

Evolutionary History of Cathepsin L (L-like) Family Genes in Vertebrates

Jin Zhou^{1,2,3§}, Yao-Yang Zhang^{4§}, Qing-Yun Li⁴, Zhong-Hua Cai^{1,2,3*}

1. The Division of Ocean Science & Technology, Graduate School at Shenzhen, Tsinghua University, Shenzhen, 518055, P. R. China
2. Shenzhen Public Platform of Screening & Application of Marine Microbial Resources, Graduate School at Shenzhen, Tsinghua University, Shenzhen, 518055, P. R. China
3. Shenzhen Key Laboratory for Coastal Ocean Dynamic and Environment, Graduate School at Shenzhen, Tsinghua University, Shenzhen, 518055, P. R. China
4. School of Life Science, Tsinghua University, Beijing, 100084, P. R. China

§: These two authors contributed equally to this work.

*Corresponding author:

Zhong-Hua Cai

E-mail address: caizh@sz.tsinghua.edu.cn

Supplementary Material

Table S1. Chromosome location of cathepsin L genes in vertebrates.

Table S2. Chromosome location and exon number of typical cathepsin L genes in vertebrates.

Table S1. Chromosome location of cathepsin L genes in vertebrates.

Sequence name	Species	Length (bp)	Chromosome location
Hsa_L1	human	333	Chromosome 9: 90,340,434-90,346,307
Hsa_L2	human	334	Chromosome 9: 99,794,940-99,801,592
Hsa_K	human	329	Chromosome 1: 150,768,684-150,780,799
Hsa_S	human	331	Chromosome 1: 150,702,672-150,738,433
Hsa_H	human	335	Chromosome 15: 79,213,400-79,237,433
Rno_L1	rat	334	Chromosome 17: 6,288,002-6,294,150
Rno_01	rat	330	Chromosome 17: 9,589,872-9,593,955
Rno_02	rat	330	Chromosome 17: 9,575,189-9,579,764
Rno_testin	rat	333	Chromosome 17: 9,532,437-9,544,923
Rno_8	rat	333	Chromosome 17: 9,182,040-9,194,211
Rno_03	rat	334	Chromosome 17: 9,274,994-9,280,494
Rno_R	rat	334	Chromosome 17: 9,235,486-9,244,244
Rno_M	rat	333	Chromosome 17: 8,931,468-8,936,876
Rno_Q	rat	343	Chromosome 17: 9,347,171-9,352,730
Rno_04	rat	331	Chromosome 17: 9,401,362-9,441,576
Rno_05	rat	343	Chromosome 17: 9,420,443-9,425,111
Rno_K	rat	329	Chromosome 2: 190,394,854-190,405,667
Rno_7	rat	331	Chromosome 17: 9,092,366-9,097,506
Rno_06	rat	331	Chromosome 2: 190,425,296-190,466,567
Rno_H	rat	333	Chromosome 8: 94,971,674-94,990,410
Mmu_L	mouse	334	Chromosome 13: 64,464,522-64,471,614
Mmu_01	mouse	331	Chromosome 13: 60,899,611-60,904,205
Mmu_02	mouse	330	Chromosome 13: 60,883,254-60,887,725
Mmu_03	mouse	333	Chromosome 13: 60,943,982-60,965,777
Mmu_M	mouse	333	Chromosome 13: 61,637,804-61,643,199
Mmu_8	mouse	333	Chromosome 13: 61,348,108-61,356,709
Mmu_6	mouse	334	Chromosome 13: 61,296,493-61,304,771
Mmu_R	mouse	334	Chromosome 13: 61,260,575-61,265,617
Mmu_J	mouse	333	Chromosome 13: 61,101,639-61,107,277
Mmu_3	mouse	332	Chromosome 13: 61,665,991-61,671,426
Mmu_Q	mouse	343	Chromosome 13: 61,136,388-61,141,992
Mmu_7	mouse	331	Chromosome 13: 61,453,822-61,459,531
Mmu_S	mouse	343	Chromosome 3: 95,330,708-95,360,322
Mmu_K	mouse	329	Chromosome 3: 95,303,208-95,313,284
Mmu_H	mouse	333	Chromosome 9: 89,949,105-89,970,927
Cfa_L1a	dog	334	Chromosome 1: 72,994,522-72,999,827
Cfa_01	dog	331	Chromosome X: 30,929,634-30,937,796
Cfa_L1b	dog	341	Chromosome 1: 75,272,527-75,275,720
Cfa_K	dog	335	Chromosome 17: 63,008,805-63,016,333

Cfa_S	dog	342	Chromosome 17: 62,973,353-62,989,357
Cfa_H	dog	334	Chromosome 3: 60,967,575-60,978,278
Sus_L1a	pig	334	Chromosome 10: 65,264,634-65,269,060
Sus_L1b	pig	334	Chromosome 10: 65,320,862-65,326,580
Sus_01	pig	332	Chromosome 10: 26,750,075-26,755,828
Sus_02	pig	332	Chromosome 10: 26,905,970-26,912,041
Sus_K	pig	330	Chromosome 4: 102,342,678-102,354,617
Sus_03	pig	342	Chromosome 4: 102,368,163-102,393,283
Sus_H	pig	335	Chromosome 7: 54,119,425-54,140,625
Bta_L1a	cow	334	Chromosome 8: 84,976,255-84,981,470
Bta_L1b	cow	333	Chromosome 8: 82,365,269-82,371,037
Bta_K	cow	334	Chromosome 3: 19,994,998-20,007,861
Bta_S	cow	331	Chromosome 3: 20,024,302-20,047,228
Bta_H	cow	335	Chromosome 21: 25,675,908-25,696,632
Oan_01	platypus	339	SuperContig Contig819: 387,673-391,951
Oan_02	platypus	338	SuperContig Contig819: 148,843-153,649
Oan_03	platypus	339	SuperContig Contig11552: 581-5,461
Oan_K	platypus	331	SuperContig Contig2231: 23,157-28,048
Oan_S	platypus	333	SuperContig Contig2231: 36,176-48,136
Oan_04	platypus	220	SuperContig Contig9554: 20,739-25,304
Mdo_01	opossum	342	Chromosome 6: 16,901,900-16,905,261
Mdo_02	opossum	335	Chromosome 6: 16,758,228-16,762,052
Mdo_03	opossum	337	Chromosome 6: 16,648,108-16,653,801
Mdo_K	opossum	337	Chromosome 2: 495,961,242-495,979,495
Mdo_S	opossum	341	Chromosome 2: 495,917,064-495,927,008
Mdo_04	opossum	344	Chromosome 1: 423,898,764-423,907,181
Mdo_05	opossum	333	Chromosome 1: 423,857,986-423,869,484
Mdo_H	opossum	333	Chromosome 1: 121,189,894-121,230,902
Aca_01	lizard	347	Chromosome 2: 40,926,589-40,932,486
Aca_S	lizard	342	Scaffold GL343862.1: 18,139-36,017
Aca_K	lizard	290	Scaffold GL343862.1: 1,635-8,701
Aca_02	lizard	212	Scaffold GL344400.1: 526-6,347
Aca_03	lizard	262	Scaffold GL343464.1: 695,102-700,742
Aca_H	lizard	285	Scaffold GL343835.1: 65,434-77,377
Xtr_L2	frog	336	Scaffold GL172758.1: 1,748,966-1,754,900
Xtr_K	frog	339	Scaffold GL172787.1: 1,815,690-1,826,066
Xtr_S	frog	333	Scaffold GL172787.1: 1,775,172-1,791,682
Xtr_01	frog	327	Scaffold GL172787.1: 1,970,633-1,975,839
Xtr_02	frog	336	Scaffold GL172718.1: 49,964-56,645
Xtr_03	frog	334	Scaffold GL172718.1: 37,778-42,450
Xtr_04	frog	323	Scaffold GL172787.1: 1,956,206-1,964,412
Xtr_05	frog	357	Scaffold GL174050.1: 56,167-64,096
Xtr_06	frog	333	Scaffold GL172718.1: 7,070-33,472
Xtr_07	frog	355	Scaffold GL172787.1: 2,008,652-2,014,793

Xtr_08	frog	295	Scaffold GL174151.1: 1,820-5,845
Xtr_09	frog	328	Scaffold GL174151.1: 11,182-17,695
Xtr_10	frog	329	Scaffold GL172718.1: 97,099-104,124
Xtr_L1	frog	334	Scaffold GL173053.1: 993,788-1,001,711
Xtr_11	frog	338	Scaffold GL180251.1: 23,068-28,614
Xtr_H	frog	331	Scaffold GL173053.1: 1,005,021-1,016,757
Ocu_L	rabbit	333	Chromosome 1: 75,697,176-75,702,184
Ocu_01	rabbit	341	Chromosome 1: 73,519,100-73,522,900
Ocu_02	rabbit	341	Chromosome 1: 73,510,454-73,513,651
Ocu_K	rabbit	329	Chromosome 13: 41,346,878-41,360,274
Ocu_S	rabbit	331	Chromosome 13: 41,370,897-41,392,664
Ocu_H	rabbit	333	Scaffold GL018746: 338,090-352,409
Cpo_L2	guinea pig	333	scaffold_42: 2,408,344-2,412,837
Cpo_K	guinea pig	329	scaffold_2: 11,816,411-11,823,540
Cpo_S	guinea pig	324	scaffold_2: 11,731,591-11,749,621
Cpo_H	guinea pig	335	scaffold_190: 394,405-408,376
Ptr_01	chimpanzee	333	Chromosome 9: 86,565,325-86,570,726
Ptr_02	chimpanzee	334	Chromosome 9: 95,803,350-95,812,986
Ptr_03	chimpanzee	329	Chromosome 1: 128,975,594-128,987,818
Ptr_04	chimpanzee	331	Chromosome 1: 128,913,629-128,947,089
Ptr_05	chimpanzee	335	Chromosome 15: 76,560,276-76,584,240
Mmul_01	macaque	333	Chromosome 15: 96,263,014-96,268,822
Mmul_02	macaque	334	Chromosome 15: 108,649,779-108,656,210
Mmul_03	macaque	332	Scaffold 1099548049343: 21,891-25,407
Mmul_K	macaque	329	Chromosome 1: 129,199,747-129,211,826
Mmul_S	macaque	331	Chromosome 1: 129,132,888-129,166,851
Mmul_04	macaque	335	Chromosome 7: 58,049,842-58,070,819
Cin_01	C.intestinalis	341	Chromosome 3: 6,084,679-6,088,543
Cin_02	C.intestinalis	335	Chromosome 12: 2,713,325-2,718,540
Cin_03	C.intestinalis	327	Chromosome 12: 2,720,429-2,723,634
Cin_04	C.intestinalis	330	Chromosome 12: 2,676,056-2,680,725
Cin_05	C.intestinalis	333	Chromosome 3: 694,722-700,309
Pma_01	Lamprey	339	Scaffold GL477089: 42,202-47,941
Pma_02	Lamprey	297	Scaffold GL499136: 454-1,347
Pma_03	Lamprey	295	Scaffold GL485125: 12,828-20,107
Pma_04	Lamprey	293	Scaffold GL482653: 10,321-33,550
Pma_05	Lamprey	347	Scaffold GL478103: 36,116-54,424
Pma_06	Lamprey	333	Scaffold GL476381: 1,843-17,023
Tru_L1	Fugu	338	scaffold_4: 2,007,591-2,010,132
Tru_01	Fugu	335	scaffold_10: 1,009,465-1,011,044
Tru_S	Fugu	335	scaffold_175: 250,039-253,308
Tru_K	Fugu	333	scaffold_175: 240,470-248,567
Tru_02	Fugu	343	scaffold_13: 1,827,107-1,829,107
Tru_03	Fugu	327	scaffold_138: 72,342-74,527

Ola_L	Medaka	338	ultracontig107: 101,405-103,709
Ola_01	Medaka	334	Chromosome 22: 5,659,178-5,661,457
Ola_K	Medaka	331	Chromosome 16: 24,479,790-24,492,011
Ola_S1	Medaka	339	Chromosome 16: 24,494,714-24,503,215
Ola_S2	Medaka	335	Chromosome 11: 6,270,370-6,272,167
Ola_02	Medaka	348	Chromosome 4: 3,384,433-3,386,091
Gac_L1	Stickleback	338	groupXIII: 19,311,958-19,315,035
Gac_01	Stickleback	332	groupXV: 7,604,544-7,607,322
Gac_K	Stickleback	329	groupXX: 16,241,460-16,251,039
Gac_S	Stickleback	339	groupXX: 16,251,960-16,257,849
Gac_02	Stickleback	343	scaffold_211: 14,095-17,474
Gac_03	Stickleback	330	groupX: 5,305,887-5,309,240
Gac_04	Stickleback	332	groupX: 5,303,181-5,305,098
Gac_H	Stickleback	331	groupII: 15,348,573-15,353,126
Dre_L1a	Zebrafish	337	Chromosome 5: 46,423,373-46,429,502
Dre_L	Zebrafish	337	Chromosome 5: 25,974,812-25,978,602
Dre_01	Zebrafish	336	Chromosome 12: 17,941,821-17,943,847
Dre_02	Zebrafish	335	Chromosome 12: 17,825,928-18,008,271
Dre_03	Zebrafish	336	Chromosome 12: 17,758,538-17,944,088
Dre_04	Zebrafish	336	Chromosome 12: 17,839,808-18,018,797
Dre_05	Zebrafish	335	Chromosome 12: 18,005,763-18,008,271
Dre_06	Zebrafish	336	Chromosome 12: 17,701,518-17,703,564
Dre_07	Zebrafish	335	Chromosome 12: 17,785,120-17,787,146
Dre_08	Zebrafish	335	Chromosome 12: 17,962,598-18,008,271
Dre_09	Zebrafish	335	Chromosome 12: 17,986,367-17,990,114
Dre_10	Zebrafish	336	Chromosome 12: 17,583,551-17,586,054
Dre_11	Zebrafish	335	Chromosome 12: 17,659,767-17,662,116
Dre_L1b	Zebrafish	336	Chromosome 12: 17,684,092-17,686,647
Dre_L1	Zebrafish	334	Chromosome 17: 27,315,894-27,320,142
Dre_K	Zebrafish	333	Chromosome 16: 31,753,228-31,762,840
Dre_Sb2	Zebrafish	330	Chromosome 16: 31,774,764-31,780,268
Dre_Sb1	Zebrafish	330	Chromosome 16: 31,769,424-31,773,725
Dre_Sa	Zebrafish	328	Chromosome 19: 320,356-327,394
Dre_12	Zebrafish	349	Chromosome 22: 16,629,168-16,638,405
Dre_H	Zebrafish	330	Chromosome 18: 26,208,613-26,215,014
Gga_01	Chicken	339	Chromosome Z: 40,911,253-40,916,030
Gga_02	Chicken	340	Chromosome 28: 4,489,297-4,491,432
Gga_03	Chicken	245	Chromosome 25: 1,787,224-1,789,851
Gga_04	Chicken	329	Chromosome 23: 2,401,585-2,403,839
Gga_05	Chicken	261	Chromosome 10: 22,098,456-22,104,465
Gga_06	Chicken	317	Chromosome 23: 2,393,838-2,398,024
Bfl_01	lancelet	327	chrUn:655788676-655792318
Bfl_02	lancelet	336	chrUn:655502356-655513866
Bfl_03	lancelet	308	chrUn:655777011-655785973

Bfl_04	lancelet	327	chrUn:655793322-655797364
Bfl_05	lancelet	307	chrUn:655811403-655814577
Bfl_06	lancelet	327	chrUn:655822049-655826562
Bfl_07	lancelet	330	chrUn:655829195-655832999
Bfl_08	lancelet	298	chrUn:198063666-198069419
Bfl_09	lancelet	295	chrUn:320778459-320781844

Table S2. Chromosome location and exon number of typical cathepsin L genes in vertebrates.

Sequence name	Species	Length of amino acids	The original sites of cdna	Positive or negative	exon number
Hsa_K	human	329	Chromosome 1: 150,768,684-150,780,799	-	8
Hsa_S	human	331	Chromosome 1: 150,702,672-150,738,433	-	8
Ptr_03	chimpanzee	329	Chromosome 1: 128,975,594-128,987,818	-	8
Ptr_04	chimpanzee	331	Chromosome 1: 128,913,629-128,947,089	-	8
Sus_K	pig	330	Chromosome 4: 102,342,678-102,354,617	+	7
Sus_03	pig	342	Chromosome 4: 102,368,163-102,393,283	+	8
Ocu_K	rabbit	329	Chromosome 13: 41,346,878-41,360,274	+	8
Ocu_S	rabbit	331	Chromosome 13: 41,370,897-41,392,664	+	7
Aca_S	lizard	342	Scaffold GL343862.1: 18,139-36,017	+	7
Aca_K	lizard	290	Scaffold GL343862.1: 1,635-8,701	+	6
Dre_K	Zebrafish	333	Chromosome 16: 31,753,228-31,762,840	+	8
Dre_Sb1	Zebrafish	330	Chromosome 16: 31,769,424-31,773,725	+	8
Ola_K	Medaka	331	Chromosome 16: 24,479,790-24,492,011	+	8
Ola_S1	Medaka	339	Chromosome 16: 24,494,714-24,503,215	+	8