

Molecular cloning and mRNA expression of a hepcidin gene from the spinyhead croaker, *Collichthys lucidus*

C. Sang¹, Y. Lin², K. Jiang³, F. Zhang³ and W. Song³

¹School of Ocean, Yantai University, Yantai, Shangdong, China
²College of Fisheries and Life Science, Shanghai Ocean University,
Shanghai, China
³East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences,
Shanghai, China

Corresponding author: W. Song E-mail: songw@ecsf.ac.cn

Genet. Mol. Res. 14 (4): 16050-16059 (2015) Received August 16, 2015 Accepted October 14, 2015 Published December 7, 2015 DOI http://dx.doi.org/10.4238/2015.December.7.18

ABSTRACT. Antimicrobial peptides are important components that participate in host innate immune activities and play crucial roles in host defense against microbial invasion. Hepcidin is an antimicrobial peptide and iron-regulatory molecule that primarily functions in the liver. In the present study, we first obtained a full-length cDNA sequence of hepcidin and its corresponding genomic DNA sequence from *Collichthys lucidus* using RT-PCR and rapid amplification of cDNA ends (RACE), and then analyzed these sequences using bioinformatics software. The results showed that *C. lucidus* hepcidin (CL-hepc) possesses two introns and three exons in the genomic DNA, with a length of 816 bp. The open reading frame was 264 bp, encoding an 87 amino acid peptide, and with high similarity (88.89%) to 83416593 *Larimichthys crocea* (ABC18307) and relatively low similarity (47.73%) to 158358729 *L. crocea* (ABY84845.1). The pre-peptide contained a signal peptide (25 amino acids), a prodomain (34 amino acids), and a mature peptide (25 amino acids). The predicted 25 amino acid hepcidin

mature peptide included 8 conserved cysteine residues. Quantitative real-time reverse transcription-PCR analysis revealed specific expression patterns of CL-hepc, with the highest expression observed in the liver, relatively low expression observed in the gill and spleen, and almost no expression detected in other tissues analyzed. In conclusion, we identified a hepcidin from *C. lucidus* that has common expression patterns with other hepcidins. However, as this hepcidin is inconsistent with two other hepcidins from *L. crocea* in terms of the phylogenetic tree, the presence of another hepcidin gene warrants further investigation.

Key words: Hepcidin; Collichthys lucidus; Gene clone; Innate immune

INTRODUCTION

Hepcidin (HEPC), a cysteine-rich antimicrobial peptide, was first discovered in human plasma ultrafiltrate (Krause et al., 2000) and urine (Park et al., 2001), and subsequently the sequences were identified from various vertebrates, including mice (Pigeon et al., 2001), dog (Fry et al., 2004), swine (Sang et al., 2006), pigeons (Fu et al., 2007), equine (Oliveira Filho et al., 2010), and zebrafish (Shike et al., 2004). In vertebrates, hepcidin is mainly involved in the innate immune system and maintenance of iron homeostasis (Park et al., 2001; Nemeth et al., 2003). Its synthesis is greatly stimulated by inflammation or iron overload. In anemia or inflammation, HEPC production is increased by approximately 100-fold, which may account for the defining feature of this condition, sequestration of iron in macrophages (Ganz, 2003). Ferroportin is an iron exporter present on the surface of absorptive enterocytes, macrophages, hepatocytes, and placental cells. HEPC binds to ferroportin to induce its degradation, thus, reducing the exportation of iron. Through this pathway, homeostasis of the iron system is regulated, i.e., iron regulates the secretion of HEPC, which in turn controls the concentration of ferroportin on the cell surface (Nemeth et al., 2004). The key role of HEPC suggests that drugs acting as HEPC agonists or antagonists may be a potential target for the treatment of iron-related diseases (Nicolas et al., 2002). In addition, studies on human HEPC have demonstrated that HEPC has a relatively wide antibacterial spectrum, including fungi, such as Candida albicans, Aspergillus fumigatus, and Aspergillus niger, and bacteria, such as Escherichia coli, Staphylococcus aureus, Staphylococcus epidermidis, and group B Streptococcus.

Due to the lack of a well-developed adaptive immune system, fish rely heavily on innate components when encountering various pathogens (Magnadottir, 2006). The sequences of HEPCs have been cloned in many fish species, including both freshwater fish, such as zebrafish (Shike et al., 2004), mud loach (Nam et al., 2011), and rice field eel (Li et al., 2011), and marine fish, including *Epinephelus coioides* (Qu et al., 2013), *Pseudosciaena crocea* (Wang et al., 2009), *Cynoglossus semilaevis* (Wang et al., 2012), rainbow trout (Alvarez et al., 2013), and *Chrysophrys major* (Chen et al., 2005). HEPC in fish has a similar structure and mRNA distribution to other species, such as mammals, although there are some differences. In zebra fish (Shike et al., 2004) and channel catfish (Bao et al., 2005), the transcripts of HEPC are mostly found in other tissues rather than in the liver. In addition, no alteration in expression level was detected in iron-deficient fish when they were submitted to either iron status modulation or bacterial infection (Rodrigues et al., 2006). HEPC (Bao et al., 2005), and synthesized peptides (Wang et al., 2009) all exhibited an effective activity against

Genetics and Molecular Research 14 (4): 16050-16059 (2015)

different species of microorganisms. Meanwhile, the antimicrobial spectrum of HEPC activity from different fish species appeared to be individually specific, suggesting that different HEPCs from different sources may have specific antimicrobial action (Wang et al., 2009). In addition, the use of HEPC antibacterial peptide from large yellow croaker as a feed additive could prevent disease and treat sea fish (Wang et al., 2011), suggesting that obtaining various HEPCs from different species may sometimes be promising for the treatment of various diseases.

Collichthys lucidus Richardson (Perciformes, Sciaenidae, Collichthys), is a commercially important near-shore species widely distributed in the South Sea and the East Sea of China in particular (Cheng et al., 2012). Based on its excellent properties and quality of meat, *C. lucidus* is highly valued by Chinese consumers. Knowing that resources of *C. lucidus* are declining as a result of increasingly serious problems, such as pollution and over-fishing, several institutions in China, including the East China Sea Fisheries Institute and Shanghai Fisheries Research Institute, have begun to study the artificial propagation of *C. lucidus*, with substantial progress being made recently. The aim of the present study was to characterize both the mRNA and genome DNA sequences of *C. lucidus* HEPC (CL-hepc) and analyze expression patterns in different tissues in an attempt to lay a preliminary foundation for promoting aquaculture of *C. lucidus* and other species.

MATERIAL AND METHODS

Determination of CL-hepc cDNA and genomic DNA sequences

Construction of the cDNA library and cloning of CL-hepc cDNA from C. lucidus

Healthy fish weighing 35 ± 5 g were captured from the East China Sea area near Shanghai, China. Tissue samples (200 ± 50 mg) for RNA extraction were obtained through vivisection on fishing boats and then stored in a 1.5 mL RNA fixer (Bioteke Corporation, Beijing, China) at -20°C. Total RNA was isolated from the hepatopancreas, gill, heart, muscle, and blood of *C. lucidus* using TRIpure reagent (Aidlab, Beijing, China) following the manufacturer protocol. Quality and concentration of RNA were checked using agarose gel electrophoresis (Universal Hood II, Bio-Rad Laboratories Inc., Hercules, CA, USA) and spectrophotometry (DU800 Nucleic acid/Protein Analyzer, Beckman Coulter, Brea, CA, USA). RNA was stored at -80°C until use.

A cDNA library was constructed successfully using a SMART[™] cDNA library construction kit (TaKaRa, Dalian, China), followed by transformation, screening, and sequencing. Random sequencing was performed after cDNA library construction. All expressed sequence tags were subjected to BLAST analysis.

A cDNA sequence comprising the CL-hepc domain was obtained and subjected to further investigation. RACE (3'-rapid amplification of cDNA ends) was carried out following the manufacturer protocol using the 3'-full RACE core set (TaKaRa) to obtain the full-length cDNA sequence. Two gene specific-primers, CL-hepc-outer (5'-ACTCGTGCTCGCCTTTATTTGC-3') as the 3'-RACE amplification outer primer and CL-hepc-inner (5'-TTCTGAGGATTCCCAGCGACGACCAA-3') as the nest PCR primer, were designed according to the obtained sequence and used for 3'-RACE. PCR was carried out following the manufacturer protocol. The PCR products were checked on a 1.0% agarose gel and purified using the Agarose Gel DNA Purification kit v2.0. (TaKaRa), and then cloned into the pMD19-T vector (TaKaRa) for sequencing.

Genetics and Molecular Research 14 (4): 16050-16059 (2015)

Genomic DNA determination of CL-hepc

Genomic DNA was isolated from the muscle of *C. lucidus* using the TIANamp Marine Animals DNA kit (TIANGEN, Beijing, China) according to the manufacturer protocol. Primers CL-hepc-g-F (5'-GAAGAAATCCTCTTGCCTGGACG-3') and CL-hepc-g-R (5'-AAGGTCATTTCCTGGGGTGAG ATAC-3') were designed based on the 5'-untranslated region (UTR) and the 3'-UTR of the cDNA sequences.

Bioinformatic analysis

General features of the CL-hepc full-length cDNA were detected by Vector NTI Advance 11.5 (http://www.lifetechnologies.com/cn/zh/home/life-science/cloning/vector-nti-software/vector-nti-advance-software/whats-new-in-vector-nti-advance.html), and the peptide sequence was deduced and verified according to other HEPCs using the NCBI Open Reading Frame (ORF) Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The biochemical features of the deduced protein were calculated using the Expert Protein Analysis System (http://web.expasy.org/). Amino acid sequences from various species were obtained from NCBI GenBank and analyzed using the MEGA software version 5.0 (Saitou and Nei, 1987). Phylogenetic analysis was performed using the neighbor-joining method using the MEGA software version 5.0 and the confidence level in the generated phylogenetic tree was obtained using 1000 bootstraps. The secondary structure of the CL-hepc protein was predicted using the Predict Protein package (available online at http://www.predictprotein.org/).

Tissue distribution analysis of CL-hepc mRNA in adults

Tissue samples, including the liver, blood, heart, muscle, gill, spleen, brain, and kidney, were collected from fish using the same method as described above, from which total RNA (approximately 1 μ g) was extracted and reverse transcribed using the quantitative real-time PCR (qRT-PCR) kit (TOYOBO, Japan) for first-strand cDNA synthesis. The products were diluted 10 times and stored at -20°C until qRT-PCR analysis.

The qRT-PCR assay was carried out in a StepOne Plus detection system (Applied Biosystems). The transcript expression profiles of CL-hepc in eight different tissues of *C. lucidus* were detected using a pair of CL-hepc specific primers (CL-hepc-RT-F: 5'-GAAACGTGGATGCCACAT TTCA-3' and CL-hepc-RT-R: 5'-GAAGCCGCAGCCTTTGTTCTT-3'), and an 18S rRNA primer pair (18S-RT-F: 5'-GCCTGAATACCGCAGCTAGGAATAA-3' and 18S-RT-R: 5'-TTTCACCTCTAGCGG CACAATACG-3') designed from the 18S rRNA sequence on NCBI GenBank (accession No. JN211725.1) used as the internal control. CL-hepc expression levels were calculated using the standard curve method (Sun et al., 2012). Amplifications were performed on a 96-well plate with a 20-µL reaction volume containing 10 µL 2X Power SYBR Green PCR Master Mix (TaKaRa, Applied Biosystems), 1.0 µL PCR forward primer (10 µM), 1.0 µL PCR reverse primer (10 µM), 2.0 µL cDNA template, and 6 µL diethylpyrocarbonate-treated water (DEPC-treated water). The reaction process for qRT-PCR was 10 min at 95°C, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min.

Genetics and Molecular Research 14 (4): 16050-16059 (2015)

RESULTS

Determination of CL-hepc cDNA and genomic DNA sequences

The full-length cDNA of CL-hepc was 816 bp in length, containing a 264-bp ORF encoding a protein of 87 amino acids, and CL-hepc included a 5'-UTR of 201 bp and a 3'-UTR of 351 bp, including a poly(A)+ tail. A typical poly-adenylation signal (AATAAA) was located 17 bp upstream of the poly(A)+ tail. The signal peptide cleavage site of the deduced CL-hepc was predicted between Ala 24 and Val 25 (Figure 1). According to the conserved motif "RQKR", the mature peptide of CL-hepc contained 25 amino acids that started with Gln 63. Thus, the mature peptide consisted of 25-amino acid residues at the C terminus. The predicted molecular weight of CL-hepc was 2876.9 Da with an isoelectric point of 8.41. Meanwhile, the mature peptide had the typical eight-Cys structure, which is a common feature of the HEPC family.

	(Exe	on 1	1)																		
1	CAC	GCAG	GGA	AGGA	AGTI	GAC	CAAC	GGG1	ICA(CCC	AAA	GAT	CTG	AAG.	AAA	TCC	тст	TGC	CTG	GACG	
61	ATC	CAC	CAT	CCAT	CAC	CTGO	AGO	CTG	AAA	AA	ΓAA.	ATT	GAA	GTT	ATT	GTG	GTG	СТС	TTT	GGTG	
121	GCC	CAT	TGA	GAAA	AAAA	GAC	2004		AGGT	CT	AAT	CTG	CAA	AGG.	ATT	TAT	TAA	СТА	AAC	CATT	
181	TTT	TG	AAA	AAGA	AAA	GCT	CAA.	AT	GAA	GGC	ATT	CAG	CAT	TGC	AGT	TGC	AGT	GAC	ACT	CGTG	
1								М	K	А	F	S	I	А	V	A	V	Т	L	V	
								-									(1	ntr	on	1)	
241	CTO	CGC	CTT	TATT	TGC	CATI	сто	GAG	GAGO	CTC	ICC	CGT	CCC	ATT	CAO	CGG	Ggt	aas	aat	ccga	
14	L	A	F	I	С	I	L	Е	S	S	A	۸V	P	F	T	G			,		
								Pro	edi	cat	ed o	l' cle	ava	ge	sit	e f	or	sie	mal	ner	tide
301	ctt	taa	acti	cact	ttca		ect	tta	tta	rct	ata	aat	gtt	ttg	tca	993	tgo	tea	gat	ccta	, or ac
				- ac			-80	(F	(voi	1 2)	aav	0			554		- 50	5		
361	aat	ot.	aca	taat	tea	+++		C	TCC		, 	тсс	100	ACC	CAC	CCA	CCA	ATC	aca	СТ	
30		ug ti	505	c aa i		1000	iac e	v	1007	I GO/	100	1001	7 F	100	1 (1 0	3 1	UNIC VI	n í	Т	
00									4		. 1				(т	, . n+r		2)		•	
410	00/	ι c τ΄	TCC		0.01		CAL		TC		201	AAC	стс	CAT	(1 (a+	1111 0 + 0	++0	2) 22+	+ ~ ~	ot as	
415	D	101	100		U	O A	IGAN T	1100	5107	v v	50A.	плс	w 10	u GAI	σgι	aig	lic	aai	lgo	cigo	
42	F	V	A	A	п	Q	E	2	3	T.	E	1	W	M							
419	atga	18 T.	t a a :	ECC 5	att	laci	21.9	2 2 5 6	- 882	атт:	aaa	αττ	taa	gτg	gat	gtg	ттт	TCC	SC C Z	000	
500									Juur											~ 55	
539	agta	atco	ctg	gtgo	ctct	tcto	cgo	caca	aag	gag	agg	atg	aaa	tgc	tgg	gtg	atg	gaa	cgc	cac	
539	agta	atco	ctg	gtgo	ctct	tcto	cg	caca	aagį	gag	agg on 3	atg 3)	aaa	tgc	tgg	gtg	atg	gaa	cgc	cac	
539 599	agta	atco	ctg: gtc	gtgo	ctct	tcto ctti	ccgo	caca	aagg (cag(agg on S CAT	atg 3) TTC	aaa AGA	tgc CAG	tgg AAG	gtg CGC	atg CAG	gaa	CAC	Cac	
539 599 56	agta	atco	ctg: gtc	gtgo tct1	tgto	tcto	ccgo tti	gcad	aag (cag(gag Ex CCA	agg on S CAT H	atg 3) TTC F	aaa AGA R	tgc CAG Q	tgg AAG K	gtg CGC R	atg CAG Q	gaa AGC S	CAC H	cac ATC I	
539 599 56	agta	atco	ctg: gtc	gtg(tgto	tete	ccgo ttts Pre	gcad gcad	aagi (cag(cte	gag Ex CCA P d p:	agg on S CAT H roc	atg 3) TTC F ess	aaa AGA R ing	tgc CAG Q si	tgg AAG K te	gtg CGC R for	atg CAG Q ma	gaa AGC S atur	CAC H rep	ATC I ept:	ide
539 599 56 659	agta gcto TCC1	atco cat:	ctg; gtc IGC	gtgo tct1 CGC1	tgto fACT	tete etti IGCI	ccgo ttts Pre	gcad edi	aagg (cag(cte	gag Ex CCA P d p IGC	agg on S CAT H roc AAG	atg 3) TTC F ess AAC	aaa AGA R ing AAA	tgc CAG Q si	tgg AAG K te TGC	gtg CGC R for GGC	CAG Q D TTC	gaa AGC S atun TGC	CAC H rep	acac ATC I ept: AGG	ide
539 599 56 659 67	agta gcto TCCI S	atco cat: TTG	ctg: gtc IGC C	gtgo teti CGC1 R	tgto fACT Y	tete etti IGCI C	ccgo ttts Pro IGC/ C	caca gcao edi AAC	cago cago cteo IGCI C	gag Ex CCA P d p CCA C	agg on S CAT H roc AAG K	atg 3) TTC F ess AAC N	aaa R R ing AAA K	tgc Q si GGC G	tgg AAG K te TGC C	gtg CGC R for GGC G	CAG Q ma TTC	gaa AGC S tur TGC C	CAC H rep TGC	ATC I ept: AGG R	ide
539 599 56 659 67 719	agta gcto TCCI S TTC <u>I</u>	atco cat; TTG L	ctg: gtc IGC0 C GGA	gtgo tett CGC1 R TTCC	tgto fACT Y CCAC	tete etti C GCG4	ecgo ttts Pro C ACG/	edi AAC	aagi (cag) cte IGC: C AAC	Ex CCA P d p CCA C C AAA	agg on 3 CAT H roc AAG K TAT	atg 3) TTC F ess AAC N TAA	aaa R ing AAA K TTT	tgc CAG Q sisi GGC G ATT	tgg AAG K te TGC C ATG	gtg CGC R for GGC G CTT	CAG Q ma TTC F TGT	agaa AGC S atur TGC C ATC	CAC H rep TGC C C TCA	ATC I ept: AGG R CCCC	ide
539 599 56 659 67 719 87	agta gcto TCCI S TTC <u>I</u> F	atco cat ITG L FGA	ctg: gtc IGC C GGA	gtgd teti CGC1 R ITCC	raci Y CCAC	tete etti C GCG4	ecgo tttg Pro IGC/ C ACG/	edi AAC	aagi cag(cte IGCI C AAC	Ex CCA P d p GCC C AAA	A gg CAT H roc AAG K TAT	atg 3) TTC F ess AAC N TAA	aaa R ing AAA K TTT	tgc Q si GGC G ATT	tgg <u>AAG</u> K te TGC C ATG	gtg CGC R for GGC G CTT	CAG Q Ma TTC F TGT	agaa AGC S itui TGC C ATC	CAC H rep TGC C TCA	ATC I action aco	ide
539 599 56 659 67 719 87 779	agta gctc TCC1 S TTC <u>1</u> F CAGC	atco atco trati L FGA SAA	ctg: gtc IGCO C GGA ATG	gtgo teti CGC1 R ITCO ACC1	raci Y CCAG	tete etti IGCI C GCGA	ccgo ttts Pro CC CC CC CC CC CC CC CC CC CC CC CC CC	gcad gcad edi AAC N ACC	cag(cag(cte IGC1 C AAC/	gag; Exc CCA P d p: IGC. C AAA	agg on S CAT H roc AAG K TAT	atg 3) TTC F ess AAC N TAA	aaa AGA R ing AAA K TTT CTC	tgc Q c si GGC G ATT CAA	tgg AAG K te TGC C ATG	gtg CGC Rfor GGC G CTT	CAG Q ma TTC F TGT	rgaa AGC S Atun TGC C TGT	CAC H Tep CTGC C CTCA	ATC I AGG R CCC	ide
539 599 56 659 67 719 87 779 839	agta gctc TCC1 S TTC <u>1</u> F CAGC CTG1	atco cat; L FGA SAAA	ctg; gtc GGGA ATG, IGT,	gtgd tct1 CGC1 R TTCC ACC1 ATCC	tgto tgto rAC1 Y CCAC	teto etti C GCCA ICTI ITCA	CCGC CTTS CC CC CC CC CC CC CC CC CC CC CC CC CC	caca gcad edi AAC N ACC	aagg (cteo IGC C AACO ITTT	gaga Exc CCA P I p: I GC. C AAA I TTT GAT	agg on S CAT H roc AAG K TAT TTC	atg 3) TTC F ess AAC N TAA TTT	aaa R ing AAA K TTT CTC AAG	CAG Q ssi GGC GATT CAA	AAG K te TGC C ATG TGC CAC	gtg CGC Rfor GGC G CTT ATT TGG	CAG CAG CAG TTC F TGT GCC AAA	rgaa AGC S Atun TGC C TGT AGA	CCAC H Tep TGC C TCA	ATC I AGG R CCCC ACA GAA	ide
539 599 56 659 67 719 87 779 839 839	agta gctc TCC1 S TTC <u>1</u> F CAGC CTG1 GCT4	atco atco cat; L <u>FGA</u> * GAAA FTT	ctg: gtc C GGA ATG. IGT. CTT	gtg (tct1 CGC1 R TTCC ACC1 ATCC	tgto tgto raci Y CCAC riti giti	teto etti IGCI C GCGA ICTI ITCA	CCGC CCCC CCC CCC CCC CCC CCC CCC CCC C	caca gcad edi AAC N ACC	aagg (cte (C AAC) ITTT TAT(ITTT	gaga Exc CCA P d p: rgC. C AAA rTT GAT	agg on 3 CAT H roc AAG. K TAT FTC CAT.	atg 3) TTC F ess AAC N TAA TTT TGG ACT	aaa AGA R Sing AAAA K TTT CTC AAG	t gc t gc CAG Q i si GGC G ATT CAA AAG GTC	AAG K te TGC C ATG C CAC CAC CCC	gtg CGC Rfor GGC CTT ATT TGG GGA	CAG CAG CAG TTC F TGT GCC AAAA	rgaa AGC S Atun TGC C ATC TGT AGA	CAC CAC H re p TGC C TGC C TCA	ATC I ept: AGG R ACC ACA GAA TTG	ide
539 599 56 659 67 719 87 779 839 839 899 959	agta gctc TCC1 S TTC <u>1</u> F CAGC CTG1 GCT4 TAA4	atco atco cat; ITG L FGA SAAA ITT ATTO	ctg; gtc IGCO C GGGA IGTI CTT CAAO	gtg tett CGC1 R TTCC ACC1 ATCC TGT/ CAC1	raci raci y ccac riti giti atti	tete etti C GCGA FCTT FTCA FATT GTAA	Pro C C C C C C C C C C C C C C C C C C C	gcad gcad edi AAC N ACC AGG GCA	aagg (cte(cte(IGC: C AAAC/ ITTT: TTT: TTT:	gag (Ex) (CA) P d p: (GC) (C AAA (CTT) (CTA)	agg on 3 CAT H roc AAG. K TAT TAT CAT. CAT.	atg 3) TTC F ess AAC N TAA TTA TGG ACT	AGA R ing AAA K TTT CTC AAG TCC CAG	CAG Q (sisi GGC G ATT CAA AAG GTC ATG	AAG K K TGC C ATG C ATG C CAC CCC TGA	cGC R for GGC G CTT ATT GGA CTG	CAG CAG CAG TTC F TGT GCC AAA AGT TTT	agaa AGC S Atun TGC C TGT AGA CTT CTT	CAC H Trep TTGC C TTCA C C C C C C C C C C C C C C C C	ATC I ept: AGG R CCC ACA GAA TTG CTA	ide

Figure 1. Nucleotide sequence of cDNA, predicted amino acid sequences, and genomic DNA organization of spinyhead croaker HEPC (CL-hepc). The exon sequence is in uppercase and the intron sequence is in lowercase. The deduced amino acid sequences are translated and shadowed. Vertical arrows show the predicted positions of cleavage sites for signal peptides and mature peptides and the box indicates the putative motif RQKR. The start codon (ATG) and stop codon (TGA) are bold and underlined, the polyadenylation signal (AATAAA) is bold.

Genetics and Molecular Research 14 (4): 16050-16059 (2015)

The CL-hepc genomic DNA sequence was amplified and sequenced and the *C. lucidus* HEPC gene was found to consist of two introns (94 and 165 bp) and three exons (Figure 1), which is identical to many other HEPCs in terms of genome structure. The first exon contained the 5'-UTR, the signal peptide, and 5-amino acid residues of the prodomain peptide. The prodomain extended from exon 1 through exon 3. Two introns were inserted into the prodomain, thus, separating the peptide into three exons. Exon 3 encoded the mature peptide and the 3'-UTR. CL-hepc genomic organization is indicated in Figure 1.

Twenty-five amino acid sequences of various species were obtained using "HEPC" in a key word search of the protein database (http://www.ncbi.nlm.nih.gov/guide/proteins/#databases). All sequences were aligned using ClustalW in the MEGA 5.0 software. All sequences contained eight Cys domains in the mature peptide located in the C-terminal (Figure 2). In addition, most sequences contained a conserved RX(K/R)R domain that could be recognized by the precursor peptidase. Signal peptides are highly conversed in fish, as they are in mammals. Of the 25 sequences obtained through the database search, HEPC from 83416593 *Larimichthys crocea* had the highest similarity (88.89%) with CL-hepc, followed by *Micropterus dolomieu* (87.78%) and *Pogonophryne scotti* (83.33%). In addition, HEPC from the *Danio rerio* isoform1 had 39.51% similarity with CL-hepc and 23.60% with *Homo sapiens*.

Species/Abbrv	*		4111	*								11111			*		*	**	**		* *	*	Ľ
1. Pogonophryne_scotti	MAF	IAV	VT-	LVLAF	ICIL	E S S A	SFI	GVQEL	EEAGAN		VAVYQ	MNE	RMM	N H I		SHI-	SIC	WCC	NCCK	GN	der	CRF	-
2. Alphestes_immaculatus	MAE	IAV	AVT-	LVLAR	ICIL	S S A	VFI	GV	EEAASN	-	AAYQ	HEME	RHM	DHV	2 0	SHL-	SIC	WCC	NCCR	GN	dg	CKF	-
3. CL hepc	MAR	IAV	AVT-I	LVLA	ICL	ESSA	VFI	GVOEL	EEAGSN	1 -1	VAAHQ	SKET	WMP	HF	Q	SHI-	SLC	YCC	NCCK	- N	de l	CRF	Í.
4. 83416593 Larimichthys crocea	MAF	IAV	AVT-I	LVLAR	ICIL	E S A	VFI	GVOEL	AGSN	-	AAHQ	HHE	WMM	N H I	Q K B Q	SHL-	SIC	WCC	NCCK	SN	de i	CRF	-
5. Micropterus salmoides	MAR	IAV	AVT-	LVLAR	ICIL	ESSA	VFI	GVQEL	EEAGSN	1 -	AAHQ	HEHE	WMM	N H I	0	SHI-	SIC	WCC	NCCR	GN	de i	CRF	-
6. Micropterus dolomieu	MVE	IAV	VT-I	LVLA	ICIL	ESSA	VFI	GVQLI	GSN	1 -1	AAHQ	HHE	WMM	NHI		SHL-	SLC	WCC	NCCK	GN	de l	CRF	-
7. Perca_fluviatilis	MTF	VAV	VA-	VLTE	ICIQ	QBAA	VA	EV	PMGI	EN-LI	AEHE	TOVD	WKM	Y N N	HKIG		F KC	VCC	GCCT	PGV-	- da i	CRF	-
8. Sebastes_schlegelii_type_I	MITE	VAV	VA-	V VL AV	ICIQ	B SA	V A	KVQEL	PISN	NPV	ADHE	TVD	L KM	Y NN	E		KC	FCC	NCCI	TG	-cove	SKRF	-
9. hepcidin_5_Epinephelus_moara	MATE	VAV	VT-	LVLAB	ICIL	ESSA	VFI	GVQEL	EEAASS	1 -	VAAYQ	HEME	RMM	DHV	Q N N Q	SHL-	SNC	WCC	NCCK	GN	CGPC	CKF	-
10. hepcidin_4_Epinephelus_moara	MITE	VAV	VT-	LVLIE	ICLO	B BBA	GSF	EV	PHNN	ES-	AAHE	KEE	WKM	Y N N	H	P	KDC	FCC	GCCP	DKS	C GP	CKY	-
11. hepcidin_3_Epinephelus_moara	MTF	VAV	AVT-I	LVLA	ICIL	E S S A	VF	GVQEL	ERALSN	- 1	VAAYQ	ME	RMM	DHV	Q	SHL-	SKC	WCC	NCCK	GN	dg r	CKF	-
12. hepcidin_2_Epinephelus_moara	MITE	VAV	AVI-I	LVLAB	ICIL	E S S A	VFI	GV	PHNN	ES-	AAHE	KEEE	WKM	YNN	H	P	KDC	FCC	GCCP	DMS	d d d s d	CIF	-
13. hepcidin_1_Epinephelus_moara	MITE	VAV	VA-	VLTE	IC Q	ESSA	GSF	EVQEL	EPMNN	ES-	AAHE	KEEE	WKM	Y N N	HKIS	P	KDC	FCC	GCCP	DHS	dg : c	CRF	-
14. Eleginops_maclovinus	MITEO	VAV	VA-	LMLTE	ICIQ	E S A	VVI	DVEL	PLIY	I I-I	TEHE	DIDD	WHI	YT-	A R B	R	RKC	FCC	GCCN	PGI-	- CQT	CTK.F	G
15. Chaenocephalus_aceratus	MPFI	NVAV.	AVA-	VLIC	ICHE	E S S A	V V I	EVPDE	EEPHSY	¥-	TEHE	PDD	WNI	YSP	AR	R	RKC	FCC	NCCS	N - I -	- CQT	CTRRF	-
16. 158358729 Larimichthys_crocea	MTF	VAV	VA-	VVLAR	ICIQ	ESSA	V A.N	EE	QQIYE	AD	P	MPVE	CKM	YYH	ENC	GSP-	ARC	FCC	RCCP	RMR	de i	C R F	-
17. isoform_2_precursor_Danio_rerio	MESI	VFL	AV-	VILTO	VOVE	QITA	VFI	Q-VQD	HHVES	EE	QENQ	LITEAR	HRL	TDPLVLF	ткр	SHL-	SLC	FCC	KCCR	N - K	CG Y	CKF	-
18. isoform_1_precursor_Danio_rerio	MELSI	VEL.	AV-	VILTO	VOVE	QITA	V FI	QQVQD	HHVES	EE	LQENQ	HLTEAR	HRL	TDPLVLE	ТÇ	SHL-	SIC	FCC	KCCR	N -	d gy(0KF	-
19. Homo_sapiens	MALS	QIW	ACL	LLLI	LASL	TGGS	VEPQ	QIGQL	ALQPO	0 R	-AGAR	SWMP	FQR		RR	THT-	PIC	FCC	GCCH	RS	C GM	CKT	-
20. Gorilla_gorilla	MALS	QIW	ACL	LILI	LASL	TGS	VEPQ	QTGQ <mark>L</mark>	ALQPO	R	-AGAR	GWTP	LQR		RR	THE-	PIC	FCC	GCCH	RS	- dax c	CKT	-
21. Ovis_aries	MALSI	QTR	TC-I	LELLV	LISL	TGS	VIPP	QIRQL	TDIQTQ	H	-AGAA	GLTP	LQ-		RR	THT-	PIC	FCC	GCCR	KGT-	- dg : c	CKI	-
22. Rattus_norvegicus	MALSI	RIQ	AC-	LLLLI	LASL	SGA	YLRQ	QTRQT	TALQPW	HG	-AESK	TDDSAT	LNL	K	R	TNE-	PIC	FCC	KCCK	NSS-	- CGL	CIT	-
23. Mus_musculus	MALST	TRIQ	AC-	LILLI	LASL	STT	YLHQ	QHRQT	TELQPL	HG	-EESR	ADIA	PHQ	K	R	TNT-	PIC	FCC	KCCN	NSQ-	- da : c	0KT	-
24. Physeter_catodon	MALNI	QIR	TC-I	LULLI	IVSL	TGGE	V_PP	QIRQL	ADLQTQ		-AGRAI	GLMS	LQ-		PR	THE-	PIC	FCC	SCCR	KGH-	- da i	CRT	-
25. Oryctolagus_cuniculus	MALS	QIQ	AC-	LLLL	LIGL	TGGS	VLQQ	QTQL	AARQPQ		-TGAQ	TSLNP	AHS	S	R	THE-	PIC	FCC	SCCR	NSK-	- CG I	CKT	-
26. Canis_lupus_familiaris	MALS	QIQ	AC-	LELLI	LASV	AVS	VLPH	QIGQI	TDIRAQ	1 1	-AGAE	GLQPT	LQLI	RRL	-RR	THE-	PIC	FCC	GCCK	TP	- da	CIT	-

Figure 2. Clustal W alignment and comparison of the amino acid sequences deduced from *Collichthys lucidus* HEPC cDNA with some similar protein sequences of HEPC. Similarities of more than 50% are labeled in different colors. The amino acid sequence deduced from spinyhead croaker is marked as CL-hepc. Other similar proteins predicted and obtained from GenBank are as follows: *Pogonophryne scotti* (ABY84821.1); *Alphestes immaculatus* (AER00228.1); 158358729 *Larimichthys crocea* (ABY84845.1); 83416593 *L. crocea* (ABC18307); *Micropterus salmoides* (ACD13027.1); *Micropterus dolomieu* (*ACD13029.1*); *Perca fluviatilis* (ABR04075.1); *Sebastes schlegelii* type I (ACD80122.1); HEPC 5 *Epinephelus moara* (ADY16665.1); HEPC 4 *E. moara* (ADY16664.1); HEPC 3 *E. moara* (ADY16663.1); HEPC 1 *E. moara* (ADY16661.1); HEPC 2 *E. moara* (ADY16662.1); *Eleginops maclovinus* (ABY84826.1); *Chaenocephalus aceratus* (ABY84828.1); HEPC isoform 2 *Danio rerio* (NP_001276723.1); HEPC 1soform 1 *D. rerio* (NP_991146.1); *Homo sapiens* (AAH20612.1); *Gorilla gorilla* (ABU75219.1); *Ovis aries* (ADK56130.1); *Rattus norvegicus* (NP_445921.1); *Mus musculus* (NP_115930.1); *Physeter catodon* (XP_007106594.1); *Oryctolagus cuniculus* (XP_008247717.1); and *Canis lupus* familiaris (AAT95397.1).

Genetics and Molecular Research 14 (4): 16050-16059 (2015)

[©]FUNPEC-RP www.funpecrp.com.br

Phylogenetic analysis carried out using the bootstrap test of phylogeny produced an NJphylogenetic tree (Figure 3). The analysis involved all 26 aligned amino acid sequences. Based on the tree, HEPCs were separated into two clusters, one for fish and the other for mammals. The fish cluster was further divided into three branches. CL-hepc has the closest relationship with 83416593 *L. crocea*, thus, supporting the traditional taxonomic relationship. Interestingly, CL-hepc has a relatively distant relationship with another type of HEPC from *L. crocea* (ABY84845.1).





Relative expression of CL-hepc in different tissues

The relative expression of CL-hepc in the liver, blood, heart, muscle, gill, spleen, brain, and kidney was investigated using quantitative real-time PCR. Expression was found to be the highest in the liver and was significantly higher than that observed in other tissues, followed by the kidney and gill. There was no or little expression in the heart, spleen, blood, and brain (Figure 4).

Genetics and Molecular Research 14 (4): 16050-16059 (2015)



Figure 4. Expression profiles of the HEPC (CL-hepc) gene in different tissues of *Collichthys lucidus* detected by quantitative real-time PCR. The quantity of CL-hepc mRNA was normalized to the 18S rRNA transcript level. The y-axis represents the relative ratio of expression levels of CL-hepc/18S rRNA.

DISCUSSION

The full-length sequence of CL-hepc is 816 bp, containing a 264-bp ORF encoding a protein of 87 amino acids. The cDNA has an 11-bp poly(A)+ tail with a typical poly-adenylation signal (AATAAA) located 17 bp upstream. The predicted signal cleavage site of CL-hepc is between Ala24 and Val25, which is similar to that in *Micropterus salmoides* (Robertson et al., 2009) and *P. crocea* (Wang et al., 2009). The amino acid sequence of CL-hepc possesses the conserved motif "RQKR" that can be catalyzed by furin (Hosaka et al., 1991), thus, the mature peptide of CL-hepc contains 25 amino acids and starts with Gln 63, and has the typical eight-Cys structure, which is a common feature of the HEPC family (Park et al., 2001). The genomic sequence of CL-hepc has two introns and three exons, which is similar to that of other identified HEPCs in mammals and fish (Park et al., 2001; Pigeon et al., 2001; Li et al., 2011; Wang et al., 2009, 2012; Qu et al., 2013; Gong et al., 2014). Above all, CL-hepc is a member of the HEPC family.

The BLAST results of the deduced amino acid sequence of CL-hepc on the NCBI website shows that it belongs to the HEPC family, with 88.89% similarity with 83416593 *L. crocea* (ABC18307), 87.78% similarity with *M. dolomieu*, 83.33% similarity with *P. scotti*, and relatively low similarity (47.73%) with 158358729 *L. crocea* (ABY84845.1). In addition, the phylogenetic tree also shows that fish HEPCs are clustered in three branches and the two HEPCs of *L. crocea* are separated. Given such a close relationship with *L. crocea* (Tian et al., 2004), the presence of an additional type of HEPC in *C. lucidus* warrants further investigation.

Antimicrobial peptides form part of the main innate immune response components of fish defense against opportunistic pathogens (Ravichandran et al., 2010). The mRNA expression pattern of a gene is closely related to its function. In this study, we examined the expression of CL-hepc in eight tissues, showing that expression was highest in the liver, which is seen as a common pattern in most HEPCs (Sang et al., 2006; Fu et al., 2007; Robertson et al., 2009; Oliveira Filho et

Genetics and Molecular Research 14 (4): 16050-16059 (2015)

al., 2010; Nam et al., 2011; Wang et al., 2012; Alvarez et al., 2013). In this sense, CL-hepc may prove to be an important component participating in innate immune activities. However, HEPCs are not expressed in the liver in channel catfish (Bao et al., 2005) or zebrafish (Shike et al., 2004), suggesting that they may be different isoforms. Based on the knowledge that the expression of PC-hepc (158358729 *L. crocea*) is high in the kidney and relatively low in the liver, we assume that CL-hepc and PC-hepc are not the same isoform. Although different, all these HEPCs play important roles in innate immune responses (Shike et al., 2004; Bao et al., 2005; Wang et al., 2009; Cai et al., 2012; Qu et al., 2013). In addition, peptides obtained from purified, recombinant, and artificial synthesis all exhibit effective activity against several pathogens (Park et al., 2001; Bao et al., 2005; Wang et al., 2009). Furthermore, it is encouraging that some drug-resistant strains, such as *Aeromonas hydrophila*, are still sensitive to the recombinant HEPC peptide (Gao et al., 2012), suggesting that a future study on HEPC is potentially promising.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the Basic Research Fund for State-Level Nonprofit Research Institutes of ESCFRI, CAFS (# Dong2012M06).

REFERENCES

- Alvarez CA, Santana PA, Guzman F, Marshall S, et al. (2013). Detection of the hepcidin prepropeptide and mature peptide in liver of rainbow trout. *Dev. Comp. Immunol.* 41: 77-81.
- Bao B, Peatman E, Li P, He C, et al. (2005). Catfish hepcidin gene is expressed in a wide range of tissues and exhibits tissuespecific upregulation after bacterial infection. *Dev. Comp. Immunol.* 29: 939-950.
- Cai L, Cai JJ, Liu HP, Fan DQ, et al. (2012). Recombinant medaka (*Oryzias melastigmus*) pro-hepcidin: Multifunctional characterization. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 161: 140-147.
- Chen SL, Xu MY, Ji XS, Yu GC, et al. (2005). Cloning, characterization, and expression analysis of hepcidin gene from red sea bream (*Chrysophrys major*). Antimicrob. Agents Chemother. 49: 1608-1612.
- Cheng J, Ma GQ, Miao ZQ, Shui BN, et al. (2012). Complete mitochondrial genome sequence of the spinyhead croaker *Collichthys lucidus* (Perciformes, Sciaenidae) with phylogenetic considerations. *Mol. Biol. Rep.* 39: 4249-4259.
- Fry MM, Liggett JL and Baek SJ (2004). Molecular cloning and expression of canine hepcidin. Vet. Clin. Pathol. 33: 223-227.
- Fu YM, Li SP, Wu YF and Chang YZ (2007). Identification and expression analysis of hepcidin-like cDNAs from pigeon (*Columba livia*). *Mol. Cell Biochem*. 305: 191-197.
- Ganz T (2003). Hepcidin, a key regulator of iron metabolism and mediator of anemia of inflammation. Blood 102: 783-788.
- Gao Y, Chen CF, Li DP, Tan JJ, et al. (2012). Cloning, expression and antimicrobial activity analysis of hepcidin from Chinese sturgeon (*Acipenser sinensis*). Acta Hydrobiol. Sin. 36: 798-803.
- Gong LC, Wang H and Deng L (2014). Molecular characterization, phylogeny and expression of a hepcidin gene in the blotched snakehead Channa maculata. Dev. Comp. Immunol. 44: 1-11.
- Hosaka M, Nagahama M, Kim WS, Watanabe T, et al. (1991). Arg-X-Lys/Arg-Arg motif as a signal for precursor cleavage catalyzed by furin within the constitutive secretory pathway. J. Biol. Chem. 266: 12127-12130.
- Krause A, Neitz S, Mägert HJ, Schulz A, et al. (2000). LEAP-1, a novel highly disulfide-bonded human peptide, exhibits antimicrobial activity. *FEBS Lett.* 480: 147-150.
- Li W, Sun WX, Tang F, Li CP, et al. (2011). Molecular characterization and expression analysis of a hepcidin gene from rice field eel (*Monopterus albus*). Afr. J. Biotechnol. 10: 7953-7961.

Magnadottir B (2006). Innate immunity of fish (overview). Fish Shellfish Immunol. 20: 137-151.

Nam YK, Cho YS, Lee SY, Kim BS, et al. (2011). Molecular characterization of hepcidin gene from mud loach (*Misgurnus mizolepis*; Cypriniformes). *Fish Shellfish Immunol.* 31: 1251-1258.

Genetics and Molecular Research 14 (4): 16050-16059 (2015)

- Nemeth E, Valore EV, Territo M, Schiller G, et al. (2003). Hepcidin, a putative mediator of anemia of inflammation, is a type II acute-phase protein. *Blood* 101: 2461-2463.
- Nemeth E, Tuttle MS, Powelson J, Vaughn MB, et al. (2004). Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. *Science* 306: 2090-2093.
- Nicolas G, Chauvet C, Viatte L, Danan JL, et al. (2002). The gene encoding the iron regulatory peptide hepcidin is regulated by anemia, hypoxia, and inflammation. J. Clin. Invest. 110: 1037-1044.
- Oliveira Filho JP, Badial PR, Cunha PHJ, Cruz TF, et al. (2010). Cloning, sequencing and expression analysis of the equine hepcidin gene by real-time PCR. *Vet. Immunol. Immunopathol.* 135: 34-42.
- Park CH, Valore EV, Waring AJ and Ganz T (2001). Hepcidin, a urinary antimicrobial peptide synthesized in the liver. J. Biol. Chem. 276: 7806-7810.
- Pigeon C, Ilyin G, Courselaud B, Leroyer P, et al. (2001). A new mouse liver-specific gene, encoding a protein homologous to human antimicrobial peptide hepcidin, is overexpressed during iron overload. J. Biol. Chem. 276: 7811-7819.
- Qu H, Chen B, Peng H and Wang K (2013). Molecular cloning, recombinant expression, and antimicrobial activity of EChepcidin3, a new four-cysteine hepcidin isoform from *Epinephelus coioides*. *Biosci. Biotechnol. Biochem*. 77: 103-110.
- Ravichandran S, Kumaravel K, Rameshkumar G and AjithKumar T (2010). Antimicrobial peptides from the marine fishes. *Res. J. Immunol.* 3: 146-156.
- Robertson LS, Iwanowicz LR and Marranca JM (2009). Identification of centrarchid hepcidins and evidence that 17betaestradiol disrupts constitutive expression of hepcidin-1 and inducible expression of hepcidin-2 in largemouth bass (*Micropterus salmoides*). *Fish Shellfish Immunol*. 26: 898-907.
- Rodrigues PN, Vazquez-Dorado S, Neves JV and Wilson JM (2006). Dual function of fish hepcidin: response to experimental iron overload and bacterial infection in sea bass (*Dicentrarchus labrax*). *Dev. Comp. Immunol.* 30: 1156-1167.
- Saitou N and Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.
- Sang Y, Ramanathan B, Minton JE, Ross CR, et al. (2006). Porcine liver-expressed antimicrobial peptides, hepcidin and LEAP-2: cloning and induction by bacterial infection. *Dev. Comp. Immunol.* 30: 357-366.
- Shike H, Lauth X, Westerman ME, Ostland VE, et al. (2002). Bass hepcidin is a novel antimicrobial peptide induced by bacterial challenge. *Eur. J. Biochem.* 269: 2232-2237.
- Shike H, Shimizu C, Lauth X and Burns JC (2004). Organization and expression analysis of the zebrafish hepcidin gene, an antimicrobial peptide gene conserved among vertebrates. *Dev. Comp. Immunol.* 28: 747-754.
- Sun M, Jiang K, Zhang F, Zhang D, et al. (2012). Effects of various salinities on Na(+)-K(+)-ATPase, Hsp70 and Hsp90 expression profiles in juvenile mitten crabs, *Eriocheir sinensis*. *Genet. Mol. Res.* 11: 978-986.
- Tian LX, Liang B, Zhang SY, Zhao S L, et al. (2004). Phylogenetic relationships of 7 Sciaenidae species based on cytochrome b gene sequences. J. Oceanogr. Taiwan Strait 23: 436-443.
- Wang KJ, Cai JJ, Cai L, Qu HD, et al. (2009). Cloning and expression of a hepcidin gene from a marine fish (*Pseudosciaena crocea*) and the antimicrobial activity of its synthetic peptide. *Peptides* 30: 638-646.
- Wang KJ, Cai JJ and Cai L (2011). New hepcidin antibacterial peptide of large yellow croaker useful as a feed additive for preventing disease and treating sea fishes. Xiamen University, Xiamen.
- Wang Y, Liu X, Ma L, Yu Y, et al. (2012). Identification and characterization of a hepcidin from half-smooth tongue sole *Cynoglossus semilaevis. Fish Shellfish Immunol.* 33: 213-219.

Genetics and Molecular Research 14 (4): 16050-16059 (2015)