยมหิดล

(ความเห็นในรายงานนี้เป็นของผู้วิจัย สกว.ไม่จำเป็นต้องเห็นด้วยเสมอไป)

Abstract

Project Code : TRG5480011

Project Title : Production of Recombinant Cathepsin L from *Paragonimus pseudoheterotremus* for diagnostic development of paragonimiasis

Investigator : Tippayarat Yoonuan

Faculty of Tropical Medicine, Mahidol University

E-mail Address : tippayarat.yoo@mahidol.edu

Project Period : 2 yers

Cathepsin L is cysteine protease that belongs to papain family. In parasitic trematodes, cathepsin L plays indispensible roles in parasite survival and host-parasite interaction. In this study, cathepsin L of lung fluke *Paragonimus pseudohetrotremus* (PpsCatL) was pioneer identified and characterized its molecular biological and immunological characteristics. The sequence analysis of PpsCatL demonstrated that this gene encodes 325 amino acid residues, which is the most similar to *P. westermani* cathepsin L. *In silico* 3D structure suggested that PpsCatL is pro-enzyme that becomes active when cleavage the pro-peptide. The recombinant pro-PpsCatL without signal peptide (rPpsCatL) was expressed in *E. coli* at the molecular weight of 35 kDa and reacted with *P. pseudohetrotremus*-infected rat sera. The native protein was detected in crude worm antigen, and excretory-secretory product and localized in the caecum with accumulation at lamellae along intestinal tract of adult parasite. Evaluation of immunodiagnostic potential using ELISA suggested

that rPpsCatL could detect paragonimiasis at high sensitivity and specificity (100% and 95.6%, respectively), which supports further development of rPpsCatL-ELISA as immunodiagnostic tool.

Keywords : *Paragonimus pseudohetrotremus*, Cathepsin L, recombinant protein, caecum, Enzyme-linked Immunosorbant Assay (ELISA)

<u>เอกสารแนบหมายเลข 2</u>

บทคัดย่อ

รหัสโครงการ: TRG5480011

ชื่อโครงการ: การสร้างรีคอมบิแนนท์คาเทปซิน แอล จาก *Paragonimus pseudoheterotremus* เพื่อพัฒนาการตรวจวินิจฉัยโรคพยาธิใบไม้ปอด

ชื่อนักวิจัย: นางทิพยรัตน์ อยู่นวล

คณะเวชศาสตร์เขตร้อน มหาวิทยาลัยมหิดล

E-mail Address : tippayarat.yoo@mahidol.edu

ระยะเวลาโครงการ: 2 ปี

คำหลัก : *Paragonimus pseudohetrotremus*, Cathepsin L, recombinant protein, caecum, Enzyme-linked Immunosorbant Assay (ELISA)

Introduction

Lung flukes in the genus *Paragonimus* spp. are causative agents for paragonimiasis, which are endemic in Asia, Africa and America (Waikagul 2007). Recently, only 7 from more than 40 species have been reported to infect human including *P. africanus*, *P. kelikotti*, *P. mexicanus*, *P. westermani*, *P. uterobilateralis*, *P. hetrotremus* and *P. pseudoheterotremus* (Intapan et al. 2012). *P. pseudohetrotremus* was firstly identified in freshwater crabs, *Larnaudia larnaudii*, from Kanchanaburi Province, Thailand (Waikagul 2007). Afterward, the first human case was reported in Thai male who admitted in the hospital with TB-like symptoms (Intapan et al. 2012). As a novel human-infected species, fundamental information especially protein database still lack and require intensive identification to elucidate the factor associated with biology and pathogenesis. In this regards, proteolytic enzyme namely cathepsin L of *P. pseudohetrotremus* (PpsCatL) was identified and characterized here the molecular biological and immunological aspects.

Cathepsin L is cysteine protease that plays important roles in entire life cycle of several trematodes (Collins et al. 2004; Dalton et al. 1996; Day et al. 1995; Park et al. 2002). In *P. westermani*, variety cathepsin L isoforms have been reported, which impact on parasite's survival and virulence (Park et al. 2002). Cathepsin Ls of *F. hepatica* and *F. gigantica* are major secretory proteases that contribute to development of immunodiagnosis (Cornelissen et al. 2001; Varghese et al. 2012) and vaccine (Villa-Mancera et al. 2014; Sansri et al. 2015). Moreover, evolutionary study of *Fasciola* cathepsin L could suggest diverged time that separated *F. hepatica* from *F. gigantica* as well as their substrate specificity (Irving et al. 2003). Using cathepsin L as a drug target was demonstrated in human African trypanosomiasis (Steverding et al. 2012). Regarding information

above, identifying and characterizing cathepsin L of *P. pseudohetrotremus* would advantage for understanding and development of muti-strategies to prevention and control this parasitic helminth.

Herein, cathepsin L-encoding gene of *P. pseudoheterotremus* was amplified from total RNA using RACE-PCR and then analyzed molecular biological features including sequence properties and *in silico* 3D structure. Transcription level of PpsCatL in different developmental stages was determined using real-time PCR and confirmed by western blot analysis. Recombinant pro-PpsCatL (rPpsCatL) was expressed in *E. coli* and used for determination of immune response. Specific location of PpsCatL was analyzed in the parasite tissue using immunolocalization. To evaluate rPpsCatL for development of immunodiagnosis, reacting this protein with paragonimiasis- and other heterologous infected human sera was performed.

Materials and methods

Ethics statement

All animal work was conducted with the approval of the Faculty of Tropical Medicine-Animal Care and Use Committee (No. FTM-ACUC 011/2012). Left-over human sera with helminthic infection and healthy used in this study were permitted by Immunodiagnosis for Helminthiasis Unit, Department of Helminthology, Faculty of Tropical Medicine under approval of The Ethics Committee of Faculty of Tropical Medicine, Mahidol University (No. MUTM 2011-056-01).

Parasite

The natural waterfall crabs (*Demanietta* sp.) collected from Tak province, Thailand were subjected to examine *P. pseudoheterotremus* infection by tissue compression technique (Sugiyama et al. 2004). To collect the metacercariae, body of *P. pseudoheterotremus*-positive crabs were separated into small pieces, homogenized in 0.85% NSS and settled in sedimental jar with several changes of NSS, respectively (Sugiyama et al. 2004). Metacercariae of *P. pseudoheterotremus* in the sediment were collected under a stereoscopic microscope (Olympus, Japan) and used for infection of definitive host. Five 8-week-old female Wistar rats purchased from National Laboratory Animal Center, Mahidol University were orally infected with five metacercariae each and then housed them at the Animal Care Unit, Faculty of Tropical Medicine, Mahidol University. After 8 week post-infection (wpi), fecal examination using simple direct smear was performed and eggs-positive rats were euthanized to obtain the adult worms from the respiratory tract. Fresh recovered worms were washed thoroughly with 0.85% NSS and stored at -80°C for further investigation. For preparation of newly excysted juvenile worms (NEJ), metacercariae were washed thoroughly with 1XPBS and then digested with 1%pepsin-HCl solution for 2 h at 37 ^{O}C. After several washing with 1XPBS, digested metacercariae were incubated with complete medium [RPMI-1640 (Gibco^{TM} Thermo Fisher Scientific, Watham, MA) supplemented with 10% Fetal Bovine Serum (FBS; Biowest SAS, Nuaillé, France) and 1X penicillin-streptomycin solution (Biowest SAS)] at 37°C for 6 h. Emerging NEJ were collected and additionally incubated in complete medium at 37°C for 3 h. Metacercaria, NEJ and adult *P. pseudoheterotremus* were stored at -80°C for further investigation.

Amplification of full-length PpsCatL cDNA

Degenerated primers of PpsCatL were designed by comparing amino acid sequences of cathepsin L orthologs using Clustal Omega program (Goujon et al. 2010). Two highly amino acid conserved motifs with the lowest degree of degeneration, GRCWAF and WGTRWG, were selected and subsequently converted to potential nucleic codons. The degenerated primers were 5'-GGI WSI TGY TGG GCI TTY-3' for forward (Fwd) and 5'-ICC CCA ICG IGT ICC CCA-3' for reverse (Rev). I, S, W and Y stand for inosine, C or G, T or A and T or C, respectively.

Total RNA of adult *P. pseudoheterotremus* was isolated using TriZol reagent (Invitrogen, Carlsbad, CA) according to manufacturer's instructions. Five microgram of total RNA was treated with 1 U of DNase I (Thermo Fisher Scientific) to eliminate genomic DNA and then converted to 1st strand cDNA using RevertAid 1st strand cDNA construction kit (Thermo Fisher Scientific). Partial cDNA region of PpsCatL was amplified by PCR in a total volume of 50 μl containing 1X Taq polymerase buffer 2 mM $MgCl_2$, 0.2 mM of each dNTP, 200 nM of each degenerated primer, and 1 U Taq polymerase (Thermo Fisher Scientific). Amplification was performed under the condition of 95 $^{\circ}$C for 5 min, 40 cycles of 95 $^{\circ}$C for 30 sec, 45 $^{\circ}$C for 30 sec, 72 $^{\circ}$C for 30 sec and final step at 72 $^{\circ}$C for 5 min. The partial cDNA sequence was subcloned into pTZ57R/T plasmid (Thermo Fisher Scientific) and then analyzed the DNA sequence by DNA sequencing (AITbiotech Pte Ltd., Singapore). The obtained sequence was used for design primers using primer 3 program (ref) to perform 5'- and 3'-Rapid Amplification of cDNA Ends (RACE)-PCR using SMARTer^{TM} RACE cDNA Amplification Kit (Clontech Laboratories, Inc., Mountain View, CA) according to manufacturer's instruction. The 3' RACE- and the 5'RACE primers were 5'-GTTTGGAATCCGAAGCTGACTATCC-

3’ and 5’-CGTGCTCCTCCTCATAAGCTCCTA-3’, respectively. The PCR products were subcloned into pTZ57R/T plasmid and then sequenced (AITbiotech Pte Ltd.).

Bioinformatics analysis

The nucleotide and predicted amino acid sequences submitted to NCBI database with an assigned accession number (-----) were used for bioinformatic analysis. The protein properties including molecular weight, isoelectric point, signal peptides and transmembrane helices were predicted using Pepstats (Rice et al. 2000), SignalP 4.1 Server (Petersen et al. 2011) and TMHMM (Krogh et al. 2001), respectively. Moreover, N- and O-glycosylation sites and disulfide bridges were determined using NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/), NetOGlyc 4.0 (Steentoft et al. 2013) and DiANNA 1.1 web server (Ferre and Clote 2006). The protein conserved motifs and catalytic triad of PpsCatL were analyzed by multiple alignment using Clustal Omega (Goujon et al. 2010). The evolutionary relationship of PpsCatL with other related orthologs was displayed by phylogenetic tree using maximum likelihood analysis (ML) with 100 bootstrap replications of MEGA program version 5 (Tamura et al. 2011). All sequences recruited in this experiment are provided in a Table 1. A three-dimensional structure of PpsCatL was *in silico* modeled by SWISS-MODEL (Biasini et al. 2014) using crystal structure of *Fasciola hepatica* cathepsin L (PDB ID: 2o6x.1.A) as a template. The optimality of the predicted structure was examined using Ramachandran plot analysis implemented on the RAMPAGE server

(http://mordred.bioc.cam.ac.uk/~rapper/rampage.php).

Table 1 Accession numbers of amino acid sequences recruited in this study

Protein name	Type	Accession no.	Species
PpsCatL	Cathepsin L	----------	*P. pseudoheterotremus*
PwCatL1	Cathepsin L	AAB93494.1	*P. westermani*
PwCatL2	Cathepsin L	AAK35220.1	*P. westermani*
PwCP1	Cathepsin L	AAF21461.1	*P. westermani*
PwCP3	Cathepsin L	AAY81942.1	*P. westermani*
PwCatL5	Cathepsin L	AAY81943.1	*P. westermani*
PwCatL6	Cathepsin L	AAY81944.1	*P. westermani*
PwCP7	Cathepsin L	AAY81945.1	*P. westermani*
PwCatL8	Cathepsin L	AAY81946.1	*P. westermani*
PwCP9	Cathepsin L	AAY81947.1	*P. westermani*
PwCP11	Cathepsin L	AAY81948.1	*P. westermani*
PwWES1	Cathepsin L	AAW28151.1	*P. westermani*
PwWES10	Cathepsin L	AAW28152.1	*P. westermani*
SmCatL	Cathepsin L	AAC46485.1	*Schistosoma mansoni*
SjCatL	Cathepsin L	CAX72171.1	*S japonicum*
ShCatL	Cathepsin L	XP_012793298.1	*S. haematobium*
FhCatL	Cathepsin L	BAA23743.1	*Fasciola hepatica*
FgCatL	Cathepsin L	AAF44677.1	*F. gigantica*
CsCatB	Cathepsin B	GAA29748.2	*Clonorchis sinensis*

Protein name	Type	Accession no.	Species
FhCatB	Cathepsin B	ABU62925.1	*F. hepatica*
FgCatB	Cathepsin B	AAO73004.1	*F. gigantica*
OvCatB	Cathepsin B	ACT99885.1	*Opisthorchis viverrini*
SmCatB	Cathepsin B	AAA29865.1	*S. mansoni*
SjCatB	Cathepsin B	CAA50305.1	*S. japonicum*

Expression of recombinant PpsCatL in *E. coli*

First-strand cDNA of *P. pseudoheterotremus* was used as a template to amplify pro-PpsCatL without signal peptide in a reaction volume of 25 μl, containing 2 μl cDNA, 1× Taq polymerase buffer, 0.2 mM of each dNTP, 2 mM $MgCl_2$, 1 U Taq polymerase, 100 nM PpsCatL Fwd- and Rev primers. The Fwd- and Rev primer sequence for amplification of pro-PpsCatL were 5'-GG<u>GGATCC</u>GTGAGCACTGTTCGGGTGC-3′ and 5′-AAA<u>AAGCTT</u>TTAACGGATGACTGCAGACG-3′, respectively. The *Bam*HI restriction site of Fwd and *Hind*III restriction site of Rev are underlined. The PCR was performed in C1000™ Thermal cycler (Bio-Rad Laboratories, Inc., Philadephia, PA) with an amplification condition of 94 °C for 5 min, followed by 35 cycles of 94 °C for 30 sec, 55 °C for 30 sec, 72 °C for 1 min, and a final step of 72 °C for 5 min. The PCR product was examined using 2% agarose gel electrophoresis and eluted from the gel using Gel/PCR DNA Fragments Extraction Kit (Geneaid Biotech Ltd., New Taipei city, Taiwan). PpsCatL cDNA and the pQE30 vector (Qiagen GmbH, Hilden, Germany) were digested with *BamH*I (Thermo Fisher Scientific) and *Hind*III (Thermo Fisher Scientific) at 37 °C for 16-18 h. The digested cDNA and vector were separated on 2% agarose gel and purified from agarose gel using Gel/PCR DNA Fragments Extraction Kit according to manufacturer's instructions (Geneaid Biotech Ltd.). PpsCatL cDNA was ligated with pQE30 using T4 DNA ligase (Thermo Fisher Scientific), transformed into *E. coli* JM109 strain using the heat-shock technique (Froger and Hall 2007), and incubated at 37 °C for 16-18 h. The correct sequence and insertion were confirmed by DNA sequencing (AITbiotech Pte Ltd.). The pQE30-PpsCatL was transformed into *E. coli* M15 strain and then expressed recombinant protein by inducing with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG; Thermo Fisher Scientific) for 4 h. Bacteria

were harvested by centrifugation at 6,000 *g* at 4°C for 30 min, and the pellet was used in purification under denaturing condition as described previously (Pakchotanon et al. 2016). Purified rPpsCatL was used for production of mouse polyclonal antibody.

Immunological analysis

Crude worm antigen (CWA) and excretory-secretory product (ES) of adult *P. pseudoheterotremus* were prepared as mentioned previously (Adisakwattana et al. 2007). CWA, ES and rPpsCatL were size-separated on 12% SDS-PAGE gel and subsequently transferred onto a PVDF membrane (Pall Corporation, Washington, NY). The membrane was cut into small strip and then performed Immunoblot analysis according to published elsewhere with some modifications (Nuamtanong et al. 2012). Briefly, membranes containing rPpsCatL, CWA or ES were blocked with blocking solution (5% skimmed milk in 1XPBS, 0.05% Tween 20) and then incubated with 1:200 *P. pseudohetertremus*-infected rat sera at 0 and 8 wpi for rPpsCatL and 1:4,000 mouse anti rPpsCatL sera for CWA and ES in blocking solution. After washing, 1:1,000 HRP-conjugated goat anti-rat IgG (KPL Inc., Gaithersburg, MD) or 1:1,000 HRP-conjugated goat anti mouse IgG (Southern Biotech, Birmingham, AL) was added and then incubated at RT for 1 h. A result was developed by incubating the membrane with colorimetric substrate; 2,6 dichloroindophenol (Sigma-Aldrich, St. Louis, MO).

Detection of PpsCatL transcript in different developmental stages using SYBR-Real time RT-PCR

Total RNA of metacercaria, NEJ and adult *P. pseudoheterotremus* were isolated using TriZol reagent according to manufacturer's instructions (Invitrogen). One microgram of total RNA was treated with Dnase I (Thermo Fisher Scientific) to eliminate contaminated DNA. DNase-free total RNA from each stage was used as a template to synthesize first strand cDNA as described above and then determine the level of PpsCatL transcription using SYBR-Real time PCR. The reaction mixture was set up in 20 µl, containing 2 µl of first strand cDNA, 1× iTaq Universal SYBR® Green (Bio-Rad Laboratories, Inc.), and 300 nM each of Fwd and Rev primers. Specific primers for PpsCatL qPCR were Fwd: 5'-CTGTTCGGGTGCCAGATAAT-3' and Rev: 5'-CTCGTTCGTGCATCTGGTAG-3'. 18S rRNA was used as internal control for normalization of gene expression. The specific primers of *P. pseudoheterotremus* 18S rRNA were Fwd: 5'-GATAACGGGTAACGGGGAAT-3' and Rev: 5'-AGCCTCTGTTGAGTCCCGTA-3'. Amplification was performed using the LightCycler® 480 II Real-Time PCR System (Roche Applied Science, Mannheim, Germany) with an amplification reaction of pre-incubation at 95 °C for 5 min, followed by 40 cycles of 95 °C for 20 sec, and 60 °C for 1 min. Melting curve analysis was performed at 65–95 °C. The level of gene expression was calculated using the $2^{-\Delta\Delta C_t}$ formula (Livak and Schmittgen 2001). Experiments were performed with three replicates.

Immunolocalization of PpsCatL in parasite tissue

Paraffin-embedded sections of adult *P. pseudoheterotremus* were prepared and performed immunolocalization as described previously with some modifications (Adisakwattana et al. 2007). In summary, 5 μm thick paraffin-embedded sections were dewaxed, retrieved antigenic epitopes and neutralized endogenous peroxidase, respectively. Non-specific binding sites were blocked with blocking solution [10% (w/v) bovine serum albumin (BSA) in PBS, pH 7.4] at RT for 20 min followed by incubation with mouse anti-rPpsCatL (1:200) or preimmune sera (1:200) in blocking solution. Consequently, the sections were incubated with HRP-conjugated goat anti-mouse IgG (1:1,000; SouthernBiotech, Birmingham, AL) and then developed the color signal using a aminoethyl carbazole (AEC) staining kit (Sigma-Aldrich) according to the manufacturer's instructions. The result was observed under light microscope.

Evaluation of immunodiagnostic potential of rPpsCatL

A total of 255 serum samples from 32 diseases were classified into three groups; group 1 were sera of 28 paragonimiasis patients, group 2 were 197 sera samples from helminths and protozoa infection patients, and group 3 were sera of 30 healthy persons. The detail was provided in the table 2.

Table 2 Parasite-infected human sera with diagnostic criteria

Diseases	No. of cases	Diagnosis			
		Stool	Immuno	Blood	Sputum
Paragonimiasis	28	+	+		+
Echinostomiasis	1	+			
Fascioliasis	3		+		
Minute intestinal fluke	10	+			
Opisthorchiasis	10	+	+		
Ascariasis	10	+			
Angiostrogyliasis	10		+		
Malayan filariasis	10		+	+	
Capillariasis	2	+			
Dirofilariasis	1		+		
Enterobiasis	1	+			
Filariasis	3		+		
Gnathostomiasis	10		+		
Hookworm infection	10	+			
Strongyloidiasis	10	+	+		
Toxocariasis	7		+		
Trichinellosis	8		+		
Trichostrongyliasis	16	+			

Diseases **No. of cases** **Diagnosis**

		Stool	Immuno	Blood	Sputum
Trichuriasis	10	+			
Bancroftian filariasis	7		+	+	
Echinococcosis	3		+		
Hymenolepiasis nana	6	+			
Neurocysticercosis	10		+		
Sparganosis	4		+		
Taeniasis	11	+			
Taeniasis solium	4	+			
Blastocystosis	4	+			
Amebiasis	5	+			
Giardiasis	2	+			
Opisthorchiasis & Minute intestinal fluke	7	+			
Creeping eruption	1		+	Clinical history	
Taeniasis & Opisthorchiasis	1	+			
Healthy	30	+			
Total	**255**				

One hundred microliter of each rPpsCatL (2.5 μg/ml) or CWA (5 μg/ml) in 0.05 M carbonate buffer, pH 9.6 were coated into each well of 96-well microtiter plate (Nunc; Thermo Scientific, Denmark) and

incubated at 37°C for 1 h and then 16-18 h at 4°C. ELISA was performed according to previously described with some modifications. In detail, non-specific binding was blocked with blocking solution (0.5% skim milk in 1X PBS) at 37°C for 1 h, followed by washing with PBST. Human sera (1:400) were added into the well and then incubated at 37°C for 1 h. After washing, HRP conjugated goat anti-human IgG (1:2,000) was added and incubated at 37°C for 1 h and subsequently visualized by adding ABTS substrate (Sigma-Aldrich). The reaction was developed in the dark at RT for 30 min and then stopped with 1% SDS. The optical density (OD) was measured using Sunrise^{TM} absorbance reader (Tecan Group Ltd., Männedorf, Switzerland) with a wavelength at 405 nm.

Result

Sequence analysis and 3D structure simulation

The full-length PpsCatL cDNA was amplified by RACE-PCR and then analyzed nucleotide composition by DNA sequencing. The result showed that it composes of 978 bp with an open reading frame of 325 amino acid residues (Fig. 1). Homology comparison of the deduced amino acid sequence with available database using protein blast program demonstrated the highest homology to *P. westermani* cathepsin L6 (accession no. AAY81944.1) at 79.69% identity. Homology of PpsCatL with cysteine protease orthologs found in Genus *Paragonimus* were compared and shown with heatmap (Fig. 2). The full-length PpsCatL cDNA and deduced amino acid sequences were submitted to NCBI database with an assigned accession no. of (-----). The protein properties were predicted using Pepstat program showed that the molecular mass of full-length PpsCatL protein was approximately 36.5 kDa with an isoelectric point (pI) of 5.4. The potential signal peptide was

predicatively identified at the amino acid positions 1-18 of N-terminus (MTLHTIRCLAFLLACACA) but undetectable transmembrane helix region. Post-translational modifications regarding N- and O-glycosylation were not observed in PpsCatL. Ten cysteine residues were found in PpsCatL amino acid sequence at positions of 8, 15, 17, 133, 136, 167, 174, 207, 265 and 313 with two potential disulfide bridges between Cys_{133}-Cys_{174} and Cys_{167}-Cys_{207}.

```
  1 ATGACACTGCATACCATAAGATGTCTTGCTTTCTTGCTCGCATGCGCCTGTGCAGTGAGCACTGTTCGGGTGCCAGATAA 80
  1 M  T  L  H  T  I  R  C  L  A  F  L  L  A  C  A  C  A  V  S  T  V  R  V  P  D  N  27

 81 TGCGCGTGATTTGTACGAACAGTTCAAACGGGACTACGGGAAGATTTATGCGAATGACGATGACGAGAAGCGATTTGCCA 160
 28  A  R  D  L  Y  E  Q  F  K  R  D  Y  G  K  I  Y  A  N  D  D  D  E  K  R  F  A  I 54

161 TCTTCAAGGACAATCTGGTGCGTGCACAAACCTACCAGATGCACGAACGAGGCACAGCTAAGTACGGTGTGACTCAGTTT 240
 55  F  K  D  N  L  V  R  A  Q  T  Y  Q  M  H  E  R  G  T  A  K  Y  G  V  T  Q  F  80

241 TTCGACTTGACACCTGAAGAATTTGCGGCCAAGTACCTGAGTTCACCAATCGACGACCAAGTGGAACACGTGCAACTGAA 320
 81 F  D  L  T  P  E  E  F  A  A  K  Y  L  S  S  P  I  D  D  Q  V  E  H  V  Q  L  N 107

321 TGATCTCAAAGCAGCTCCCGAACGTATTGACTGGCGAGAGAAGGGTGCTGTAGCACCGGTTGAAGATCAAGGCTGGTGTG 400
108  D  L  K  A ▲A  P  E  R  I  D  W  R  E  K  G  A  V  A  P  V  E  D  Q  G  W [C] G 134

401 GTTCGTGTTGGGCATTTTCGGTAGCAGGAAATATTGAAGGTCAATGGTTTCTGAAGACCGGCCAGCTTGTCAGTCTGAGC 480
135  S  C  W  A  F  S  V  A  G  N  I  E  G  Q  W  F  L  K  T  G  Q  L  V  S  L  S  160

481 AAACAGCAATTGGTCGATTGTGACACGGTGGACAGCGGATGTAATGGTGGATGGCCACCATTAACATACGGCGAGATCAA 560
161 K  Q  Q  L  V  D [C] D  T  V  D  S  G [C] N  G  G  W  P  P  L  T  Y  G  E  I  K 187

561 ACGTCTGGGTGGTTTGGAGGCGCAACGAGACTATCCCTATGTTGGCAGAGAGCAAACGTGTAGATTGGATAAGTCGAAGT 640
188  R  L  G  G  L  E  A  Q  R  D  Y  P  Y  V  G  R  E  Q  T [C] R  L  D  K  S  K  L 214

641 TGTTAGCCAAAATCGACGGATCGGTTGTTTTGGAAAGAGATGAGTATAAACAGGCAGCTTGGCTCGCAGAACACGGACCA 720
215  L  A  K  I  D  G  S  V  V  L  E  R  D  E  Y  K  Q  A  A  W  L  A  E  H  G  P  240

721 ATGGCTTCGGCCCTAAATGCTAATTATTTTCAGTACTACCGATCCGGAATCAGTCATCCGTCCAGTTATGAGTGTAATCC 800
241 M  A  S  A  L  N  A  N  Y  F  Q  Y  Y  R  S  G  I  S  H  P  S  S  Y  E  C  N  P 267

801 AGATGGATTGAACCACGCCGTATTGACTGTTGGTTACGGCACGGAAAATGGTATTCCTTACTGGACTATCAAAAATAGTT 880
268  D  G  L  N  H  A  V  L  T  V  G  Y  G  T  E  N  G  I  P  Y  W  T  I  K  N  S  W 294

881 GGGGTACCGGTTGGGGCGAGAACGGCTACTTCCGACTCTACCGTGGTGATGGAACATGTGGAATCGAAAAAGTTGTTTCG 960
295  G  T  G  W  G  E  N  G  Y  F  R  L  Y  R  G  D  G  T  C  G  I  E  K  V  V  S  320

961 TCTGCAGTCATCCGTTAA 978
321 S  A  V  I  R  *  325
```

Fig. 1 Nucleotide sequence of full-length PpsCatL (accession no.) comprises of 978 nucleotides that encodes 325 amino acid residues. Signal peptide identified at amino acid residues 1-18 is underlined. Cleavage site between the pro-peptide and mature protein is indicated with arrowhead. Potential disulfide bonds are in the box.

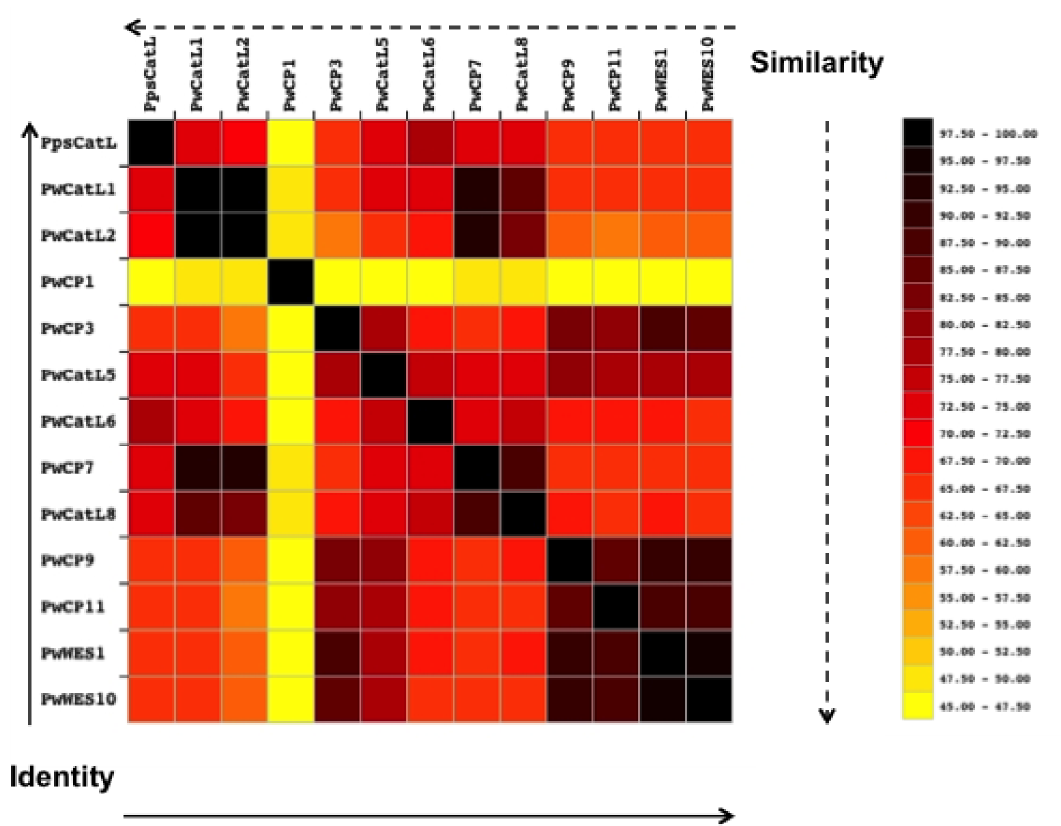


Fig. 2 Homology heatmap of cathepsin L found in *Paragonimus* spp. Percentage identity are indicated with dash arrow (lower left) and percentage similarity are indicated with closed arrow (upper right).

Multiple sequence alignment of PpsCatL with orthologs identified pro-region cleavage site between the Ala_{111} and Ala_{112}. A mature protease is composed of 214 amino acid residues with a calculated molecular mass approximately 23.7 kDa and a theoretical pI of 5.3. A conserved ERFNIN and GNFD motif, signature of cathepsin L, were observed in a pro-region of PpsCatL at amino acid positions of 48-63 and 76-82, respectively. A conserved structural motif, GCNGG, was presented in PpsCatL at positions 173-177. A catalytic triad was located at Cys_{136}, His_{272} and Asn_{292} and S_1 subsite of the protease active site were identified at Gln_{130}, Cys_{136}, His_{272}, Asn_{292}, and Trp_{294}. The S_2 subsite of the active site contains Leu_{192}, Glu_{193}, Ala_{244}, and Leu_{270} (Fig. 3).

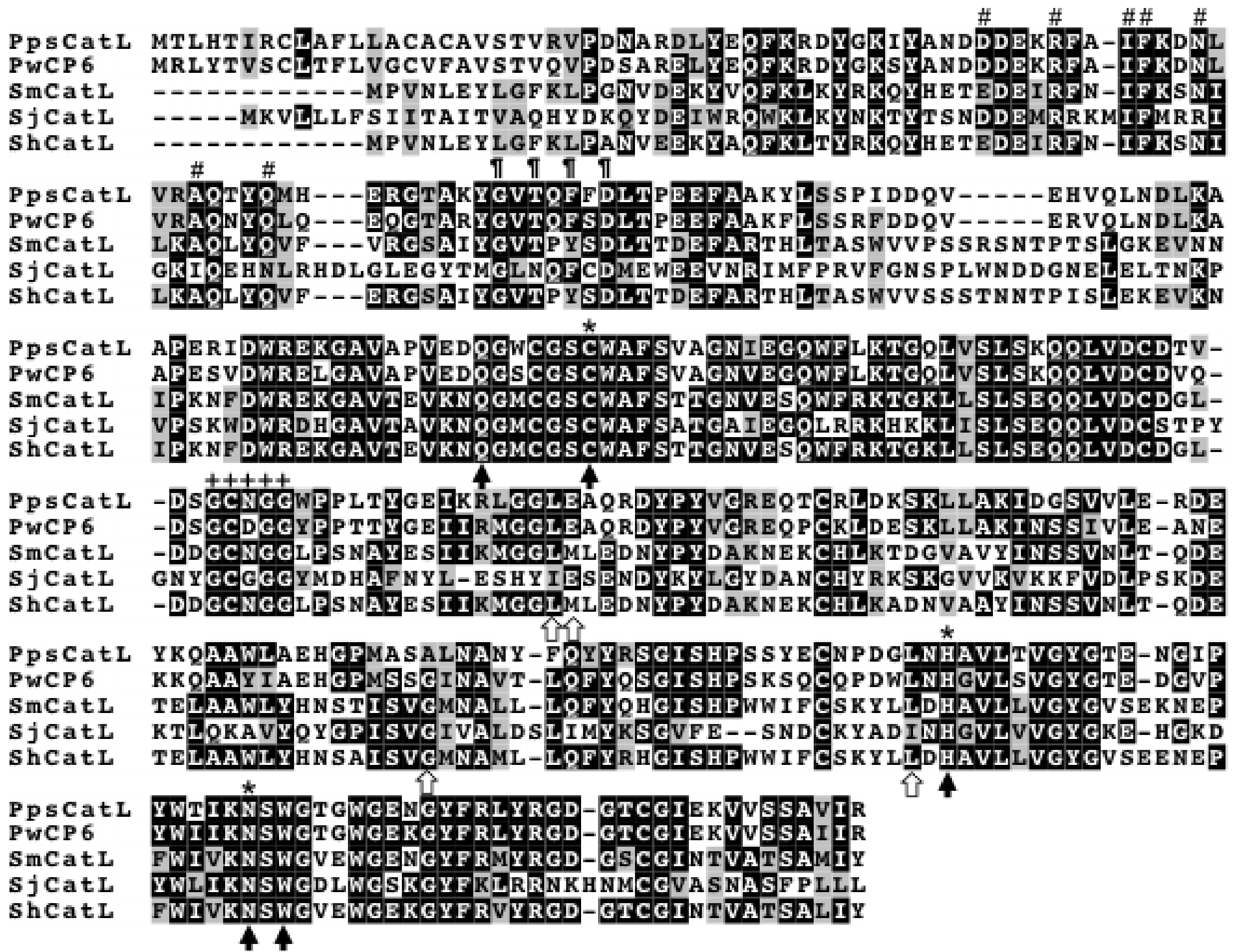


Fig. 3 Multiple alignment of deduced amino acid sequences indicates conserved motifs within orthologous cathepsin L. Identical and similar residues are indicated by black and grey shading, respectively. Gaps are indicated by a dash (-). Putative amino acid residues of the catalytic triad [Cys (C), His (H), Asn (N)] are indicated with asterisk (*) above the residues. ERFNIN, GNFD, and GCNGG motifs are labeled above the alignment with #, ¶, and +, respectively. S_1 subsites are indicated with black arrow below the alignment and S_2 subsites are indicated with white arrow below the alignment.

An evolutionary relationship of PpsCatL with orthogs was analyzed using maximum likelihood method (ML) with 100 bootstrap replications showed that PpsCatL was clustered in the same branch with *P. westermani* cathepsin L, which exhibited the closest relationship with PwCatL6. Relationship of Paragonimus cathepsin Ls with *Schistosoma*- was closer than with *Fasciola* cathepsin L. The trematode cathepsin B were distinctly separated from cathepsin L as shown in the phylogenetic tree (Fig. 4).

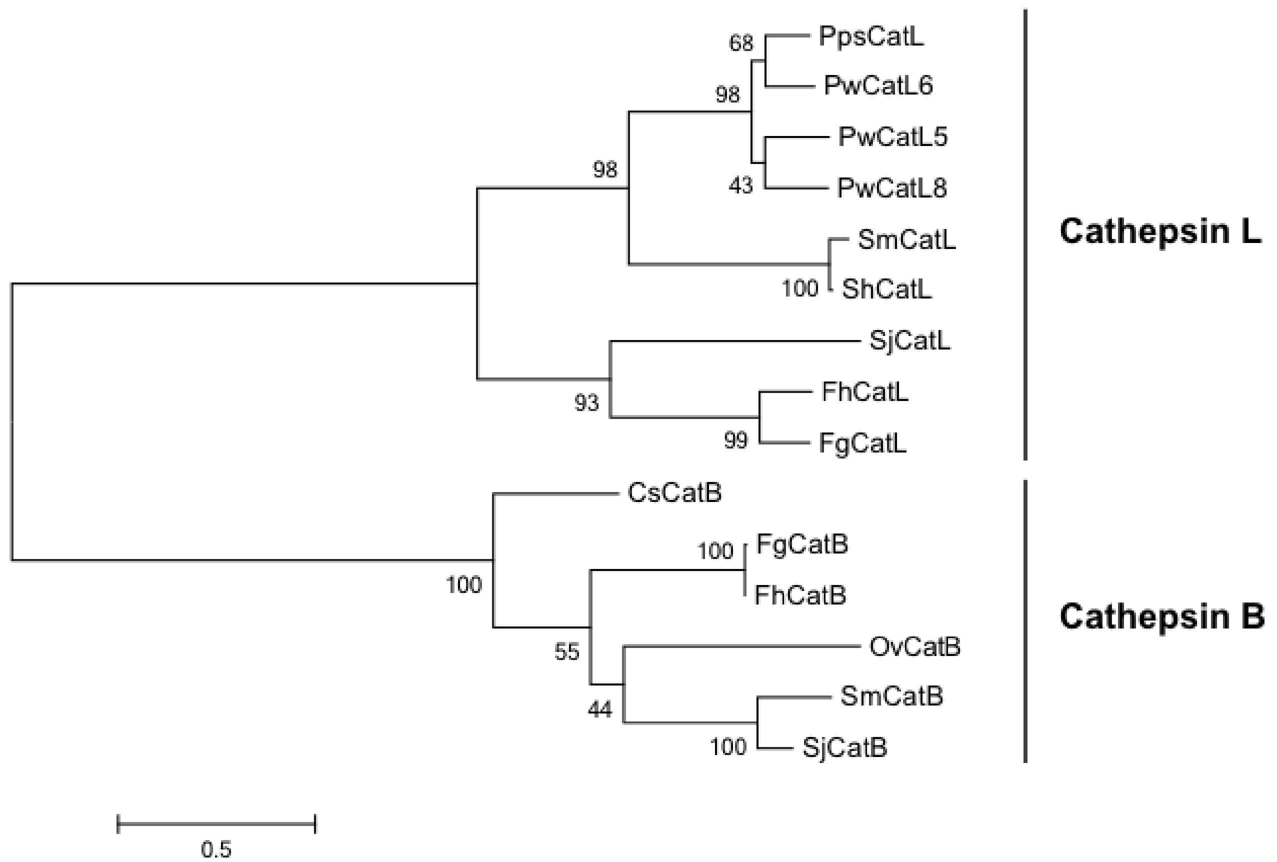


Fig. 4 Phylogenetic tree demonstrating that PpsCatL is closely related with *P. westermani* cathepsin L

3D structures of pro- and mature PpsCatL were simulated using crystal structure of *F. hepatica* catthepsin L (PDB ID: 2o6x.1.A) as a template. The pro-PpsCatL 3D structure reveals that pro-peptide region folds on the cathepsin L surface to hinder the active site from substrates (Fig. 5A). After removing pro-region (amino acid residues 19-111), the structure of PpsCatL becomes mature pretease (amino acid residues 112-325) that exhibits its active site to expose with the substrate (Fig. 5B). The structure of mature PpsCatL is composed of two domains (left and right), which the left domain contains three α-helices and the right domain contains β-sheet with a front helix forming a coiled structure. At the top of the two domains interacting creates the active site cleft (Fig. 5B). Ramachandran plotting suggested that the structural model was optimal (data not shown).

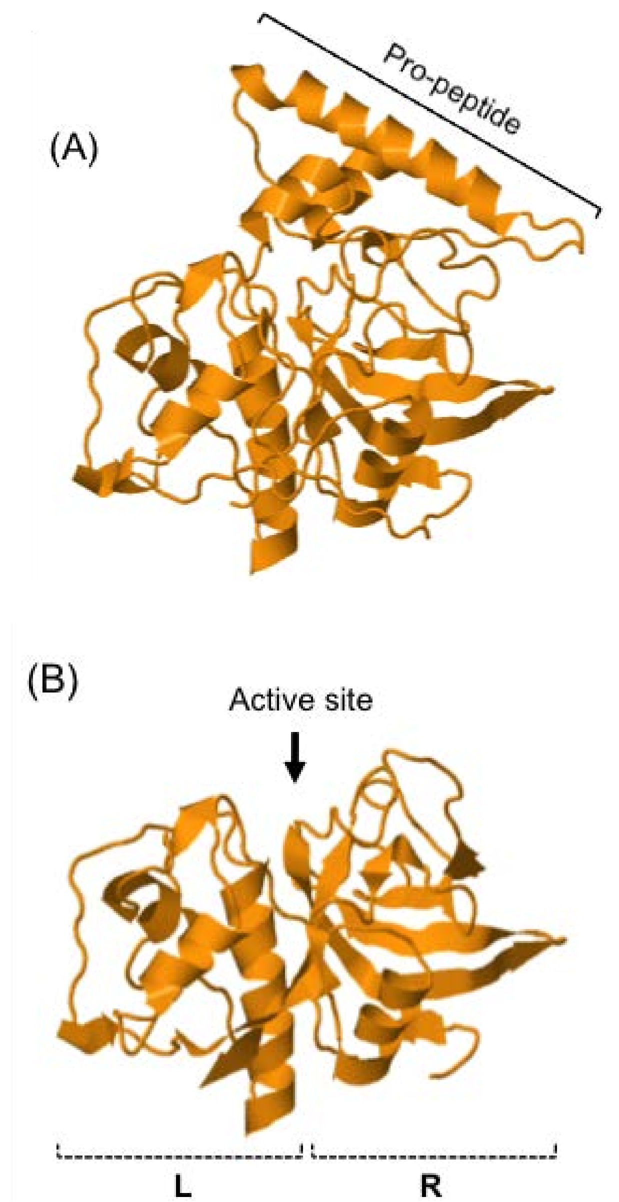


Fig. 5 *In silico* 3D structure of pro-PpsCatL (A) and mature-PpsCatL (B). Pro-peptide hinders the active site of protease to inhibit substrate exposure. Left (L) and right (R) domains contain three α-helices and β-sheet, respectively.

Expression and purification of rPpsCatL

After inducing protein expression with IPTG, rPpsCatL was predominantly expressed at an approximately molecular size of 35 kDa (Fig. 6A). The water-soluble property was determined and showed that rPpsCatL was expressed as an insoluble form in *E. coli*, which was detected with anti-His Tag antibody (Fig. 6A). rPpsCatL was purified under denaturing condition, which presented a predominant bands at 35 kDa were eluted (Fig. 6B). Purified rPpsCatL reacted with anti-His tag antibody showed positive result (data not shown).

Detection of PpsCatL in native parasite antigens and determination of immune response against rPpsCatL

Presence of PpsCatL in CWA and ES of *P. pseudoheterotremus* was detected with western blot analysis probing with mouse anti-rPpsCatL immune sera. In CWA, the anti-rPpsCatL sera could detect native protein at the molecular weight approximately 35, 28 and 25 kDa, respectively. In ES, a protein band at a molecular weight of 25 kDa was predominantly reacted with anti-rPpsCatL sera. Moreover, protein bands at range of 10-18 kDa in ES, which strongly reacted at 14 kDa (Fig. 6C).

Immune response of *P. pseudoheterotremus* infected rats against rPpsCatL was observed with western blot analysis. The result indicated that all five infected rat sera at 8 wpi could react with rPpsCatL at 35 kDa and other truncated proteins (Fig. 6D). Normal rat- and mouse sera were used as a negative control.

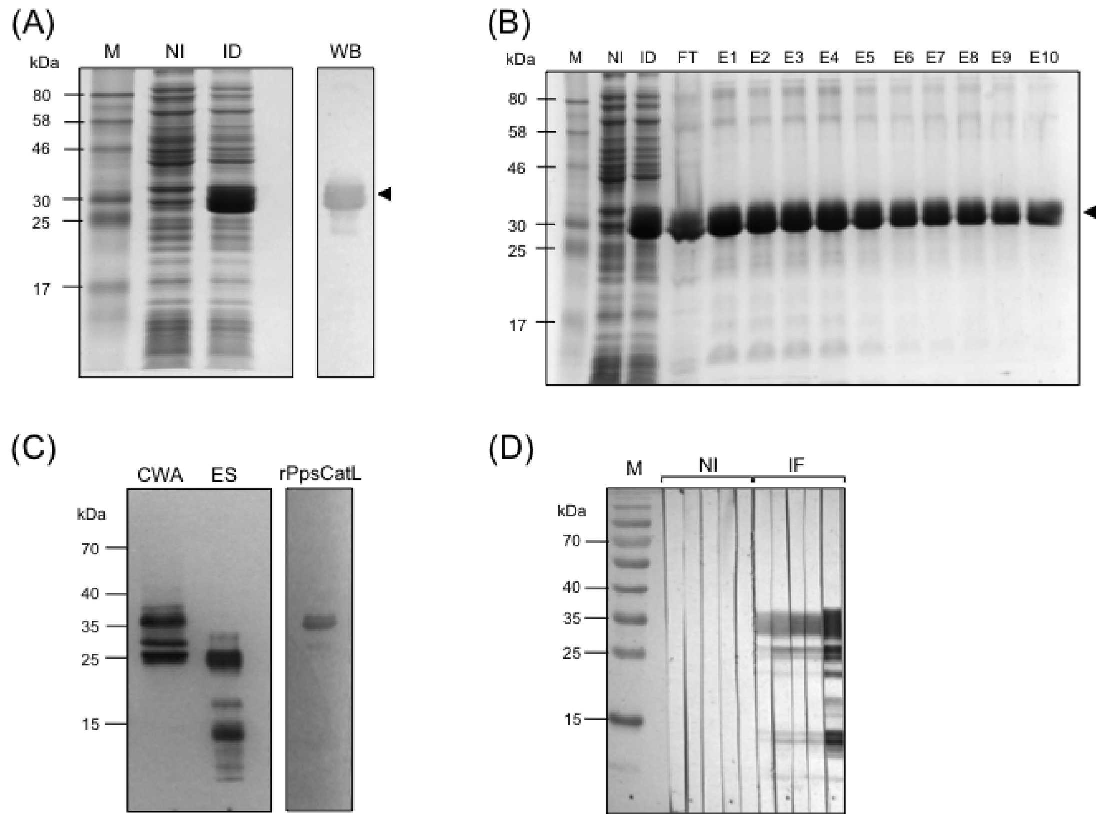


Fig. 6 rPpsCatL was heterologously expressed in *E. coli* (A) and then purified using Co^{2+} affinity column (B). M: broad range prestained protein marker, NI: noninducing, ID: inducing, WB: western blot analysis with anti-His tag antibody, FT: flow-through, E: elution. C, Western blot analysis detecting native PpsCatL in CWA and ES using mouse anti-rPpsCatL sera. CWA: crude worm antigen, ES: excretory-secretory product, rPpsCatL: recombinant *P. pseudohetrotremus* cathepsin L. D, Detection of rPpsCatL with *P. pseudohetrotremus*-infected rat sera. M: broad range prestained protein marker, NI: non-infected sera, IF: infected sera.

Determination of PpsCatL transcript in different stage of *P. pseudoheterotremus*

Transcription level of PpsCatL in different developmental stages (metacercaria, NEJ, and adult) was performed to define stage specific expression. The result showed that PpsCatL was transcribed at the highest level in adult parasite with approximately 7.7 folds over metacercaria and NEJ. Expression of mRNA level of metacercaria and NEJ was not significantly different (Fig. 7).

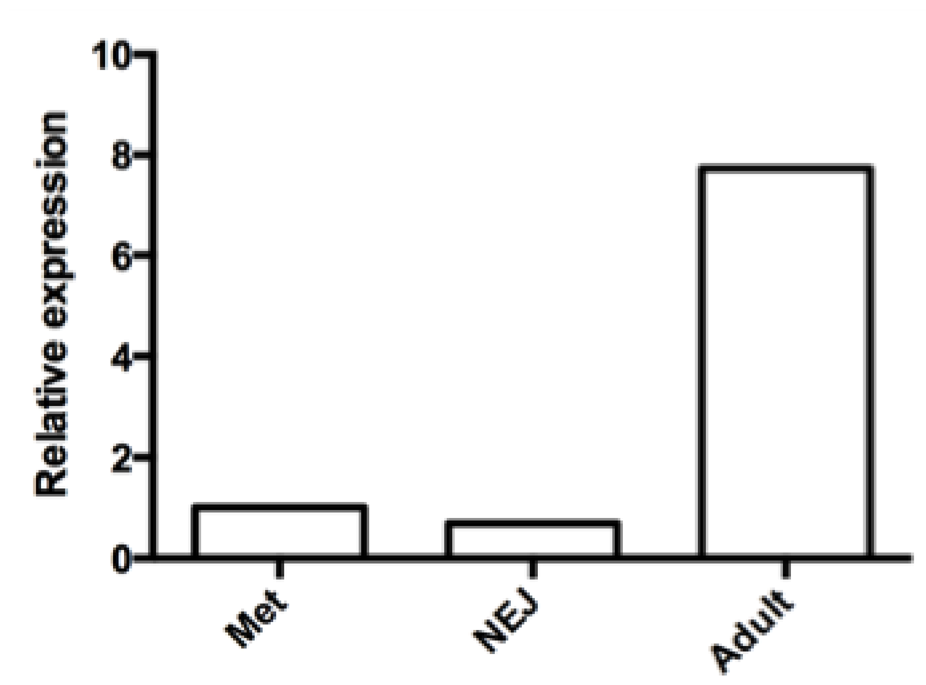


Fig. 7 Stage-specific transcription of PpsCatL was monitored in different developmental stages of *P. pseudohetrotremus* using SYBR realtime RT-PCR. Met: metacercaria, NEJ: newly excysted juvenile, Adult: adult worm. Relative expression level was calculated by comparison with transcription level of metacercaria.

Tissue-specific localization of PpsCatL in adult *P. pseudoheterotremus*

The immunolocalization of PpsCatL was performed in adult stage according to an evidence of the highest transcription of this gene. PpsCatL was specifically expressed in digestive system, including esophagus and intestine but not other tissues. The localization of PpsCatL was intensively accumulated at lamellae lining in esophagus and along the intestine of whole parasite body (Fig. 8A,B). Immunolocalization with mouse pre-immunized sera showed negative result (Fig. 8C).

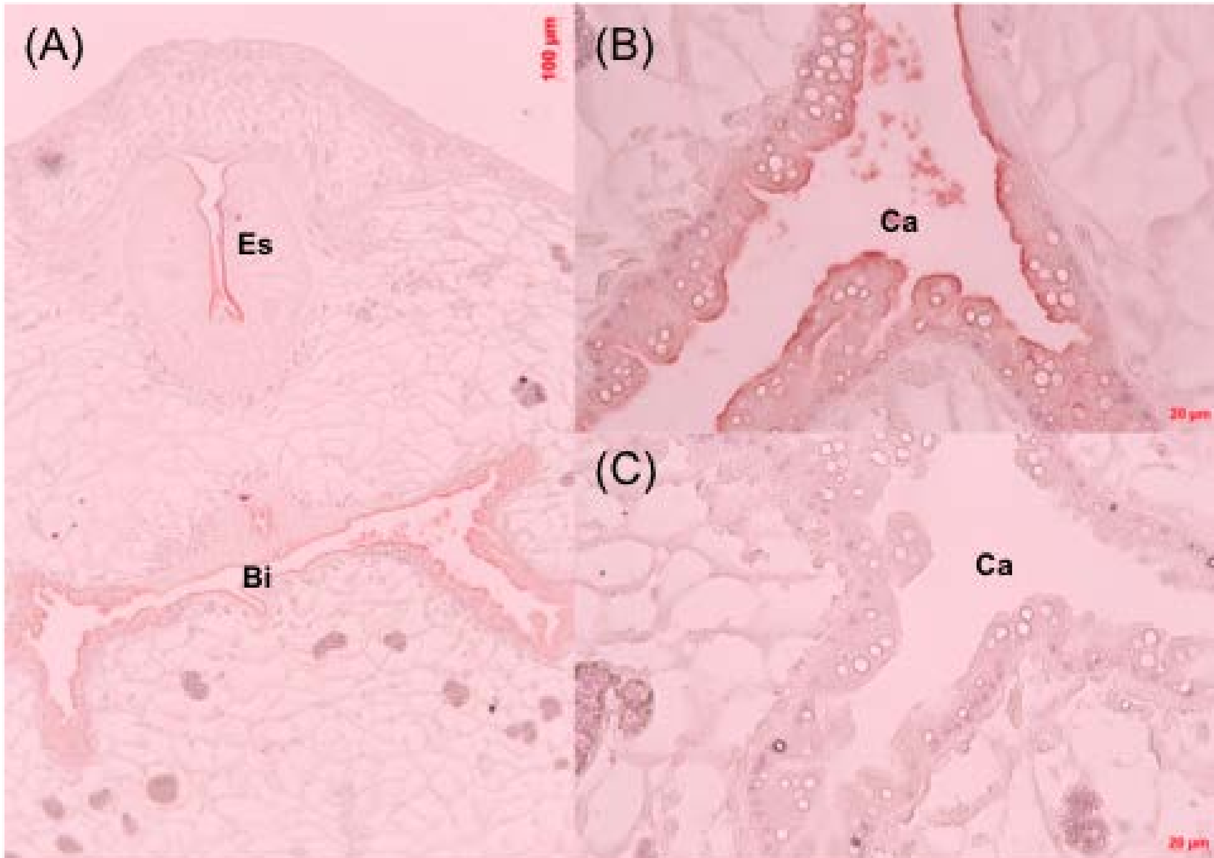
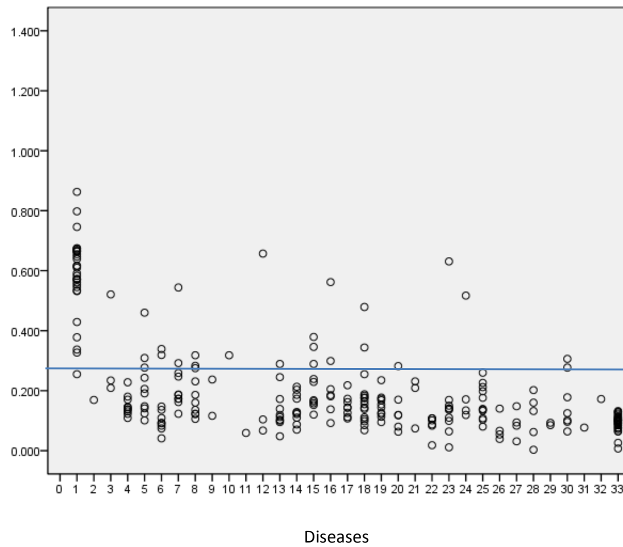


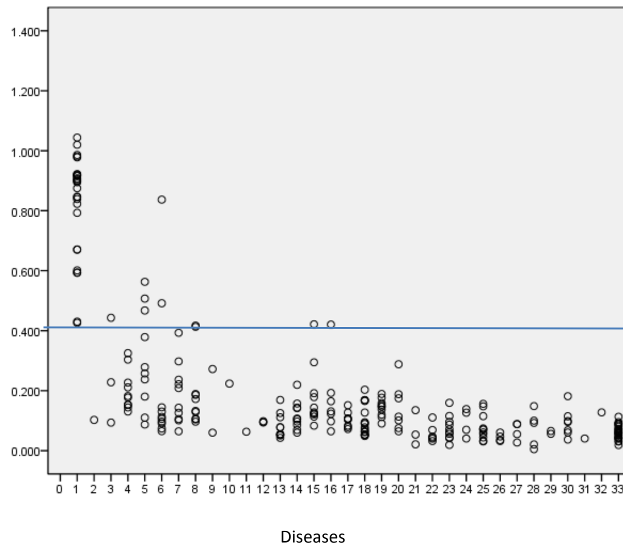
Fig. 8 Immunolocalization of PpsCatL in tissue of adult *P. pseudohetrotremus* (A-C). A, low magnification of adult tissue showed localization of PpsCatL at intestinal tissue. Es: esophagus, Bi: bifurcated intestine. B, High magnification of parasite caecum; Ca: caecum. C, Control reacted with mouse preimmune sera.

rPpsCatL-ELISA

Suitable dilution of human sera and secondary antibody was optimized using checkerboard titration (data not shown). The optimal dilution of human sera and secondary antibody were 1:400 and 1:2,000, respectively. In this study, ELISA using rPpsCatL and CWA as antigen were performed to compared sensitivity and specificity between them. The cut-off points of CWA- and rPpsCatL-ELISA were 0.232 (mean+5SD) and 0.412 (mean+6SD), respectively. Frequency distribution of negative, positive and samples was plotted to indicate the number of true positive, true negative, false positive, and false negative (Fig. 9A left and 9B left). The sensitivity and specificity of CWA were 100% and 84.1%, respectively, and of rPpsCatL were 100% and 95.6%, respectively (Fig. 9A right and 9B right).



CWA-ELISA	STANDARD: Positive	STANDARD: Negative
Positive	28	36
Negative	0	191
	Sensitivity 100%	**Specificity 84.1%**



PpsCatL-ELISA	STANDARD: Positive	STANDARD: Negative
Positive	28	10
Negative	0	217
	Sensitivity 100%	**Specificity 95.6%**

Fig. 9 Evaluation of CWA-ELISA (A) and rPpsCatL-ELISA (B). Scatter plot indicates absorbance value of the parasite-infected sera (A left and B left). Dashed lines indicate cut-off values of CWA-ELISA and rPpsCatL-ELISA at 0.232 and 0.412, respectively. Sensitivity and specificity were calculated by 2X2 table shows that sensitivity and specificity of CWA-ELISA are 100 and 84.1%, respectively, and of rPpsCatL are 100 and 95.6%, respectively (A right and B right).

Discussion

In parasitic trematode, cathepsin L is indispensible cysteine protease that involves in variety activities through their life cycle including encystation, excystation, migration, digestion, maturation, fertilization, immune-evasion, *etc* (Grams et al. 2001; Shin et al. 2001; Collins et al. 2004; Na et al. 2006; Chung et al. 2008). Previously, several cathepsin L isoforms were identified in a medical important lung fluke, *P. westermani*, but not other *Pargonimus* spp. (Park et al. 2002). Herein, cathepsin L of zoonotic paragonimiasis, *P. pseudoheterotremus*, was firstly identified and characterized using bioinformatics and molecular biology approaches. Identification of PpsCatL using primer degenerated RT-PCR combination with RACE-PCR created full-length gene, which showed high homology to *P. westermani* cathepsin L, especially PwCatL6 and PwCatL5, respectively. PwCatL5 was well-characterized cysteine protease found in *P. westermani* adult worm, which was classified in cathepsin L family (Park et al. 2002). Amino acid properties of PpsCatL demonstrated that the protein has a molecular size of 36 kDa containing signal peptide but not transmembrane helix, which indicate excretory-secretory potential. These may suggest the extracellular roles of this protein involving digestion, migration, invasion, moreover, host-parasite interaction (Day et al. 1995; Dalton et al. 1996; Collins et al. 2004). Presence of ten cysteine residues with two potential disulfide bonds in PpsCatL possibly facilitates tertiary structure forming and intermolecular interaction creating dimer as shown in other cathepsin L orthologs. Cathepsin L found in embryos and larvae of brine shrimp (*Artemia franciscana*) forms heterodimer with cathepsin L-associated protein (CLAP), which more active and stable than cathepsin L monomer (Warner et al. 2004). Multiple alignment of PpsCatL with orthologs indicates ERFNIN and GNFD motifs at the propeptide of PpsCatL, which is

described as a signature of cathepsin L-like but not B-like proteases (Kollien et al. 2004; Pandey et al. 2009). The properties of ERFNIN and GNFD motifs were described to facilitate inhibition of falcipain 2 propeptide, which was confirmed by deletion of both motifs impaired inhibitory activity against falcipain 2 (Pandey et al. 2009). GCNGG is another conserved motif found at amino acid residues 173-177 where located in a mature PpsCatL. This motif was previously reported to be associated with protein globular formation relied upon disulfide bridge (Karrer et al. 1993). The catalytic triad of mature PpsCatL composed of Cys_{136}, His_{272} and Asn_{292}, which conserve within the papain family (clan CA) (Pandey and Dixit 2012). S_1 and S_2 subsites were predicted at the active site of PpsCatL, which conserve with other orthologs. The S_1 and S_2 subsites of cathepsin L, especially S2 contribute to substrate preference (Alves et al. 2001; Lecaille et al. 2007). Mutation of S2 pocket (L67Y/A205L) in human cathepsin L affected a change of their enzymatic specificity both amino acid preference in peptidyl substrate and small selective inhibitor (Lecaille et al. 2007).

Phylogenetic tree demonstrated that PpsCatL was closely related with cathepsin L of *P. westermani*, PwCatL6, PwCatL5 and PwCatL8, respectively that might suggest some biological and biochemical associations of this protein in genus *Paragonimus*. *In silico* tertiary structure of mature PpsCatL introduced (reviewed by Turk et al. 2012). However, crystallography of PpsCatL as well as other *Paragonimus* cathepsin L should be performed to clearly understand the structure of this protein in genus *Paragonimus*.

Expression of pro-PpsCatL without signal peptide in prokaryotic expression system found that the protein was produced as an insoluble product at an approximately molecular size of 35 kDa, which was similar to calculated molecular weight of pro-PpsCatL (34.5 kDa). Detection of PpsCatL

in CWA and ES using mouse anti-rPpsCatL sera indicated that pro-PpsCatL produced by *E. coli* was the same size as found in CWA. These supported an evidence of non-glycosylated PpsCatL, which was consistent with no potential glycosylation site predicted in this protein. Cathepsin L without potential glycosylation site was previously observed in *F. hepatica* procathepsin L1 but was still detectable in ES of adult parasite (Collins et al. 2004). Lacking of glycosylation in human and mice procathepsin L did not disturb structural folding, stability and secretory property of this protease (Kane 1993; Smith et al. 1989). In *P. pseudoheterotremus*, PpsCatL was presented in ES at predominant molecular weight of 25 kDa that was equal to predicted mature PpsCatL. In *F. hepatica*, mature cathepsin L1 (FhCatL1) and L2 (FhCatL2) but not full length were detected in ES of adult parasite (Collins et al. 2004). Ladder-liked pattern found in CW and ES of adult *P. pseudoheterotremus* emerged from catalytic of pro-enzyme and degradation, respectively. Immune response against PpsCatL was studied by reacting with *P. pseudoheterotremus*-infected rats. The strong signal was observed in all infected animals suggests a potential of this protein for immunodiagnostic development. Using *F. hepatica* cathepsin L for immunodiagnosis was successfully detected fascioliasis in both human and ruminants (Cornelissen et al. 2001; Gonzales Santana et al. 2013).

Transcriptional up-regulation of PpsCatL was mainly detected in adult stage of *P. pseudoheterotremus* but less in other developmental stages. Differential expression of cathepsin L isoforms in different developmental stages was frequently reported in *Fasciola* spp. (Dowd et al. 1994). FgCatL1G was mainly expressed in metacercaria and NEJ of *F. gigantica* but not in adult (Sansri et al. 2013). The previous study showed that biochemical properties between (Norbury et al. 2011). Expression of PpsCatL in adult *P. pseudoheterotremus* tissue was located at gastrodermal

tissue along the parasite gut, especially gut lamellae. Moreover, secretory PpsCatL was determined in the gut lumen and ES of the parasite. These suggest important roles of PpsCatL in the parasite itself and host environment. Secreted cathepsin L1 from *F. hepatica* was proposed as a major virulence that played essential roles in helminth pathogenicity such as nutrient acquisition by host protein digestion, host extracellular matrix cleavage lead to parasite migration, and impairing host immunity by destroying immunoglobulin and suppressing Th1 immune response (Collins et al. 2004).

rPpsCatL-ELISA was highly sensitive and specific to diagnose human paragonimiasis. However, the result here could not determine species-specificity according to lack of species-confirmed sera. Cross-reaction of cathepsin L in the same genus was occurred in *F. hepatica* and *F. gigantica*-infected ruminants. Detection of *F. gigantica* with FhCatL-ELISA elicited similar outcome as shown in *F. hepatica* (Cornelissen et al. 2001). Thus, positive result shown in this experiment may come from diverse *Paragonimus* spp. infection such as *P. westermani*, *P. heterotremus*, *P. pseudoheterotremus etc*. Another evidence regarding cross-reaction was suggested by high percentage identity of cathepsin L amino acid sequences, more than 45%, within *Paragonimus* spp. However, validation of rPpsCatL-ELISA with confirmed species-specific pargonimiasis sera, especially *P. heterotremus*, should be performed to clarify this issue. *P. heterotremus* and *P. pseudohetrotremus* are sister species, which have high genetic similarity (Intapan et al. 2012). Therefore, high homology of cathepsin L occurring in both species might affect cross-reactivity of rPpsCatL-ELISA. Nonetheless, cathepsin L has not yet been identified in *P. heterotremus* that should be further discovered for comparison fundamental function and property of these sibling species including sequence homology, protein structure, biochemistry, immune response, *etc*. 10 cases of

heterologous sera exhibited false positive with rPpsCatL-ELISA including *O. viverrini* (3 cases), *A. lumbricoides* (2 cases), *B. malayi* (2 cases), *S. stercoralis* (1 case), *Toxocara* spp. (1 case), and *Fasciola* spp. (1 case). According to the result, almost all heterologous sera (9 cases) cross-reacted with rPpsCatL located nearby cutoff line, which might be due to intense antibody titer against those diseases affect antibody specificity. As demonstrated previously, false positive of HIV test (Murex HIV Ag/Ab combination EIA) was significantly associated with increasing level of *S. haematobium* IgG1 (Everett et al. 2010). Moreover, one case of *A. lumbricoides* infection highly reacted with rPpsCatL was observed in our study. This is possibly caused by previous history of paragonomiasis or other closely related helminthic infections.

High sensitivity (100%) and specificity (95.6%) introduced in rPpsCatL-ELISA suggests the potential to use this test for screening paragonimiasis. The higher specificity of rPpsCatL-ELISA (95.6%) over CWA-ELISA (84.1%) emphasizes replacement of CWA with rPpsCatL in clinical immunodiagnosis. However, evaluation of rPpsCatL-ELISA with non-parasitic infectious diseases, *e.g.* viruses, bacteria, fungi, and non-infectious diseases, *e.g.* lung cancer, lung diseases need to be included in the further study.

In summary, a novel cathepsin L-like protease was firstly identified in *P. pseudohetrotremus*, which was predominantly expressed in adult stage. This cathepsin L was found in ES product and specifically localized at intestinal brush border of the parasite. Recombinant protein expressed in *E. coli* was reacted with rat-infected sera and could differential diagnose human paragonimiasis with high sensitivity and specificity. Further studies focused on evaluation of rPpsCatL-ELISA with

species-specific paragonimiasis as well as non-parasitic infection and some non-infected lung diseases will advantage for development of rPpsCatL-based immunodiagnosis in the near future.

References

Adisakwattana P, Viyanant V, Chaicumpa W, Vichasri-Grams S, Hofmann A, Korge G, Sobhon P, Grams R (2007) Comparative molecular analysis of two asparaginyl endopeptidases and encoding genes from *Fasciola gigantica*. Mol Biochem Parasitol 156(2):102-16 doi:S0166-6851(07)00209-5 [pii] 10.1016/j.molbiopara.2007.07.006

Alves LC, Melo RL, Cezari MH, Sanderson SJ, Mottram JC, Coombs GH, Juliano L, Juliano MA (2001) Analysis of the S(2) subsite specificities of the recombinant cysteine proteinases CPB of *Leishmania mexicana*, and cruzain of *Trypanosoma cruzi*, using fluorescent substrates containing non-natural basic amino acids. Mol Biochem Parasitol 117(2):137-43 doi:S0166685101003401 [pii]

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo Cassarino T, Bertoni M, Bordoli L, Schwede T (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42(Web Server issue):W252-8 doi:gku340 [pii] 10.1093/nar/gku340

Chung YB, Kita H, Shin MH (2008) A 27 kDa cysteine protease secreted by newly excysted *Paragonimus westermani* metacercariae induces superoxide anion production and degranulation of human eosinophils. Korean J Parasitol 46(2):95-9 doi:200806095 [pii] 10.3347/kjp.2008.46.2.95

Collins PR, Stack CM, O'Neill SM, Doyle S, Ryan T, Brennan GP, Mousley A, Stewart M, Maule AG, Dalton JP, Donnelly S (2004) Cathepsin L1, the major protease involved in liver fluke

(*Fasciola hepatica*) virulence: propetide cleavage sites and autoactivation of the zymogen secreted from gastrodermal cells. J Biol Chem 279(17):17038-46 doi:10.1074/jbc.M308831200 M308831200 [pii]

Cornelissen JB, Gaasenbeek CP, Borgsteede FH, Holland WG, Harmsen MM, Boersma WJ (2001) Early immunodiagnosis of fasciolosis in ruminants using recombinant *Fasciola hepatica* cathepsin L-like protease. Int J Parasitol 31(7):728-37 doi:S0020-7519(01)00175-8 [pii]

Dalton JP, Clough KA, Jones MK, Brindley PJ (1996) Characterization of the cathepsin-like cysteine proteinases of *Schistosoma mansoni*. Infect Immun 64(4):1328-34

Day SR, Dalton JP, Clough KA, Leonardo L, Tiu WU, Brindley PJ (1995) Characterization and cloning of the cathepsin L proteinases of *Schistosoma japonicum*. Biochem Biophys Res Commun 217(1):1-9 doi:S0006291X85727374 [pii]

Dowd AJ, Smith AM, McGonigle S, Dalton JP (1994) Purification and characterisation of a second cathepsin L proteinase secreted by the parasitic trematode *Fasciola hepatica*. Eur J Biochem 223(1):91-8

Everett DB, Baisely KJ, McNerney R, Hambleton I, Chirwa T, Ross DA, Changalucha J, Watson-Jones D, Helmby H, Dunne DW, Mabey D, Hayes RJ (2010) Association of schistosomiasis with false-positive HIV test results in an African adolescent population. J Clin Microbiol 48(5):1570-7 doi:JCM.02264-09 [pii] 10.1128/JCM.02264-09

Ferre F, Clote P (2006) DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. Nucleic Acids Res 34(Web Server issue):W182-5 doi:34/suppl_2/W182 [pii] 10.1093/nar/gkl189

Froger A, Hall JE (2007) Transformation of plasmid DNA into *E. coli* using the heat shock method. J Vis Exp(6):253 doi:10.3791/253

Gonzales Santana B, Dalton JP, Vasquez Camargo F, Parkinson M, Ndao M (2013) The diagnosis of human fascioliasis by enzyme-linked immunosorbent assay (ELISA) using recombinant cathepsin L protease. PLoS Negl Trop Dis 7(9):e2414 doi:10.1371/journal.pntd.0002414 PNTD-D-13-00668 [pii]

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38(Web Server issue):W695-9 doi:gkq313 [pii] 10.1093/nar/gkq313

Grams R, Vichasri-Grams S, Sobhon P, Upatham ES, Viyanant V (2001) Molecular cloning and characterization of cathepsin L encoding genes from *Fasciola gigantica*. Parasitol Int 50(2):105-14 doi:S138357690100068X [pii]

Intapan PM, Sanpool O, Thanchomnang T, Imtawil K, Pongchaiyakul C, Nawa Y, Maleewong W (2012) Molecular identification of a case of *Paragonimus pseudoheterotremus* infection in Thailand. Am J Trop Med Hyg 87(4):706-9 doi:ajtmh.2012.12-0235 [pii] 10.4269/ajtmh.2012.12-0235

Irving JA, Spithill TW, Pike RN, Whisstock JC, Smooker PM (2003) The evolution of enzyme specificity in *Fasciola* spp. J Mol Evol 57(1):1-15 doi:10.1007/s00239-002-2434-x

Kane SE (1993) Mouse procathepsin L lacking a functional glycosylation site is properly folded, stable, and secreted by NIH 3T3 cells. J Biol Chem 268(15):11456-62

Karrer KM, Peiffer SL, DiTomas ME (1993) Two distinct gene subfamilies within the family of cysteine protease genes. Proc Natl Acad Sci U S A 90(7):3063-7

Kollien AH, Waniek PJ, Nisbet AJ, Billingsley PF, Schaub GA (2004) Activity and sequence characterization of two cysteine proteases in the digestive tract of the reduviid bug *Triatoma infestans*. Insect Mol Biol 13(6):569-79 doi:IMB504 [pii] 10.1111/j.0962-1075.2004.00504.x

Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305(3):567-80 doi:10.1006/jmbi.2000.4315 S0022-2836(00)94315-8 [pii]

Lecaille F, Chowdhury S, Purisima E, Bromme D, Lalmanach G (2007) The S2 subsites of cathepsins K and L and their contribution to collagen degradation. Protein Sci 16(4):662-70 doi:16/4/662 [pii] 10.1110/ps.062666607

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25(4):402-8 doi:10.1006/meth.2001.1262 S1046-2023(01)91262-9 [pii]

Na BK, Kim SH, Lee EG, Kim TS, Bae YA, Kang I, Yu JR, Sohn WM, Cho SY, Kong Y (2006) Critical roles for excretory-secretory cysteine proteases during tissue invasion of *Paragonimus westermani* newly excysted metacercariae. Cell Microbiol 8(6):1034-46 doi:CMI685 [pii] 10.1111/j.1462-5822.2006.00685.x

Norbury LJ, Beckham S, Pike RN, Grams R, Spithill TW, Fecondo JV, Smooker PM (2011) Adult and juvenile *Fasciola* cathepsin L proteases: different enzymes for different roles. Biochimie 93(3):604-11 doi:S0300-9084(10)00444-X [pii] 10.1016/j.biochi.2010.12.004

Nuamtanong S, Dekumyoy P, Adisakwattana P (2012) Evaluation of recombinant serine protease inhibitor from *Trichinella spiralis* for immunodiagnosis of swine trichinosis. Southeast Asian J Trop Med Public Health 43(5):1094-104

Pakchotanon P, Molee P, Nuamtanong S, Limpanont Y, Chusongsang P, Limsomboon J, Chusongsang Y, Maneewatchararangsri S, Chaisri U, Adisakwattana P (2016) Molecular characterization of serine protease inhibitor isoform 3, SmSPI, from *Schistosoma mansoni*. Parasitol Res doi:10.1007/s00436-016-5053-y 10.1007/s00436-016-5053-y [pii]

Pandey KC, Barkan DT, Sali A, Rosenthal PJ (2009) Regulatory elements within the prodomain of Falcipain-2, a cysteine protease of the malaria parasite *Plasmodium falciparum*. PLoS One 4(5):e5694 doi:10.1371/journal.pone.0005694

Pandey KC, Dixit R (2012) Structure-function of falcipains: malarial cysteine proteases. J Trop Med 2012:345195 doi:10.1155/2012/345195

Park H, Kim SI, Hong KM, Kim MJ, Shin CH, Ryu JS, Min DY, Lee JB, Hwang UW (2002) Characterization and classification of five cysteine proteinases expressed by *Paragonimus westermani* adult worm. Exp Parasitol 102(3-4):143-9

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8(10):785-6 doi:nmeth.1701 [pii] 10.1038/nmeth.1701

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16(6):276-7 doi:S0168-9525(00)02024-2 [pii]

Sansri V, Changklungmoa N, Chaichanasak P, Sobhon P, Meemon K (2013) Molecular cloning, characterization and functional analysis of a novel juvenile-specific cathepsin L of *Fasciola gigantica*. Acta Trop 128(1):76-84 doi:S0001-706X(13)00168-X [pii] 10.1016/j.actatropica.2013.06.013

Sansri V, Meemon K, Changklungmoa N, Kueakhai P, Chantree P, Chaichanasak P, Lorsuwannarat N, Itagaki T, Sobhon P (2015) Protection against *Fasciola gigantica* infection in mice by vaccination with recombinant juvenile-specific cathepsin L. Vaccine 33(13):1596-601 doi:S0264-410X(15)00176-0 [pii] 10.1016/j.vaccine.2015.02.010

Shin MH, Kita H, Park HY, Seoh JY (2001) Cysteine protease secreted by *Paragonimus westermani* attenuates effector functions of human eosinophils stimulated with immunoglobulin G. Infect Immun 69(3):1599-604 doi:10.1128/IAI.69.3.1599-1604.2001

Smith SM, Kane SE, Gal S, Mason RW, Gottesman MM (1989) Glycosylation of procathepsin L does not account for species molecular-mass differences and is not required for proteolytic activity. Biochem J 262(3):931-8

Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, Wandall HH, Levery SB, Clausen H (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J 32(10):1478-88 doi:emboj201379 [pii] 10.1038/emboj.2013.79

Steverding D, Sexton DW, Wang X, Gehrke SS, Wagner GK, Caffrey CR (2012) *Trypanosoma brucei*: chemical evidence that cathepsin L is essential for survival and a relevant drug target. Int J Parasitol 42(5):481-8 doi:S0020-7519(12)00080-X [pii] 10.1016/j.ijpara.2012.03.009

Sugiyama H, Morishima Y, Kameoka Y, Arakawa K, Kawanaka M (2004) *Paragonimus ohirai* metacercariae in crabs collected along the Arakawa River in Tokyo, Japan. J Vet Med Sci 66(8):927-31 doi:JST.JSTAGE/jvms/66.927 [pii]

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28(10):2731-9 doi:msr121 [pii] 10.1093/molbev/msr121

Turk V, Stoka V, Vasiljeva O, Renko M, Sun T, Turk B, Turk D (2012) Cysteine cathepsins: from structure, function and regulation to new frontiers. Biochim Biophys Acta 1824(1):68-88 doi:S1570-9639(11)00270-6 [pii] 10.1016/j.bbapap.2011.10.002

Varghese A, Raina OK, Nagar G, Garg R, Banerjee PS, Maharana BR, Kollannur JD (2012) Development of cathepsin-L cysteine proteinase based Dot-enzyme-linked immunosorbent assay for the diagnosis of *Fasciola gigantica* infection in buffaloes. Vet Parasitol 183(3-4):382-5 doi:S0304-4017(11)00501-2 [pii] 10.1016/j.vetpar.2011.07.032

Villa-Mancera A, Reynoso-Palomar A, Utrera-Quintana F, Carreon-Luna L (2014) Cathepsin L1 mimotopes with adjuvant Quil A induces a Th1/Th2 immune response and confers significant protection against *Fasciola hepatica* infection in goats. Parasitol Res 113(1):243-50 doi:10.1007/s00436-013-3650-6

Waikagul J (2007) A new species of Paragonimus (Trematoda: Troglotrematidae) from a cat infected with metacercariae from mountain crabs *Larnaudia larnaudii.* J Parasitol 93(6):1496-500 doi:10.1645/GE-1054.1

Warner AH, Pullumbi E, Amons R, Liu L (2004) Characterization of a cathepsin L-associated protein in *Artemia* and its relationship to the FAS-I family of cell adhesion proteins. Eur J Biochem 271(20):4014-25 doi:EJB4338 [pii] 10.1111/j.1432-1033.2004.04338.x

Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.

1. ผลงานตีพิมพ์ในวารสารวิชาการนานาชาติ

ชื่อผู้แต่ง : Tippayarat Yoonuan, Supaporn Nuamtanong, Paron Dekumyoy, Orawan Phupisut, Poom Adisakwattana*

ชื่อเรื่อง : Molecular and immunological characterization of cathepsin L-like cysteine protease of *Paragonimus pseudoheterotremus*

ส่งตีพิมพ์ในวารสาร Parasitology Research

และทางวารสารได้ส่งเรื่องกลับมาแล้ว ขณะนี้อยู่ในระหว่างการตอบคำถาม และ ปรับปรุงแก้ไข ตามข้อเสนอแนะจาก reviewers

2. การนำผลงานวิจัยไปใช้ประโยชน์

จะมีการทำวิจัยเพิ่มเติมในเรื่องของการวินิจฉัยโรคด้วยวิธี ELISA โดยใช้ cathepsin L-like cysteine protease of *Paragonimus pseudoheterotremus* เป็นแอนติเจนโดยทดสอบกับ IgG subclass ซึ่งหากได้ผลดีจะสามารถนำมาใช้ในหน่วยวินิจฉัยโรคพยาธิของภาควิชาฯ ทดแทนการใช้ crude antigen จาก *P. heterotremus*