Immunogenetics (2019) 71:421–432
https://doi.org/10.1007/s00251-019-01117-w

ORIGINAL ARTICLE

Identification and characterization of the lamprey cathepsin genes

Dong Wang[1,2] · **Peng Su**[1,2] · **Xuefeng Wang**[1] · **Kai Liu**[1,2] · **Changzhi Li**[1,2] · **Xingxing Gao**[1,2] · **Jiali Lu**[1,2] · **Feng Sun**[1,2] · **Qingwei Li**[1,2] · **Yue Pang**[1,2]

Received: 2 February 2019 / Accepted: 10 April 2019 /

Abstract

Cathepsins are key mammalian proteases that play an important role in the immune response. Several studies have revealed the versatile and critical functions of cathepsins. Here, we obtained ten kinds of cathepsin homologs and identified seven homologs with complete coding sequences. Phylogenetic analysis verified their identities and supported the classification of cathepsins into seven families, which is similar to other vertebrates. Tissue-specific expression analysis showed that all lamprey cathepsins (L-cathepsins) are present in the supraneural body (SB), kidney, gill, intestine, brain, heart, and liver, but their relative abundance varied among tissues. Additionally, we focused on the lamprey cathepsin L (L-cathepsin L) and used recombinant L-cathepsin L protein (rL-cathepsin L) to prepare anti rL-cathepsin L polyclonal antibodies, which were used to detect its distribution in lamprey tissues. The L-cathepsin L protein was primarily detected in the SB, kidney, gill, intestine, brain, and liver via western blot and immunohistochemistry assays. Importantly, quantitative real-time PCR (RT-PCR) revealed that the expression level of L-cathepsins mRNA significantly increased after exposure to three different stimuli (poly I:C, *Staphylococcus aureus* (S.a) and *Vibro anguilarum* (V.an)). This suggested that L-cathepsins may participate in defense processes. These results revealed that L-cathepsins may play key roles in the immune response to exogenous stimuli. The findings provide important information for future studies aiming to understand the molecular mechanisms underlying the immune response to pathogen invasion in lamprey.

Keywords Lamprey · Cathepsins · Evolution · Immune

Introduction

Cathepsin is derived from the Greek kathepsein (to digest), and it is active in a slightly acidic environment (Willstätter and Bamann 1929). Researchers have already identified 15 kinds of cathepsins in humans to date. According to different catalytic active site residues, they have been classified into three different families: serine cathepsins (cathepsins A and G), aspartate cathepsins (cathepsins D and E), and cysteine cathepsins (cathepsins B, C, F, H, K, L, O, S, V, X, and W) (Stoka et al. 2016). Cathepsins are present in almost all species, and they widely evolved into several classes of proteases. All of the cathepsins are synthesized as the inactive form of zymogens that are further processed by active proteases or by autocatalysis which is regulated by endogenous protein inhibitors known as cystatins, stefins, thyropins, and serpins (Turk et al. 2002; Conus and Simon 2010). Cathepsins execute important functions in the cytosol, nucleus, and extracellular space; thus, they play key roles in physiological and pathological processes (Cocchiaro et al. 2017). Importantly, cysteine cathepsin proteases, which are normally found in lysosomes and belong to the papain superfamily of cysteine proteases, are a class of enzymes that are evolutionarily conserved as far as the prokaryotes. All of the cysteine cathepsins share high sequence homology and an identical peptidase domain, which provides a convenient way to analyze their functions (Olson and Joyce 2015), including intracellular protein degradation, autophagy, and inflammatory and immune responses (Rawlings et al. 2016). Cysteine cathepsins are a

---

Dong Wang, Peng Su, Xuefeng Wang and Kai Liu contributed equally to this work. Qingwei Li and Yue Pang are corresponding author.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00251-019-01117-w) contains supplementary material, which is available to authorized users.

✉ Qingwei Li
liqw@263.net

✉ Yue Pang
pangyue01@163.com

[1] College of Life Science, Liaoning Normal University, Dalian 116081, China

[2] Lamprey Research Center, Liaoning Normal University, Dalian 116081, China

group of endosomal proteases involved in antigen processing and preparation of MHC class II molecules for antigen presentation. After cleavage by cathepsins, the invariant chain—which blocks the antigen-presenting site during maturation of MHC molecules—is replaced by antigenic peptides (Hsing and Rydensky 2005; Dijkstra and Yamaguchi 2019). Additionally, conversation of cathepsin L and conserved orthologs with invariant chain residues suggest their long co-evolution in the antigen presentation pathway (Criscitiello et al. 2012). As a member of the cysteine cathepsin protease family, cathepsin L has been implicated in antigen processing (Watts 1997). Sea urchin cathepsin L is upregulated in LPS stimulated coelomocytes (Nair et al. 2005). The potent endoprotease activities of cathepsin L may be important in revealing and releasing antigens (Riese and Chapman 2000), which may play key roles in the immunization process.

Lampreys are known as cyclostomes, the most basal extant group of vertebrates which is thought to have existed largely unchanged for more than 500 million years (Nikitina et al. 2009; Gustafsson et al. 2008; Smith et al. 2013). Over 20 kinds of cathepsins have been identified in vertebrates, and their specific roles in metabolism have also been investigated; thus, the importance of cathepsins has recently been emphasized in invertebrates (Uinuk-Ool et al. 2003). In a previous study, we reported some basic information about L-cathepsin D in lampreys, including that it degrades hemoglobin, fibrinogen, and serum albumin (Xiao et al. 2015). Other family members of cathepsins have recently been identified through third-generation sequencing methods. After identification and classification, we confirm that there are approximately ten kinds of cathepsin homologs in lamprey: cathepsins A, B, D, K, L (cathepsin, cathepsin 1, cathepsin 2), O, and Z. In this study, the sequences of cathepsin family members from lamprey are first reported and analyzed, and the relative expression of *L-cathepsin* mRNA is detected. Importantly, analysis of the immune response of lamprey cathepsins (L-cathepsins) will help us to understand how lampreys defend pathogen invasion. Moreover, we focus on L-cathepsin L, a member of the cysteine cathepsin family with unique synthesis and transportation mechanisms (Reiser et al. 2010; Ishidoh and Kominami 2002). We examine the localization and distribution of L-cathepsin L in lamprey tissues by immunohistochemistry and western blot. This research lays the foundation for studying the origin and evolution of the innate immune system in lampreys.

Materials and methods

Animals and immunization

Adult lampreys (*Lampetra japonica*), which were captured from the Tong River of China, were kept in 4 °C water in a glass tank. Lamprey handling and all of the experimental procedures were approved by the Animal Welfare and Research Ethics Committee of the Institute of Dalian Medical University (Permit Number: SYXK2004-0029). We performed the animal experiments at the Institute of Dalian Medical University. The lampreys (three in each group) were intraperitoneally injected with phosphate-buffered saline (PBS) (0.1 mL/each), *Staphylococcus aureus* (S.a) and *Vibro anguilarum* (V.an) (approximately 1×10^8/each in 0.1 mL of PBS), and poly I:C (100 mM in 0.1 mL of PBS).

Cloning and analyzing the full-length cDNA of *L-cathepsin L* genes

The cDNA sequence was obtained from the third generation sequencing result. It was identified with the Basic Local Alignment Search Tool X (BLASTX) from the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/). Total RNA was isolated from the liver of lampreys. The procedure was conducted according to the manufacturer's instructions (TaKaRa, Dalian, China). The total RNA was dissolved in RNase-free water and stored at − 80 °C. Then, cDNA was synthesized from the total RNA samples by reverse transcription using PrimeScript™ II First Strand cDNA Synthesis Kit (TaKaRa, Dalian, China) according to the manufacturer's protocol. The specific PCR primers in the open reading frame (ORF) were designed by PRIMER 5 and were based on the lamprey cDNA library our laboratory established for L-cathepsin L. The ORF of L-cathepsin L was amplified and cloned into a pMD19-T Simple Vector for sequence confirmation (TaKaRa, Dalian, China). The full-length cDNA sequence of L-cathepsin L was spliced by Sequencher 4.2 software, and the deduced amino acids were analyzed with DNAMAN V6 and DNASTAR 5.0 software.

Phylogenetic tree and genomic synteny construction

Sequences of additional cathepsin family members from other species were obtained from NCBI (http://www.ncbi.nlm.nih.gov/). The multiple sequence alignments of cathepsins were performed by Bioedit 7.0 software using default settings. A neighbor-joining (NJ) tree was constructed by MEGA 7.0 software based on the pairwise deletion of gaps/missing data and a Poisson model matrix of an amino acid model with 1000 bootstrapped replicates. The genomic synteny was obtained from Genomics (http://www.genomicus.biologie.ens.fr) and Stowers Institute for Medical Research base (SIMRbase, https://genomes.stowers.org/).

Expression and purification of lamprey cathepsin L

The ORF of L-cathepsin L was subcloned into the pCold I vector with a His-tag, and the recombinant protein was expressed in *E. coli* BL21 (DE3) which was induced by 0.1 mM IPTG for 24 h at 16 °C. The recombinant protein was purified by Ni affinity chromatography (GE Healthcare). The protein concentration was measured with the BCA Protein Quantitation Kit (Solarbio, Beijing, China) and was analyzed with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Production of anti-L-cathepsin L polyclonal antibodies

New Zealand rabbits were used to generate polyclonal antibodies against the recombinant L-cathepsin L protein (rL-cathepsin L). The rabbits were immunized with multipoint intradermal injections four times every 2 weeks. For the first immunization, 500 mg of the rL-cathepsin L antigen in 500 mL of PBS mix was mixed with an equal volume of Freund's complete adjuvant (Sigma). For the rest of immunizations, the dose was halved and Freund's complete adjuvant (Sigma) was replaced with Freund's incomplete adjuvant (Sigma). The sera of the immunized rabbits were collected, and the polyclonal antibodies were purified by affinity chromatography in a protein G column (GE Healthcare). The titer of the purified antibody was monitored with ELISA.

Western blot analysis

The SB cells were isolated from lamprey SB tissues by 0.25% trypsin treatment; other tissues and cells were obtained the same way. The protein was extracted with RIPA Lysis Buffer. After measuring the concentration, protein samples were separated by 12% SDS-PAGE and transferred onto polyvinylidene difluoride (PVDF) membranes. Membranes were blocked with 5% skim milk for 3 h and subsequently probed overnight at 4 °C with the primary rabbit anti-rL-cathepsin L antibody. After washing 3 times with PBS-T, the membranes were incubated with HRP-conjugated goat anti-rabbit IgG (1:5000). After washing 3 times with PBS-T, the membrane was developed using ECL substrate (Beyotime, China).

Immunohistochemical staining

The paraffin sections were deparaffinized with xylene and rehydrated with descending concentrations of ethanol. Endogenous peroxidase was blocked by treatment with 3% H_2O_2 in methanol for 20 min at room temperature. After blocking the sections with 10% normal donkey serum for 3 h at room temperature, they were incubated with the anti-rL-cathepsin L rabbit polyclonal antibody at 4 °C overnight. Next, they were incubated with HRP-conjugated goat anti-rabbit IgG (1:100) and rinsed with PBS-T. Normal rabbit IgG was used as a negative control. Subsequently, the slides were stained with diaminobenzidine (DAB), and the time of coloration depended on the real-time observation of coloration. The coloration was terminated with distilled water, and the sections were counterstained with hematoxylin. Following dehydration, the sections were passed through xylene two times for 15 min and then mounted in neutral resin.

Immunofluorescence

Lamprey SB cells were isolated from lamprey by 0.25% trypsin treatment. Then, they were plated in wells and fixed for 25 min in 4% paraformaldehyde in PBS at room temperature. After washed twice with PBS, the cells were fixed and

Table 1 Gene information for entire cathepsins identified in *Lampetra japonica* and their primer sequences for quantitative Real-time PCR

No.	Classes of cathepsins	Protease type	Length (aa)	pI	MW (kDa)	RT-PCR primer
1	Cathepsin A	Serine Cathepsin	474	5.92	52885.31	F:5'GTGGACATGGCCTGCAACTT-3' R:5'GAGGAAGGCAAGGTTGGAGAA-3'
2	Cathepsin B	Cysteine Cathepsin	333	5.75	35947.55	F:5'-CCGCCCCTACTCCATCAAG-3' R:5'-TCGATGTTGTACGCCTTCTTACC-3'
3	Cathepsin D	Aspartic Cathepsin	397	6.23	42772.47	F:5'-ACCACAAGTACGACTCCTCCAAGT-3' R:5'-CTCCGAACACCTGCTTCTTCAC-3'
4	Cathepsin K	Cysteine Cathepsin	184	5.56	20370.46	F:5'-GCCAGCCTCAACTCCTTCAC-3' R:5'-TCGTTCCCCAGCTGTTCTTG-3'
5	Cathepsin	Cysteine Cathepsin	329	6.07	37157.65	
6	Cathepsin L	Cysteine Cathepsin	157	4.34	16879.54	F:5'-CCTCGGCAAGCACTCCTACA-3' R:5'-TCTGTCCACTTCTTCACCTCGTT-3'
7	Cathepsin 1	Cysteine Cathepsin	799	5.92	85877.43	
8	Cathepsin 2	Cysteine Cathepsin	347	5.48	38686.22	
9	Cathepsin O	Cysteine Cathepsin	240	5.95	25927.34	F:5'-AGGAGGTGGGTTGCTCAGAA-3' R:5'-AAATTGAGGGCAGAACCATGTC-3'
10	Cathepsin Z	Cysteine Cathepsin	242	5.54	26749.76	F:5'-ACATGGGAGTGTGGAAATTCG-3' R:5'-CGGGTGAAGTTCTTGATGTTGA-3'

Table 2 Accession numbers of L-cathepsins

Gene Name	Accession number
Cathepsin A	MK676073
Cathepsin B	MK676074
Cathepsin K	MK676075
Cathepsin L	MK676076
Cathepsin	MK676077
Cathepsin 1	MK676078
Cathepsin 2	MK676079
Cathepsin O	MK676080
Cathepsin Z	MK676081

permeabilized for 10 min using 0.1% Triton X-100. They were blocked with normal goat serum for 30 min and incubated with the rabbit anti-rL-cathepsin L primary antibody (1:200) at 4 °C overnight. Next, the cells were washed with PBS twice and then incubated with Alexa Fluor 488-conjugated goat anti-rabbit IgG antibody (1:400). Following three more washes with PBS, the cells were stained with DAPI (1:1000). After two washes with PBS, the coverslips were mounted on glass slides with one drop of antifade solution. The immunofluorescence was visualized and captured with a Zeiss LSM 780 inverted microscope (Carl Zeiss, Inc).

FACS analysis

Lamprey SB cells were plated in wells and fixed for 20 min in 90% methanol in PBS at room temperature. Then, the cells were washed three times with PBS, blocked with normal goat serum for 30 min, and incubated with rabbit anti-rL-cathepsin L antibody (1:200) in PBS for 1 h at room temperature. After three washes, the cells were incubated with FITC-conjugated donkey anti-rabbit IgG (1:500) for 45 min at room temperature protected from light. Then, they were washed three times with PBS. The cells were resuspended with PBS and analyzed on a FACSAria flow cytometer (BD Biosciences). Cells incubated with FITC-conjugated goat anti-rabbit IgG were used as isotype controls. Data analysis was performed using Flowjo software (Tree Star).

Quantitative RT-PCR

The immunized lampreys were anesthetized with 0.05% tricaine methanesulfonate (MS-222; 3-aminobenzoic acid ethyl ester, Sigma), and then their SB cells were collected. According to the manufacturer's instructions, total RNA was extracted with the MiniBEST Universal RNA Extraction kit (TaKaRa, Dalian, China). Reverse transcription was performed as previously described. The cDNA was used as template to determine the mRNA expression of *L-cathepsins*. The sequences of primers used for real-time PCR (RT-PCR) are listed in Table 1. The PCRs were carried out in a TaKaRa PCR Thermal Cycler Dice Real Time System using a TaKaRa SYBR® PrimeScript™ RT-PCR Kit. Each reaction was performed in triplicate, and the data were normalized to L-GAPDH, which was the internal control to normalize the starting quantity of RNA. Additionally, the heatmap was constructed by TB tools.

Statistical analysis

All of the data are presented as the means ± SE based on separate experiments. Differences between treatment groups were analyzed by Student's *t* tests. $P < 0.05$ was set as the threshold for significance ($^{*}P < 0.05$, $^{**}P < 0.01$).

Results

Cloning, sequence analysis of *L-cathepsins*, and phylogenetic analysis

The sequences of cathepsins were obtained from our library and were sent to GenBank (Table 2). The open reading frame (ORF) sequence of cathepsins from lamprey was analyzed along with human cathepsins (Fig. 1). Additionally, the number of amino acids, molecular weight (MW), and isoelectric point (pI) for each *L-cathepsin* was calculated and predicted, as shown in Table 1. Functional domain analysis demonstrated that *L-cathepsins* contain similar domains to human cathepsins (Fig. 1). To further determine the phylogenetic relationship of the cathepsin family in lamprey, we constructed a neighbor-joining tree (NJ tree) using the amino acid sequence of cathepsin family members. The tree suggested that all members of the cathepsin family in lamprey were classified into corresponding clusters. Interestingly, four sequences were classified into the *cathepsin L* cluster. According to their sequence similarity, they were named *L-cathepsin*, *L-cathepsin 1*, *L-cathepsin 2*, and *L-cathepsin L*. The phylogenetic analysis indicated that each cathepsin in lamprey agreed with the species unique characteristic and followed a lower to higher evolutionary relationship according to their evolutionary history (Fig. 2).

Genomic synteny analyses of the lamprey cathepsin family

Genomic synteny analyses can be useful not only for establishing gene homology relationships but also in providing clues about the mechanistic origin of new genes. To better understand the evolution of the cathepsin gene family during vertebrate expansion, the neighboring gene environment of lamprey cathepsins was compared between fish and mammals. There are four species in the genomic synteny analysis: human, zebrafish, hagfish and lamprey (Fig. 3). In the

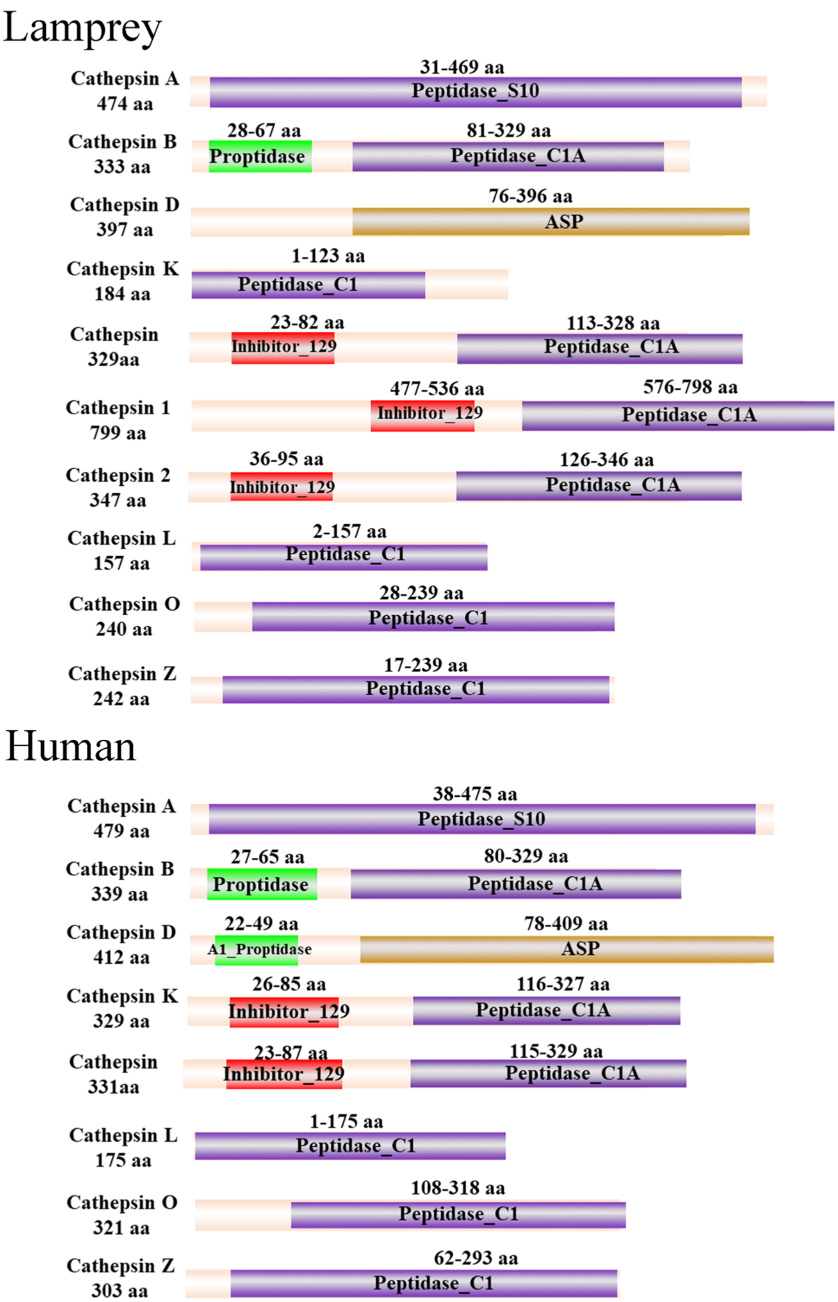

Fig. 1 Schematic of lamprey and human cathepsins proteins as well as their conserved domains. Similar domains are color-coded with the same color; the "peptidase" domain is purple, the "proptidase" domain is green, and the "inhibitor" domain is red. The length of the domains is presented above the diagram. The amino acid (aa) lengths are presented on the left of the diagram

lamprey genomes, the cathepsin genes, although detectable, could not be examined for syntenic relationship with mammalian cathepsin family genes because of the incompleteness of the genome sequence; the hagfish faced the same problem (Fig. 3). Strong syntenic relationships among cathepsin A, B, and Z orthologs were easily detected in the human and zebrafish genome sequences that we examined. The data reveals a high degree of evolutionary lability within cathepsin gene clusters. Of note, the TBX20 gene is present in both lamprey and zebrafish (Fig. 3).

Expression of L-cathepsin L recombinant protein and preparation of anti-rL-cathepsin L rabbit polyclonal antibody

The recombinant protein was expressed in *E. coli* BL21 (DE3) after induction with 0.1 mM IPTG for 24 h at 16 °C, as detected in the SDS-PAGE analysis (Supplementary 1a). The purified recombinant protein migrated as a single bind on the SDS-PAGE gel with a molecular mass of approximately 21 kDa. This was agrees

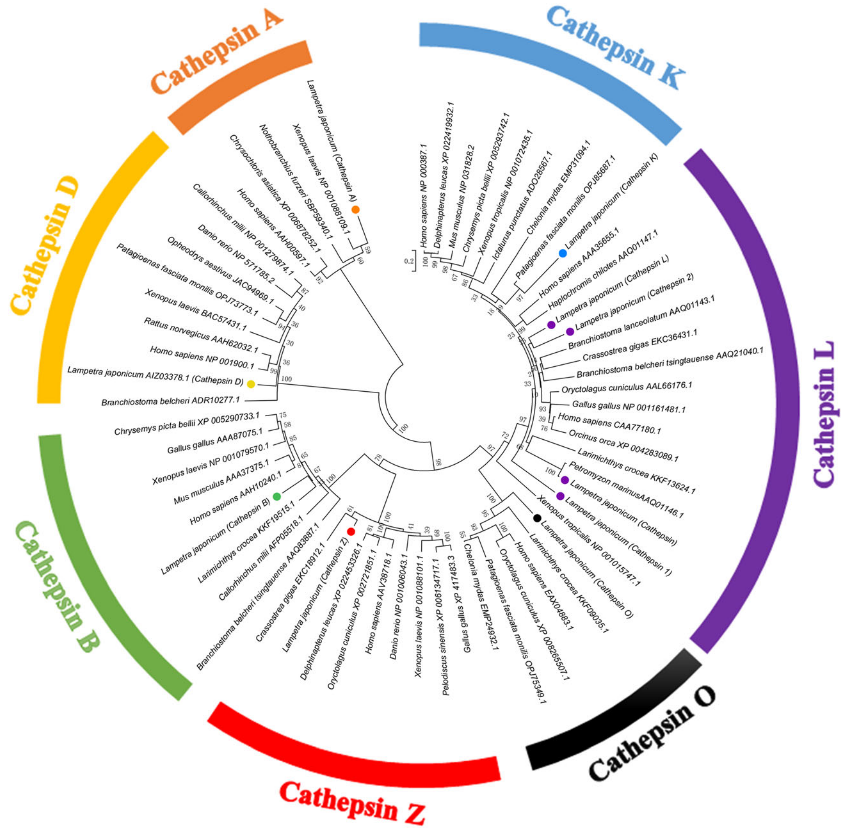

Fig. 2 Evolutionary relationships of L-cathepsin proteins. Phylogenetic analysis was conducted in MEGA 7.0. The evolutionary history was inferred using the NJ method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset

with the molecular mass predicted from the DNA sequence (Supplementary 1b). The polyclonal antibodies against rL-cathepsin L protein were purified from rabbit serum by protein GSepharose affinity chromatography (Supplementary 1d). The titer at which the ELISA used the rabbit anti-rL-cathepsin L polyclonal antibodies was higher than 1:320,000 (Supplementary 1c), and preimmunized rabbit IgG was used as a negative control.

Localization and distribution of L-cathepsins in lamprey tissues

Expression of *L-cathepsins* mRNA in various tissues was determined by semi-quantitative RT-PCR using GAPDH as a control (Fig. 4a, b). *L-cathepsin*, *A*, *D*, and *L* showed widespread expression, and the rest of *L-cathepsin* family members were expressed in specific tissues. On the whole, the expression of *L-cathepsin K* and *O* were lower than that of the other family members. Interestingly, the results revealed that all of the L-cathepsins were expressed in the SB. Importantly, cathepsin L is a member of the cysteine cathepsin family and has a unique synthesis and transportation mechanism (Reiser et al. 2010; Ishidoh and Kominami 2002); therefore, we further investigated L-cathepsin L. The tissue distribution of L-cathepsin L was determined via immunohistochemistry (Fig. 5a) and western blot analysis (Fig. 5b) using a specific anti-rL-cathepsin L antibody. The results revealed that L-cathepsin L protein is expressed in the SB, liver, kidney, intestine, and gill. The highest expression level was in the SB, followed by the kidney, gill, and liver. Lower expression levels were observed in the brain and intestine, while the protein was not detected in the heart or oral glands. To investigate the intracellular localization of L-cathepsin L in lamprey SB cells, we used immunofluorescence. The results demonstrated that L-cathepsin L mainly localizes in the cytoplasm of SB cells (Fig. 5c).

Relative expression of *L-cathepsins* mRNA in stimulated lamprey SB

Functionally, SB processes constantly survey the environment and act as important responders to inflammation. Additionally, the lamprey SB contains all of the L-cathepsins (Fig. 4a, b). To assess the response of *L-cathepsins* to different antigens, we performed RT-PCR with total RNA extracted from the SB of lampreys immunized with gram-positive bacteria (S.a), gram-negative bacteria (V.an), and a virus (poly I:C) for 1–3 days. When lampreys were challenged with virus (poly I:C), *L-cathepsin D* and *L* expression was significantly elevated and their expression peaked at 24 h compared to the other time points. *L-cathepsin K* and *Z* expression peaked at 72 h in response to virus, but there were no significant change in the expression of *L-cathepsins A*, *B*, and *L*. When lampreys were stimulated with gram-negative

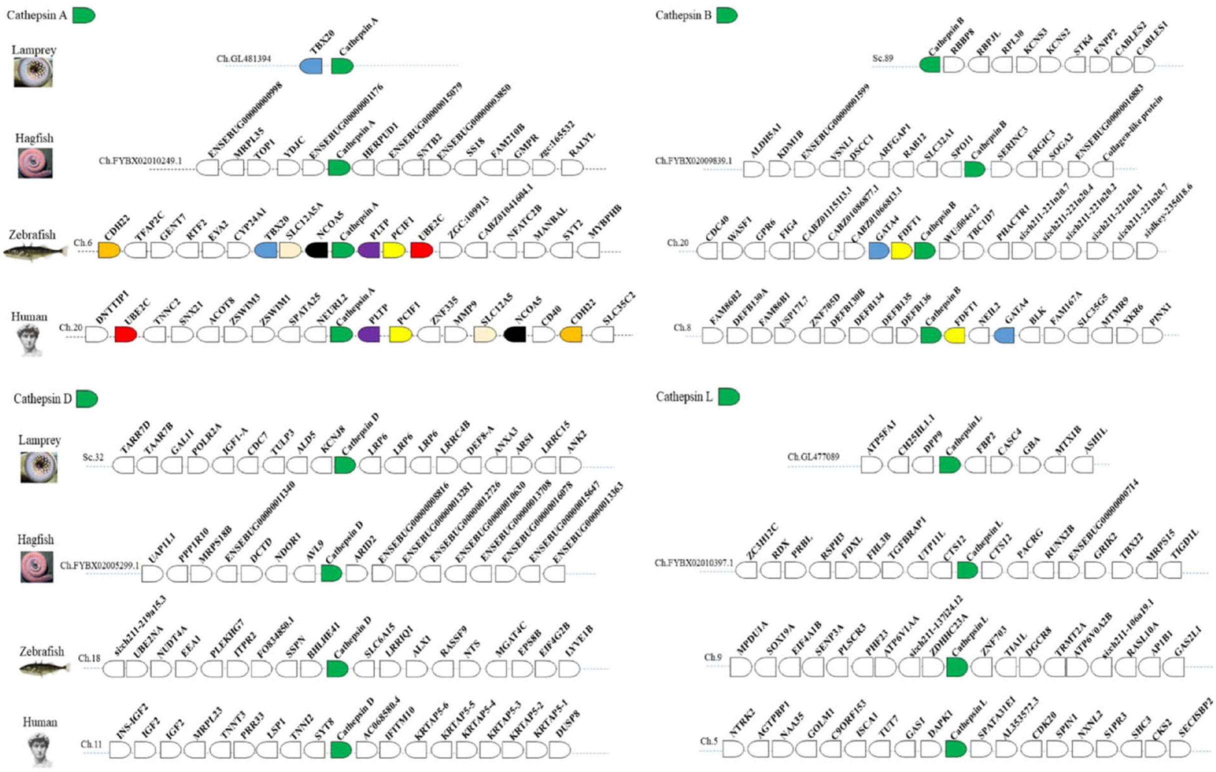

Fig. 3 Genomic environments of cathepsin family genes. Arrows or arrowheads in opposite directions indicate genes located on opposite strands. Ch and Sc denote chromosome and scaffold, respectively. Orthologous genes contributing to conserved synteny are similarly color-coded, and cathepsin clusters are green

bacteria (V.an), only *L-cathepsin A* did not change significantly of the targets tested. As the time from V.an exposure increased, the expression of *L-cathepsins D*, *K*, *L*, and *Z* increased and, importantly, the expression of *L-cathepsin B* and *O* were significantly upregulated at 72 h. The findings show that *L-cathepsin B* and *O* may exert their main function in defending lamprey from gram-negative bacteria (V.an) compared to other members of *L-cathepsins*. In contrast, the expression of *L-cathepsin A* increased compared to other cathepsins when lampreys were stimulated with gram-positive bacteria (S.a). As shown in Fig. 6, we found a substantial upregulation of *L-cathepsin* expression over the 3 days, regardless of which antigen the lamprey was stimulated with. Additionally, the same *L-cathepsins* presented different responses for different antigens. The results indicate that L-cathepsins play potential roles in many aspects of lamprey immune activities, and different members of L-cathepsins may exert different functions in antigen invasion.

Discussion

Cathepsins are widely expressed throughout the animal and plant kingdoms and have also been identified in viruses and bacteria (Lecaille et al. 2002). On the basis of the lamprey cDNA library our laboratory established, ten kinds of cathepsins have been detected in the lamprey genome; they are named cathepsins A, B, D, K, L (cathepsin, cathepsin 1, cathepsin 2), O, and Z. Cathepsin A belongs to the serine protease family; cathepsin D belongs to the aspartic protease family; and cathepsins B, K, L (cathepsin, cathepsin1, cathepsin2), O, and Z belong to the cysteine protease family. In particular, *L-cathepsin* is 100% similar to the *Petromyzon marinus* ortholog (AAQ01146.1), and the NJ tree revealed that it belongs to the *cathepsin L* family. All *L-cathepsins* in lamprey contain a conserved peptidase domain except for *L-cathepsin D*, which contains the ASP domain consistent with human *cathepsin D*. Although cathepsins are conserved proteins, there are some differences. For example, *L-cathepsin D* in lamprey does not contain the propeptide domain which is a potent inhibitor of its cognate enzyme in vitro (Turk et al. 2002; Caglic et al. 2007). Moreover, genomic synteny reveals a high degree of evolutionary lability within cathepsin gene clusters.

To investigate the potential function of cathepsins in lampreys, the tissue distribution pattern was detected by semi-quantitative RT-PCR. The tissue-specific expression was determined by calculating the grayscale value (Fig. 4a, b). In

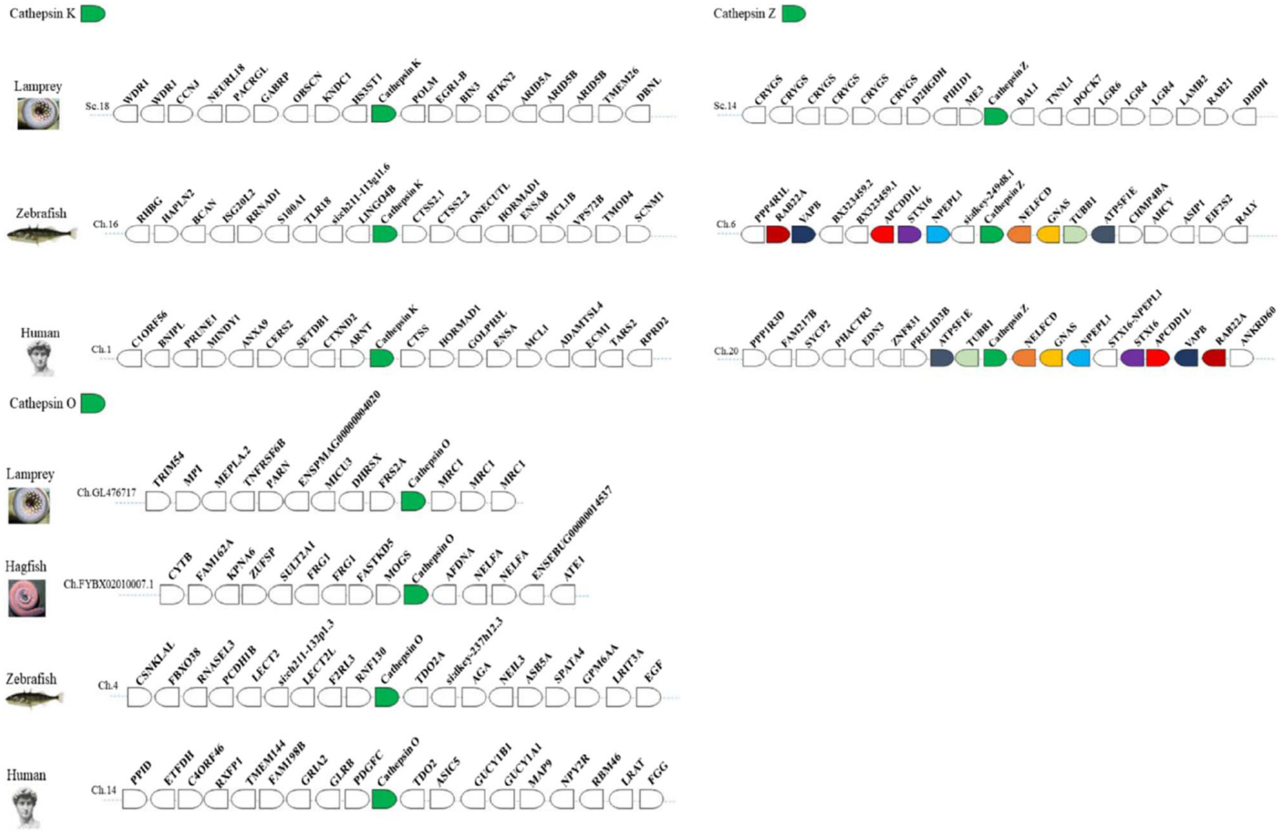

Fig. 3 continued.

humans, cathepsins are widely expressed in most tissues and cell lines (Rossi et al. 2004). However, not all cathepsins are expressed at the same levels in different tissues (Rossi et al. 2004; Unanue et al. 2016). Some of the cathepsins, such as cathepsins B, L, and O, are ubiquitous and highly to moderately expressed in various tissues and cells including the brain. This suggests that these enzymes play a role in general protein degradation and turnover. Such diverse localization and specific expression of some cysteine cathepsins suggest that they have more specific cellular functions (Cook et al. 2014; Duncan et al. 2008; Goulet et al. 2004). However, the tissue distribution of cathepsins has not been well studied in fish species. We discovered that all of the L-cathepsins are expressed in the liver, kidney, SB, leukocytes, and oral glands, even though the levels of expression are different. Meanwhile, L-cathepsins L and D expressed in all of the tissues tested. Moreover, all L-cathepsins are detectable in immune-relevant tissues, suggesting their potential function in defending infections. The different expression patterns in different species suggest that cathepsins might play different roles in different species. Further studies are necessary to correlate and compare the results of cathepsin tissue distribution patterns in lamprey and other organisms. Moreover, as an important immune tissue of the immune defense system, the lamprey SB plays an important role in the native and adaptive immune systems (Pang et al. 2015). Our western blot and immunohistochemical staining showed that the L-cathepsin L protein is abundant in the SB, implying that L-cathepsins may play a crucial role in the SB in pathogen defense. In recent years, cathepsins have been shown to have important roles in processing and/or degradation of several important neuronal proteins, thereby exerting either neuroprotective or harmful effects (Stoka et al. 2016). L-cathepsin A, D, and L have high expression in the lamprey brain, suggesting that they might be involved in neuroinflammation. This hypothesis needs further investigation.

Cathepsins are key acid hydrolases within the lysosome, and they represent the principal effectors of protein catabolism and autophagy (Pan et al. 2012; Dennemärker et al. 2010a, b). Intracellular cathepsins are also crucial components of signal transduction pathways that drive inflammation (Navab et al. 2008; Bruchard et al. 2013; Dennemärker et al. 2010a, b; Park et al. 2008). Moreover, cathepsins are a group of peptidases involved in different processes of immune responses, including apoptosis, inflammation, and antigen processing (Zavasnik-Bergant and Turk 2006). However, their detailed immune roles against pathogen infection are still limited to a handful of fish species. For example, *V. parahaemolyticus* and *Listeria monocytogenesthe* challenge in the disk abalone (*Haliotis discus*) results in high expression of cathepsin Z in

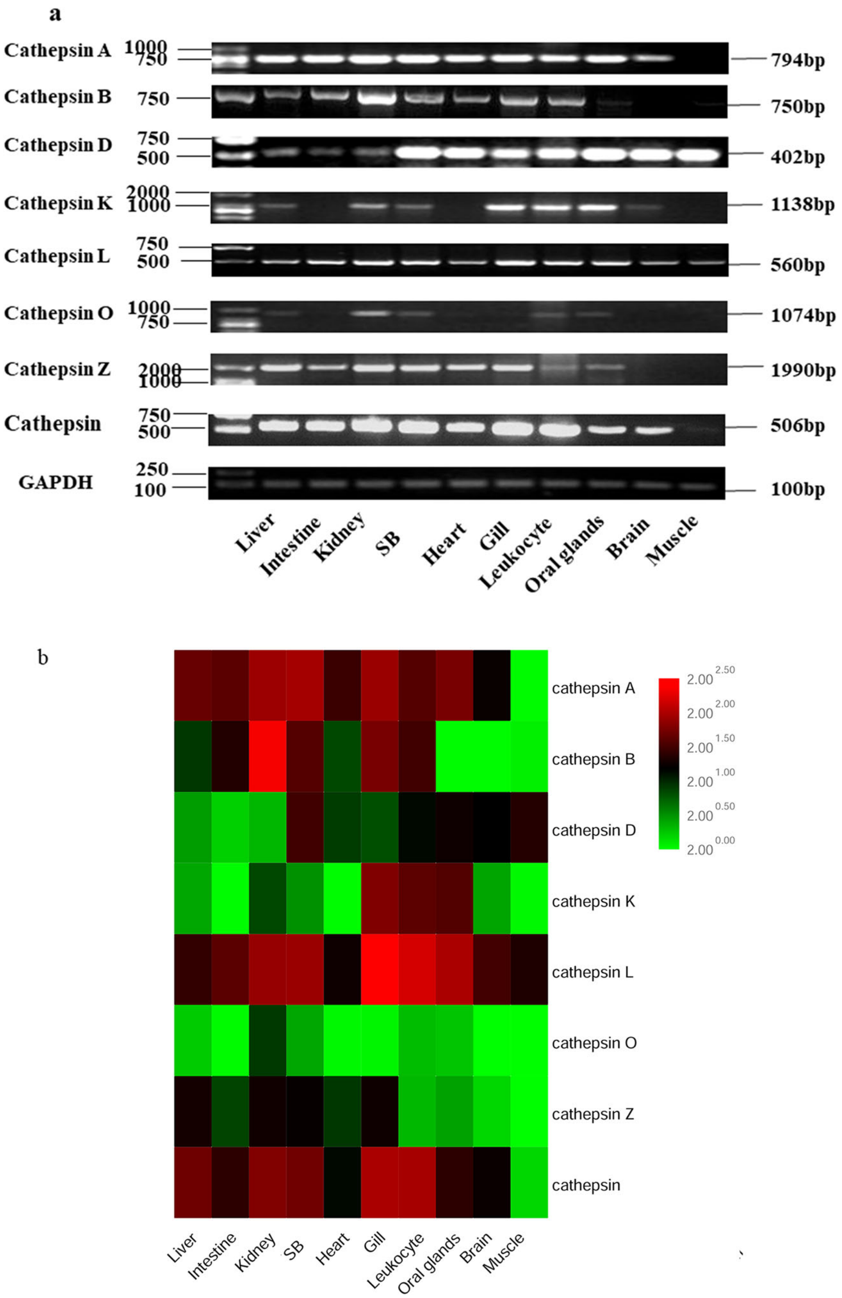

Fig. 4 Expression of L-cathepsins in lamprey tissues. **a** The mRNA expression level of *L-cathepsins* in lamprey tissues detected by semi-quantitative RT-PCR. **b** Heatmap of mRNA expression levels of *L-cathepsins* in lamprey tissues. Green and red represent low-to-high expression levels, and the color scales correspond to the expression values of the microarray

the gill and hemocytes, which indicates that cathepsin Z is involved in host innate immune responses (Godahewa et al. 2017). Our RT-PCR analysis revealed that L-cathepsins participate in the immune defense process for various antigens such as S.a, V.an and poly I:C. The expression levels of *L-cathepsin D* and *L* were upregulated 48 h after virus (poly I:C) stimulation, and the levels of *L-cathepsin Z* were upregulated significantly after 72 h. This implies that L-cathepsins A, B, and L may not contribute to the virus immune response of the lamprey, while others play key roles in this process. In response to gram-negative bacteria (V.an), the expression levels of *L-cathepsin K* were upregulated at 48 h, and the levels of *L-cathepsins B* and *O* were upregulated significantly at 72 h. When the lampreys were challenged with gram-positive bacteria (S.a), the expression levels of *L-cathepsin A* were upregulated significantly at 72 h. These findings indicate that the lamprey may have different immune mechanisms to defend various antigens, and different L-cathepsins exert different functions in lamprey immune responses. We predict that L-cathepsins participate in the immune response by

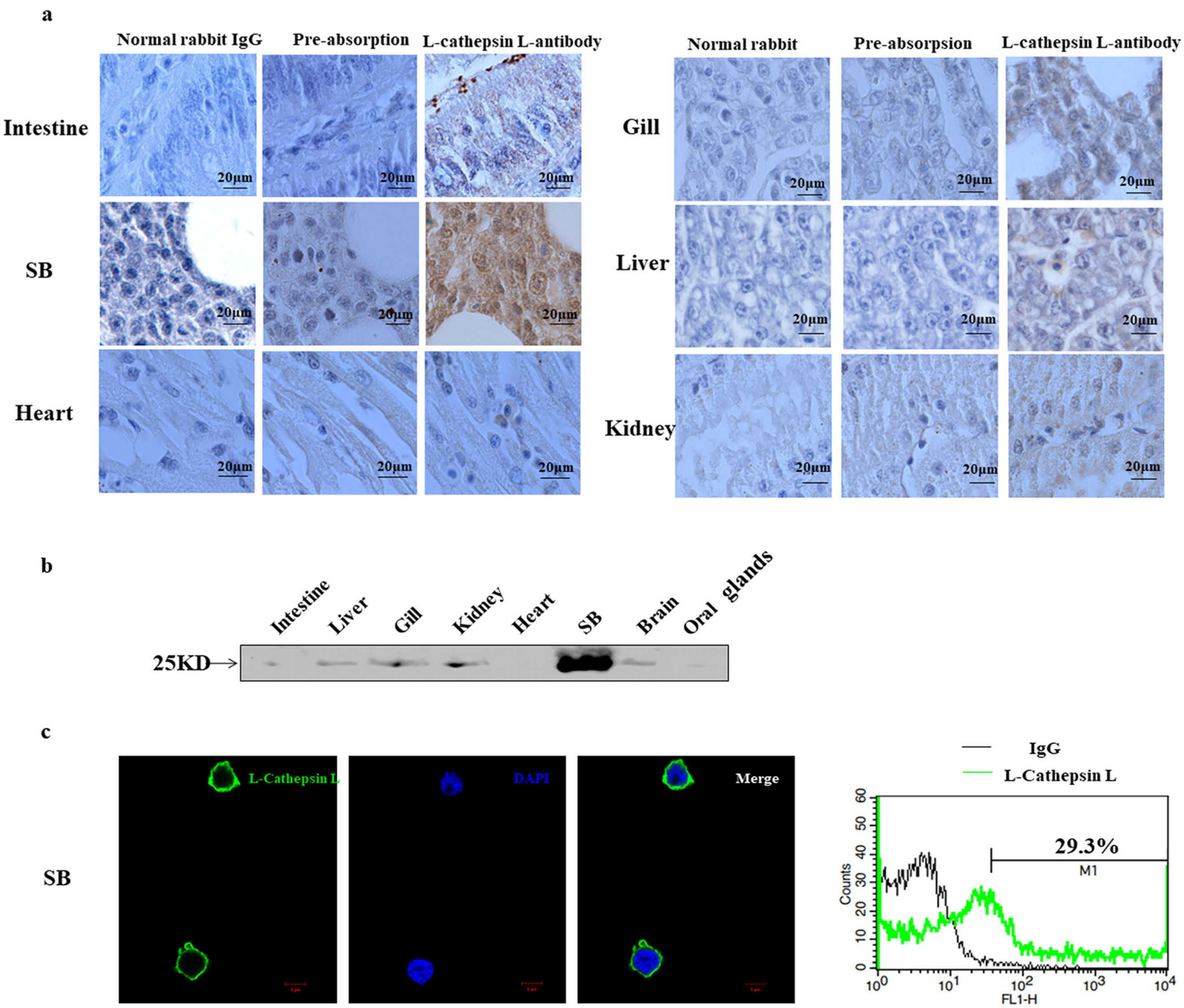

Fig. 5 Localization and distribution of L-cathepsins in lamprey tissues. **a** Protein in lamprey tissues, as observed via immunohistochemical staining. The tissue sections were incubated with the rabbit anti-rL-cathepsin L antibody (1:2000). Preabsorption of the antibody with recombinant L-cathepsin L abolished the signal, which supports the specificity of the signal. Normal rabbit IgG antibody (1:5000) also served as a negative control. Scale bars, 20 mm. **b** Protein expression level of L-cathepsin L in lamprey tissues detected with western blot. Lane 1: intestine; Lane 2: liver; Lane 3: gill; Lane 4: kidney; Lane 5: heart; Lane 6: SB; Lane 7: brain; Lane 8: oral glands. **c** Immunofluorescence detection of L-cathepsin L in SB cells. The primary antibody was rabbit anti-rL-cathepsin L (1:200) and the secondary antibody was an Alexa Fluor 488-conjugated goat anti-rabbit IgG (1:400). The SB cells were treated with an isotype-matched IgG antibody and then with FITC-labeled goat anti-rabbit IgG antibody as a control. The cells were stained with DAPI. The immunofluorescence was visualized and captured using a fluorescence microscope (Carl Zeiss, Inc). Scale bars, 5 μm. FACS analysis of L-cathepsin L in SB cells. The SB cells that were bound to the rabbit anti-rL-cathepsin L polyclonal antibodies were incubated with FITC-conjugated anti-rabbit IgG secondary antibodies followed by FACS analysis (green). The cells were selected first by light scattering (FSC, forward scatter; SSC, side scatter). The SB cells were first treated with isotype-matched IgG antibody and then with FITC-labeled goat anti-rabbit IgG antibody (black)

presenting the antigen, which is similar to the role in humans (Palermo and Joyce 2008); however, this notion needs further study for validation. Meanwhile, exposure to different antigens leads to different responses in lampreys. This suggests that lamprey cathepsins may exert different functions in different immune responses, and it may provide some clues to explore the evolution of the innate immune system in lampreys.

In conclusion, we first detected and identified the cathepsin family members in lampreys, including cathepsins A, B, D, K, L (cathepsin, cathepsin 1, cathepsin 2), O, and Z. Additionally, the tissue-specific expression of L-cathepsins was described. Importantly, we discovered that the cathepsins may play key roles in the lamprey immune response and that lamprey cathepsins may take part in different immune response mechanisms to protect the species from foreign antigen invasion.

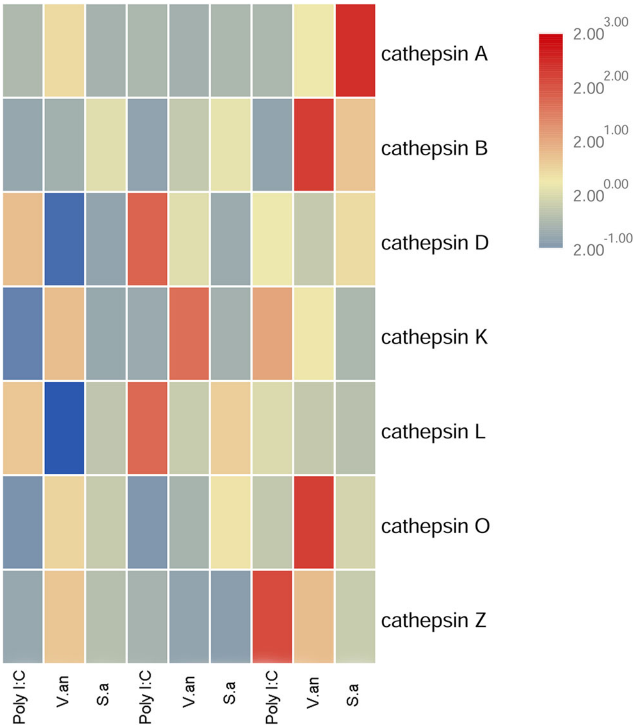

Fig. 6 Different antigens upregulated the expression of L-cathepsins. The heatmap of L-cathepsin expression levels, which were determined via RT-PCR in lamprey SB cells after stimulation with poly I:C, *Vibro anguilarum* (V.an) and *Staphylococcus aureus* (S.a) at 24 h, 48 h, and 72 h. Blue and red colors represent low to high expression levels, and the color scales correspond to the expression values of the microarray

Funding information This work was funded by the Chinese Major State Basic Research Development Program (973 Program; Grant2013CB835304), the Marine Public Welfare Project of the State Oceanic Administration (No.201305016), Chinese National Natural Science Foundation Grants (No. 31170353, No. 31202020, No. 31772884, and No. 31801973), Science and Technology Innovation Fund Research Project (No. 2018J12SN079), and the project of Department of Ocean and Fisheries of Liaoning Province (No. 201805).

Compliance with ethical standards

Competing interests The authors declare that they have no conflicts of interest.

References

Bruchard M, Mignot G, Derangère V, Chalmin F, Chevriaux A, Végran F, Boireau W, Simon B, Ryffel B, Connat JL, Kanellopoulos J, Martin F, Rébé C, Apetoh L, Ghiringhelli F (2013) Chemotherapy-triggered cathepsin B release in myeloid-derived suppressor cells activates the Nlrp3 inflammasome and promotes tumor growth. Nat Med 19(1): 57–64

Caglic D, Pungercar JR, Pejler G, Turk V, Turk B (2007) Glycosaminoglycans facilitate procathepsin B activation through disruption of propeptide-mature enzyme interactions. J Biol Chem 282(45):33076–33085

Cocchiaro P, De Pasquale V, Della Morte R, Tafuri S, Avallone L, Pizard A, Moles A, Pavone LM (2017) The multifaceted role of the lysosomal protease cathepsins in kidney disease. Front Cell Dev Biol 5:114

Conus S, Simon HU (2010) Cathepsins and their involvement in immune responses. Swiss Med Wkly 140:w13042

Cook C, Stankowski JN, Carlomagno Y, Stetler C, Petrucelli L (2014) Acetylation: a new key to unlock tau's role in neurodegeneration. Alzheimers Res Ther 6(3):29

Criscitiello MF, Ohta Y, Graham MD, Eubanks JO, Chen PL, Flajnik MF (2012) Shark class II invariant chain reveals ancient conserved relationships with cathepsins and MHC class II. Dev Comp Immunol 36(3):521–533

Dennemärker J, Lohmüller T, Mayerle J, Tacke M, Lerch MM, Coussens LM, Peters C, Reinheckel T (2010a) Deficiency for the cysteine protease cathepsin L promotes tumor progression in mouse epidermis. Oncogene 29(11):1611–1621

Dennemärker J, Lohmüller T, Müller S, Aguilar SV, Tobin DJ, Peters C, Reinheckel T (2010b) Impaired turnover of autophagolysosomes in cathepsin L deficiency. Biol Chem 391(8):913–922

Dijkstra JM, Yamaguchi (2019) Ancient features of the MHC class II presentation pathway, and a model for the possible origin of MHC molecules. Immunogenetics 71(3):233–249

Duncan EM, Muratore-Schroeder TL, Cook RG, Garcia BA, Shabanowitz J, Hunt DF, Allis CD (2008) Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation. Cell 135(2):284–294

Godahewa GI, Perera NCN, Lee S, Kim MJ, Lee J (2017) A cysteine protease (cathepsin Z) from disk abalone, Haliotis discus discus: genomic characterization and transcriptional profiling during bacterial infections. Gene 627:500–507

Goulet B, Baruch A, Moon NS, Poirier M, Sansregret LL, Erickson A, Bogyo M, Nepveu A (2004) A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor. Mol Cell 14(2):207–219

Gustafsson OS, Collin SP, Kröger RH (2008) Early evolution of multifocal optics for well-focused colour vision in vertebrates. J Exp Biol 211(Pt 10):1559–1564

Hsing LC, Rydensky AY (2005) The lysosomal cysteine proteases in MHC class II antigen presentation. Immunol Rev 207:229–241

Ishidoh K, Kominami E (2002) Processing and activation of lysosomal proteinases. Biol Chem 383(12):1827–1831

Lecaille F, Kaleta J, Brömme D (2002) Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. Chem Rev 102(12):4459–4488

Nair SV, Del Valle H, Gross PS, Terwilliger DP, Smith LC (2005) Macroarray analysis of coelomocyte gene expression in response to LPS in the sea urchin. Identification of unexpected immune diversity in an invertebrate. Physiol Genomics 22(1):33–47

Navab R, Pedraza C, Fallavollita L, Wang N, Chevet E, Auguste P, Jenna S, You Z, Bikfalvi A, Hu J, O'Connor R, Erickson A, Mort JS, Brodt P (2008) Loss of responsiveness to IGF-I in cells with reduced cathepsin L expression levels. Oncogene 27(37):4973–4985

Nikitina N, Bronner-Fraser M, Sauka-Spengler T (2009) The sea lamprey *Petromyzon marinus*: a model for evolutionary and developmental biology. Cold Spring Harb Protoc 2009(1):pdb.emo113

Olson OC, Joyce JA (2015) Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. Nat Rev Cancer 15(12):712–729

Palermo C, Joyce JA (2008) Cysteine cathepsin proteases as pharmacological targets in cancer. Trends Pharmacol Sci 29(1):22–28

Pan L, Li Y, Jia L, Qin Y, Qi G, Cheng J, Qi Y, Li H, Du J (2012) Cathepsin S deficiency results in abnormal accumulation of autophagosomes in macrophages and enhances Ang II-induced cardiac inflammation. PLoS One 7(4):e35315

Pang Y, Wang S, Ba W, Li Q (2015) Cell secretion from the adult lamprey supraneural body tissues possesses cytocidal activity against tumor cells. Springerplus 4:569

Park B, Brinkmann MM, Spooner E, Lee CC, Kim YM, Ploegh HL (2008) Proteolytic cleavage in an endolysosomal compartment is required for activation of Toll-like receptor 9. Nat Immunol 9(12): 1407–1414

Rawlings ND, Barrett AJ, Finn R (2016) Twenty years of the MEROPS database of proteolyticenzymes, their substrates and inhibitor. Nucleic Acids Res 44(D1):D343–D350

Reiser J, Adair B, Reinheckel T (2010) Specialized roles for cysteine cathepsins in health and disease. J Clin Invest 120(10):3421–3431

Riese RJ, Chapman HA (2000) Cathepsins and compartmentalization in antigen presentation. Curr Opin Immunol 12(1):107–113

Rossi A, Deveraux Q, Turk B, Sali A (2004) Comprehensive search for cysteine cathepsins in the human genome. Biol Chem 385:363–372

Smith JJ, Kuraku S, Holt C, Sauka-Spengler T, Li W et al (2013) Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. Nat Genet 45(4):415–421 421e1-412

Stoka V, Turk V, Turk B (2016) Lysosomal cathepsins and their regulation in aging and neurodegeneration. Ageing Res Rev 32:22–37

Turk B, Turk D, Salvesen GS (2002) Regulating cysteine protease activity: essential role of protease inhibitors as guardians and regulators. Curr Pharm Des 8(18):1623–1637

Uinuk-Ool TS, Takezaki N, Kuroda N, Figueroa F, Sato A, Samonte IE, Mayer WE, Klein J (2003) Phylogeny of antigen-processing enzymes: cathepsins of a cephalochordate, an agnathan and a bony fish. Scand J Immunol 58(4):436–448

Unanue ER, Turk V, Neefjes J (2016) Variations in MHC class II antigen processing and presentation in health and disease. Annu Rev Immunol 34:265–297

Watts C (1997) Capture and processing of exogenous antigens for presentation on MHC molecules. Annu Rev Immunol 15:821–850

Willstätter R, Bamann E (1929) Über die Proteasen der Magenschleimhaut. Erste Abhandlung über die Enzyme der Leukocyten. Hoppe Seylers Z Physiol Chem 180:127–143

Xiao R, Zhang Z, Wang H, Han Y, Gou M, Li B, Duan D, Wang J, Liu X, Li Q (2015) Identification and characterization of a cathepsin D homologue from lampreys (Lampetra japonica). Dev Comp Immunol 49(1):149–156

Zavasnik-Bergant T, Turk B (2006) Cysteine cathepsins in the immune response. Tissue Antigens 67(5):349–355