



# HHS Public Access

Author manuscript

*Vet Parasitol.* Author manuscript; available in PMC 2020 March 01.

Published in final edited form as:

*Vet Parasitol.* 2019 March ; 267: 9–16. doi:10.1016/j.vetpar.2019.01.004.

## Characterization of a novel cathepsin L-like protease from *Taenia solium* metacestodes for the immunodiagnosis of porcine cysticercosis.

Nancy León-Janampa<sup>a</sup>, Ruddy Liendo<sup>a</sup>, Robert H. Gilman<sup>b</sup>, Carlos Padilla<sup>c</sup>, Hector H. García<sup>d</sup>, Armando Gonzales<sup>e</sup>, Patricia Sheen<sup>a</sup>, Mónica J. Pajuelo<sup>a,b</sup>, Mirko Zimic<sup>a,\*</sup>  
Cysticercosis Working Group in Perú

<sup>a</sup>Laboratorio de Bioinformática y Biología Molecular. Facultad de Ciencias y Filosofía. Universidad Peruana Cayetano Heredia, Lima, Peru

<sup>b</sup>Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, U.S.A.

<sup>c</sup>Instituto Nacional de Salud, Lima, Peru

<sup>d</sup>Cysticercosis Unit, Instituto de Ciencias Neurológicas, Lima, Peru

<sup>e</sup>School of Veterinary Medicine, Universidad Nacional Mayor de San Marcos, Lima, Peru

### Abstract

Porcine cysticercosis is an endemic parasitic disease caused by infection with *Taenia solium* that is found predominantly in developing countries. In order to aid in the development of simple diagnostic approaches, identification and characterization of potential new antigens for immunodiagnostic purposes is desired. The cysteine protease family has previously been found to have important immunodiagnostic properties. These proteases are expressed as zymogens which contain a signal peptide, pro-peptide, and an active domain. Subsequent catalytic cleavage of the pro-peptide converts these zymogens into enzymes. With the use of bioinformatic tools we

\*Corresponding author: Laboratory of Bioinformatics and Molecular Biology, Faculty of Science and Philosophy Universidad Peruana Cayetano Heredia. Av. Honorio Delgado 430, San Martín de Porres, Lima, Perú. mirko.zimic@upch.pe.

Author Informations:

Nancy León Janampa, Laboratorio de Bioinformática y Biología Molecular. Facultad de Ciencias y Filosofía. Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430, Urb. Ingeniería, San Martín de Porres. Lima, Peru.

Ruddy Liendo, Laboratorio de Bioinformática y Biología Molecular. Facultad de Ciencias y Filosofía. Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430, Urb. Ingeniería, San Martín de Porres. Lima, Peru.

Robert H. Gilman, Department of International Health, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, Maryland, USA.

Carlos Padilla, Laboratorio de Biotecnología y Biología molecular, Instituto Nacional de Salud, Cápac Yupanqui 1400 - Jesús María. Lima, Peru.

Hector H. García, Cysticercosis Unit, Instituto de Ciencias Neurológicas, Jirón Ancash 1271, Cercado de Lima 15003. Lima, Peru.  
Armando Gonzales, School of Veterinary Medicine, Universidad Nacional Mayor de San Marcos, Av. Circunvalación cdra. 28, San Bojiya. Lima, Peru.

Patricia Sheen, Laboratorio de Bioinformática y Biología Molecular. Facultad de Ciencias y Filosofía. Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430, Urb. Ingeniería, San Martín de Porres. Lima, Peru.

Mónica J. Pajuelo, Laboratorio de Bioinformática y Biología Molecular. Facultad de Ciencias y Filosofía. Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430, Urb. Ingeniería, San Martín de Porres. Lima, Peru.

Mirko Zimic, Laboratorio de Bioinformática y Biología Molecular. Facultad de Ciencias y Filosofía. Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430, Urb. Ingeniería, San Martín de Porres. Lima, Peru.

### Disclosure:

The authors have no conflict of interest to declare.

identified an active domain of a novel cathepsin L-like cysteine protease (TsolCL) in the *T. solium* genome. The TsolCL gene includes 705 nucleotides (nt) within a single intron and a 633 nt exonic sequence encoding an active protein of 211 amino acids. Sequence alignment and phylogenetic analysis suggest that the TsolCL gene is closely related to genes found in *Echinococcus granulosus* and *E. multilocularis*. In addition, TsolCL was found to have a 61.9% – 99.0% similarity to other cathepsin L proteins found in other helminths and mammals.

We cloned, expressed, purified, and characterized the recombinant active TsolCL (27 kDa) using the baculovirus-insect cell expression system. TsolCL showed cysteine protease enzymatic activity with the capacity to hydrolyze the Z-Phe-Arg-AMC substrate as well as bovine serum albumin. However, TsolCL was not able to hydrolyze human immunoglobulin. In addition, TsolCL has cathepsin L conserved amino acid residues in the catalytic site (Gln8, Cys14, His159, Asn179 and Trp181) and the motif GCNGG. Using ELISA, TsolCL was able to distinguish circulating IgG antibodies between healthy animals and naturally infected pigs with cysticercosis, showing a moderate sensitivity of 83.33% (40/48; 95% CI: [69.8% - 92.5 %]), and a specificity of 83.78% (31/37; 95% CI: [67.9% - 93.8%]).

In conclusion, a novel cathepsin L-like cysteine protease from a *T. solium* metacestode was expressed successfully in Baculovirus system and was evaluated as a candidate antigen to diagnose porcine cysticercosis using the ELISA immunoassay.

## Keywords

cathepsin L; cysticercosis; immunodiagnosis; *Taenia solium*; Baculovirus Expression Vector System

## 1. INTRODUCTION

*Taenia solium* is a cestode (flatworm) and a zoonotic parasite whose life cycle involves both humans and pigs. Humans are the definitive host because they are infected by adult tapeworms (taeniasis), while pigs, the intermediate host, are infected by larvae (cysticercosis).

Taeniosis/cysticercosis is distributed around the world, and is endemic in many regions with developing countries: Latin America, Sub-Saharan Africa, India, parts of China, and South East Asia (García et al., 2003; Winkler, 2012; Garcia et al., 2014a;). Porcine cysticercosis leads to economic losses for farmers, as infected meat cannot be officially sold. However, in rural areas infected pig carcasses may be illegally distributed and sold at a discounted price, an activity that contributes to the continued survival of the parasite (The Cysticercosis Working Group in Perú, 1993; Jayashi et al., 2012).

Parasitic transmission occurs via the fecal-oral route. When meat from pigs containing cysts are ingested by humans, the larvae attach to the mucosa of the small intestine. The parasite then uses the human's nutrition to grow until becoming an adult parasite, at which point they produce eggs containing an oncosphere, which are excreted in human feces. A parasite embryos, released from infective eggs, is developed into larva (cyst) in the body of the pig that previously ingested the eggs of the parasite. However, humans can also become

accidental intermediate hosts if they ingest food contaminated by feces and develop cysticercosis. The larval stage of *T. solium* typically becomes established within muscles, however, the more serious condition occurs when the larvae invade the central nervous system causing neurocysticercosis (NCC) (Garcia and Brutto, 2005).

The neurologic features of NCC are diverse and are dependent on various features including the number and location of cysts, as well the immune system of the host (Montano et al., 2005; Garcia et al., 2014b). NCC can cause headaches, epilepsy, hydrocephalus, psychiatric disorders, as well as death. It is possible to prevent human disease by diagnosing and vaccinating pigs due they form part of the life cycle of *T. solium* (Lightowlers, 2010). This is because by reducing cysticercosis in pigs, the contagion of *T. solium* in humans is also reduced.

In the literature there are many epidemiological studies of porcine cysticercosis reporting the use of different screening and serodiagnostic techniques. Examples include enzyme-linked immunoelectrotransfer Blot (EITB), tongue inspection, and Enzyme-linked Immunosorbent assay (ELISA) (Jayashi et al., 2017). Although EITB has been shown to be superior for diagnosis, tongue inspection is still the chosen technique for epidemiological surveys of porcine cysticercosis given that it is inexpensive with adequate sensitivity and a high specificity (Gonzalez et al., 1990).

Gonzalez et al. (1990), have previously evaluated the use of EITB for porcine cysticercosis serodiagnosis. They also compared EITB with both ELISA and tongue inspection. In this study tongue inspection was reported to have a sensitivity of 70% and a specificity of 100% whilst ELISA was found to have a sensitivity of 79% and a specificity of 75%. However, in contrast to the aforementioned results, Sciutto et al. (1998) later reported that antigen and antibody ELISA, EITB, and tongue inspection in fact had a lower sensitivity and specificity when used to test rurally reared infected pigs, especially in cases with a low burden of cysts.

The nature of antigens used in immunodiagnostic tests is diverse. In the last few years recombinant antigens as well as synthetic peptides that are glycoproteic in nature have been used successfully (Ferrer et al., 2012; Lescano et al., 2007). Cysteine proteases have an important role in the development and survival of parasites since it facilitates the penetration of the parasite in the host. Li et al. (2006) and Zimic et al. (2009) have proposed cathepsin L-like enzymes, from *T. solium*, as potential immunodiagnostic antigens and candidate vaccine targets.

Cysteine protease is usually synthesized as a precursor protein (zymogen), containing a signal peptide, pro-peptide, and mature catalytic domain. The signal peptide is cleaved off by signal peptidases following transportation. The pro-peptide acts as an intramolecular chaperone which in turn catalyzes the folding of the catalytic domain of the protease (Dutta et al., 2016). These pro-peptides contain highly conserved elements such as the consensus sequence ERFNIN motif that is present in the  $\alpha 2$  helix of cathepsin L in numerous species. GxNx $\alpha$ FD heptapeptide or GNFD is another conserved motif that is located at the kink of the  $\beta$ -sheet (Karrer et al., 1993). The Asp residue in the GNFD motif is believed to be essential for the correct processing of protease precursors (Vernet et al., 1995).

In a previous study, a partially purified cathepsin L-like fraction obtained from *T. solium* cyst fluid was tested in a Western blot and found to have a sensitivity of 96% and a specificity of 98% in NCC patients with multiple cysts. The ELISA assay using this same antigen was reported to be 92.7% specific and 98.0% sensitive in NCC patients with multiple cysts. This same method was demonstrated to have a sensitivity of 84.0% for the detection of NCC associated with a single cyst (Zimic et al., 2009).

Likewise, the first recombinant L-like cathepsin from *T. solium*, reported by Li et al. (2006), showed catalytic activity and antigenicity against the sera of patients with cysticercosis, sparganosis or fasciolosis, but weak or no antigenicity against the sera of patients with paragonimiasis or clonorchiosis, in western blot testing.

In this study we characterize and evaluate a novel active recombinant cathepsin L-like protease expressed in insect cells as a new potential immunodiagnostic antigen for porcine cysticercosis using samples from experimentally infected pigs.

## 2. MATERIALS AND METHODS

### 2.1. Parasites

*T. solium* (Tsol) metacestodes from a naturally infected pig were obtained and donated by The Cysticercosis Working Group (CWGP). This pig was from an endemic area in the central highlands of Peru and was confirmed to have cysticercosis via tongue inspection (Gonzalez et al., 1990). The metacestodes were collected after being washed in phosphates buffered saline (PBS, pH 7.4) and were immediately stored in RNA later solution (Ambion, Carlsbad, USA) at  $-70^{\circ}\text{C}$ .

### 2.2. Sequence analysis of TsoIcL

The first cathepsin L-like from *T. solium* was reported by Li et al. (2006). In this study, we evaluate a new CL for this parasite. With the use of the Blast2GO 4.1.9 bioinformatic tool, a TsoIcL sequence was identified in the *T. solium* genome sequence (Project ID: PRJNA183343)(Pajuelo et al., 2017). This sequence has been submitted to the GenBank<sup>TM</sup> database using Accession N° KX906691.1 (protein ID: AQQ11627.1). Alignment of the TsoIcL amino acid sequence with other cathepsin L-like cysteine protease sequences of other species was determined using MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0 and the online NCBI-Blast program (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Similarity percentage was determined using SMS software ([http://www.bioinformatics.org/sms2/ident\\_sim.html](http://www.bioinformatics.org/sms2/ident_sim.html)). Phylogenetic analysis of the aligned sequences was performed with the maximum likelihood method using MEGA6 with a 1000 replicates bootstrap. Conserved domains and active sites were found using Conserved Domain Search Service (CD Search) (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) (Marchler-bauer et al., 2011). Putative N-glycosylation sites were predicted with the NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>).

### 2.3. TsolCL amplification and cloning in a pFastBac HTA plasmid

Primers were designed using Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>), Primer Blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>), and OligoAnalyzer 3.1 (<https://eu.idtdna.com/calc/analyzer>) software; and *Nco* I and *Xho* I restriction sites (*forward* 5'-ATGCCATGGTAACAGGGGTGAAAGATCAG-3' and *reverse* 5'-CCGCTCGAGGATCAGAGGAT AGGAAGCTG-3') were included. Total RNA was extracted from the *T. solium* cysts using TRIzol (Invitrogen™ Life Technologies, Carlsbad, CA) according to the manufacturer's protocol. The spectrophotometer Nanodrop ND 2000 (Thermo Scientific, Delaware, USA) was used to determine RNA quality and concentration. cDNA was synthesized by reverse transcription using oligo d(T)s and RT-PCR amplification kits (Applied Biosystems, Foster, USA). The TsolCL cDNA was amplified by a conventional PCR reaction (1x buffer PCR, 0,2 mM dNTPs; 2 mM MgCl<sub>2</sub>; 10% dimetil sulphoxide (DMSO) (Merck, Darmstadt, Germany); primers (0.6 μM); Taq polymerase (Invitrogen™ Life Technologies) 0.025 U/μl and cDNA 50 ng/μl). The cycling PCR program profile comprised 35 cycles with denaturation at 94 °C for 30 s, annealing at 55 °C for 60 s, and extension at 72 °C for 90 s.

The amplified TsolCL coding sequence was purified from the gel using QIAquick® Gel Extraction Kit (QIAGEN, Center Mainz, Germany). TsolCL coding sequence and the pFastBac™ HT A vector were digested with *Nco* I and *Xho* I restriction endonucleases (RE) and ligated with T4 DNA ligase (New England Biolab, Ipswich, England). The *E. coli* DH5α competent cells were transformed with the ligation mixture and plated on LB agar containing ampicillin (100 μg/ml). The transforming positive clones (pFastBac HTA-TsolCL) were confirmed using PCR and double digestion with restriction enzymes.

### 2.4. Construction of the recombinant baculovirus genome-shuttle vector

*E. coli* DH10Bac™ cells (Invitrogen, Carlsbad, USA), which carry a plasmid containing the baculovirus genome (bacmid) and a plasmid encoding a transposase enzyme were transformed using pFastBac HTA-TsolCL to generate a recombinant bacmid DNA (baculovirus shuttle vector) using transposition. Transforming positive clones were selected on a plate of LB agar containing kanamycin (50 μg/ml), tetracycline (10 μg/ml), gentamicin (7 μg/ml), isopropyl b-D-1-thiogalactopyranoside (IPTG; 40 μg/ml), and X-gal (100 μg/ml). The plate was incubated at 37 °C for 48 hours. As part of this process the gene of interest was transposed into the Bacmid through lacZ gene disruption. The TsolCL sequence was inserted via Tn7 transposable elements immediately downstream of the polyhedrin promoter in the baculoviral DNA. The *E. coli* DH10Bac™ transformants containing the recombinant bacmid were determined by the blue/white-color colony screening method on a plate. To verify the presence of the gene of interest after transposition, high molecular weight DNA (bacmid-TsolCL) was isolated from an overnight culture and amplified by PCRs with pUC/M13 primers. These hybridize to sites flanking the mini-attTn7 site within the lacZα – complementation region, where the gene of interest is inserted. This means that PCR products from the transposition of recombinant bacmid-TsolCL generate a band of approximately 3.05 Kb which consists of 2420 bp from a pFastBac HTA vector and 636 bp from the TsolCL DNA sequence.

## 2.5. Transfection of *Sf9* cells and production of the recombinant baculovirus

Purified bacmid-TsolCL was used to transfect *Spodoptera frugiperda* (*Sf9*) insect cells. In a six-well plate *Sf9* cells were distributed with 90% confluency ( $1 \times 10^6$  cells/well) and incubated for one hour to allow cell attachment to the plate. The Cellfectin reagent and the bacmid-TsolCL DNA were diluted separately into 100  $\mu$ l of SF-900 SFM without antibiotics and combined to form lipid-DNA complexes, which were diluted down to 1 ml with SFM and laid over the washed *Sf9* cells. The cells were incubated for five hours at 27 °C, rinsed, and then incubated for another six to nine days. The culture medium was collected and stored as generation P1 viral stock. Another two rounds of viral stock amplifications were carried out to generate a high-titer of P3 viral stock. After three successive rounds of amplifications the virus titer (pfu/ml) was determined using an endpoint dilution plate assay with Neutral Red stain (Sigma-Aldrich, Saint Louis, USA). The recombinant baculovirus were amplified in cultures of *Sf9* cells at Multiplicity Optimal Infection (MOI) of 1 (one pfu to one *Sf9* insect cell). The amplified viral stock (P3) was used in subsequent steps of protein expression.

## 2.6. Recombinant TsolCL expression in *Sf9* insect cells and enzymatic activity

TsolCL expression was performed in a 200 ml suspension of *Sf9* insect cells ( $2 \times 10^6$  cells/ml) in a glass spinner (Coming, Lowell, USA) at 27 °C for six to nine days using a MOI of 20. Infected and uninfected *Sf9* cells were morphologically distinguishable. Uninfected cells continue dividing and formed a confluent monolayer whilst infected cells stopped dividing and were enlarged.

As the expressed recombinant TsolCL contained a 6-histidine tag at the N-terminal it was purified using HisTrap HP histidine-tagged protein purification columns (GE Healthcare Life Sciences, Uppsala, Sweden) according to the manufacturer's instructions. The recombinant TsolCL expressed by infected cells was confirmed using Western blot with mouse anti-histidine monoclonal antibodies.

The enzymatic activity of recombinant TsolCL (0.33 nmol) expressed in insect cells was evaluated fluorometrically using a commercial peptide substrate coupled with 7-amino-4-metilcoumarin (AMC) and Z-phe-Arg-AMC (MP Biomedicals Aurora, Ohio, USA). The peptide hydrolysis was conducted in citrate buffer (10 mM sodium citrate, pH 4.5) and 5 mM DTT incubated at 37 °C for 1, 2, and 24 h. To confirm this specific activity, inhibition of peptidase activity was tested with trans-Epoxy succinyl-L- Leucylamido-(4-Guanidino) Butano (E64). fluorescence was detected in the TD-700 fluorometer (Turner Designs, Sunnyvale, California, USA) at wavelengths of 360 nm (excitation) and 460 nm (emission).

In addition, the recombinant TsolCL was tested for its capacity to degrade human immunoglobulin G (IgG) and bovine serum albumin (BSA) such as Li et al. (2006) previously reported in the first cathepsin L-like (rTsCL-1) from *T. solium*. This reaction was conducted in different buffers: citrate buffer, phosphate buffer (10 mM sodium phosphate pH 6.5; 200 mM NaCl, 2 mM EDTA, 5 mM DTT), and acetate buffer (100 mM sodium acetate pH 4.5, 5 mM DTT), and incubated for 24 h at 37 °C. As a positive control each reaction was repeated with commercial papain cysteine protease (Sigma-Aldrich, St Louis, USA).



Negative control reactions were also undertaken by using nuclease free PCR water in place of the TsoIcL solution. Evidence of proteolytic activity was detected using SDS-PAGE electrophoresis.

## 2.7. Antigenicity of recombinant TsoIcL using sera from infected pigs with cysticercosis

Ninety two pig serum samples were provided by The Cysticercosis Working Group (CWGP) repository, but only 85 serum pig samples had complete necropsy and LLGP-EITB data: 48 positive samples were obtained from experimentally and naturally infected pigs that were LLGP-EITB positive; and 37 negative samples were obtained from pigs from industrialized farms from a non-endemic area that were LLGP-EITB negative (Vargas-Calla et al., 2016). For the infected pigs, the necropsy report included details about the total number of cysts and viable cysts found in the carcasses. As the serum samples used in this study were provided by the CWGP, the infection of the pigs, LLGP-EITB of sera and necropsy protocols are described in more details in previous publications (Gomez-Puerta et al., 2018; Moreno et al., 2012; Vargas-Calla et al., 2016).

ELISA immunoassays were performed for duplicated, using the porcine serum samples. Recombinant TsoIcL (1.2 µg) diluted in 0.2 M sodium carbonate-bicarbonate buffer (pH 9.2), was fixed in wells of Nunc MaxiSorp ELISA (Sigma-Aldrich, St Louis, USA) for 12 h at 4 °C. Then the wells were blocked with 3% BSA for 1 h at 37 °C. After, titres of 1/100, 1/200, 1/400, and 1/800 of pools of sera from infected pigs were added and incubated for 12 h at 4 °C. The immune complex antigen-antibody was recognized with 1/2500 dilution of rabbit anti-Pig IgG (H+L) secondary antibody, HRP (Thermo Lisher Scientific, Rockford, USA), and revealed with o-phenylenediamine dihydrochloride substrate (1 mg/ml). The reaction was stopped with 0.5 M HCl after five minutes and absorbance was read at a wavelength of 492 nm. Six negative samples and two strong positive samples were included in each plate for validation purposes. In order to allow comparability between plates, a percent positivity (PP) was calculated for each sample by dividing its mean optic density (OD) by the mean OD of the strong positive controls of the plate x 100 (Wright et al., 1993; Zimic et al., 2009).

## 2.8. Statistical analysis

With the PP of the tested sera used as a predictor, a logistic regression was performed to model whether or not the pig was infected. Sensitivity, specificity, and a Receiver Operating Curve (ROC) were then calculated for different cutoffs for the given logistic regression predictor (PP). Different combinations of sensitivity and specificity are available. The best combination was selected based on the maximization of the Youden's Index (sensitivity + specificity – 1). The correlation between parasite load (number of total or viable cysts) among positive pigs and PP was evaluated by calculating the Spearman's rank correlation coefficient. All analysis was performed using Stata 14 statistical software (Stata Corp., College Station, TX).

### 3. RESULTS

#### 3.1. Sequence analysis of TsolCL

The TsolCL nucleotide sequence was 712 bp in length, comprising one intron of 79 bp and a coding sequence of 633 bp which translates into a protein consisting of 211 amino acids. Two putative N-glycosylation sites, Asn 90 and Asn 96, were predicted. CD Search showed that our protein belongs the peptidase Superfamily and the Papain family cysteine protease (2-210 aminoacids sequence). The sequence alignment confirmed a GCNGG motif that was found within the mature functional domain (Figure 1). The typical ERFNIN and GNFD conserved domains found in other cathepsin L proteins were not identified in our amino acid sequence because they are only present in the pro-peptide form which was not evaluated in this study. The amino acid residues crucial for active site formation were found in the TsolCL sequence. These include the nucleophile CysX, the proton donor His159, Gln8 that forms the oxi-anionic site that is related to the tetrahedral intermediate, Asn179 that is involved in the orientation of the His159 ring, and Trp181. Another important and conserved domain found in TsolCL was the S2 which is related to the specificity of protease hydrolysis. The S2 domain includes mainly hydrophobic residues (58 Leu, 59 Met, 129 Ala, 157 Leu, 160 Gly, and 205 Ala).

Multiple sequence alignment of the TsolCL mature domain with the cathepsin L-like cysteine protease of other helminths and mammals revealed a similarity of 61.9% – 99.0% (Figures 1). Similarity with the previously reported cathepsin L-like sequences from *F. hepatica* (77.8%), *T. solium* (75.5%), and *T. pisiformis* (75.9%) was found; as well as a marked similarity with the cathepsin L-like sequences from other cestodes such as *E. granulosus* (99%), *E. multilocularis* (98.6%), *Hymenolepis microstoma* (91.9%), *Clonorchis sinensis* (83%), and *Schistosoma japonicum* (82%).

Phylogenetic analysis showed that TsolCL shares a more inclusive clade with cathepsin L-like cysteine proteases of cestode and trematode parasites such as *E. granulosus*, *E. multilocularis*, *H. microstoma*, *C. sinensis*, and *S. japonicum* than with those from nematodes or mammals. Interestingly, the phylogenetic tree also shows that the TsolCL reported in this study is evolutionarily distant from the cathepsins L-like proteases reported for other taenias. This creates a distinct clade together with other cestodes confirming that this TsolCL is a novel cathepsin L-like protease of *T. solium* (Figure 2).

#### 3.2. Expression and enzymatic activity of TsolCL

According to the plate assay the amount of baculovirus produced by infected cells was  $6 \times 10^7$  pfu/ml viral titre after amplifying P3 virus stock in a *Sf9* suspension cell culture.

The molecular weight of the purified TsolCL was approximately 27 kDa (Figure S1 A, B and C). The largest amount of recombinant TsolCL was found in the cytoplasm of infected insect cells. After purification, 565 µg/ml of recombinant TsolCL expressed in 200 ml of infected insect cells (>2,5 mg/1L cells culture) was obtained.

We confirmed that the purified recombinant TsolCL had catalytic peptidase activity. Papain and TsolCL were able to hydrolyze the specific Z-Phe-Arg-AMC fluorogenic peptide



substrate. The enzymatic activity was highest in citrate buffer at a pH of 5.2 and was substantially inhibited by 0.1 mM E64. Non-infected insect cells were not evaluated because their catalytic activity was determined from a purified recombinant TsolCL using affinity chromatography.

TsolCL was able to degrade BSA in phosphate buffer at a pH of 6.5 and citrate buffer at pH of 6.5. The catalytic reaction was evidenced by SDS PAGE (Figure 3 A, B and C). TsolCL was not able to degrade human immunoglobulins. The proteolytic activity of commercial papain (positive control cysteine protease) was observed against BSA and IgG in a citrate buffer with a pH of 5.2, an acetate buffer with a pH of 4.5, and a phosphate buffer with a pH of 6.5.

### 3.3. Antigenicity of recombinant TsolCL with sera from pigs infected with cysticercosis

For immunological tests, 85 sera from pigs were evaluated: 48 positive and 37 negative sera. The positive sera used in this study belonged to pigs that harbored between 1 - 4069 cysts: 25% of pigs (12/48) with degenerated cysts, 14.6% (7/48) with viable cysts and 60.4% (9/48) with viable and degenerated cysts. It was observed that there was not a direct relationship between the parasitic load and the antibody titer against cysticercosis. This means that a positive serum sample, from a pig with a high parasitic burden, did not result in a high PP in ELISA test.

In the ELISA (TsolCL ELISA), the highest titer of pooled positive pig sera that differentiated healthy animals from those infected with cysticercosis was 1/400 (Figure 4). A cutoff in the PP of 90.47% resulted in a classification of 46 positive samples and 39 negative samples. The resulting test sensitivity was 83.33% (40 of 48; 95% CI: [69.78% - 92.52%]), and a specificity of 83.78% (31 of 37; 95% CI: [67.99% - 93.81%]). The area under the ROC was 0.90 (95% CI: [0.83–0.96]) (Figure 5, Table S1). Among positive pigs, no correlation was found between PP and total parasite load ( $\rho = -0.10$ ,  $p=0.525$ ) or between PP and the number of viable cysts ( $\rho = 0.11$ ,  $p=0.482$ ).

## 4. DISCUSSION

In this study we report a novel cathepsin L-like cysteine protease, TsolCL, that is found in mature *T. solium* metacestodes. This TsolCL exhibits the ability to hydrolyze the highly specific Z-Phe-Arg-AMC peptide with its activity being inhibited by E64. In addition, the reported TsolCL was able to hydrolyze BSA but not human IgG. Furthermore, TsolCL was found to have promising features for use as an antigen in the immunodiagnosis of porcine cysticercosis.

The literature reports the presence of cysteine proteases in multiple parasite species (Caffrey et al., 2018). These proteins are involved in physiological processes related to the development and survival of these parasites (Grote et al., 2018). Examples of these processes include immunoevasion, enzyme activation, virulence, as well as tissue and cellular invasion. Due to the important role of cysteine proteases, they have been considered as potential targets for chemotherapy and immunoprophylaxis (Sajid and Mckerrow, 2002; Li et al., 2006; Cortez et al., 2009; Zimic et al., 2009; Wang et al., 2013).

Within cells cysteine protease is translated, producing a zymogen. Here it consists of a signal peptide, a pro-peptide, and a mature catalytic domain. The pro-peptide has an important role in the correct folding and spatio-temporal regulation of the enzyme's proteolytic activity (Dutta et al., 2016). In addition, pro-peptides contain the highly conserved consensus sequences of ERFNIN (Karrer et al., 1993) and GNFD. In contrast, here we report a TsoLCL that only contains the mature catalytic domain. Thus, as a result of the absence of the pro-peptide in the amino acid sequence of TsoLCL, the conserved ERFIN and CGNFD motifs are not present. According to Vernet et al. (1995) the Asp residue in the GNFD motif appears to be essential for the correct processing of the protease precursor. However, only the GCNGG motif was present in the catalytic domain of TsoLCL.

In a previous study, Li et al. (2006) reported the first recombinant cathepsin L-like cysteine protease from the cestode parasite *T. solium* (TsCL-1, Genbank ID: AAS00027.1). The cathepsin L-like cysteine protease, TsoLCL, reported by us in this study is the second enzyme of its class to be reported for *T. solium*. We have found that TsoLCL has a 58.5% identity match and 75.5% similarity with the protease reported by Li et al. Therefore, it can be concluded that these two sequences constitute paralogous genes. The TsCL-1 identified by Li et al. has 216 amino acids, a molecular weight of 28 kDa, and is reported to be involved in cysticerci host invasion. This recombinant protein was expressed first in an *E. coli* system, however, it did not possess proteolytic enzyme activity. Therefore, Li et al. expressed the cathepsin L-like cysteine protease in a *P. pastoris* eukaryotic expression system. In our study, however, the recombinant TsoLCL was expressed in a Baculovirus-Insect cell System (BEVS).

The fact that *T. solium* is a eukaryotic organism means that post-translation modifications are important for the proper enzymatic function and immunogenicity of active *T. solium* proteases. The BEVS enables the rapid and efficient expression of recombinant glycosylated proteins in insect cells (Jarvis, 2009). This system includes a modified baculovirus genome (~130 kb/bacmid), derived from wild-type *Autographa californica* multicapsid nuclear polyhedrosis virus (AcMNPV), where a foreign gene can be introduced by site-specific Tn7 transposition (Trowitzsch et al., 2010) in DH10BAC *E. coli*.

According to the prediction generated by SnapGene<sup>R</sup> software ([https://www.snapgene.com/products/snapgene/free\\_trial/](https://www.snapgene.com/products/snapgene/free_trial/)), the molecular weight of linear TsoLCL was calculated to be 22.5 kDa. However, following the expression of TsoLCL in insect cells, we found that purified TsoLCL had a molecular weight of approximately 27 kDa. This difference could be attributed to glycosylation that may occur as part of post-translational modifications. There are multiple studies in support of this theory, reporting that post-translational modifications of recombinant proteins obtained via expression in insect cells using BEVS are frequently related to glycosylation (Demain and Vaishnav, 2009; Jarvis, 2009). These modifications do not appear to alter the function of the recombinant protein from that of the native protein, which is an important observation as the glycosylation of recombinant proteins can alter their activity, function, and antigenicity (Brooks, 2004). A considerable benefit of the BEVS is that it is a non-pathogenic system for vertebrates and plants which consequently ameliorates biosafety issues (Jarvis, 2009).

As a result of the high similarity of the TsolCL amino acid sequence to the L-like cathepsins of *E. granulosus*, *E. multilocularis*, *S. japonicum*, and *F. hepatica*, cross-reactions are likely to occur in immunological tests where antibody detection is performed using serum of pigs infected by any one of these parasites (Iburg et al., 2007). *E. granulosus* infection is typical in dogs and sheep, and *E. multilocularis* (Conraths and Deplazes, 2015) in red foxes. However, there is a broad spectrum of accidental hosts for these parasites including pigs. For example, Sánchez et al. (2012) identified the presence of secondary hydatid cysts from *E. granulosus* in alpacas and pigs. This highlights the importance of cross-reaction evaluation. Furthermore, additional markers (candidate antigens from *T. solium*) are needed in addition to TsolCL in order to increase the specificity of the proposed immunodiagnostic test. However, the markers would need to be carefully selected in order to preserve a reasonable sensitivity.

The similarity of TsolCL with cathepsin L-like cysteine proteases from other cestodes, such as *E. granulosus*, suggests that TsolCL may be homologous to these enzymes or that both enzymes belong to the same family of cathepsins. Analyzing the phylogenetic relationships, Robinson et al. (2008) classified cathepsin L-like cysteine protease from *F. hepatica* into five separate families (L3 and L4 in early infective larvae and L1, L2, and L5 in adult parasites). A similar analysis could be done with the cathepsin L-like cysteine protease of *T. solium*, however, there are currently not enough reported sequences. Nonetheless, it is clear that the TsolCL we report in our study belongs to a family that is distinct to that of the *T. solium* cathepsin L-like cysteine protease reported by Li et al. (2006).

The TsolCL we expressed in insect cells showed proteolytic activity and was able to degrade the Z-Phe-Arg-AMC substrate and BSA, but not immunoglobulin, and was also found to be inhibited by E64, confirming the cysteine/cathepsin L-like nature of the recombinant protein. The observed limited catalytic activity of TsolCL could be due to inadequate folding, or because TsolCL could have a different role during the parasitic invasion of host tissue. Further studies are required to clarify the cause of this observed limited activity.

We have previously reported a partially purified fraction from the fluid of *T. solium* cysts with cathepsin L activity (Zimic et al., 2009). This fraction included dominant proteins of 53 and 25 kDa and revealed a sensitivity and specificity of 96% and 98%, respectively, for the diagnosis of human cysticercosis using a Western immunoblot. In addition a sensitivity of 98.0% and a specificity of 92.7% were recorded for an ELISA and dot-ELISA assay of these same proteins. The recombinant TsolCL we have identified was found to have antigenicity against the sera from pigs infected with cysticercosis with a sensitivity of 83.3% and a specificity of 83.7% when tested using an ELISA assay. Even though the aforementioned marker did not prove to be better than existing antigens, we propose their combined use could improve sensitivity and specificity of immunological tests. However, further studies are still required to evaluate cross reactions with sera from pigs infected with other parasites such as *T. asiatica*, *E. granulosus* and *E. multilocularis* due to the similarities they share with *T. solium*.

## 5. CONCLUSION:

In summary, we report the identification and expression of a novel cathepsin L-like cysteine protease, TsolCL, with proteolytic activity, from a *T. solium* metacestode. TsolCL is a candidate for use in combination with other antigens for the improvement of porcine cysticercosis immunodiagnosis.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements:

We would like to thank Sebastian Loli (BSc) for identifying the TsolCL sequence in the *T. solium* genome and for designing TsolCL primers. We would also like to thank Nicholas Rich, Dr Nicolas Rey de Castro and Emma Carter for proofreading and assisting with the language editing of this manuscript.

### Financial Support:

This study was supported by the Programa Nacional de Innovación para la Competitividad y Productividad - Innóvate Peru under the contract 181-FINCYT-IB-2013, The World Academy of Science (TWAS) No.14-233RG-BIO-LA UNESCO FR324028597 award, NIH training grant TW001140, D43TW010074 and TMRC U19AI129909 and by the cysti TMRC grant.

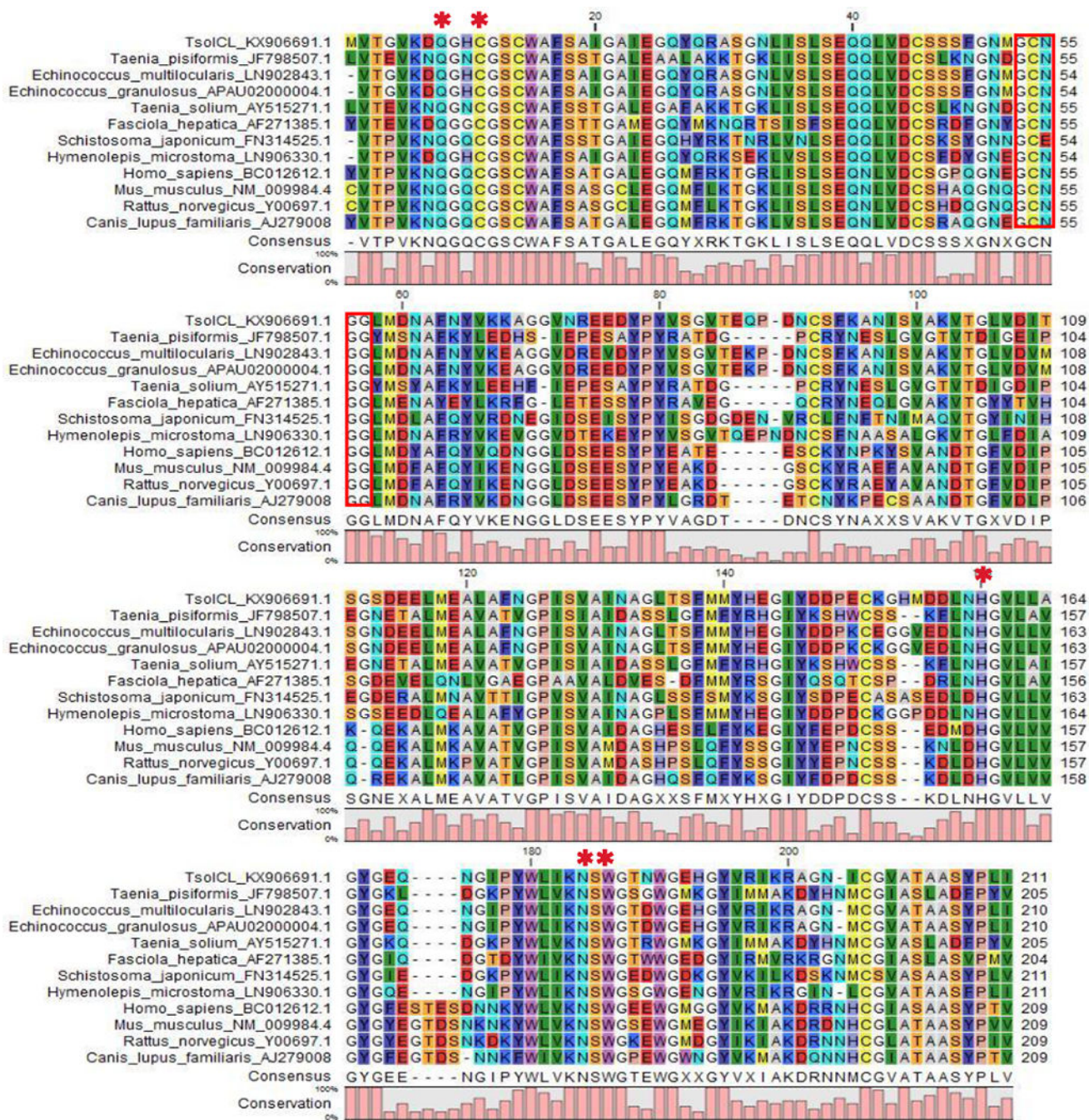
## REFERENCES

- Brooks SA, 2004 Appropriate glycosylation of recombinant proteins for human use: implications of choice of expression system. *Mol. Biotechnol* 28, 241–255. [PubMed: 15542924]
- Caffrey CR, Goupil L, Rebello KM, Dalton JP, Smith D, 2018 Cysteine proteases as digestive enzymes in parasitic helminths 1–20.
- Conraths FJ, Deplazes P, 2015 *Echinococcus multilocularis*: Epidemiology, surveillance and state-of-the-art diagnostics from a veterinary public health perspective. *Vet. Parasitol* 213, 149–161. [PubMed: 26298509]
- Cortez AP, Rodrigues AC, Garcia HA, Neves L, Batista JS, Bengaly Z, Paiva F, Teixeira MMG, 2009 Cathepsin L-like genes of *Trypanosoma vivax* from Africa and South America – characterization, relationships and diagnostic implications. *Mol. Cell. Probes* 23, 44–51. [PubMed: 19063960]
- Demain AL, Vaishnav P, 2009 Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv* 27, 297–306. [PubMed: 19500547]
- Dutta S, Choudhury D, Roy S, Dattagupta JK, Biswas S, 2016 Mutation in the pro-peptide region of a cysteine protease leads to altered activity and specificity — A structural and biochemical approach. *PLoS One* 11, 1–19.
- Ferrer E, Sánchez J, Milano A, Alvarez S, La R, Lares M, Miguel L, Milagros M, Davila I, Harrison LJS, Parkhouse RME, Garate T, 2012 Diagnostic epitope variability within *Taenia solium* 8 kDa antigen family: Implications for cysticercosis immunodetection. *Exp. Parasitol* 130, 78–85. [PubMed: 22075212]
- Garcia HH, Brutto O.H. Del, 2005 Neurocysticercosis: updated concepts about an old disease. *Lancet Neurol* 4, 653–661. [PubMed: 16168934]
- García HH, Gilman RH, Gonzales A, Verastegui M, Rodriguez S, Gavidia CM, Tsang V, Falcon N, Lescano A, Moulton LH, Bernal T, Tovar M, Perú TCWG in, 2003 Hyperendemic human and porcine *Taenia solium* infection in Perú. *Am J Trop Med Hyg* 68, 268–275.
- Garcia HH, Nash TE, Brutto O.H. Del, 2014a Clinical symptoms, diagnosis, and treatment of neurocysticercosis. *Lancet Neurol.* 13, 1202–1215. [PubMed: 25453460]

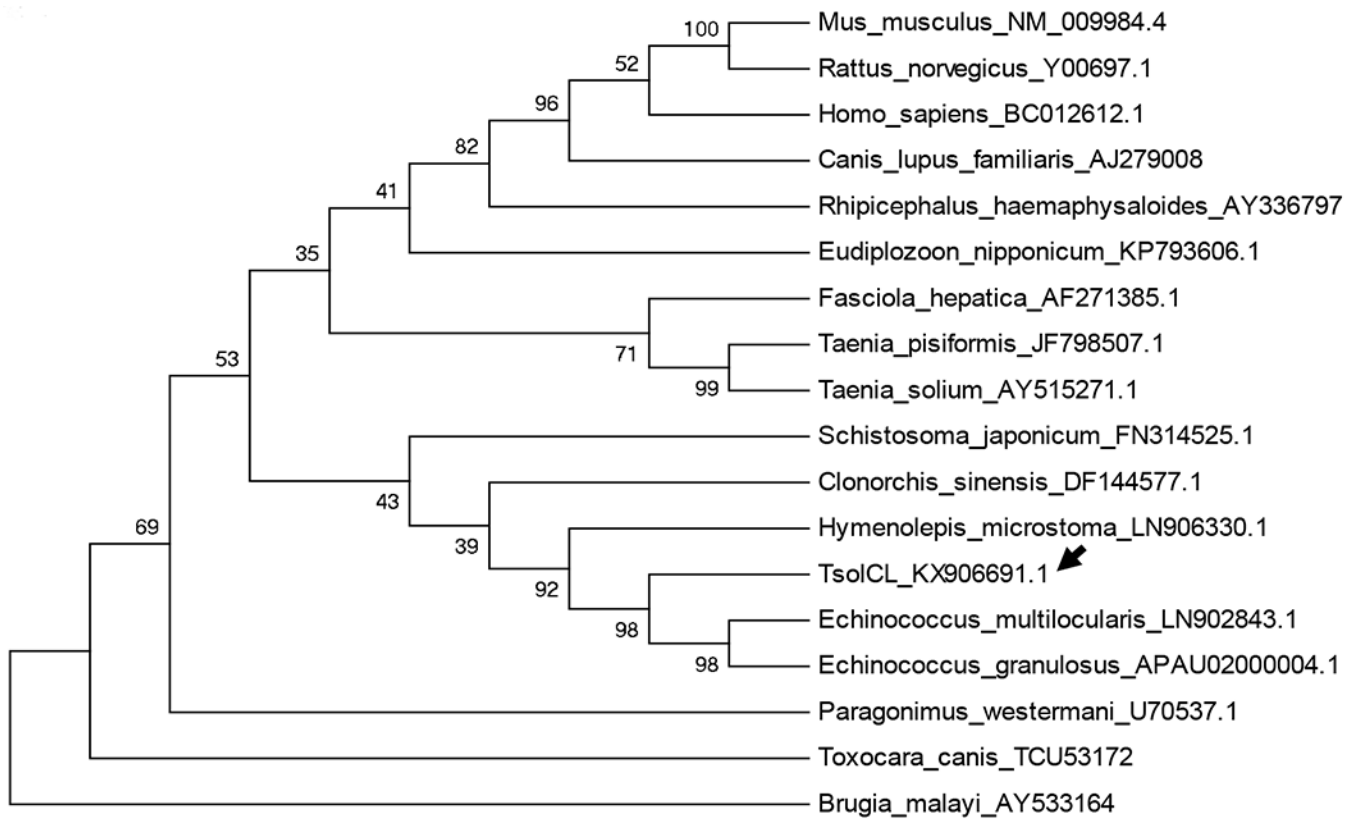
- Garcia HH, Rodriguez S, Friedland JS, The Cysticercosis Working Group in Perú, 2014b Immunology of *Taenia solium* Taeniasis and human cysticercosis. *parasite Immunol.* 36, 388–396. [PubMed: 24962350]
- Gomez-Puerta LA, Garcia HH, Gonzalez AE, 2018 Experimental porcine cysticercosis using infected beetles with *Taenia solium* eggs. *Acta Trop.* 183, 92–94.
- Gonzalez A, Cama V, Gilman R, Tsang V, Pilcher J, Chavera A, Castro M, Montenegro T, Verastegui M, Miranda E, 1990 Prevalence and comparison of serologic assays, necropsy, and tongue examination for the diagnosis of porcine cysticercosis in Peru. *Am J Trop Med Hyg* 43, 194–199. [PubMed: 2389823]
- Grote A, Caffrey CR, Rebello KM, Smith D, Dalton JP, Lustigman S, 2018 Cysteine proteases during larval migration and development of helminths in their final host. *PLoS Negl. Trop. Dis*
- Iburg T, Johansen MV, Leifsson PS, Willingham AL, Lindberg R, 2007 Hepatic Changes in Congenital *Schistosoma japonicum* Infections in Pigs. *J comp Path* 136, 250–255. [PubMed: 17459408]
- Jarvis D, 2009 Baculovirus-insect cell expression systems. *Methods Enzym.* 463, 191–222.
- Jayashi CM, Gonzalez AE, Castillo R, Rodríguez S, Garcia HH, Lightowlers MW, Cysticercosis Working Group in Peru, 2017 Validity of the Enzyme-linked Immunoelctrotransfer Blot (EITB) for naturally acquired porcine cysticercosis. *Vet. Parasitol* 199, 42–49.
- Jayashi M, Arroyo G, Lightowlers MW, Gonzalez AE, 2012 Seroprevalence and risk factors for *Taenia solium* cysticercosis in rural pigs of northern Peru. *PLoS Negl. Trop. Dis* 6, 1–5.
- Karrer KM, Peiffert SL, Ditomas ME, 1993 Two distinct gene subfamilies within the family of cysteine protease genes. *Proc. Natl. Acad. Sci. U. S. A* 90, 3063–3067. [PubMed: 8464925]
- Lescano A, Garcia H, Gilman R, Guezala M, Tsang V, Gavidia CM, Rodriguez S, Moulton LH, Green JA, 2007 Swine cysticercosis hotspots surrounding *Taenia solium* Tapeworm Carriers. *Am. J. Trop. Med. Hyg* 76, 376–383. [PubMed: 17297051]
- Li A, Moon S, Park Y, Na B, Hwang M, Oh C, Cho S, Kong Y, Kim T-S, Chung P-R, 2006 Identification and characterization of a cathepsin L-like cysteine protease from *Taenia solium* metacestode. *Vet. Parasitol* 141, 251–259. [PubMed: 16872751]
- Lightowlers MW, 2010 Eradication of *Taenia solium* cysticercosis : A role for vaccination of pigs. *Int. J. Parasitol* 40, 1183–1192. [PubMed: 20470777]
- Marchler-bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, Deweese-scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH, 2011 CDD : a Conserved domain database for the functional annotation of proteins. *Nucleic Acids Res.* 39, 225–229. [PubMed: 20823090]
- Montano S, Villaran M, Ylquimiche L, Figueroa J, Rodriguez S, Bautista C, Gonzalez A, Tsang V, Gilman R, Garcia EL, Group CW, 2005 Neurocysticercosis association between seizures, serology, and brain CT in rural Peru. *Neurology* 65, 229–233. [PubMed: 16043791]
- Moreno L, Lopez-Urbina MT, Farias C, Domingue G, Donadeu M, Dungu B, García HH, Gomez-Puerta LA, Lanusse C, González AE, 2012 A high oxfendazole dose to control porcine cysticercosis: Pharmacokinetics and tissue residue profdes. *Food Chem. Toxicol* 50, 3819–3825. [PubMed: 22841955]
- Sajid M, Mckerrow JH, 2002 Cysteine proteases of parasitic organisms. *Mol. Biochem. Parasitol* 120, 1–21. [PubMed: 11849701]
- The Cysticercosis Working Group in Perú, 1993 The marketing of cysticercotic pigs in the Sierra of Peru. *Bull. World Health Organ.* 71, 223–228. [PubMed: 8490986]
- Trowitzsch S, Bieniossek C, Nie Y, Garzoni F, Berger I, 2010 New baculovirus expression tools for recombinant protein complex production. *J. Struct. Biol* 172, 45–54. [PubMed: 20178849]
- Vargas-Calla A, Gomez-Puerta LA, Calcina J, Gonzales-Viera O, Gavidia C, Lopez-Urbina MT, Garcia HH, Gonzalez AE, 2016 Evaluation of activity of triclabendazole against *Taenia solium* metacestode in naturally infected pigs. *Asian Pac. J. Trap. Med* 9, 23–26.
- Vernet T, Bertis P, de Montigny C, Musil R, Tessier D, Ménard R, Magny M, Storer A, Thomas D, 1995 Processing of the papain precursor. The ionization state of a conserved amino acid motif within the Pro region participates in the regulation of intramolecular processing. *J. Biol. Chem* 270, 10838–10846. [PubMed: 7738022]

- Wang Q, Zhang S, Luo X, Hou J, Zhu X, Cai X, 2013 Cloning and characterization of a cathepsin L-like cysteine protease from *Taenia pisiformis*. *Vet. Parasitol* 194, 26–34. [PubMed: 23411373]
- Winkler AS, 2012 Neurocysticercosis in sub-Saharan Africa: a review of prevalence, clinical characteristics, diagnosis, and management. *Pathog. Glob. Heal* 106, 261–274.
- Wright P, Nilsson E, Van Rooij E, Lelenta M, Jeggo M, 1993 Standardisation and validation of enzyme-linked immunosorbent assay techniques for the detection of antibody in infectious disease diagnosis. *Rev Sci Tech.* 12, 435–450. [PubMed: 8400384]
- Zimic M, Pajuelo M, Rueda D, Lopez C, Arana Y, Castillo Y, Calderón M, Rodríguez S, Sheen P, Vinetz JM, Gonzalez AE, Garcia HH, Gilman RH, 2009 Utility of a Protein Fraction with Cathepsin L-Like Activity Purified from Cysticercus Fluid of *Taenia solium* in the Diagnosis of Human Cysticercosis. *Am. J. Trop. Med. Hyg* 80, 964–970.



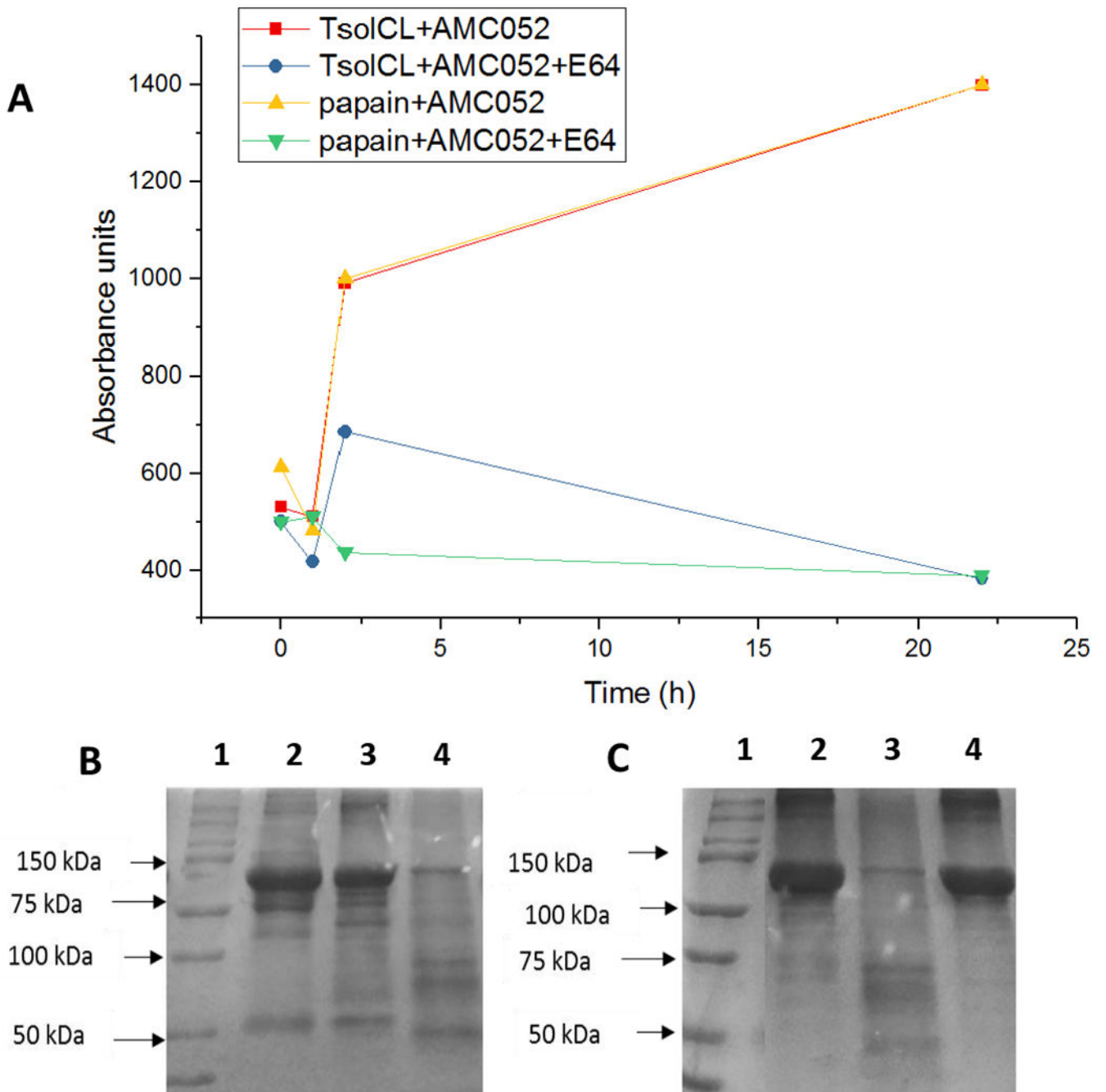


**Figure 1:** Multiple sequence alignment of the Tso1CL amino acid sequence with cathepsin L cysteine proteases from other species using CLC Sequence Viewer software. The red box shows the conserved motif GCNNGG from cathepsin L-like cysteine protease. Asterisks (\*) represent the residues from the protease active site (Gln 8, Cys 14, His 159, Asn 179, and Trp 181).



**Figure 2:**

Phylogenetic tree based on the deduced amino acid sequence of TsoLCL and cathepsin L-like cysteine protease from other species. The unrooted maximum likelihood tree with 1000 bootstrap replicates was constructed using the MEGA6 program. Each sequence presents its accession number together with the name of the species. The arrow indicates the TsoLCL protease reported in this study.



**Figure 3:**

Catalytic activity of purified TsolCL expressed in insect cells. (A) TsolCL catalytic activity curves on the specific substrate AMC052. It is observed that the catalytic activity of TsolCL and papain increases over time. However, this is inhibited when using E64. TsolCL activity using BSA, in (B) citrate buffer with a pH of 5.2 (Lanes 1: protein molecular weight marker, 2: BSA + TsolCL + E64 reaction, 3: BSA, and 4: BSA + TsolCL reaction) and (C) phosphate buffer pH 6.5 (Lanes 1: protein molecular weight marker; 2: BSA, 3: BSA + TsolCL reaction, and 4: BSA + TsolCL + E64 reaction). It is seen that TsolCL can degrade



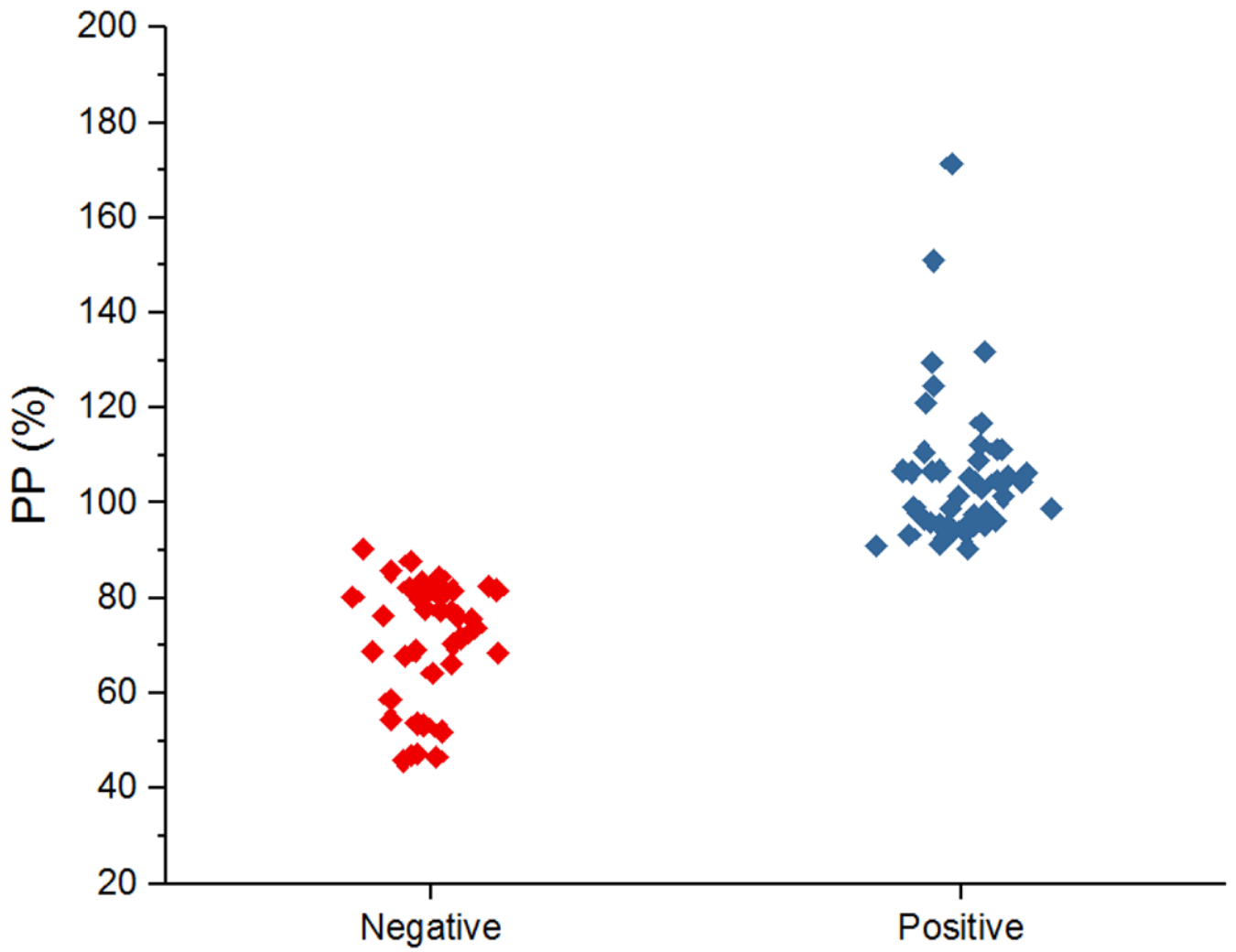
BSA in citrate buffer and phosphate buffer at acidic pHs. However, note that TsoLCL was unable to degrade human IgG (data not shown).

Author Manuscript

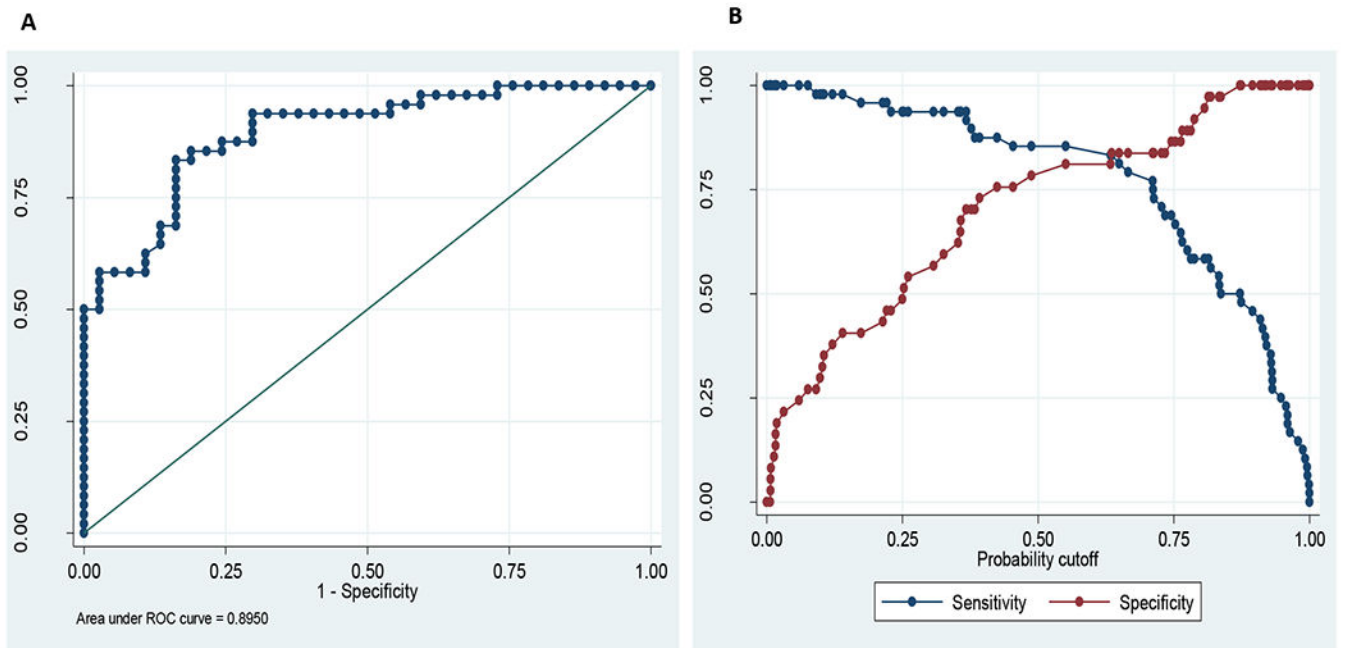
Author Manuscript

Author Manuscript

Author Manuscript



**Figure 4:** Plot representing the PP (cut-off: 90.4%) of positive and negative results, obtained from TsolCL ELISA test. Using sera from non-infected and infected pigs with cysticercosis in a 1/400 dilution.



**Figure 5:** Statistical analysis. (A) ROC curve for TsoLCL ELISA test, ratio of true positive versus false positive rate with area under ROC curve of 0.895. (B) Probability cutoff between sensitivity (83.33%) and specificity (83.78%) of TsoLCL ELISA test.