e-ISSN 1734-9168
http://www.isez.pan.krakow.pl/en/folia-biologica.html

Folia Biologica (Kraków), vol. 69 (2021), No 1
https://doi.org/10.3409/fb_69-1.05

Molecular Characterization, Tissue Distribution, and Expression Profiling of the *CTSD* Gene during Goose Ovarian Follicle Development

Jiaran ZHU, Shenqiang HU, Yao LU, Yujing RONG, Enhua QING, Liang LI, and Jiwen WANG

Accepted March 23, 2021 *Published online March 31, 2021* *Issue online March 31, 2021*

Original article

ZHU J., HU S., LU Y., RONG Y., QING E., LI L., WANG J. 2021. Molecular characterization, tissue distribution, and expression profiling of the *CTSD* gene during goose ovarian follicle development. Folia Biologica (Kraków) **69**: 39-48.

Cathepsin D (*CTSD*) is known to be crucial for the degradation and utilization of yolk precursors in ovarian follicles. However, little is known about its expression profiles and physiological actions in avian ovarian cells. In this study, the intact coding sequence of the *CTSD* gene in geese was cloned for the first time, with a length of 1197 bp. It encoded a polypeptide of 398 amino acids (AA) consisting of a signal peptide and two conserved functional domains (i.e., A1_Propeptide and Cathepsin_D2). The AA sequence of goose *CTSD* had > 96% similarities with the homologs of turkeys, chickens, and ducks. Results from real-time PCR showed that goose *CTSD* mRNA was present in all tissues examined, with higher levels in the adrenal gland, liver, heart, and reproductive organs. Furthermore, levels of *CTSD* mRNA were much higher in goose granulosa layers than in the theca layers in any follicular category. Significantly, its expression remained almost unchanged in the theca layers throughout follicle development, while it increased gradually in the granulosa layers from 2-4 mm to F5 follicles but declined there after. These results suggested that *CTSD* may regulate goose ovarian follicle development through its actions on both the degradation and absorption of yolk precursors and granulosa cell apoptosis.

Key words: Goose, *CTSD*, molecular cloning, granulosa cell, follicle development.

Jiaran ZHU, Shenqiang HU✉, Yao LU, Yujing RONG, Enhua QING, Liang LI, Jiwen WANG✉, Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Sichuan Agricultural University, Chengdu 611130, China.
E-mail: wjw2886166@163.com; shenqiang.hu@sicau.edu.cn

In birds, ovarian follicles committed to ovulation are usually arranged in a strict hierarchy, comprising hundreds of visible follicles at different developmental stages (HERNANDEZ & BAHR 2003; HU & ZADWORNY 2017). During the reproductive lifespan of birds, only a very small number of follicles will be ultimately ovulated while most of them will undergo apoptosis which could take place at any developmental stage (JOHNSON *et al.* 1996; JOHNSON & WOODS 2009). In particular, compared to ducks and chickens with annual egg production numbers of 250, domesticated geese (*Anser cygnoides domestica*) show poor egg production performance, usually laying 60 eggs per year, which severely impedes the rapid development of the goose industry. Because the fate of ovarian follicles determines goose lifetime reproductive performance and depends on a well-organized follicular hierarchy, it is of great theoretic and practical value to identify the factors regulating goose follicle development as well as the underlying mechanisms.

It is known that avian follicle development is not only characterized by the differentiation of follicular cells, including granulosa and theca cells, but also accompanied by the deposition and absorption of liver-synthesized yolk precursors, *i.e.,* very low density lipoprotein (VLDL) and vitellogenin (VTG) (ROMANO *et al.* 2004). Over the past two decades, a lot of studies have been conducted to explore the molecular mechanisms controlling the differentiation of follicular cells (JOHNSON 2015), whereas little is known about the

factors regulating the degradation and utilization of yolk precursors as well as their actions in follicular cells. Our previously published studies have suggested that the VLDL receptor (VLDLR) plays a critical role in the transport of yolk precursorsinto goose ovarian follicles and that the VLDLR-mediated endocytosis of lipids as well as endogenously synthesized triglycerides are important for maintaining normal follicle development (HU *et al.* 2014b; HU *et al.* 2020). Regarding the reported roles of lysosomal proteinases in lipid degradation (THELEN & ZONCU 2017), it is proposed that they may be responsible for the utilization of yolk precursors in birds. Indeed, cathepsin D (*CTSD*), a lysosomal aspartyl protease, has been demonstrated to be involved in the degradation and absorption of yolk precursors, thereby providing essential energy and nutrients for oocyte and embryonic development in fish and chickens (RETZEK *et al.* 1992; KWON *et al.* 2002; TINGAUD-SEQUEIRA & CERDÀ 2007; PALOMINO *et al.* 2017). In fish, *CTSD* mediates the processing of VTG and VLDL, and its derivative yolk proteins and lipids are subsequently utilized for embryonic development (CARNEVALI *et al.* 1999a). Furthermore, PALOMINO *et al.* (2017) and CARNEVALI *et al.* (1999b) have indicated that both the expression level and enzymatic activity of *CTSD* are key determinants of egg quality in several fish species and could be considered as the molecular markers for egg quality. In addition, *CTSD* can influence cellular activities, including proliferation, autophagy, and apoptosis, in a range of cell types. It was reported that in human neuroblastoma cells *CTSD* exerted an anti-apoptotic effect and attenuated doxorubicin-induced apoptosis (VITALIA *et al.* 2008) while in mouse embryonic fibroblasts cytoplasmic *CTSD* levels regulated the balance between cell survival and cell death (KIM *et al.* 2013). In the tadpoles of a tropical toad, *CTSD* was involved in programmed cell death in the epidermis, muscles, spinal cord, and blood cells during tail resorption (MAHAPATRA & MAHAPATRA 2012), and in lepidopteran insects, the differential expression and autophagy-mediated maturation of *CTSD* determined its roles in cell proliferation and apoptosis, eventually determining the cell fates of tissues during metamorphosis (DI *et al*. 2020). Thus, we hypothesized that *CTSD* may influence avian ovarian follicle development not only by mediating the degradation and absorption of yolk precursors but also by regulating follicular cell activities.

The main objectives of the present study were to first clone and characterize the full-length coding sequence (CDS) of the *CTSD* gene in geese and to subsequently examine its mRNA expression in a wide range of tissues as well as in the granulosa and theca layers of goose follicles at different developmental stages. These data are expected to provide a basis for further clarification of the exact roles of *CTSD* in avian ovarian follicular cells.

Materials and Methods

Ethics statement

All experimental procedures involving the manipulation of birds were conducted in concordance with the "Guidelines for Experimental Animals" of the Ministry of Science and Technology (Beijing, China). This study has been reviewed and approved by the Sichuan Agricultural University Animal Ethical and Welfare Committee (Approval No.: 20180034).

Experimental birds and tissue collection

Females from the maternal line of Tianfu meat geese (*Anser cygnoide*), 35-45 weeks of age and laying in regular sequences of at least 2-3 eggs, were used in all studies described. The geese were kept under natural conditions of light and temperature and were provided with free access to feed and water at the Waterfowl Breeding Experimental Farm of Sichuan Agricultural University (Ya'an campus, Sichuan, China). Among them, three healthy geese having similar body weights were sacrificed by exsanguination, and thereafter, a total of thirteen types of tissue, including heart, liver, spleen, kidney, intestine, muscle, adipose, brain, hypothalamus, pituitary, adrenal gland, ovarian stroma, and oviduct tissues, were collected from each individual. The follicles within each ovary were classified as previously described (HU *et al.* 2014a; HU *et al.* 2020) into the prehierarchical 2-4-, 4-6-, 6-8-, and 8-10 mm in diameter follicles as well as the preovulatory F5-F1 (F5<F4<F3<F2<F1 in diameter) follicles. Furthermore, the granulosa layers were separated from the theca layers for all follicular categories according to the method introduced by HU *et al.* (2014a). All tissue samples were rapidly frozen in liquid nitrogen and finally stored at -80^{o}C until RNA isolation.

RNA isolation and cDNA synthesis

Total RNA was isolated from each sample using Trizol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. Its purity and concentration was determined using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and its integrity was assessed by visualization of the 28S/18S rRNA ratio after electrophoresis on 1.5% agarose gels. The cDNA was then synthesized from 1 μg RNA using the PrimeScript™ RT reagent Kit with gDNA Eraser (Takara Biotechnology Co., Ltd., Dalian, China) according to the manufacturer's instructions.

Primer design

According to the reported cDNA sequences of the *CTSD* gene in chickens, swine, humans, and mice

(GenBank accession number: NM_205177.1, NM_001037721.1, NM_001909.4, and NM_009983.2, respectively), three pairs of primers (namely D1-D3) were designed to amplify three overlapping fragments spanning the full-length CDS of the goose *CTSD* gene. Then, based on the obtained intact CDS, the primer named D4 was designed to examine the relative mRNA expression of the goose *CTSD* gene in multiple tissues. Meanwhile, the primers named E1, based on the goose *ß-ACTIN* gene (GenBank accession number: M26111), and E2, based on the goose *GAPDH* gene (GenBank accession number: MG674174), were also designed and used as internal controls. All of the above primers were designed using Primer Premier 5 software (Primer Biosoft International, Palo Alto, CA, USA)and are listed in Table 1.

RT-PCR

Using the cDNA transcribed from the goose ovary as the template, three overlapping fragments of goose *CTSD* cDNA were amplified using corresponding primer pairs by reverse transcription (RT)-polymerase chain reaction (PCR). The PCR reactions were performed in a Bio-Rad thermal cycler (Bio-Rad, USA) with the following program: pre-denaturation at 95°C for 5 min, followed by 35 cycles of denaturation at 95°C for 30 s, annealing at the primer-specific annealing temperature for 30 s, and extension at 72°C for 1 min, and a final extension at 72°C for 10 min. The PCR products were evaluated by DNA electrophoresis on 1.5% agarose gels and purified using a gel extraction kit (Watson Biomedical Inc., Shanghai, China). Target cDNAs were then cloned into the pMD19-T vector (Invitrogen, USA) and sent to Invitrogen Corporation (Applied Invitrogen, Shanghai, China) for sequencing.

Real-time PCR

The mRNA expression levels of the goose *CTSD* gene in various tissues as well as in the theca and granulosa layers of follicles at different stages of development were measured using real-time PCR. The PCR reactions were performed on the CFX96™ Real-Time PCR Detection System (Bio-Rad, USA) using the SYBR Premix Ex Taq™ II (Takara Biotechnology Co., Ltd., Dalian, China). Reactions were conducted under the following conditions: pre-denaturation at 95°C for 5 min, followed by 40 cycles of denaturation at 95°C for 15 s and annealing/extension at the corresponding temperature of each primer set for 30 s. The no-template controls and negative controls without reverse transcriptase were also included in all qPCR runs. Target specificity for each primer set was validated by melting curve analyses, and the identity of all amplicons was verified by sequencing. All samples were amplified in triplicate and the relative mRNA expression levels of the goose *CTSD* gene were normalized to the reference genes *ß-ACTIN* and *GAPDH* using the comparative Cq method (ΔΔCq) (SCHMITTGEN & LIVAK 2008). The primers used for real-time PCR are listed in Table 1.

Table 1

Primer pairs used in this study

Primer ID	Primer name	Primer sequence (5' to 3')	Amplicon (bp)	Purpose
D1	*CTSD*-1F	CTGCGCTTCTGCTTTAGG	485	RT-PCR
	CTSD-1R	GGCTCCCAGTCCCATAGT		
D2	*CTSD*-2F	CATGGATGCCCAGTATTATG	834	RT-PCR
	CTSD-2R	TCTCCTTGTGCAGAAACCTT		
D3	*CTSD*-3F	CCAATGGGCTGACTCTTT	454	RT-PCR
	CTSD-3R	GTGTATGCAAGTGTTTGTGT		
D4	*CTSD*-4F	ACTACAGCGGCGACTTCAGC	91	Real-time PCR
	CTSD-4R	GTCAGCCCATTGGCAACATC		
E1	*ß-ACTIN*-F	CAACGAGCGGTTCAGGTGT	92	Real-time PCR
	ß-ACTIN-R	TGGAGTTGAAGGTGGTCTCGT		
E2	*GAPDH*-F	GCTGATGCTCCCATGTTCGTGAT	86	Real-time PCR
	GAPDH-R	GTGGTGCAAGAGGCATTGCTGAC		

Note: F, forward primer; R, reverse primer.

Bioinformatics analysis

All amplified fragments of goose *CTSD* cDNA were assembled using the DNAMAN software. The open reading frame (ORF) Finder tool on the NCBI website was used to predict the ORF of the goose *CTSD* gene. Homology analyses of both the nucleotide and AA sequences were performed using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) (SIEVERS and HIGGINS 2014). Multiple alignments of the *CTSD* AA sequences among different species were performed using the DNAMAN software. A phylogenetic tree was constructed using the MEGA 7.0 software. The conserved functional domains and the aspartyl catalytic residues within the AA sequence of the goose *CTSD* gene were predicted using the CD-search tool on the NCBI website. The putative N-glycosylation site and signal peptide were predicted using the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/) and SignalP-5.0 Server (http://www.cbs.dtu.dk/services/SignalP/), respectively.

Statistical analysis

Data from the real-time PCR experiments were presented as the mean ±SD of three individuals and were subjected to an analysis of variance (ANOVA) followed by Duncan's Multiple Range Test using SAS 9.4 software (SAS Institute, Cary, USA). P-values below 0.05 were considered statistically significant.

Results

Molecular characterization and sequence analysis of goose *CTSD* cDNA

Three overlapping fragments were cloned using corresponding primer pairs and assembled to successfully obtain the full-length CDS of the goose *CTSD* gene (GenBank accession number: MT812700). As shown in Figure 1, the intact CDS of the goose *CTSD* gene had a length of 1197 bp, which encoded a polypeptide of 398 AA. Results from multiple alignments suggested that there were two active aspartyl residues (located at AA 96 and 283) and two potential N-glycosylation sites (located at AA 133 and 251) present in the AA sequence of the goose *CTSD* gene, and more importantly, these residues were largely conserved between geese and other species (Fig. 2). Figure 2 also showed that the AA sequence of the goose *CTSD* gene mainly comprised a signal peptide of 20 AA and two conserved functional domains, including the A1_Propeptide of 37 AA and Cathepsin_D2 of 325 AA. As shown in Table 2, a homology analysis revealed that both the nucleotide and AA sequences of the goose *CTSD* gene shared much higher similarities with corresponding homologs from the other avian species compared to those of mammals and fish, particularly with more than 90% similarities at the nucleotide level and more than 96% similarities at the AA level with those of ducks, chickens, and turkeys. These results strongly indicated that the cDNA

```
1       ATGGCGCCCCGCGGCCTTCTCGTCCTGCTGCTCCTCGCCCTGGTGGGGCCCTGCGCGGCACTCATCAGGATCCCCCTCACCAAATTCACCTCCACGCGCCGCATGCTGACCGAGGTGGGC
1        M  A  P  R  G  L  L  V  L  L  L  L  A  L  V  G  P  C  A  A  L  I  R  I  P  L  T  K  F  T  S  T  R  R  M  L  T  E  V  G

121     AGCGAGATCCCTGACATGAACGCCATCACCCAGTTCCTCAAGTTCAAGCTGGGTTTTGCTGACCTGGCTGAGCCCACCCCGGAAATCCTCAAGAATTACATGGATGCCCAGTATTATGGT
41       S  E  I  P  D  M  N  A  I  T  Q  F  L  K  F  K  L  G  F  A  D  L  A  E  P  T  P  E  I  L  K  N  Y  M  D  A  Q  Y  Y  G

241     GAGATTGGCATTGGGACCCCCCCACAGAAGTTCACTGTGGTCTTTGACACGGGCTCCTCCAACCTCTGGGTGCCGTCAGTGCACTGTCACCTGCTAGACATCGCCTGTTTGCTACACCAC
81       E  I  G  I  G  T  P  P  Q  K  F  T  V  V  F  D  T  G  S  S  N  L  W  V  P  S  V  H  C  H  L  L  D  I  A  C  L  L  H  H

361     AAGTATGATGCATCCAAATCTAGCACCTATGTGGAGAATGGCACTGAGTTTGCCATCCACTATGGGACAGGGAGCCTCTCAGGATACCTGAGCCAGGACACGGTCACGCTCGGGAACTTG
121      K  Y  D  A  S  K  S  S  T  Y  V  E  N  G  T  E  F  A  I  H  Y  G  T  G  S  L  S  G  Y  L  S  Q  D  T  V  T  L  G  N  L

481     AAAATCAAGAACCAGATCTTCGGGGAGGCCGTGAAGCAGCCGGGCATCACCTTCATTGCTGCCAAGTTCGATGGCATCCTTGGCATGGCATTCCCCAGGATCTCTGTGGACAAGGTCACT
161      K  I  K  N  Q  I  F  G  E  A  V  K  Q  P  G  I  T  F  I  A  A  K  F  D  G  I  L  G  M  A  F  P  R  I  S  V  D  K  V  T

601     CCTTTTTTTTGATAACGTCATGCAGCAGAAGCTGATTGAGAAAAACATCTTCTCCTTCTACCTGAACAGAGACCCCACAGCTCAGCCAGGCGGCGAGCTGCTGCTTGGGGGCACTGACCCC
201      P  F  F  D  N  V  M  Q  Q  K  L  I  E  K  N  I  F  S  F  Y  L  N  R  D  P  T  A  Q  P  G  G  E  L  L  L  G  G  T  D  P

721     AAATACTACAGCGGCGACTTCAGCTGGGTGAACGTCACACGCAAAGCCTACTGGCAGGTCCACATGGATGCGGTGGATGTTGCCAATGGGCTGACTCTGTGTAAGGGGGGCTGTGAGGCC
241      K  Y  Y  S  G  D  F  S  W  V  N  V  T  R  K  A  Y  W  Q  V  H  M  D  A  V  D  V  A  N  G  L  T  L  C  K  G  G  C  E  A

841     ATCGTGGACACGGGCACCTCGCTCATCACCGGCCCCACCAAGGAAGTGAAGGAGCTGCAGACAGCAATTGGTGCGAAACCACTCATCAAGGGCCAGTATGTGATCCCCTGTGAGAAGGTG
281      I  V  D  T  G  T  S  L  I  T  G  P  T  K  E  V  K  E  L  Q  T  A  I  G  A  K  P  L  I  K  G  Q  Y  V  I  P  C  E  K  V

961     TCATCTCTGCCTGTTGTCACGCTCACGTTAGGAGGGAAGCCCTACCAGCTCACCGGAGAGCAGTATGTCTTCAAGGTTTCTGCACAAGGAGAGACCATCTGCCTGAGTGGGTTTTCAGGC
321      S  S  L  P  V  V  T  L  T  L  G  G  K  P  Y  Q  L  T  G  E  Q  Y  V  F  K  V  S  A  Q  G  E  T  I  C  L  S  G  F  S  G

1081    CTGGACGTCCCACCTCCTGGTGGCCCGCTCTGGATCCTGGGGGATGTCTTCATTGGTCCCTACTACACTGTCTTTGACCGTGATAATGACTCTGTTGGCTTCGCCAAATGTGTTTAA
361      L  D  V  P  P  P  G  G  P  L  W  I  L  G  D  V  F  I  G  P  Y  Y  T  V  F  D  R  D  N  D  S  V  G  F  A  K  C  V  *
```

Fig. 1. The intact coding sequence and deduced amino acid sequence of the goose *CTSD* gene. Nucleotides and amino acids were numbered on the left side. The TAA stop codon was marked with an asterisk.

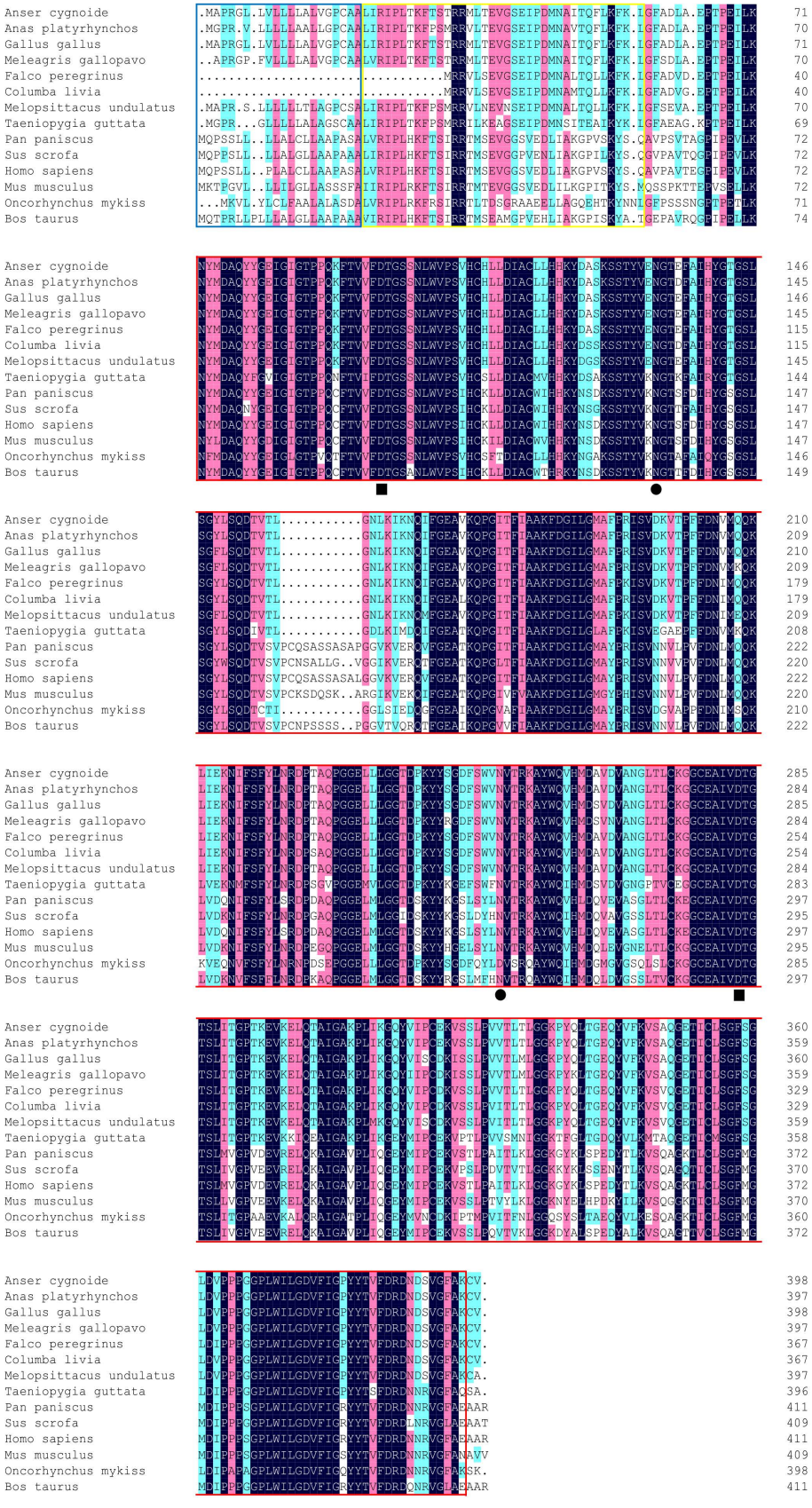

Fig. 2. Multiple alignments of the deduced amino acid sequences of the *CTSD* genes from geese and other species. One putative signal peptide and two conserved functional domains (*i.e.,* A1_Propeptide and Cathpsin_D2) were boxed in blue, yellow, and red, respectively. ■ – indicates active aspartyl residues located at AA 96 and 283. ● – indicates the potential N-glycosylation residues located at AA 133 and 251.The GenBank accession numbers of the *CTSD* genes from other species were shown in Table 2.

Table 2

Homology analysis of goose *CTSD* gene CDS with other species

Species (Latin name)	Nucleotide (%)	GenBank Accession No.	Amino acid (%)	GenBank Accession No.
Duck (*Anas platyrhynchos*)	92.78	XM_027457884.1	97.48	XP_027313685.1
Chicken (*Gallus gallus*)	92.16	NM_205177.1	98.49	NP_990508.1
Turkey (*Meleagris gallopavo*)	90.28	XM_010711093.2	96.47	XP_005149264.1
Tiercel (*Falco peregrinus*)	88.59	XM_005237329.3	96.73	XP_005237386.1
Pigeon (*Columba livia*)	87.43	XM_005499231.2	95.91	XP_005499288.1
Parrot (*Melopsittacus undulatus*)	86.41	XM_005149207.3	93.20	XP_005149264.1
Zebra finch (*Taeniopygia guttata*)	79.21	XM_002199057.5	79.04	XP_002199093.1
Human (*Homo sapiens*)	69.04	NM_001909.5	70.35	NP_001900.1
Bonobo (*Pan paniscus*)	69.36	XM_008969110.2	70.35	XP_008967358.1
Cattle (*Bos taurus*)	67.68	XM_005227296.3	67.59	XP_005227353.1
Swine (*Sus scrofa*)	69.84	NM_001037721.1	70.60	NP_001032810.1
Mouse (*Mus musculus*)	68.23	NM_009983.3	68.84	NP_034113.1
Fish (*Oncorhynchus mykiss*)	66.29	NM_001124711.1	67.68	NP_001118183

isolated from the goose ovary in this study encodes the true goose *CTSD* gene and is largely conserved with the homologs of other avian species.

Phylogenetic analysis of *CTSD*s from geese and other species

Based on the AA sequences of *CTSDs* from geese and another 13 species, a phylogenetic tree was constructed (Fig. 3). The overall topology of this tree suggested that there were obviously three clades among all examined species, including birds, fish, and mammals. In particular, the goose *CTSD* gene was clustered into the clade of birds and was evolutionarily closest to that of duck *CTSD*, which was in good agreement with the above results from the homology analysis. Noticeably, in the clade of birds, the zebra finch *CTSD* gene had a relatively longer evolutionary distance from the others.

Distribution of goose *CTSD* mRNA in various tissues

The levels of *CTSD* mRNA were examined in thirteen different tissues of laying geese using real-time PCR. As shown in Figure 4, *CTSD* mRNA was ubiquitously present in all of the examined tissues but displayed different expression levels. Specifically, the highest levels of *CTSD* mRNA were detected in the adrenal gland ($p<0.05$), followed by the liver and heart ($p<0.05$). Moderate levels of *CTSD* mRNA were detected in the reproduction-related organs (including ovarian stroma, oviduct, hypothalamus, and pituitary), spleen, and kidneys, while the lowest levels were detected in the intestine, brain, muscle, and adipose tissue.

Expression profiling of goose *CTSD* mRNA in the granulosa and theca layers during follicle development

Changes in levels of goose *CTSD* mRNA in the granulosa and theca layers during follicle development were examined using real-time PCR. As shown in Figure 5, goose *CTSD* mRNA was present in both the granulosa and theca layers of all follicular categories but exhibited different levels between the granulosa and theca layers of the same sized follicles and different expression profiles in either the granulosa or theca layers throughout follicle development. Specifically, *CTSD* mRNA was much more abundant in the granulosa layer than in the theca layer at all stages of follicle development. Furthermore, the expression of *CTSD* mRNA was maintained at relatively low and stable levels in the theca layers throughout follicle development ($p>0.05$). In contrast, as the follicles matured, the mRNA levels of *CTSD* in the granulosa layers increased gradually from the 2-4 mm to the F5 follicles and reachedmaximum in the F5 follicles ($p>0.05$), followed by a decline in the F4 follicles and were finally maintained at comparably stable levels in the F4-F1 follicles ($p>0.05$).

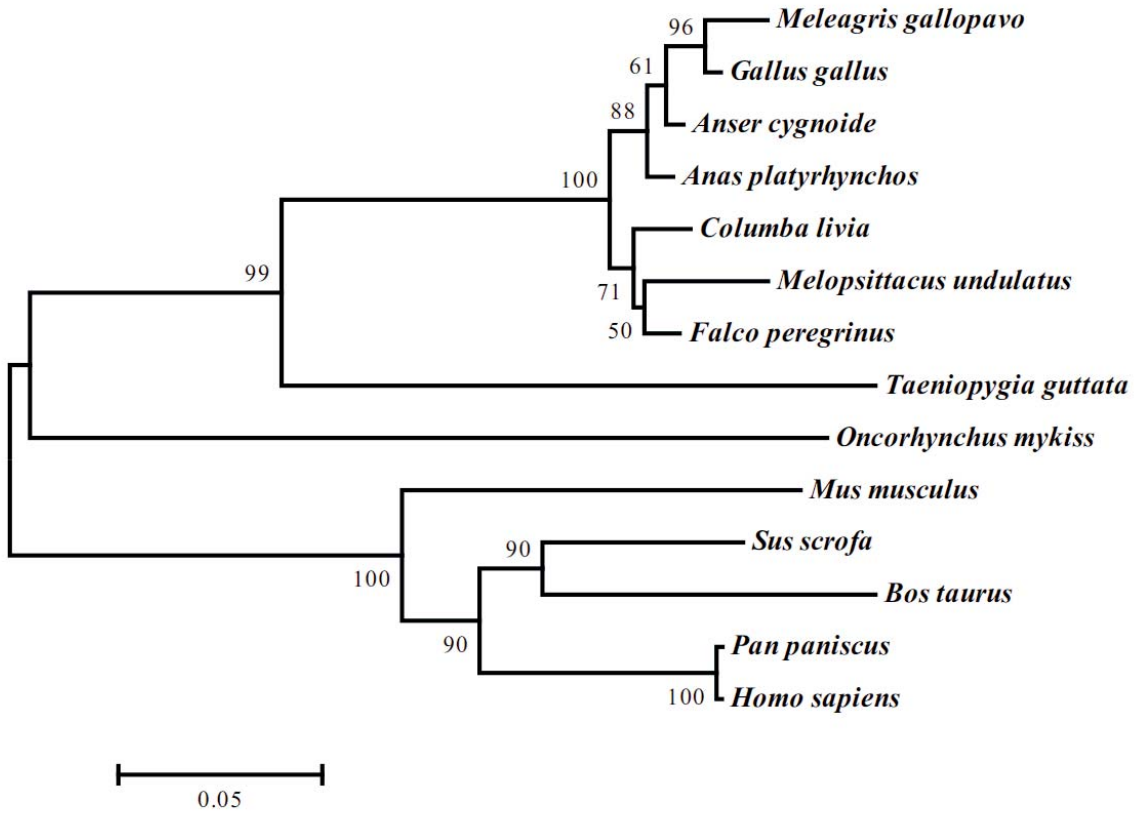

Fig. 3. A phylogenetic tree of *CTSD*s from geese and other species constructed with the neighbor-joining method using MEGA 7.0. A 1000 bootstrap replication was used to test the reliability of each branch, and bootstrap values were indicated as numbers at the branch nodes. The GenBank accession numbers of the *CTSD* genes from other species were shown in Table 2.

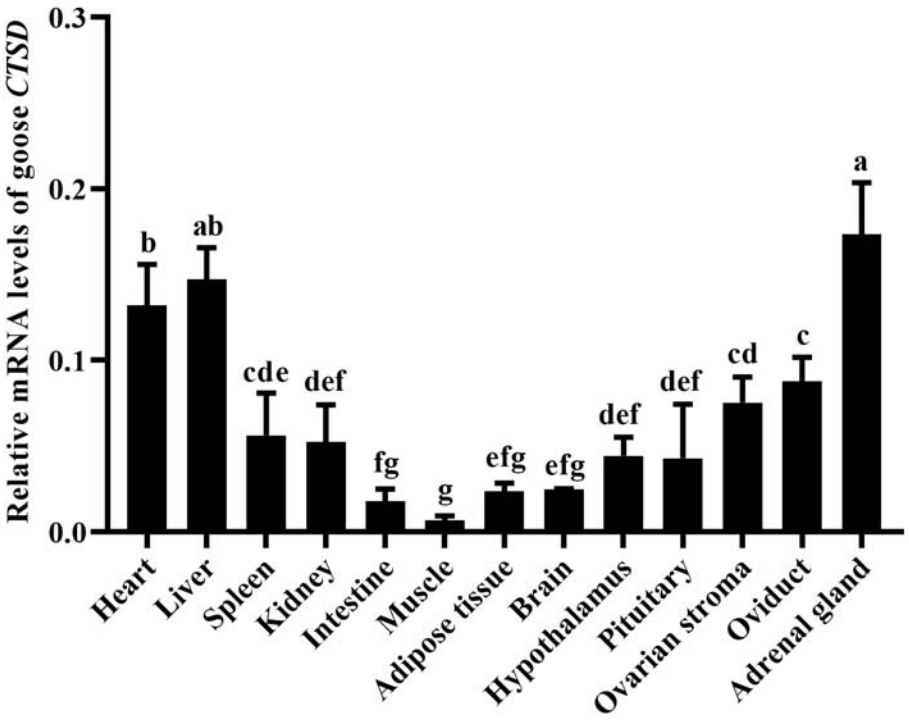

Fig. 4. Relative mRNA expression levels of *CTSD* in various tissues of laying geese. Bars with different lowercase letters indicated significant differences at $p<0.05$ among different tissues.

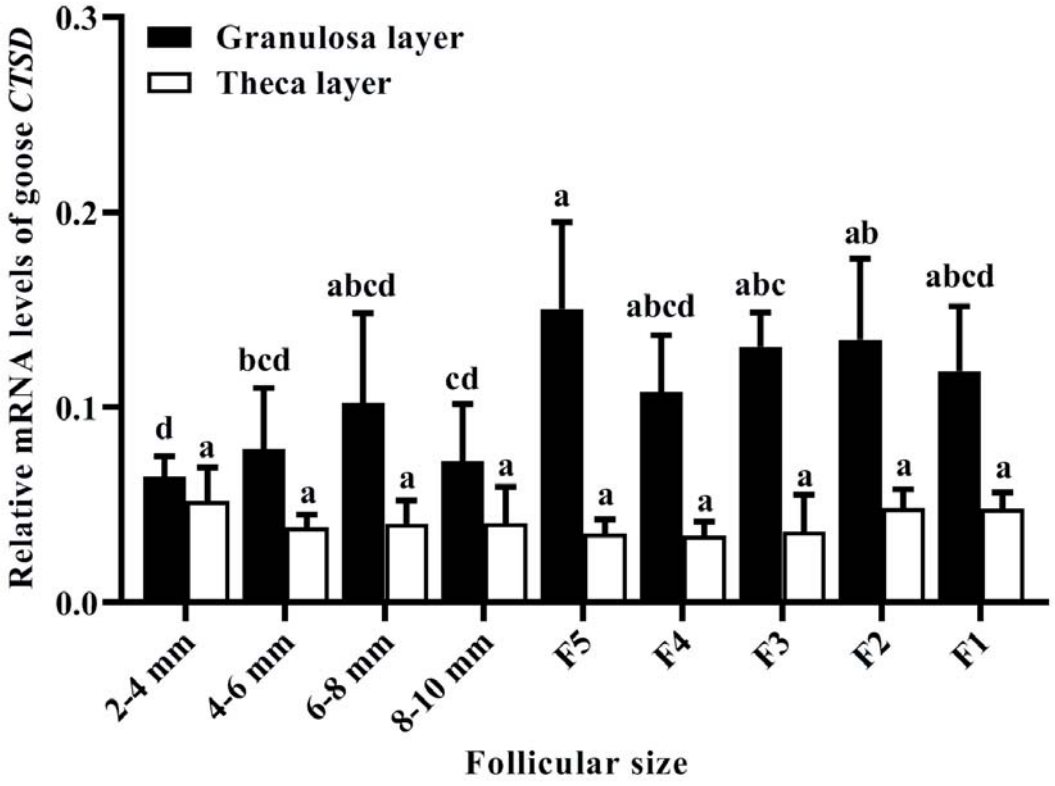

Fig. 5. Relative mRNA expression levels of *CTSD* in the granulosa and theca layers isolated from different sized follicles of laying geese. Bars with different lowercase letters indicate significant differences at $p<0.05$ among either the granulosa or theca layers of different sized follicles.

Discussion

In the present study, the intact CDS of the goose *CTSD* gene with a length of 1197bp encoding 398 AA was obtained for the first time. Similar to other reported species such as chickens, swine, and rats (BIRCH & LOH 1990; MEI *et al.* 2008; SHENG *et al.* 2013), the deduced AA sequence of the goose *CTSD* gene mainly comprised a signal peptide of 20 AA and two conserved functional domains including the A1_Propeptide of 37 AA and Cathepsin_D2 of 325 AA, which are known as the common features of the lysosomal aspartic proteinases (DUNN 1991), suggesting that goose *CTSD* functions as a secreted aspartic protease and could have similar physiological actions to the homologs of other vertebrate species. Meanwhile, two potential N-glycosylation sites were identified in the AA sequence of goose *CTSD*, which further supported that *CTSD* is a secreted protease because proteins that enter the secretory pathways are generally modified by N-glycans (ANELLI & SITIA 2008). Like other aspartic proteinases, *CTSD* is a bilobed molecule, and each of the two evolutionary related lobes contributes to one active-site aspartic acid residue (DUNN 1991). It was predicted that the Cathepsin_D2 domain of goose *CTSD* contained two strongly conserved aspartic acid residues, whose locations were similar to their counterparts in other vertebrate species, suggesting that these residues could be essential for substrate binding and catalysis. Besides, results from homology and phylogenetic analyses suggested that goose *CTSD* shared relatively higher AA sequence similarities with the homologs from other avian species, especially more than 96% similarities with those of ducks, chickens, and turkeys and was evolutionarily clustered into the avian clade. These results strongly demonstrated that the cDNA isolated from the goose ovary in this study encodes the true goose *CTSD* gene and suggested functional conservation of avian *CTSDs*.

To preliminarily explore the physiological actions of the goose *CTSD* gene, its mRNA expression levels in various tissues were determined.It was observed that goose *CTSD* mRNA was ubiquitously expressed in all of the examined tissues, indicating that *CTSD* is a versatile protease that regulates the development and functions of almost all tissues and organs. Specifically, goose *CTSD* mRNA was much more abundant in the adrenal gland, liver, and heart, which is indicative of a potential role for *CTSD* in maintaining

metabolic processes such as lipid homeostasis (KHURANA *et al.* 2019). Notably, moderate levels of goose *CTSD* mRNA were detected in the reproductive organs, including the ovarian stroma, oviduct, hypothalamus, and pituitary, implying its important roles in regulation of female reproductive activities because the hypothalamic-pituitary-gonadal/adrenal (HPG/A) axes are known as the central regulators of the development and function of the reproductive system (WINGFIELD & SAPOLSKY 2003). In view of this, expression of goose *CTSD* mRNA in the granulosa and theca layers of ovarian follicles at different developmental stages was further determined. The ubiquity of goose *CTSD* mRNA in both the granulosa and theca layers of all follicular categories potentiated its roles in regulating follicular cell functions. Considering that in oviparous species vitellogenesis is essential for oocyte and embryonic development by providing them with energy and essential nutrients such as amino acids, lipids, carbohydrates, vitamins, phosphorus, and sulfur (ROMANO *et al.* 2004) and that *CTSD* is recognized as a lysosomal aspartic protease crucial for protein degradation and lipid metabolism (HAIDAR *et al.* 2006), *CTSD* produced by the granulosa and theca cells was proposed to have a role in the degradation and absorption of liver-synthesized yolk precursors during avian follicle development. In support of this, it was evidenced that *CTSD* is a key enzyme responsible for yolk formation in chicken ovaries (RETZEK *et al.* 1992). Furthermore, levels of goose *CTSD* mRNA were much higher in the granulosa layers than in the theca layers throughout follicle development, and moreover, its levels remained almost constant in the theca layers but changed significantly in the granulosa layers as the follicles matured, potentially suggesting a predominant role for *CTSD* in regulating granulosa cell activities. It is widely accepted that granulosa cell apoptosis is the principal cellular mechanism of avian ovarian follicular atresia and the susceptibility of granulosa cells to apoptosis depends on the stages of follicle development (JOHNSON *et al.* 1996; JOHNSON 2003). Moreover, *CTSD* has been demonstrated to be involved in regulating multiple cellular activities including apoptosis (MINAROWSKA *et al.* 2007). Thus, it was postulated that the varying mRNA levels of the goose *CTSD* gene in the granulosa layers could be closely related to the fate of granulosa cells and thereby the process of follicle development. Of note, the mRNA levels of the goose *CTSD* gene in the granulosa layers reached a maximum in the F5 follicles and were significantly higher than in the 8-10 mm follicles, which suggested that *CTSD* may have a role in follicle selection by regulating granulosa cell activities, a process through which a single follicle among the 8-10 mm cohort is selected to enter the preovulatory hierarchy while the non-selected follicle will cease growth and eventually undergo atresia (JOHNSON & WOODS 2009; JOHNSON 2015; HU *et al.* 2020).

In conclusion, we have successfully cloned the full-length CDS of the goose *CTSD* gene (GenBank accession No.: MT812700) and described its structural characteristics, primarily consisting of a signal peptide and two conserved functional domains including the A1_Propeptide and Cathepsin_D2. The high sequence similarities of the *CTSD* genes in avian species suggested functional conservation of avian *CTSDs*, and the varying levels of *CTSD* in the granulosa layers could be closely related to the process of follicle development possibly by regulating both the degradation and absorption of yolk precursors and granulosa cell apoptosis. However, the exact roles of the *CTSD* gene in avian follicular cells as well as its underlying mechanisms remain to be further elucidated.

Author Contributions

Research concept and design: S.H.; Collection and/or assembly of data: J.Z., Y.L., Y.R., E.Q, L.L.; Data analysis and interpretation: J.Z., S.H., Y.L., E.Q.; Writing the article: J.Z., S.H.; Critical revision of the article: S.H., J.W.; Final approval of article: J.Z., S.H., Y.L., Y.R., E.Q., L.L., J.W.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (31802064) and the China Agricultural Research System (CARS-42-4).

Declaration of interest

The authors declare no conflict of interest.

References

ANELLI T., SITIA R. 2008. Protein quality control in the early secretory pathway. EMBO. J. **27**: 315-327. https://doi.org/10.1038/sj.emboj.7601974

BIRCH N.P., LOH Y.P. 1990. Cloning, sequence and expression of rat cathepsin D. Nucleic. Acids. Res. **18**: 6445-6446. https://doi.org/10.1093/nar/18.21.6445

CARNEVALI O., CARLETTA R., CAMBI A., VITA A., BROMAGE N. 1999a. Yolk formation and degradation during oocyte maturation in seabream sparus aurata: involvement of two lysosomal proteinases. Biol. Reprod. **60**: 140-146. https://doi.org/10.1095/biolreprod60.1.140

CARNEVALI O., CENTONZE F., BROOKS S., MAROTA I., SUMPTER J. P. 1999b. Molecular cloning and expression of ovarian cathepsin d in seabream, *Sparus aurata*. Biol. Reprod. **61**: 785-791. https://doi.org/10.1095/biolreprod61.3.785

DI Y.Q., HAN X.L., KANG X.L., WANG D., CHEN C.H., WANG J.X., ZHAO X.F. 2020. Autophagy triggers *CTSD* (cathepsin D) maturation and localization inside cells to promote apoptosis. Autophagy. 1-23. https://doi.org/10.1080/15548627.2020.1752497

DUNN B.M. 1991. Structure and Function of the Aspartic Proteinases: Genetics, Structures, and Mechanisms (1st edition). Plenum Press, New York. Pp.143-147.

HAIDAR B., KISS R.S., SAROV-BLAT L., BRUNET R., HARDER C., MCPHERSON R., MARCEL Y.L. 2006. Cathepsin D, a lysosomal protease, regulates ABCA1-mediated lipid efflux. J. Biol. Chem. **281**: 39971-39981. https://doi.org/10.1074/jbc.M605095200

HERNANDEZ A., BAHR J. 2003. Role of FSH and epidermal growth factor (EGF) in the initiation of steroidogenesis in granulosa cells associated with follicular selection in chicken ovaries. Reproduction **125**: 683-691.

HU S., GAN C., WEN R., XIAO Q., GOU H., LIU H., ZHANG Y., LI L., WANG J. 2014a. Role of leptin in the regulation of sterol/steroid biosynthesis in goose granulosa cells. Theriogenology. **82**: 677-685. https://doi.org/10.1016/j.theriogenology.2014.05.025

HU S., GAO S., ZHU J., GAN X., CHEN X., HE H., LIANG L., HU B., HU J., LIU H., HAN C., KANG B., XIA L., WANG J. 2020. Differential actions of diacylglycerol acyltransferase (DGAT) 1 and 2 in regulating lipid metabolism and progesterone secretion of goose granulosa cells. J. Steroid. Biochem. **202**: 105721. https://doi.org/10.1016/j.jsbmb.2020.105721

HU S., LIU H., PAN Z., XIA L., DONG X., LI L., XU F., HE H., WANG J. 2014b. Molecular cloning, expression profile and transcriptional modulation of two splice variants of very low density lipoprotein receptor during ovarian follicle development in geese (*Anser cygnoide*). Anim. Reprod. Sci. **149**: 281-296. https://doi.org/10.1016/j.anireprosci.2014.06.024

HU S., ZADWORNY D. 2017. Effects of nonglycosylated and glycosylated prolactin on basal and gonadotropin-stimulated steroidogenesis in chicken ovarian follicles. Domest. Anim. Endocrinol. **61**: 27-38. https://doi.org/10.1016/j.domaniend.2017.05.002

JOHNSON A. 2003. Intracellular mechanisms regulating cell survival in ovarian follicles. Anim. Reprod. Sci. **78**: 185-201. https://doi.org/10.1016/S0378-4320(03)00090-3

JOHNSON A. 2015. Ovarian follicle selection and granulosa cell differentiation. Poult. Sci. **94**: 781-785. https://doi.org/10.3382/ps/peu008

JOHNSON A., BRIDGHAM J., WITTY J., TILLY J. 1996. Susceptibility of avian ovarian granulosa cells to apoptosis is dependent upon stage of follicle development and is related to endogenous levels of bcl-xlong gene expression. Endocrinology **137**: 2059-2066. https://doi.org/10.1210/en.137.5.2059

JOHNSON A., WOODS D.C. 2009. Dynamics of avian ovarian follicle development: cellular mechanisms of granulosa cell differentiation. Gen. Comp. Endocrinol. **163**: 12-17. https://doi.org/10.1016/j.ygcen.2008.11.012

KHURANA P., YADATI T., GOYAL S., DOLAS A., HOUBEN T., OLIGSCHLAEGER Y., AGARWAL A.K., KULKARNI A., SHIRI-SVERDLOV R. 2019. Inhibiting extracellular cathepsin D reduces hepatic steatosis in Sprague–Dawley rats. Biomolecules. **9**: 171. https://doi.org/10.3390/biom9050171

KIM S.J., KIM K.H., AHN E.R., YOO B.C., KIM S.Y. 2013. Depletion of cathepsin D by transglutaminase 2 through protein cross-linking promotes cell survival. Amino. Acids. **44**: 73-80. https://doi.org/10.1007/s00726-011-1089-6

KWON J.Y., PRAT F., RANDALL C., TYLER C.R. 2002. Molecular characterization of putative yolk processing enzymes and their expression during oogenesis and embryogenesis in rainbow trout (*Oncorhynchus mykiss*). Biol. Reprod. **65**: 1701-1709. https://doi.org/10.1095/biolreprod65.6.1701

MAHAPATRA C., MAHAPATRA, P.K. 2012. Immunohistochemical localization of cathepsin D and a possible role of melanocytes during tail resorption in tadpoles of a tropical toad. J. Morphol. **273**: 791-800. https://doi.org/10.1002/jmor.20027

MEI Y., CHEN Y., LI J., GAO P., WANG C., ZHANG H., LING F., LI Y., XIE S., LI S. 2008. Sequence identification, tissue distribution and polymorphism of the porcine cathepsin d (*CTSD*) gene. Anim. Biotechnol. **19**: 144-158. https://doi.org/10.1080/10495390802072088

MINAROWSKA A., MINAROWSKI L., KARWOWSKA A., GACKO M. 2007. Regulatory role of cathepsin D in apoptosis. Folia. Histochem. Cytobiol. **45**: 159-163.

PALOMINO J., HERRERA G., TORRES-FUENTES J., DETTLEFF P., PATEL A., MARTÍNEZ V. 2017. Assessment of cathepsin mRNA expression and enzymatic activity during early embryonic development in the yellowtail kingfish *Seriola lalandi*. Anim. Reprod. Sci. **180**: 23-29. https://doi.org/10.1016/j.anireprosci.2017.02.009

RETZEK H., STEYRER E., SANDERS E.J., NIMPF J., SCHNEIDER W.J. 1992. Molecular cloning and functional characterization of chicken cathepsin D, a key enzyme for yolk formation. DNA. Cell. Biol. **11**: 661-672. https://doi.org/10.1089/dna.1992.11.661

ROMANO M., ROSANOVA P., ANTEO C., LIMATOLA E. 2004. Vertebrate yolk proteins: a review. Mol. Reprod. Dev. **69**: 109-116. https://doi.org/10.1002/mrd.20146

SCHMITTGEN T.D., LIVAK K.J. 2008. Analyzing real-time PCR data by the comparative C(T) method. Nat. Protoc. **3**: 1101-1108. https://doi.org/10.1038/nprot.2008.73

SHENG Q., CAO D., ZHOU Y., LEI Q., HAN H., LI F., LU Y., WANG C. 2013. Detection of SNPs in the cathepsin D gene and their association with yolk traits in chickens. PLoS One. **8**: e56656. https://doi.org/10.1371/journal.pone.0056656

SIEVERS F., HIGGINS D.G. 2014. Clustal omega. Curr. Protoc. Bioinformatics. **48**: 3-13. https://doi.org/10.1002/0471250953.bi0313s48

THELEN A.M., ZONCU R. 2017. Emerging roles for the lysosome in lipid metabolism. Trends. Cell. Biol. **27**: 833-850. https://doi.org/10.1016/j.tcb.2017.07.006

TINGAUD-SEQUEIRA A., CERDÀ J. 2007. Phylogenetic relationships and gene expression pattern of three different cathepsin L (CTSL) isoforms in zebrafish: CTSL A is the putative yolk processing enzyme. Gene. **386**: 0-106. https://doi.org/10.1016/j.gene.2006.08.018

VITALIA S., DANIEL M., EVGENY S., TOBIAS P., MANFRED S., FRANK W. 2008. Cathepsin D protects human neuroblastoma cells from doxorubicin-induced cell death. Carcinogenesis **29**: 1869-1877. https://doi.org/10.1093/carcin/bgn147

WINGFIELD J., SAPOLSKY R. 2003. Reproduction and resistance to stress: when and how. J. Neuroendocrinol. **15**: 711-724. https://doi.org/10.1046/j.1365-2826.2003.01033.x