Parasitology Research (2019) 118:1581–1592
https://doi.org/10.1007/s00436-019-06271-4

PROTOZOOLOGY - ORIGINAL PAPER

Cathepsin L—a novel cysteine protease from *Haemaphysalis flava* Neumann, 1897

Yali Sun[1,2] · Lan He[1,2] · Long Yu[1,2] · Jiaying Guo[1,2] · Zheng Nie[1,2] · Qin Liu[1,2] · Junlong Zhao[1,2]

Received: 25 April 2018 / Accepted: 14 February 2019 / Published online: 2 March 2019

Abstract

Ixodid ticks are ectoparasites responsible for the transmission of a large number of bacterial, viral, and protozoan pathogens to animals and humans. As long-term blood-pool feeders, the digestion of host blood is critical to their development as well as to the establishment of the sexual cycle of hemoparasites such as *Babesia* parasites, the agents of human and animal babesiosis. Previous studies have demonstrated that cysteine proteases are involved in blood digestion, embryogenesis, and pathogen transmission in other species of ticks, but their characteristics and functions are still unidentified in *Haemaphysalis flava*. Here, we describe the characterization of a cysteine protease *HfCL* from *H. flava*. We show that *HfCL* belongs to the L-like papain family of proteases, exhibits high expression in nymphs and adults, and localizes to both the midgut and salivary glands. Biochemical assays using purified recombinant enzyme reveal that *rHfCL* can hydrolyze the fluorogenic substrate Z-phe-Arg-MCA with optimal activity detected at pH 6. Furthermore, the short-term growth assay indicates that *rHfCL* can inhibit the intraerythrocytic development of *Babesia microti* and *Babesia gibsoni* in vitro.

Keywords Cathepsin L · Papain · *Haemaphysalis flava* · *Babesia microti* · *Babesia gibsoni*

Introduction

The hard tick *Haemaphysalis flava* is an obligate hematophagous ectoparasite that belongs to arthropoda, Ixodidae. During its long-term blood feeding on mammals, the tick can cause skin and vascular damage to its hosts and can trigger strong inflammatory and immune responses (WonKoo et al. 1997; Xu et al. 2016). It is also during this process that the tick can transmit serious public pathogens such as *Rickettsia*, *Ehrlichia*, *Borrelia*, and protozoans (Fournier et al. 2002; Otsuka 1976; Rar et al. 2010; Moon et al. 2013; Yoon et al. 1996), which can cause severe losses in animal production and health problems for humans (Fuente et al. 2017; Kern et al. 2011). In contrast to short-term blood feeders like mosquitoes, ticks are long-term blood feeders with continuous blood feeding for several days as well as a complex life cycle, thus posing great challenges to the prevention and control of tick-borne diseases (Akov 1982; Aksenova et al. 1976).

As ticks depend on host blood as the primary food source for survival, their midgut is considered to be involved in blood degradation (Smit et al. 1977). Similar to that of other blood-sucking endoparasites like *Schistosoma*, *Haemonchus*, and *Plasmodium* (Delcroix et al. 2006; Rosenthal 2004; Williamson et al. 2003), the digestive mechanism of the midgut begins with the fusion of primary lysosomes (digestive vesicles) by endocytosis to form secondary lysosomes, various free dipeptides, and amino acids. Therefore, the digestive proteolytic system during blood feeding is dependent on both individual hemoglobin enzymes and multienzyme networks regulated by related cysteine proteinases (Franta et al. 2010; Horn et al. 2009).

Previous studies have shown that cysteine proteases not only affect the survival of ticks and their blood digestion rate

Section Editor: Norbert Becker

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00436-019-06271-4) contains supplementary material, which is available to authorized users.

✉ Lan He
helan@mail.hzau.edu.cn

✉ Junlong Zhao
zhaojunlong@mail.hzau.edu.cn

[1] State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine Huazhong Agricultural University, Wuhan 430070, Hubei, China

[2] Key Laboratory of Animal Epidemical Disease and Infectious Zoonoses, Ministry of Agricultural, Huazhong Agricultural University, Wuhan 430070, Hubei, China

(Yamaji et al. 2013) but also play a pivotal role in the development of the embryo as well as the molting, egg-laying, and growth of ticks (Daniel Sojka et al. 2013; Grandjean 1984). Additionally, the pathogens acquired by ticks from blood meals may alter the expression of midgut proteins and migrate to salivary glands to maintain their proliferation, followed by transmission to the next host by blood feeding (Francischetti et al. 2009; Hovius et al. 2007). This suggests the close relationship of salivary glands to the interaction between host adaptive immunity and pathogen replication (Ibelli et al. 2014; Tirloni et al. 2016). Furthermore, cathepsin G isolated from *Ixodes ricinus* salivary gland has been demonstrated to be involved in platelet aggregation and anti-inflammatory activity of host animal (Chmelar et al. 2011).

However, to our knowledge, no report is available so far about the function and characterization of the cysteine protease in *H. flava* ticks. Here, an assembled contig from the transcriptome database of *H. flava* was found to be homologous to the cysteine protease (unpublished data). According to the contig assembly sequence information in the in-house data, we isolated a novel cathepsin L cDNA from *H. flava* and named it as *HfCL*. The expression of *HfCL* was found to be up-regulated at the starvation stage of nymphs and adults. Additionally, *HfCL* was expressed in both mid-gut and salivary glands as detected by immunolocalization. Functional verification assays demonstrated that *HfCL* might be involved in the survival and proliferation of *Babesia* parasites.

Materials and methods

Experimental animals and parasite materials

H. flava was maintained in our laboratory. New Zealand white rabbits were used as the host animal during the blood-sucking stage of ticks. Fully engorged ticks were collected from rabbit and incubated in plates at 28 ± 1 °C and 88% relative humidity to obtain the next-generation of ticks. Anti-immune serum was prepared using five-week-old Kunming mice. *Babesia gibsoni* and *Babesia microti* parasites were used to infect 1-year-old Beagles and 5-week-old BALB/c mice, respectively. At 5% parasitemia, parasites were collected from host animals for in vitro culture. All the animals were obtained from the Laboratory Animal Research Center of Hubei Center for Animal Disease Control and Prevention, Hubei, PR China.

RACE and cloning of the *HfCL* coding region

After washing three times with diethyl-pyrocarbonate (DEPC) water, the total RNA was extracted from unfed ticks using TRIzol Reagent (Invitrogen, NY, USA) and the RNA quality was checked by 1× TBE (tris base-boric acid-EDTA) agarose gel electrophoresis. The 5′-end and 3′-end cDNA were prepared by using Revert Aid First Strand cDNA Synthesis Kit (Fermentas, Germany) and specific primers designed by clone manager software (http://clone-manager.software.informer.com/). The synthesis of all the primers and DNA sequencing were performed by Sangon (Shanghai, China).

In this study, we first verified an assembled contig sequence from a transcriptome database (unpublished data) by PCR using the B144F1 and B144R1 primers (Table S1) under the conditions of 94 °C for 2 min, followed by 35 cycles (at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min), and finally at 72 °C for 10 min. According to the sequencing results of the intermediate fragments, specific-primers were designed by using Clone Manager software (version 9.0; Sci-Ed Software, Cary).

Next, the 5′-end and 3′-end cDNA were amplified using the SMARTer™ rapid amplification of cDNA ends (RACE) cDNA Amplification Kit (Clontech, CA, USA) according to the manufacturer's instructions. Specifically, the 5′-RACE PCR was performed using the specific internal primer (B144-2) (Table S1) and universal primer at 94 °C for 2 min, followed by 35 cycles at 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min. Then, with the first round of PCR products used as the template, the nest-PCR was carried out using the specific outer primer (B144-3) and universal primer as described above. Meanwhile, the 3′-end cDNA was amplified using total 3′-ends cDNA as the template and the specific internal primer (C097-1) with a universal primer as described above. Similarly, with the first round of PCR products used as the template, the nest-PCR was carried out using a specific outer primer (C097-2) with a universal primer at 94 °C for 2 min, followed by 35 cycles at 95 °C for 30 s, 57 °C for 30 s, and 72 °C for 1 min.

Finally, the sequencing results of 5′-RACE and 3′-RACE were merged to the identified intermediate sequence region to generate the full-length *HfCL* cDNA information. The entire open reading frame was verified by PCR using a pair of specific primers (QCB144F1 and QCB144R1) designed based on the full-length cDNA sequence information under the conditions of 94 °C for 3 min, followed by 35 cycles (at 94 °C for 30 s, 62 °C for 30 s, and 72 °C for 2 min), and finally at 72 °C for 10 min. The full-length ORF was cloned to pMD 19-T (Takara, Dalian, China) and then sequenced for subsequent experiments. The full *HfCL* cDNA sequence has been deposited to GenBank with the accession number MG914066.

Bioinformatics and phylogenetic analyses

After predicting the conceptual translation of *HfCL* cDNA into amino acid (aa) sequences, the aa sequences of orthologs were aligned using the Bioedit software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html#downloads), and the enzyme activity sites were predicted by software (http://myhits.isb-sib.ch/cgi-bin/motif_scan). To clarify the genetic identity of

HfCL, the phylogenetic tree was constructed using the MEGA 7.0 software (http://www.megasoftware.net/index.php) based on the amino acid sequences of 27 orthologs obtained from the GenBank database. Confidence limits were assessed by bootstrapping using 1000 pseudo-replicates for neighbor joining (NJ), and other parameters were set at the default values in MEGA 5.0. A 50% cutoff value was used for constructing the consensus tree.

The three-dimensional structure (3D) model of *HfCL* was predicted based on the human procathepsin L (PDB code: 1CS8_A) crystal structure model (https://swissmodel.expasy.org/). The crystal structure of cathepsin from *H. flava* with a 45% identity with the human procathepsin L. 3D model was generated by PyMOL software (http://www.pymol.org/).

Real-time PCR

Total RNAs were extracted separately from eggs, larvae, nymphs, adult females, and males. Specifically, 2-day eggs were collected after the production of female ticks; larvae were collected after eggs developed into larvae with high activity; nymphs and adults were collected at the starvation stage. The cDNA was synthesized using the PrimeScript® RT Reagent Kit with the gDNA Eraser (Takara, Dalian, China), and PCR amplification was performed using the SYBR Green PCR Master Mix kit according to the manufacturer's instructions (Takara, Dalian, China), with the β-actin gene used as an internal standard control (Xu et al. 2016), and double distilled water as a negative control.

The primers were designed using the GenScript Real-time PCR primer Design tool (https://www.genscript.com/tools/real-time-pcr-tagman-primer-design-tool), and the sequences of qRT-PCR primers (*HfCL*-F/*HfCL*-R and actin-F/actin-R) are shown in Table S1. The PCR reactions were performed at 50 °C for 2 min, and 94 °C for 30 s, followed by 30 cycles (of 95 °C for 15 s, then 56 °C for 15 s, and 72 °C for 20 s), and finally at 72 °C for 10 min. Data were processed using the formula $2^{-\triangle\triangle ct}$, and larvae were used as a calibrator relative to the ticks at the other different stages.

Expression and purification of the *rHfCL* in E. coli

The coding sequence of *HfCL* was sub-cloned into the expression vector pET-28a by the homologous recombination method using the primers (*HfCL* expressed F/R). After confirmation by the restriction enzyme (BamHI/XhoI) double digestion and sequencing analysis, the recombinant plasmid pET-28a-*HfCL* was transformed into the *Escherichia coli* BL21 strain (*Transetta*, Transgene, China) for prokaryotic protein expression in vitro. The soluble His-fusion protein expression was induced at 25 °C and 100 rpm for 20 h in 1:1000 diluted IPTG (isopropyl-β-D-thiogalactoside). Next, the recombinant protein was purified using His-Gravitrap Ni-NTA Resin (GE Healthcare) according to the manufacturer's instructions. Finally, the protein concentration was determined using the BCA protein assay kit (Beyotime, China).

Preparation of anti-immune serum and Western blot

After collecting the serum of healthy 5-week-old Kunming mice as a negative control, the mice were immunized five times subcutaneously at an interval of 7 days with 100 μg of the *rHfCL* emulsified by Freund complete adjuvant (Sigma, Shanghai, China). After the last immunization, the anti-immune serum was collected for detecting the antibody titer by ELISA (enzyme-linked immunosorbent assay).

The quality of the immune serum was checked by SDS-polyacrylamide gel and Western blotting. Briefly, the native protein lysates of the adult ticks (*HfCL*) were subjected to 12% SDS-PAGE, transformed onto the NC membrane (nitrocellulose, Millipore, USA), washed three times with TBST (Tris buffer Saline and Tween 20), and blocked with 1% BSA (Bull serum albumin) at 4 °C overnight. Next, the NC membranes were incubated separately with anti-*rHfCL* immune serum (diluted 1:400) and pre-immune mouse serum (negative control) for 3 h at room temperature, followed by three washes with TBST (tris-buffered saline and Tween 20), and incubation with goat anti mouse IgG (diluted 1:1000) conjugated to horseradish peroxidase (Beyotime, Shanghai, China). Western immunoblotting results were visualized using ECL chemiluminescent detection system according to the manufacturer's instructions (Thermo Scientific, MA, USA).

Immunolocalization of *HfCL* in tick tissues

After dissecting and collecting the mid-gut and salivary glands of adult ticks under an anatomical microscope, the tissues were washed three times in 1× PBS (phosphate buffered solution) and fixed separately in the Bouin's fixation fluid. Meanwhile, the Bouin's fixation fluid was prepared with 15% saturated picric acid, 4% paraformaldehyde, 1% glacial acetic acid, and 79% ddH_2O. After soaking separately in wax and Tissue-Tek O.C.T medium (Torrance, USA) at − 20 °C, the ultra-thin gut and salivary glands (0.4 μm) were fixed on glass slides for HE (hematoxylin–eosin) staining according to a standard protocol.

Indirect immunofluorescence-stained sections were incubated for 4–6 h with 1% BSA (Bull serum albumin) in PBS-Tween (Tween 20, Sigma, USA), followed by three washes with 1× PBS, permeabilization for 20 min with 0.1% Triton™ X-100 (Sigma, USA), another three washes with 1× PBS, and then incubation separately with anti-*rHfCL* antibody serum and pre-immune mouse serum (diluted 1:500) for 3 h at 37 °C. Finally, the tissue sections were washed three times with 1× PBS and incubated with goat anti-mouse IgG (diluted 1:1000) (Alexa 488 Streptavidin, Invitrogen, USA). DAPI (4′,

6-diamidino-2-phenylindole) was used for nuclear staining, and images were collected by using a fluorescence microscope.

Protease activity assays

The *rHfCL* activity was measured by hydrolyzing 0.001 M fluorogenic substrate (Z-phe-Arg-MCA, Sigma, USA) as a special substrate for cathepsin L-like family protease. The optimum protease concentration and pH were tested separately by reaction with substrates, with the protease diluted at 0.5, 1, 2, 4, 8, 16, and 32 μM using CSP (citrate–sodium phosphate) buffer (pH 3 to 10) containing 100 mM NaAC, 100 mM NaCl, 1 mM EDTA, 0.005% Triton™ X-100, and 4 mM DTT. The activity assays were performed in triplicate at the optimal protein concentration and buffer pH, using E-64 (L-trans-epoxysuccinyl-leucylamido 4-guanidino butane) as a special inhibitor for the cathepsin L-like family protease. The optical density (OD) was read at 450 nm on a Bio Tek ELISA reader.

In vitro growth inhibition assay of *Babesia* merozoites

Babesia merozoites were in vitro cultivated in host erythrocytes in HL-20 media (Gibco Life Technologies) supplemented with 1% Albuman I (Gibco Life Technologies), 2% HB101 (Irvine Scientific), 1% L-glutamine (ATLANTA Biologicals), 2% antibiotic/antimycotic 100× (Corning), and 20% healthy dog or mouse serum at 37 °C in microaerophilous stationary phase (5% CO_2, 2% O_2, 93% N_2). Meanwhile, the *Babesia*-infected erythrocytes were diluted with the healthy erythrocytes till the parasitemia reached 2%, followed by the culture of *B. microti* and *B. gibsoni* merozoites separately in the mice and dog erythrocytes.

For the growth inhibition assay of *Babesia* merozoite, the *rHfCL* protease at a concentration of 500, 250, 100, 50, and 10 μg/mL was added separately to *B. microti* and *B. gibsoni*

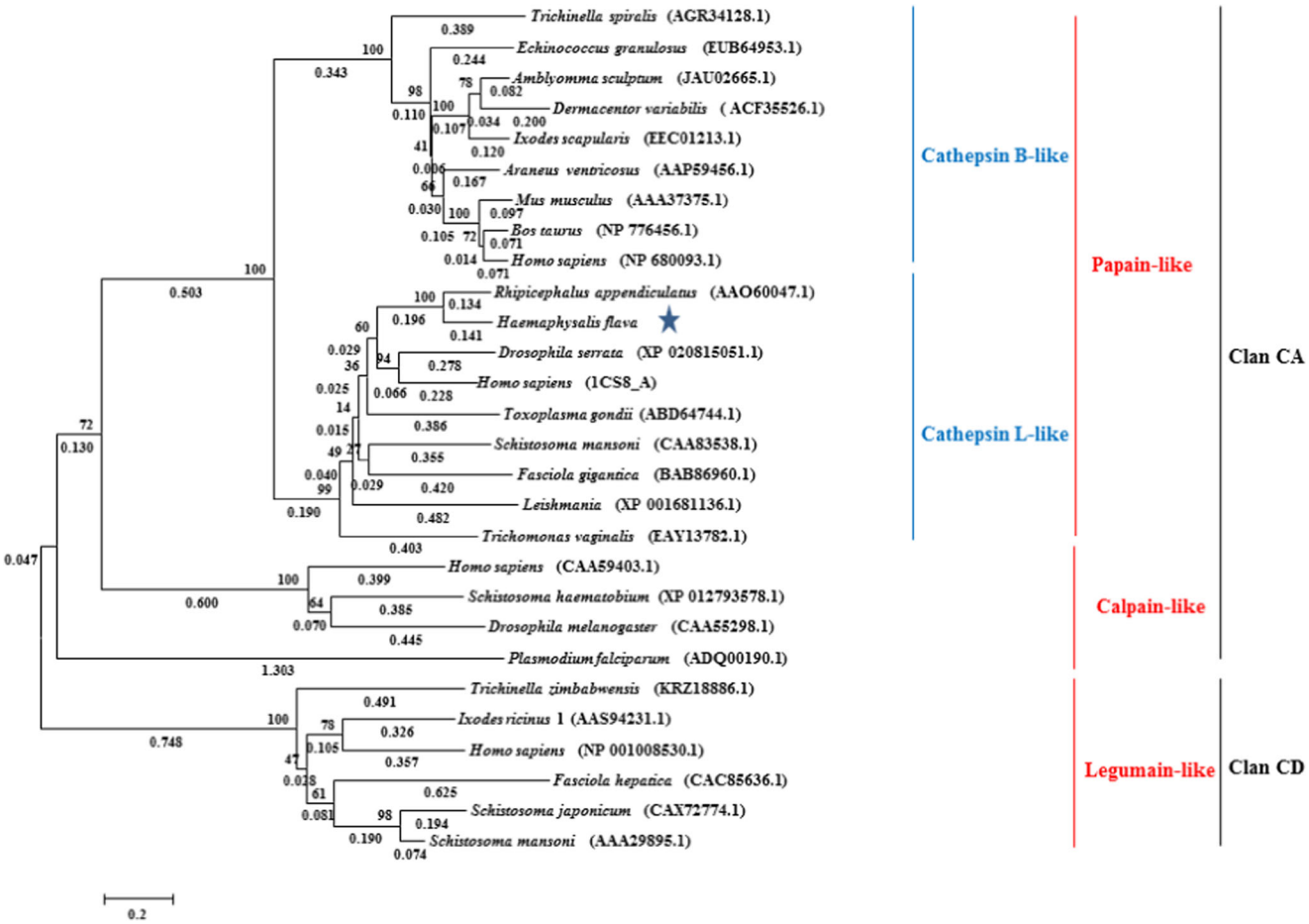

Fig. 1 Phylogenetic tree of *HfCL* with amino acid sequences of 27 orthologs. Blue braces represent the cathepsin B-like and cathepsin L-like families. Red braces indicate the papain-like, calpain-like and legumain-like families. Black braces represent clan CA and clan CD superfamilies. Accession numbers of the NCBI protein database for the various sequences are given next to each species name. The bootstrap values are presented above or below the tree branches, and the scale bar displays the branch length as a standard of the total tree length

culture wells at 0 h, using 1% BSA used as a negative control and the normal *Babesia* in vitro culture group without added protease or 1% BSA as the blank control. Next, the plates were incubated for 96 h, with the culture medium and the protein renewed every 12 h, and the parasitemia of each well was determined every 24 h by microscopic observation on the basis of approximately 3000 RBCs in Giemsa-stained blood smear.

Statistical analysis

Real-time PCR, *rHfCL* activity assay and in vitro growth inhibition assay of *Babesia* merozoites were carried out in triplicate for each group. Statistical analyses were performed by GraphPad Prism version 6.0 (http://www.graphpad.com/) using a one-way analysis of variance (ANOVA), followed by Tukey's multiple comparison test. The results were reported as the mean ± standard deviation of three independent experiments with significant differences at $P < 0.05$.

Results

Sequencing and phylogenetic analysis of *HfCL*

The transcriptome of *H. flava* was sequenced (unpublished data), and data analysis revealed an assembled contig (C290670_g1) with a 78% similarity to the *Rhipicephalus appendiculatus* midgut cysteine proteinase (AY208823). The cloning and sequencing of PCR products obtained from nested 5′ and 3′ RACE amplification revealed a 1216-bp long sequence, containing a 1041-bp open reading frame (ORF) encoding 346 amino acids, a 38-bp 5′-untranslated region (UTR), and a 137-bp 3′-UTR including the poly-A tail preceded by a non-canonical poly-adenylation signal (AAAAA) 59 bp. We named this new transcript *H. flava* cysteine proteinase (*HfCL*) gene.

Phylogenetic analysis of the amino acid sequences of the cysteine proteinase from 27 orthologs showed that *HfCL* grouped together with orthologs from *R. appendiculatus*, *Drosophila serrate*, *Schistosoma mansoni*, *Fasciola gigantica*, *Leishmania*, *Trichomonas vaginalis*, and *Homo sapiens* (Fig. 1). Based on sequence homology, the biological classification of cysteine proteases can be mainly divided into four groups, clans CA, CB, CC, and CD, and most viral proteases are the members of CB and CC clans. The phylogenetic tree of the cysteine proteases in parasites mainly covered clans CA and CD, including family C1 (papain-family), C2 (calpain-like), and C13 (legumain-like) (Sajid and Mckerrow 2002). Among the seven selected species, *HfCL* is closely related to *R. appendiculatus* (AAO60047.1) and belongs to the papain-like superfamily of the cathepsin L-like family. Figure 2 shows the alignment of the amino acid sequences of *HfCL* with the other seven cathepsin L orthologs, including *R. appendiculatus* (71%), *D. serrate* (44%), *S. mansoni* (38%), *F. gigantica* (38%), *Leishmania* (36%), *T. vaginalis* (39%), and *Homo sapiens* (45%). As demonstrated by the active motif sites (yellow shadows in the figure), the C-terminal domain of the protease is highly conserved in the orthologs (positions 320–360).

Modeling of the 3D structure of *HfCL* based on human procathepsin L crystal structure (Fig. 3) revealed that the

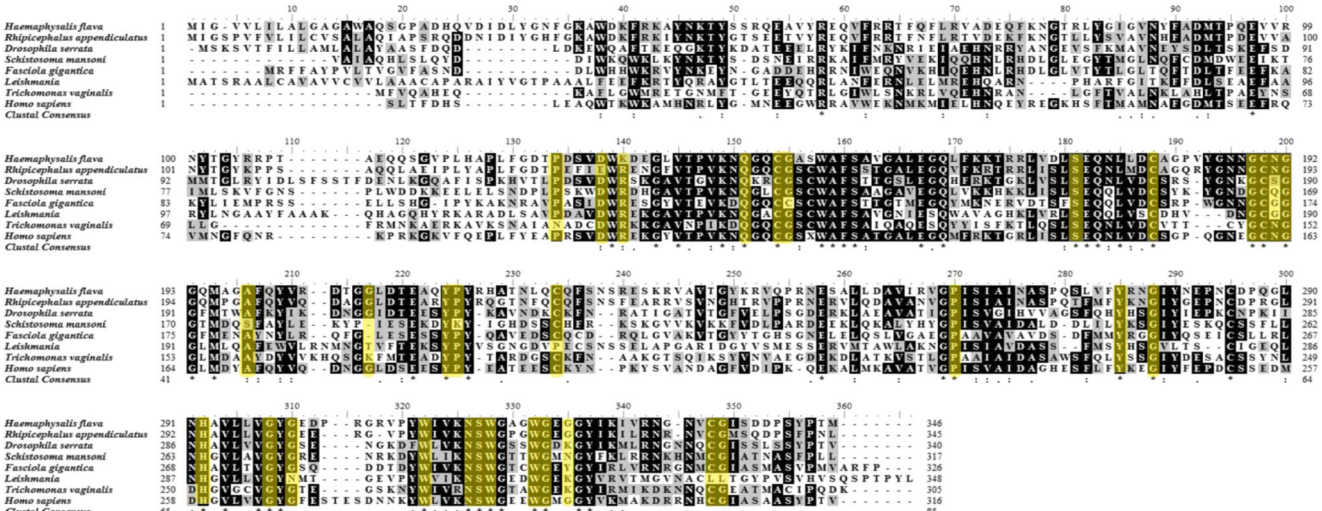

Fig. 2 Alignment of the inferred amino acid sequences of *HfCL* with related cathepsin L orthologs. The seven selected species are *Rhipicephalus appendiculatus* (AAO60047.1), *Drosophila serrate* (XP 020815051.1), *Schistosoma mansoni* (CAA83538.1), *Fasciola gigantica* (BAB86960.1), *Leishmania* (XP 001681136.1), *Trichomonas vaginalis* (EAY13782.1), and *Homo sapiens* (1CS8_A). Asterisk represents amino acid domain with complete similarity. Colon represents amino acid domain with high similarity. Period represents low similarity, and no symbol indicates no similarity among domains. Numbers represent the sequence length above the amino acid domain. The active motif sites of amino acids are indicated by yellow shadows

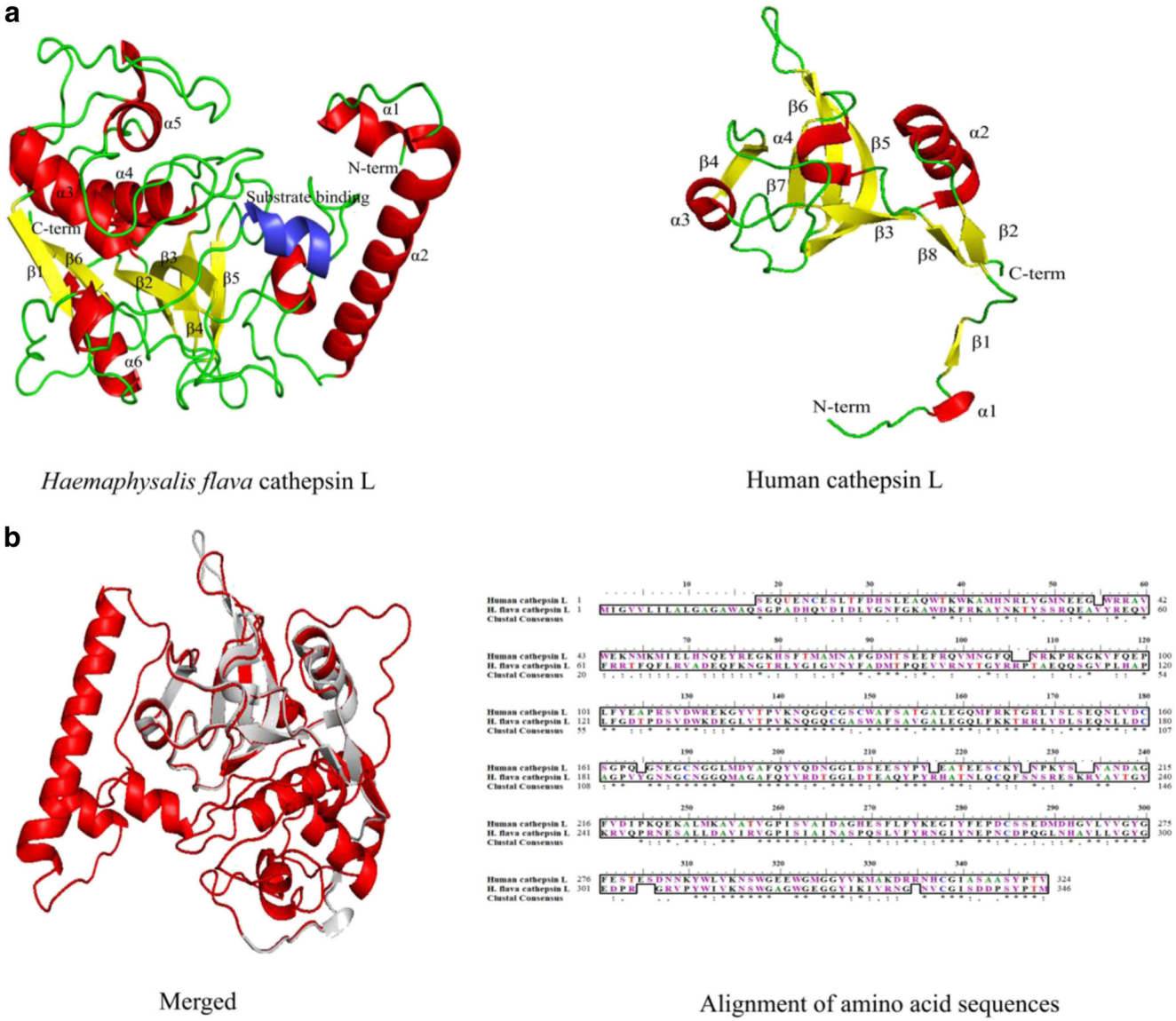

Fig. 3 3D structural model of *HfCL*. **a** The homology model of *HfCL* obtained from human cathepsin L. The modeled structure of *HfCL* shows different main structural elements: α-helices (red), β-sheet (yellow), substrate binding cleft (blue), connecting loops (green), N-terminus, and C-terminus. **b** The third structural representation of *H. flava* cathepsin L merged with human cathepsin L. Red indicates *H. flava* cathepsin L and gray indicates human cathepsin L. The secondary structural elements for each protein are structures

HfCL protease encompasses the N-terminal segment, C-terminal segment, six α-helices, six β-sheets, and a substrate binding site. The initial protein model had an N-terminal conformation within a four-turn α-helix similar to that of human cysteine proteases. In the human procathepsin L crystal structure, there are extensive contacts between cathepsin L and the prosegment within the substrate binding cleft involving hydrolysis residues. The obvious substrate binding site can be observed from the surface of the enzyme structural modeling of *HfCL*, which is wedged between the α6-helix and β4-sheet close to the C-terminal segment. However, the human prosegment emerges from the substrate binding cleft and follows in an extended conformation towards the N-terminus of cathepsin L, with its side chains spreading on both sides of the backbone along the surface of the enzyme. Compared with the results of amino acid arrangement, the substrate binding site of *HfCL* contains hydrolyzed residues and is consistent with that of the isogenous cathepsin L.

Transcriptional profiles of *HfCL*

The *HfCL* transcript level in the four different development stages of *H. flava* was analyzed by qRT-PCR with larvae used as a calibrator and β-actin as an internal standard. As shown in

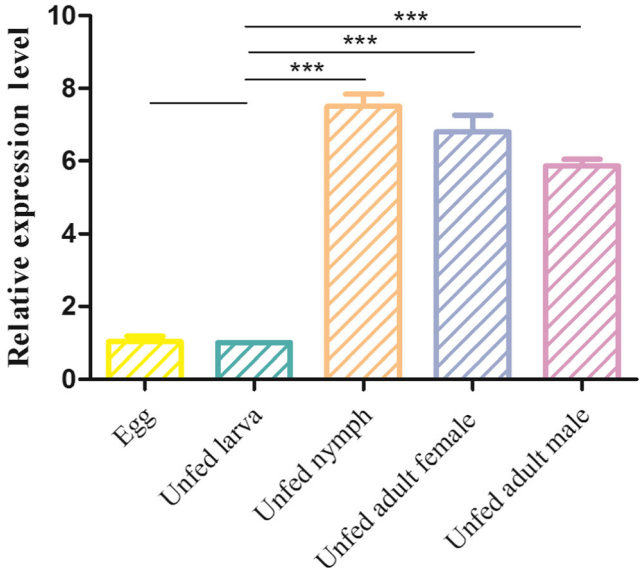

Fig. 4 Transcriptional profile of *HfCL* in four life stages (egg, larva, nymph, adult female, and adult male). Transcript abundances were determined in biological triplicate by qRT-PCR. Error bars showed 95% confidence intervals. Data processing was performed by using the formula $2^{-\Delta\Delta ct}$. Larva was compared separately with egg, nymph, adult female, and adult male, with β-actin gene used as an internal standard. Statistically significant differences are shown as $^{*}P < 0.05$, $^{**}P < 0.001$, $^{***}P < 0.0001$, and no symbol indicates no significant difference

Fig. 4, *HfCL* was most abundantly transcribed in nymphs and adults ($P < 0.0001$) and no significant difference was observed between eggs and larvae in the transcriptional level ($P > 0.05$). However, the transcript level was significantly higher in nymphs, adult females, and males than in larvae ($P < 0.0001$).

Localization of *HfCL* in the tick's tissue

The native *HfCL* in *H. flava* from the lysates of adult ticks was investigated using anti-*rHfCL* serum, and a specific 38-kDa band appeared on the membrane (Fig. 5), indicating that the adult *H. flava* contains *HfCL*. It is known that intestinal gut and salivary glands are involved in the digestion of host blood and storage of pathogens in ticks, which was well supported by the finding of this study. In Fig. 6, immunolocalization results revealed that *HfCL* is localized in the gut epithelium, gut lumen, and the vesicles of salivary gland, while no fluorescence staining signal was detectable in either the internal or external tissues treated with pre-immune serum when compared with the HE staining control group (normal tick gut and salivary glands).

Protease activity assays

The enzyme activity of the recombinant *HfCL* (*rHfCL*) was tested by hydrolyzing the fluorogenic substrate Z-phe-Arg-MCA (Fig. 7). *rHfCL* maximum activity was obtained at 16 μM and pH 6 and the minimum activity at 0.5 μM and pH 10 (Fig. 7c). In the inhibition assay, 0.1 mM E-64 inhibitor reduced the *rHfCL* activity of hydrolyzing Z-phe-Arg-MCA relative to the normal hydrolysis reaction without inhibitor (Fig. 7d). The overall results demonstrated that *rHfCL* has good activity.

In vitro growth inhibition assay of *Babesia* parasites

We further examined the inhibitory effect of *rHfCL* protease on the growth of *B. microti* and *B. gibsoni* in vitro using a short-term in vitro culture assay. As indicated in Fig. 8a, *rHfCL* could inhibit the growth of *Babesia* merozoites in a concentration-dependent manner. Specifically, *B. microti* growth was significantly inhibited at 100 μg/mL of *rHfCL* ($P < 0.05$), while the other concentrations showed an effective inhibition after 72 h of treatment. Meanwhile, *B. gibsoni* growth was significantly inhibited at 250 μg/mL of *rHfCL* ($P < 0.05$), while the other concentrations showed an effective inhibition after 96 h of treatment relative to the control (Fig. 8b). However, under the treatment of 1% BSA, parasitemias exhibited normal growth, similar to the treatment without protease. Furthermore, no morphological changes were observed in the erythrocytes in the presence of *rHfCL* at any concentration.

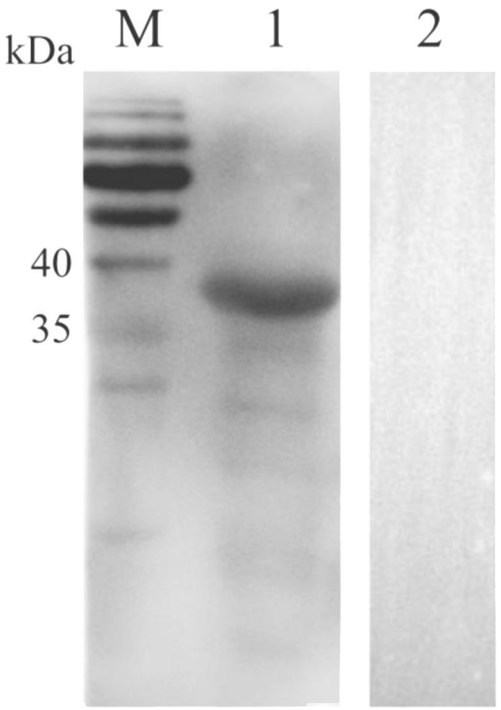

Fig. 5 Immunoblot analysis of anti-*HfCL* serum. The native protein from lysates of adult ticks was probed separately with anti-*HfCL* serum (lane 1) and pre-immune mice serum (lane 2). Goat anti-mouse IgG conjugated to horseradish peroxidase was used as the second antibody. M: Protein molecular weight standard (kDa)

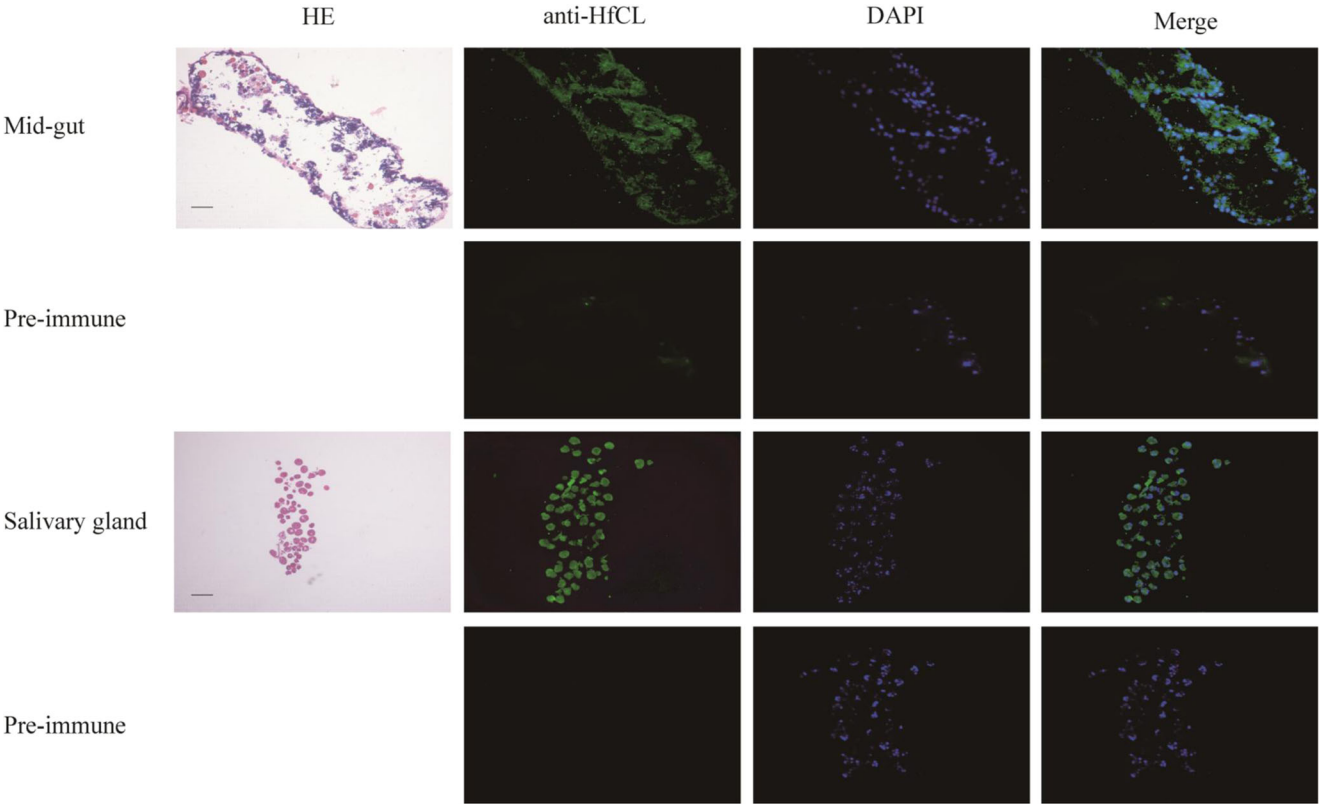

Fig. 6 Immunolocalization of *HfCL* in tick tissue. HE staining presents the normal tissue morphology of mid-gut and salivary gland. The endogenous *HfCL* was incubated with mouse anti-*HfCL* serum. The second antibody used was anti-mice IgG Alexa 488 (Green). The nucleus was stained with DAPI. The mid-gut and salivary gland were incubated with pre-immune mouse serum as a negative control

Discussion

Babesia parasites, the causative agent of babesiosis, are a major threat to the health of livestock and humans (Kjemtrup and Conrad 2000; Hunfeld et al. 2008), implying the necessity to find effective strategies to control parasite transmission. This can be achieved by targeting the cysteine proteases of ticks due to their important roles in hemoglobin digestion and development of the arthropod vector (Cho et al. 1999; Uchida et al. 2001). A previous study has shown that a cysteine proteinase from *H. longicornis* can mediate direct killing of *B. equi*, but no report is available about the interaction of tick cysteine with *Babesia* parasites (Tsuji et al. 2008). Here, the full-length cDNA of cathepsin L was cloned and characterized from the tick *H. flava*, which was named *HfCL* and found to possess functionality of inhibiting the proliferation of *Babesia* merozoites.

Proteinases can be classified into cysteine protease, serine protease, aspartate protease, threonine protease, and metalloproteinase based on proteolytic activity (Beers et al. 2004), but different families based on structure, peptide loops, and peptide substrates. In mammals, the members of the papain superfamily can be designated as cathepsin L, B, S, P, and K (Barrett 1994; Sajid and Mckerrow 2002), and in hematophagous invertebrates, the peptidases mainly consist of cysteine CA and CD clan members based on the proteolytic machinery. In this study, the isolated cysteine protease was confirmed to be L-like papain by phylogenetic analysis and alignment of the amino acid sequence with other orthologous sequences. The first crystal structure of cysteine proteases from the papain superfamily was described in 1968 (Drenth et al. 1968). The 3D structure of *HfCL* was predicted based on the human procathepsin L crystal structure (Coulombe et al. 1996), and it was observed to have the same essential cysteine residues in the active site for hydrolysis. However, little is known about *HfCL* except for its basic structural elements such as α-helices, β-sheet, and the active site for hydrolysis. The conservative domain was identified to be the C-terminal structure of protease from multiple sequence alignment, which may be related to the hydrolysis mechanism of the papain superfamily (Cygler et al. 1996). To enhance the understanding of the protein crystal structure, further research should focus on the specific elements of the *HfCL* structure.

The transcriptional profiles of cathepsin L were examined at different developmental stages in other parasites (Brady et al. 2000; Liang et al. 2014). For instance, the cathepsin L mRNA expression and activity have been found to vary with the duration of blood-sucking process in *I. ricinus* and knockdown cathepsin gene in ticks, leading to an obvious effect on body weight, egg-laying rate and egg hatchability (Sojka et al.

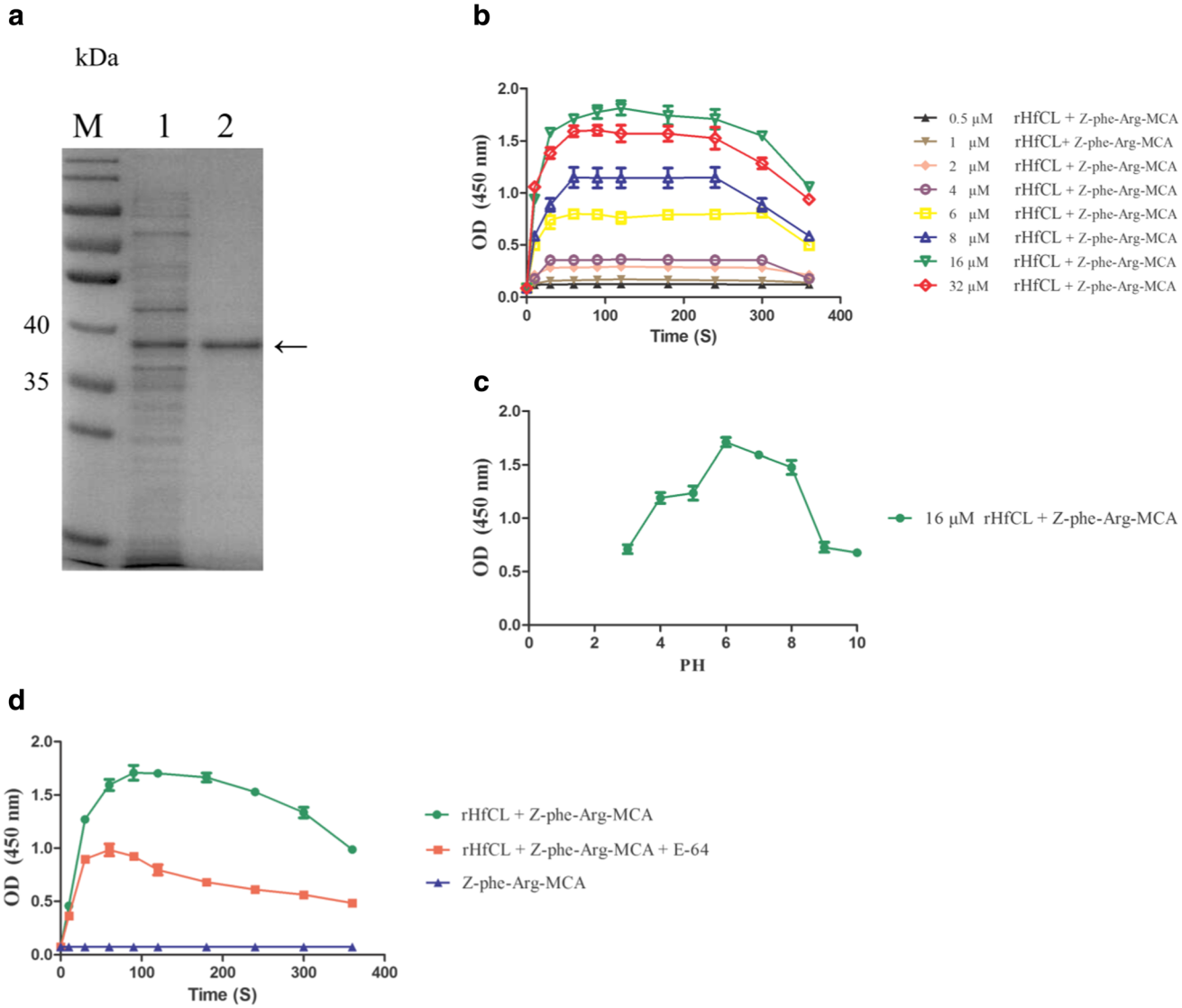

Fig. 7 Protease activity assays. **a** *rHfCL* was expressed in *E. coli*. Lane 1 indicates IPTG-induced bacterial solution, and lane 2 indicates the purified *rHfCL* by 12% SDS-PAGE. **b** Different concentrations of *rHfCL* were hydrolyzed separately with the substrate (Z-phe-Arg-MCA). **c** Enzyme activity was tested using 16 μM *rHfCL* with substrate (Z-phe-Arg-MCA) at a different pH. **d** Inhibition assay of *rHfCL* using E-64 inhibitor

2012). *H. flava* (Ixodid ticks), a three host tick, has four development stages (egg, larva, nymph, and adult), and the transcriptional profiling indicated that cathepsin L is transcribed in all stage of *H. flava*, with higher transcriptional levels detected in nymphs and adults than in larva, implying that *HfCL* contributes to growth and development of *H. flava* ticks. Unlike the host blood digestive machinery, the cathepsin L might help the survival of starving ticks by providing amino acids for energy catabolism, enabling the survival of tick at every stage until the availability of the first blood meal.

Based on immunofluorescence localization and mRNA expression level, the endogenous cathepsin L was also expressed at both midgut and salivary glands, and further analysis revealed that cathepsin L is located on both midgut epithelium and gut lumen, implicating that gut epithelium is passively secreted into the lumen during energy catabolism. Additionally, HE staining of the control group showed that the histology structure of salivary glands mainly contains acinus and acinar duct, and immunofluorescence was detected in the individual acinus rather than acinar duct. This finding was consistent with previous studies reporting that the salivary gland as a kind of hyperosmotic secretion can absorb water from outside air during long starvation periods (Kim et al. 2017), indicating the potential contribution of *HfCL* to salivary gland in absorbing water to maintain body moisture.

Furthermore, most *Babesia* spp. are tick-borne protozoan parasites and perform propagation by transovarial transmission (Boldbaatar et al. 2008), and the salivary glands have

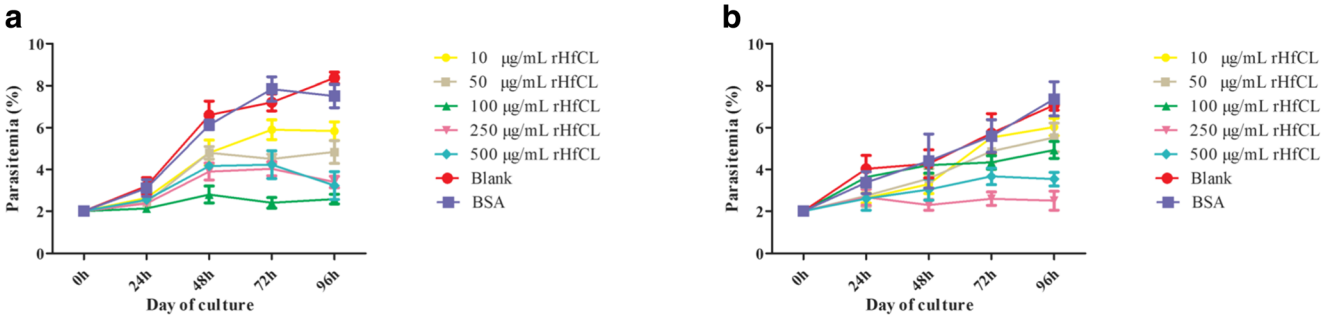

Fig. 8 In vitro inhibition assay of *Babesia* parasite erythrocytes. **a** Different concentrations of *rHfCL* were incubated separately with *B. microti*. **b** Different concentrations of *rHfCL* were incubated separately with *B. gibsoni*. *Babesia* parasites incubated with 1% BSA were considered as a negative control. The blank control consisted of only *Babesia* parasites with HL-20 medium. The infection rate of initial *Babesia* parasites was 2% in host erythrocytes

been reported to store pathogens that would be transmitted to ovary when the tick entered the next development stage (AM et al. 2013; Büscher et al. 1988; Moltmann et al. 1982; Santos et al. 2000). In the present study, the immunofluorescence results suggest that cathepsin L is not only involved in energy metabolism during the starvation period but also probably involved in the transmission of *Babesia* parasites in ticks. Unfortunately, due to difficulty in ovarian anatomy and tissue dissection, we failed to detect cathepsin L in the tick ovaries in this study.

Generally, the optimal activity of cysteine protease was stabilized in the acidic and lysosomal microenvironment. In vitro enzymatic functional assays showed that *rHfCL* has the maximum activity at pH 6 and 16 μM in hydrolyzing Z-phe-Arg-MCA (substrate for cathepsin L/B) and its enzyme activity declines sharply above pH 9. This result is different from a previous study on an asparaginyl endopeptidase from *I. ricinus*, which showed the inactivation of enzyme activity at a pH value over 6 (Sojka et al. 2007), probably due to pH variations during hydrolysis as shown in humans (Chen et al. 1997). Moreover, the proteolytic activity of purified *rHfCL* was effectively inhibited in the presence of E-64 inhibitors at the indicated concentration. Overall, the substrate specificity and inhibitor sensitivity indicate that *HfCL* is a typical cathepsin L.

Previous studies have shown that the serine protease of the mosquito has the ability of regulating the malaria parasite to invade the intestinal epithelium (Abraham and Jacobslorena 2004), and tick intestine-related proteolytic enzymes have a major impact on the survival of pathogens (Narasimhan et al. 2014; Villar et al. 2013). In our study, on the Giemsa-stained blood smear, free *Babesia* merozoites were observed outside the erythrocytes and died in the medium during cultivation, implying that *rHfCL* could effectively inhibit the proliferation of merozoites, which is consistent with the study reporting that longipain kills *B. equi* in vitro (Tsuji et al. 2008).

Collectively, our results demonstrate that *HfCL* is a multifunctional papain with a pivotal role in the digestion of blood meal, providing energy nutrition in starvation stage and regulating the vectorial capacity for *Babesia* parasites. Further elucidation of the functional properties of *HfCL* will contribute to evaluate this enzyme as a drug target or vaccine candidate for controlling ticks and tick-borne babesiosis.

Conclusion

In this study, a novel cathepsin L was identified from *H. flava* by cloning and characterizing *HfCL* cDNA. Evolutionary analysis indicated that the identified *HfCL* belongs to the L-like papain family, and multiple sequence alignment demonstrated a relatively high similarity between *H. flava* and *R. appendiculatus*. Additionally, the cathepsin L mRNA transcriptional levels were found significantly higher at the nymph and adult stages than at the egg and larval stages. Furthermore, Western blotting identified native *HfCL* protein (38 kDa) in the tick, which was found to be mainly localized at the gut and salivary gland by immunolocalization using the specific anti-*HfCL* serum as the probe. Moreover, *rHfCL* showed the maximum activity at pH 6 and 16 μM, with obvious inhibition on the growth of *B. microti* and *B. gibsoni* merozoites by co-incubation. Our integrated data suggest that *HfCL* is probably involved in the growth of *Babesia* parasites in ticks.

Acknowledgments We thank all the research team members in our lab for assistance. The authors would like to thank Professor Choukri Ben Mamoun (Yale University, USA) for revising manuscript.

Author contributions YS performed the study and wrote the manuscript. JZ and LH conceived the study. ZN, JG, and QL collected samples. LH critically revised the manuscript. YS and LY performed the experiments and data analyses. All authors have read and approved the final manuscript.

Funding This study was supported by the National Basic Science Research Program (973 program) of China (Grant No. 2015CB150300), the National Key Research and Development Program of China (Grant No. 2017YFD0501201), the National Natural

Science Foundation of China (Grant No. 31772729), and the Natural Science Foundation of Hubei Province (Grant No. 2017CFA020).

Compliance with ethical standards

Ethics approval and consent to participate The care and maintenance of experimental rabbit, mice, and dog in this study was approved by the Institutional Animal Care and Use Committee of Huazhong Agricultural University, and all experiments were performed according to the regulations for the administration of affairs concerning experimental animals of Hubei Province, P.R. China consent for publication.

Competing interests The authors declare that they have no competing interests.

Consent for publication Not applicable.

Data availability *HfCL* mRNA sequence was available in GenBank (accession no. MG914066).

Abbreviations *H. flava*, *Haemaphysalis flava*; *B. microti*, *Babesia microti*; *B. gibsoni*, *Babesia gibsoni*; *HfCL*, *Hf*-cathepsin L; *rHfCL*, recombinant *Hf*-cathepsin; ORF, open reading frame; aa, amino acid; DEPC, diethyl-pyrocarbonate; TBE, tris base-boric acid-EDTA; 3D, three-dimensional structure; RACE, rapid-amplification of cDNA ends; CDS, coding sequence; qRT-PCR, quantitative real-time PCR; PCR, polymerase chain reaction; IPTG, Isopropyl-β-D-thiogalactoside; BSA, bull serum albumin; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; PBS, phosphate buffered solution; DAPI, 4′ 6-diamidino-2-phenylindole; ECL, enhanced chemiluminescent; TBST, tris buffer saline Tween 20; RBCs, red blood cells; NC, nitrocellulose membrane; CSP, citrate–sodium phosphate; E-64, L-trans-epoxysuccinyl-leucylamido 4-guanidino butane; OD, optical density; IFA, indirect immunofluorescence assay

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Abraham EG, Jacobslorena M (2004) Mosquito midgut barriers to malaria parasite development. Insect Biochem Mol Biol 34(7):667–671

Akov S (1982) Blood digestion in ticks. Physiology of Ticks 1982:197-211

Aksenova AS, Kupriianova ES, Anufrieva VN, Ermakova RM, Shlenova MF (1976) Comparative characteristics of blood digestion stages in the mosquitoes Anopheles Meig., Culex pipiens L. and four species of Aedes Meig. (Aedes cantans Meig., Aedes caspius Pall., Aedes cinereus Meig., and Aedes communis De Geer). Meditsinskaia Parazitologiia I Parazitarnye Bolezni

AM H et al (2013) The ovarian transcriptome of the cattle tick, *Rhipicephalus* (*Boophilus*) *microplus*, feeding upon a bovine host infected with *Babesia bovis*. Parasit Vectors 6(1):276

Barrett AJ (1994) Classification of peptidases. Methods Enzymol 244(1): 1–15

Beers EP, Jones AM, Dickerman AW (2004) The S8 serine, C1A cysteine and A1 aspartic protease families in Arabidopsis. Phytochemistry 65(1):43–58

Boldbaatar D, Battsetseg B, Matsuo T, Hatta T, Umemiya-Shirafuji R, Xuan X, Fujisaki K (2008) Tick vitellogenin receptor reveals critical role in oocyte development and transovarial transmission of *Babesia* parasite. Biochem Cell Biol-biochim Et Biol Cell 86(4):331–344

Brady CP, Brindley PJ, Dowd AJ, Dalton JP (2000) *Schistosoma mansoni*: differential expression of cathepsins L1 and L2 suggests discrete biological functions for each enzyme. Exp Parasitol 94(2): 75–83

Büscher G, Friedhoff KT, Elallawy TA (1988) Quantitative description of the development of *Babesia ovis* in *Rhipicephalus bursa* (hemolymph, ovary, eggs). Parasitol Res 74(4):331–339

Chmelar J, Oliveira CJ, Rezacova P, Francischetti IMB, Kovarova Z, Pejler G, Kopacek P, Ribeiro JMC, Mares M, Kopecky J, Kotsyfakis M (2011) A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. Blood 117(2):736–744

Cho WL, Tsao SM, Hays AR, Walter R, Chen JS, Snigirevskaya ES, Raikhel AS (1999) Mosquito cathepsin B-like protease involved in embryonic degradation of vitellin is produced as a latent extraovarian precursor. J Biol Chem 274(19):13311–13321

Coulombe R, Grochulski P, Sivaraman J, Ménard R, Mort JS, Cygler M (1996) Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. EMBO J 15(20):5492–5503

Delcroix M, Sajid M, Caffrey CR, Lim KC, Dvořák J, Hsieh I, Bahgat M, Dissous C, McKerrow JH (2006) A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. J Biol Chem 281(51):39316–39329

Drenth J, Jansonius JN, Koekoek R, Swen HM, Wolthers BG (1968) Structure of papain. Nature 218(5145):929–932

Fournier PE, Fujita H, Takada N, Raoult D (2002) Genetic identification of rickettsiae isolated from ticks in Japan. J Clin Microbiol 40(6): 2176–2181

Francischetti IM, Sanunes A, Mans BJ, Santos IM, Ribeiro JM (2009) The role of saliva in tick feeding. Front Biosci 14(14):2051–2088

Franta Z et al (2010) Dynamics of digestive proteolytic system during blood feeding of the hard tick Ixodes ricinus. Parasit Vectors 3(1):1–11

Fuente JDL et al (2017) Tick-pathogen interactions and vector competence: identification of molecular drivers for tick-borne diseases. Front Cell Infect Microbiol 7:114

Grandjean O (1984) Blood digestion in Ornithodorus moubata Murray sensu stricto Walton (Ixodoidea: Argasidae) females. I. Biochemical changes in the midgut lumen and ultrastructure of the midgut cell, related to intracellular digestion. Acarologia 25:147–165

Horn M, Nussbaumerová M, Šanda M, Kovářová Z, Srba J, Franta Z, Sojka D, Bogyo M, Caffrey CR, Kopáček P, Mareš M (2009) Hemoglobin digestion in blood-feeding ticks: mapping a multipeptidase pathway by functional proteomics. Chem Biol 16(10):1053–1063. https://doi.org/10.1016/j.chembiol.2009.09.009

Hovius JW, van Dam AP, Fikrig E (2007) Tick-host-pathogen interactions in Lyme borreliosis. Trends Parasitol 23(9):434–438

Hunfeld KP, Hildebrandt A, Gray JS (2008) Babesiosis: recent insights into an ancient disease. Int J Parasitol 38(11):1219–1237

Ibelli AMG, Kim TK, Hill CC, Lewis LA, Bakshi M, Miller S, Porter L, Mulenga A (2014) A blood meal-induced *Ixodes scapularis* tick saliva serpin inhibits trypsin and thrombin, and interferes with platelet aggregation and blood clotting. Int J Parasitol 44(6):369–379

Kern A, Collin E, Barthel C, Michel C, Jaulhac B, Boulanger N (2011) Tick saliva represses innate immunity and cutaneous inflammation in a murine model of Lyme disease. Vector Borne Zoonotic Dis 11(10):1343–1350

Kim D, Maldonado-Ruiz P, Zurek L, Park Y (2017) Water absorption through salivary gland type I acini in the blacklegged tick, *Ixodes scapularis*. Peerj 5(2):e3984

Kjemtrup AM, Conrad PA (2000) Human babesiosis: an emerging tick-borne disease. Int J Parasitol 30(12–13):1323–1337

Liang P, He L, Xu Y, Chen X, Huang Y, Ren M, Liang C, Li X, Xu J, Lu G, Yu X (2014) Identification, immunolocalization, and characterization analyses of an exopeptidase of papain superfamily, (cathepsin C) from Clonorchis sinensis. Parasitol Res 113(10):3621–3629

Moltmann UG, Mehlhorn H, Friedhoff KT (1982) Ultrastructural study of the development of *Babesia ovis* (Piroplasmia) in the ovary of the vector tick Rhipicephalus bursa. J Eukaryot Microbiol 29(1):30–38

Moon S et al (2013) Autochthonous lyme borreliosis in humans and ticks in Korea. Osong Public Health Res Perspect 4(1):52

Narasimhan S, Rajeevan N, Liu L, Zhao YO, Heisig J, Pan J, Eppler-Epstein R, DePonte K, Fish D, Fikrig E (2014) Gut microbiota of the tick vector *Ixodes scapularis* modulate colonization of the Lyme disease spirochete. Cell Host Microbe 15(1):58–71

Otsuka H (1976) Studies on transmission of *Babesia gibsoni* Patton (1910) by *Haemaphysalis flava* Neumann (1897). Bulletin of the Faculty of Agriculture—Miyazaki University (Japan)

Rar VA, Livanova NN, Panov VV, Doroschenko EK, Pukhovskaya NM, Vysochina NP, Ivanov LI (2010) Genetic diversity of Anaplasma and Ehrlichia in the Asian part of Russia. Ticks Tick-borne Dis 1(1):57–65

Rosenthal PJ (2004) Cysteine proteases of malaria parasites. Int J Parasitol 34(13–14):1489–1499

Sajid M, Mckerrow JH (2002) Cysteine proteases of parasitic organisms. Mol Biochem Parasitol 120(1):1–21

Santos TRBD, Gonzales JC, Chies JM, Rosa Farias NAD (2000) *Babesia bovis* transovarian transmission in *Boophilus microplus*: obtention of a Babesia free tick strain. Ciênc Rural 30(3):455–459

Smit JD, Grandjean O, Guggenheim R, Winterhalter KH (1977) Haemoglobin crystals in the midgut of the tick Ornithodorus moubata Murray. Nature 266(5602):536–538

Sojka D, Hajdušek O, Dvořák J, Sajid M, Franta Z, Schneider EL, Craik CS, Vancová M, Burešová V, Bogyo M, Sexton KB, McKerrow JH, Caffrey CR, Kopáček P (2007) IrAE: an asparaginyl endopeptidase (legumain) in the gut of the hard tick *Ixodes ricinus*. Int J Parasitol 37(7):713–724

Sojka D, Franta Z, Frantová H, Bartošová P, Horn M, Váchová J, O'Donoghue AJ, Eroy-Reveles AA, Craik CS, Knudsen GM, Caffrey CR, McKerrow JH, Mareš M, Kopáček P (2012) Characterization of gut-associated cathepsin D hemoglobinase from tick *Ixodes ricinus* (IrCD1). J Biol Chem 287(25):21152–21163

Sojka D, Franta Z, Horn M, Caffrey CR, Mareš M, Kopáček P (2013) New insights into the machinery of blood digestion by ticks. Trends Parasitol 29(6):276–285

Tirloni L, Kim TK, Coutinho ML, Ali A, Seixas A, Termignoni C, Mulenga A, da Silva Vaz I Jr (2016) The putative role of *Rhipicephalus microplus* salivary serpins in the tick-host relationship. Insect Biochem Mol Biol 71:12–28

Tsuji N, Miyoshi T, Battsetseg B, etal (2008) A cysteine protease is critical for *Babesia spp*. Transmission in *Haemaphysalis Ticks PLoS* Pathogens 4(5):e1000062

Uchida K, Ohmori D, Ueno T, Nishizuka M, Eshita Y, Fukunaga A, Kominami E (2001) Preoviposition activation of cathepsin-like proteinases in degenerating ovarian follicles of the mosquito Culex pipiens pallens. Dev Biol 237(1):68–78

Villar M et al (2013) Identification and characterization of Anaplasma phagocytophilum proteins involved in infection of the tick vector, Ixodes scapularis. Infection of immature *Ixodes scapularis* (Acari: Ixodidae) by membrane feeding. Two Anaplasma phagocytophilum strain. J Clin Microbiol 51(3):954–958

Williamson AL, Brindley PJ, Knox DP, Hotez PJ, Loukas A (2003) Digestive proteases of blood-feeding nematodes. Trends Parasitol 19(9):417–423

WonKoo L, JaeWon L, SoYoung L, InYong L (1997) Redescription of *Haemaphysalis flava* and *Ixodes tanuki* collected from a raccoon dog in Korea. Korean J Parasitol 35(1):1

Xu XL, Cheng TY, Yang H, Liao ZH (2016) De novo assembly and analysis of midgut transcriptome of *Haemaphysalis flava* and identification of genes involved in blood digestion, feeding and defending from pathogens. Infect Genet Evol 38(1):62–72

Yamaji K, Miyoshi T, Hatta T, Matsubayashi M, Alim MA, Anisuzzaman, Kushibiki S, Fujisaki K, Tsuji N (2013) HlCPL-A, a cathepsin L-like cysteine protease from the *ixodid* tick *Haemaphysalis longicornis*, modulated midgut proteolytic enzymes and their inhibitors during blood meal digestion. Infect Genet Evol 16(2):206–211

Yoon SY, et al. (1996) A case of tick bite caused by *Haemaphysalis flava* 34:326–330