Adaptive Evolution of Hepcidin Genes in Antarctic Notothenioid Fishes

Qianghua Xu,‡[1] Chi-Hing Christina Cheng,†[1] Peng Hu,*[1] Hua Ye,* Zuozhou Chen,* Lixue Cao,* Lei Chen,* Yu Shen,* and Liangbiao Chen**

*Key Laboratory of Molecular and Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, People's Republic of China; †Department of Animal Biology, University of Illinois at Urbana-Champaign; and ‡College of Marine Sciences and Shanghai International Marine Bio-resources Research Center, Shanghai Fisheries University, Shanghai, China

Hepcidin is a small bioactive peptide with dual roles as an antimicrobial peptide and as the principal hormonal regulator of iron homeostasis in human and mouse. Hepcidin homologs of very similar structures are found in lower vertebrates, all comprise ~20–25 amino acids with 8 highly conserved cysteines forming 4 intramolecular disulfide bonds, giving hepcidin a hairpin structure. Hepcidins are particularly diverse in teleost fishes, which may be related to the diversity of aquatic environments with varying degree of pathogen challenge, oxygenation, and iron concentration, factors known to alter hepcidin expression in mammals. We characterized the diversity of hepcidin genes of the Antarctic notothenioid fishes that are endemic to the world's coldest and most oxygen-rich marine water. Notothenioid fishes have at least 4 hepcidin variants, in 2 distinctive structural types. Type I hepcidins comprise 3 distinct variants that are homologs of the widespread 8-cysteine hepcidins. Type II is a novel 4-cysteine variant and therefore only 2 possible disulfide bonds, highly expressed in hematopoietic tissues. Analyses of d_N/d_S substitution rate ratios and likelihood ratio test under site-specific models detected significant signal of positive Darwinian selection on the mature hepcidin–coding sequence, suggesting adaptive evolution of notothenioid hepcidins. Genomic polymerase chain reaction and Southern hybridization showed that the novel type II hepcidin occurs exclusively in lineages of the Antarctic notothenioid radiation but not in the basal non-Antarctic taxa, and lineage-specific positive selection was detected on the branch leading to the type II hepcidin clade under branch-site models, suggesting adaptive evolution of the reduced cysteine variant in response to the polar environment. We also isolated a structurally distinct 4-cysteine (4cys) hepcidin from an Antarctic eelpout that is unrelated to the notothenioids but inhabits the same freezing water. Neighbor-Joining (NJ) analyses of teleost hepcidins showed that the eelpout 4cys variant arose independently from the notothenioid version, which lends support to adaptive evolution of reduced cysteine hepcidin variants on cold selection. The NJ tree also showed taxonomic-specific expansions of hepcidin variants, indicating that duplication and diversification of hepcidin genes play important roles in evolutionary response to diverse ecological conditions.

Introduction

The Southern Ocean, characterized by chronic icy, freezing conditions (Hunt et al. 2003; DeVries and Cheng 2005), is the harshest marine environment in the world for ectothermic fish. The Antarctic notothenioid fish underwent an adaptive radiation in this isolated, freezing environment, and the Notothenioidei now dominates the Antarctic fish fauna in species (~45%) and biomass (>90%), representing the only known marine species flock (Eastman 2005). The ecological success of the Antarctic notothenioids had been enabled by evolutionary adaptations to cold including antifreeze glycoproteins that prevent freezing (DeVries 1971; Chen et al. 1997) and cold-stable biochemical and physiological functions (Fields and Somero 1998; Detrich et al. 2000). Recently, we carried out multitissue transcriptome analyses of the notothenioids to elucidate genome-wide adaptive changes that contributed to cold adaptation (Chen, Cheng, Zhang et al. in preparation). In the process, we discovered a highly expressed, novel hepcidin variant that differs significantly from all known vertebrate hepcidins in the conserved cysteine content, leading to a full characterization of the diversity and evolution of hepcidin genes in Notothenioidei.

Hepcidin is initially identified as antimicrobial peptide from human (Krause et al. 2000; Park et al. 2001). Shortly after, the peptide was discovered to be the principal hormone that regulates iron homeostasis (Fleming and Sly 2001; Nicolas et al. 2001; Weinstein et al. 2002). Iron is an integral element in many proteins and enzymes and plays vital roles in cellular processes including O_2 binding, mitochondrial energy production, and cell proliferation (Aisen et al. 2001). Iron metabolism is tightly regulated in all life forms, and conceivably for this critical role, the iron-regulatory hormone hepcidin is ubiquitously present in all vertebrates from fish to mammal (Shi and Camus 2006).

All known hepcidin genes contain 3 exons and 2 introns and encode a prepropeptide that is posttranslationally processed to form the bioactive mature hepcidin (Thompson et al. 1994; Pigeon et al. 2001; Chen, Xu et al. 2005). The length and composition of the mature peptide hepcidins vary between variants and species, however, all contain 8 cysteine residues in conserved positions throughout vertebrate evolution (Ganz and Nemeth 2006; Shi and Camus 2006). The 8 conserved cysteines form 4 intramolecular disulfide bonds that stabilize a hairpin-like structure, important for the proper binding of the hepcidin to its receptor, the cellular iron exporter, ferroportin (Hunter et al. 2002; Lauth et al. 2005). Hepcidin in human is primarily expressed in liver (Park et al. 2001) and is regulated by many biotic and abiotic factors. Hepcidin expression is upregulated upon inflammation (Frazer et al. 2004; Nemeth, Rivera, et al. 2004), anemia (Latunde-Dada et al. 2004), and high iron concentration (Nemeth, Tuttle, et al. 2004) and is downregulated under hypoxic conditions (Nicolas et al. 2002; Leung et al. 2005). The secreted hepcidin is transported via blood stream and bound to ferroportins on the cell surface of enterocytes, macrophages, and hepatocytes,

[1] Equal contribution to this work.

Key words: hepcidin, iron regulation, cold adaptation, positive selection, gene duplication.

E-mail: lbchen@genetics.ac.cn

Mol. Biol. Evol. 25(6):1099–1112. 2008
doi:10.1093/molbev/msn056
Advance Access publication February 29, 2008

causing cellular internalization and the degradation of the iron exporter, thereby reducing iron absorption from the diet and decreased iron efflux from the macrophages and hepatocytes so that plasma iron homeostasis is restored (Nemeth, Tuttle, et al. 2004). In addition, hepcidins in both the mammals and fishes are recognized as antimicrobial agents that function to inhibit the growth of invading microbes (Krause et al. 2000; Park et al. 2001; Shi and Camus 2006).

Many of the above factors that regulate the hepcidin expression in mammals could be encountered by the teleost fishes constantly or periodically in their respective environments and might serve as selection pressures in hepcidin evolution. In correlation with this, the teleost fishes in contrast have much greater hepcidin diversity than human and mice. For example, 5 hepcidin genes were identified from the winter flounder (Douglas et al. 2003) and 7 from the black sea bream (Yang et al. 2007). The diversity of teleost hepcidin genes may indicate elevated quantitative as well as various functional needs for the hepcidin products in coping with their diverse aquatic environments.

The novel notothenioid hepcidin variant discovered in our transcriptome analyses contains only 4 cysteines instead of the widespread 8-cysteine (8cys) hepcidins, with only 2 possible disulfide bonds and predicted greater structural flexibility in cold temperatures. To investigate whether this novel 4-cysteine (4cys) hepcidin as well as its 8cys homologs in Antarctic notothenioids evolved adaptively under selection from the freezing polar environment, we thoroughly characterized the diversity of the hepcidin genes and their evolutionary pattern across Notothenioidei.

Materials and Methods

Specimen and Tissue Collection

The endemic Antarctic notothenioid species (Perciformes: Notothenioidei) representing the 5 Antarctic families were collected as follows: *Dissostichus mawsoni* and *Pagothenia borchgrevinki* (Nototheniidae), and *Gymnodraco acuticeps* (Bathydraconidae) were collected from McMurdo Sound, *Harpagifer antarcticus* (Harpagiferidae) from South Georgia (gift from Tony North of the British Antarctic Survey), *Pogonophryne cerebropogon* and *Pogonophryne scotti* (Artedidraconidae) from the Ross Sea, and *Notothenia coriiceps* (Nototheniidae) and *Chaenocephalus aceratus* (Channichthyidae) from the Antarctic Peninsula near Anvers Island. The cool-temperate water nototheniids, *Dissostichus eleginoides* and *Notothenia angustata*, were collected in northern Ross Sea and Otago Harbor of South Island, New Zealand, respectively. Specimens representing 2 basal non-Antarctic notothenioid families, *Bovichtus variegatus* (Bovichtidae) and *Eleginops maclovinus* (Eleginopidae), were collected from Otago Harbor, New Zealand and Punta Arenas, Chile, respectively. The Antarctic eelpouts *Lycodichthys dearborni* (Perciformes: Zoarcidae) was collected from the deep water of McMurdo Sound. The phylogenetic relationship and geographic distribution of the species used in the study were shown in supplementary figure S1 (Supplementary Material online). Tissues from fish specimens were dissected, flash frozen in liquid nitrogen, and stored at −80 °C until use.

Construction of cDNA Libraries

A cDNA library each was constructed with polyA+ RNA isolated from the liver and head kidney of the Antarctic notothenid *D. mawsoni* and from the liver of the New Zealand notothenid *N. angustata*. Total RNA was first prepared with the Ultraspec II RNA Isolation Kit (Bio-Tex Inc, Houston, TX), and the polyA$^+$ RNA fraction was isolated using Oligo-dT cellulose (Omega Bio-Tek Inc, Norcross, GA) following a batch chromatography protocol. About 7 microgram of polyA$^+$ RNA was utilized for each cDNA library construction using the pCMV-Script XR cDNA Library Construction Kit (Stratagene, CA) following the manufacturer's instructions.

Notothenioid Hepcidin cDNA Sequences

Hepcidin expressed sequence tags (ESTs) were identified from a much larger transcriptome sequence data set generated from an ongoing sequencing effort of the aforesaid cDNA libraries in our laboratory. Plasmid DNA of cDNA clones was isolated in 96-well plate format using the AxyPrep Easy-96 Plasmid DNA Kit (Axygen Biosciences, Union City, CA), and sequencing reactions were performed using BigDye V.3.1 (Applied Biosystems, Foster City, CA) and resolved on an ABI 3730 sequencer. Automated base calling was carried out with PHRED (Ewing et al. 1998) and a stringent error allowance of 0.1%. The EST sequences were trimmed of vector and linker sequences and assembled using CAP3 (Huang and Madan 1999), allowing 1 mismatch in 50 nt if present. An in-house pipeline (Chen, Xue, et al. 2005) was used to batch-send assembled contigs and singlets as BlastX queries to search for protein homologs in the National Center for Biotechnology Information nonredundant, Trembl, and Swissprot databases, generate the association between query sequences and database hits, and calculate EST frequencies. Contigs and singlets with significant similarities to database entries annotated as hepcidin were manually extracted for further analysis. To authenticate the assembled contig sequence, a cDNA clone with the longest insert from each contig was sequenced in its entirety. The collection of notothenioid preprohepcidin cDNA sequences thus obtained was aligned with BioEdit (Hall 1999), and putative cleavage sites of the signal peptide and prodomain were identified using SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen et al. 2004) and by comparison to known vertebrate hepcidin sequences.

Polymerase Chain Reaction Amplification of Notothenioid Hepcidin Genes

Notothenioid hepcidin cDNA sequence analysis revealed 2 distinct types of hepcidin—type I with 8cys (with 3 apparent variants) and type II with 4cys in the mature hepcidin peptide (see Results). To preferentially polymerase chain reaction (PCR) amplify the gene of each of these hepcidin paralogs from select species from each notothenioid family, primer pairs were designed to sufficiently distinct sequences in the 5′ untranslated regions (UTR) and 3′ UTR in the cDNA obtained from the transcriptome data (supplementary table 1 [Supplementary Material online];

primer pairs 1–4). About 200 ng of genomic DNA was used in each PCR amplification reaction. Amplification conditions were an initial denaturation at 95 °C for 5 min, followed by 35 cycles of 50 s denaturation at 95 °C, 45 s annealing at 52 °C or 55 °C, and 1 min extension at 72 °C. PCR product of expected size was purified and cloned into the T-vector (Takara, Dalian, China) or pGemT$_{easy}$ (Promega, Madison, WI) and sequenced. The intron/exon junctions of the gene sequences were delineated by referencing to the cDNA sequence of the respective type of hepcidin gene.

Northern Analysis of Tissue Expression of Hepcidin Messenger Ribonucleic Acid

About 15 μg each of total RNA from the pyloric ceca, liver, head kidney, mesenteric tissue (contains exocrine pancreas), testis, and ovary of *D. mawsoni* and from the liver of *N. angustata* and *L. dearborni* were electrophoresed on denaturing (2.2 M formaldehyde) agarose gel and vacuum transferred onto Hybond-N nylon membrane (Amersham Biosciences, Piscataway, NJ). The blot was hybridized sequentially to 2 notothenioid hepcidin cDNA fragments—the mature peptide–coding sequence (corresponding to the third exon) of 8cys hepcidin variant2 and of 4cys hepcidin. The probe templates were generated by PCR amplification of the respective full-length hepcidin cDNA from *D. mawsoni* using specific primer pairs 5 (8cys hepcidin v2) and 6 (4cys hepcidin) (primer sequences in supplementary table 1 [Supplementary Material online]). Each probe was labeled with ^{32}P-deoxycytidine triphosphate with the Random Primer DNA Labeling Kit V.2 (Takara) and hybridized to the blot in PerfectHyb solution (Sigma, St. Louis, MO) at 56 °C. The blot was washed with 0.1× standard saline citrate/0.5% sodium dodecyl sulfate (SDS) at 56 °C for 30 min and autoradiographed on X-OMAT X-ray film (Kodak, Rochester, NY) for 12–36 h depending on signal intensity. The hybridized probe was stripped with 0.1% SDS at 100 °C before the blot was hybridized to the second probe.

PCR Amplification and Southern Blot Survey of Hepcidin Genes in Teleost Lineages

About 100–200 ng of genomic DNA from 10 notothenioid species representing 7 of the 8 families of Notothenioidei—*B. variegatus*, *E. maclovinus*, *D. mawsoni*, *D. eleginoides*, *N. coriiceps*, *P. borchgrevinki*, *H. antarcticus*, *P. cerebropogon*, *G. acuticeps*, and *C. aceratus*—and from the Antarctic eelpout *L. dearborni*, as well as from non-Antarctic teleosts, zebrafish *Danio rerio* and the puffer fish *Tetraodon nigroviridis*, were used in PCR with primer pairs 5 and 6 (supplementary table 1, Supplementary Material online) that bracket the coding sequences of the mature hepcidin paralogs (third exon) to amplify the specific hepcidin gene of interest. PCR conditions were similar to those described above. The PCR products were resolved on agarose gels, vacuum transferred to nylon membrane, hybridized to each of the 3 hepcidin probes that target the hepcidin third exon (coding sequences [cds] of mature hepcidin), washed, and autoradiographed as described for Northern blot analysis above.

Hepcidin Genes and cDNA from Antarctic Zoarcid fish

Hepcidin genes from the Antarctic eelpout *L. dearborni* (family Zoarcidae) were amplified from genomic DNA using primer pair 7 (supplementary table 1, Supplementary Material online) designed to the conserved 5′ end of the signal peptide–coding sequence (forward primer) and a conserved 3′ UTR site (reverse primer) based on the alignments of known teleost hepcidin sequences and notothenioid sequences obtained in this study. Conditions for PCR amplification of genomic DNA and the cloning of the PCR product were similar to those described for notothenioid fishes. Sufficient number of clones were randomly picked to obtain both 8cys and 4cys hepcidin genes and sequenced. The intron/exon boundaries were delineated by cDNA cloning and also comparison to notothenioid hepcidin genes in this study and other known vertebrate hepcidin genes.

Phylogenetic Analyses of Notothenioid Hepcidins

A total of 20 full-length notothenioid hepcidin gene sequences encompassing hepcidin cDNAs from the transcriptome and genomic hepcidin sequences (spliced protein-coding regions) amplified from select species across Notothenioidei were aligned with ClustalX (Thompson et al. 1997). The nucleotide alignment (supplementary fig. S2, Supplementary Material online) was codon-constrained preserving the positions of the start and stop codons, prodomain cleavage site, and conserved residues in the variable mature hepcidin–coding region. The Hasegawa–Kishino–Yano (HKY) model plus a proportion of invariable sites parameter (I) was selected as the best-fit nucleotide substitution model by both Akaike information criterion and Bayesian information criterion in Modeltest v.3.8 (Posada 2006) implemented on the web-based Modeltest Server v.1 (http://darwin.uvigo.es/), using likelihood scores computed for the hepcidin data set in PAUP4.0b10 (Swofford 2002). The selected model was used in Neighbor-Joining (NJ) tree reconstruction of the hepcidin sequences in PAUP4.0b10 (Swofford 2002), and supports for tree nodes were evaluated with 1,000 bootstrap pseudoreplicates.

Tests for Positive Selection in Notothenioid Hepcidins

The topology of the NJ tree obtained from phylogenetic analysis of the 20 notothenioid hepcidins (preceding section) was used as the input tree to test for the presence of positive selection. Nonsynonymous/synonymous substitution rate ratio ($\omega = d_N/d_S$) tests for positive selection in amino acid sites of the notothenioid hepcidins, as well as in the branch leading to the reduced cysteine (4cys) hepcidin clade, were carried using site models (Yang 2002; Wong and Nielsen 2004) and branch-site models (Yang and Nielsen 2002; Zhang et al. 2005), respectively, using the CODEML program implemented in PAML version 3.15 (Yang 1997). Amino acid site selection was evaluated for the prodomain and the mature hepcidin peptide separately, and parameter estimates (ω, and proportion of each class of sites) and likelihood scores were calculated for 6 models of sequence

evolution—M0 (1-ratio model), M1a (nearly neutral model), M2a (positive selection model), M3 (discrete model), M7 (beta), and M8 (beta and ω) (Yang 2002; Wong and Nielsen 2004). These 6 models comprise 3 nested pairs of a null (no sites under selection) versus a more general model (a proportion of sites under selection)—M0 versus M3, M1a versus M2a, and M7 versus M8. Likelihood ratio test (LRT) was conducted to compare the 2 models in each of these 3 nested pairs to test for presence of positively selected amino acid sites. To determine if lineage-specific positive selection has operated on the reduced cysteine (4cys) hepcidin genes, the branch to the 4cys hepcidin clade was marked in the NJ tree as foreground branch and all other branches as background branches. LRT statistics were estimated to compare model A that allows positive selection on the foreground branch with the null model M1a that does not allow such positive selection (Zhang et al. 2005) and model B (allows selections) with the null model M3 (discrete) (Yang and Nielsen 2002).

Phylogenetic Analysis of Hepcidins across Teleost Lineages

The deduced amino acids of 8 notothenioid hepcidins (a subset of the 20 sequences above) representing all identified types, 2 hepcidins from the Antarctic eelpout from this study and 24 non-Antarctic teleost hepcidins from databases were aligned with ClustalW. Teleostei taxonomic lineages represented in this alignment included Cypriniformes (zebrafish), Beloniformes (medaka), Perciformes (black and red seabreams, Japanese black seabass, 3-spined stickleback, notothenioids, and eelpout), and Pleuronectiformes (olive and winter flounders). The NJ tree reconstruction of the aligned amino acid sequences was performed using the *P* distance model (Saitou and Nei 1987), pairwise deletion of gaps, and uniform rates distribution as implemented in MEGA4.0 (Tamura et al. 2007). Support for tree nodes was evaluated with 1,000 bootstrap pseudoreplicates.

Results

Hepcidin cDNA and Gene from Notothenioid Fish

Hepcidin cDNAs from the Antarctic notothenioid *D. mawsoni* and from *N. angustata* a closely related (both members of the family Nototheniidae) cool-temperate New Zealand species were identified from EST sequences of unnormalized cDNA libraries. From ~9,000 liver and ~5,000 head kidney cDNA clones of *D. mawsoni*, BlastX searches identified a total of 282 ESTs coding for hepcidin (predominantly from the head kidney), whereas 23 hepcidin ESTs were identified from about 4,300 *N. angustata* liver cDNA clones. Assembly of these ESTs produced 3 distinct contigs in *D. mawsoni* and 6 in *N. angustata*, representing different hepcidin variants shown in the alignment of the translated sequences, along with their respective EST frequencies, in fig. 1*A*. The lengths of the preprohepcidins range from 84 to 90 residues and the 3 regions—signal peptide (predomain), propeptide (prodomain), and the mature peptide are demarcated in the protein translation based on characterized vertebrate preprohepcidins (fig. 1*A*). The notothenioid hepcidin variants can be further grouped into 2 distinct types based on a striking sequence divergence feature in the mature hepcidin peptide—8 versus 4 cysteines (fig. 1*A*). Most characterized or annotated putative vertebrate hepcidins have 8 cysteines in the mature peptide at highly conserved positions (Shi and Camus 2006), thus we designated the notothenioid 8cys variants as type I hepcidin. The 4cys hepcidins with the drastic reduction in cysteine residues thus represent a novel type, which we designate as type II hepcidin (fig. 1*A*). Analyses of the genomic sequences of hepcidin variants—3 variants of 8cys hepcidin (variant 1 or v1, v2, and v3) and the 4cys hepcidin, obtained from PCR amplification of genomic DNA from *D. mawsoni* and *N. angustata* (and a number of other species; subsequent Results) show that the gene structures of type I and type II hepcidins are conserved and have the same 3-exon/2-intron structure and exon/intron junctions as other vertebrate hepcidin genes (fig. 1*B*). In exon 3, the conserved residues RX(R/K)R are likely the cleavage sequence of the propeptide (fig. 1), using known vertebrate hepcidins as reference. The mature hepcidin peptide encoded in the remainder of exon 3, in contrast to the pre- and propeptides, shows significant sequence divergence between variants, most notably the 8cys versus 4cys variation, which is also the primary cause for the variation in the total number of amino acids in the mature peptide (fig. 1). The positions of missing cysteine residues in type II hepcidins are conserved, corresponding to the second, fourth, fifth, and seventh cysteines in the mature peptide of the 8cys hepcidins (fig. 1*A*).

The 4cys type II hepcidins in *D. mawsoni* dominate the head kidney transcriptome; *Dm*_4cysHep_a singularly represented 5.56% of total transcripts and *Dm*_4cysHep_b (identical to *Dm*_4cysHep_a except for 2 fewer residues in the propeptide) represents 0.02%, for a total of 5.58% (fig. 1*A*). The 8cys hepcidin (variant 2; see Results on phylogenetic analysis) was expressed at much lower level, at a 0.02% of total transcripts in *D. mawsoni* liver as indicated by the EST frequency. In *N. angustata* liver, 3 variants of 8cys hepcidin and the 4cys hepcidin are found at various low EST frequencies (fig. 1*A*). The three 8cys hepcidin variants differ in the propeptide sequence and more pronouncedly in the mature peptide, indicative of distinct paralogs, whereas the 4cys hepcidin is near 100% identical in sequence to its counterpart in *D. mawsoni*, indicating that the type II hepcidin genes in the 2 species are orthologous (fig. 1*A*; also see Results on phylogenetic analysis).

Northern Analyses of Hepcidin Expression in Notothenioids

Expression level of hepcidin variants was assessed in various tissues of *D. mawsoni* and *N. angustata* by Northern blot analyses (fig. 2*A*). The *D. mawsoni* type I hepcidin probe (mature peptide cds or third exon of *Dm*_c8cysHep_v2) and type II hepcidin probe (mature peptide cds or third exon of *Dm*_c4cysHep_a) detected hepcidin transcripts of ~450 nt and ~600 nt, respectively (fig. 2), consistent with their full-length cDNA sequences from EST assembly and in a tissue-dependent manner.

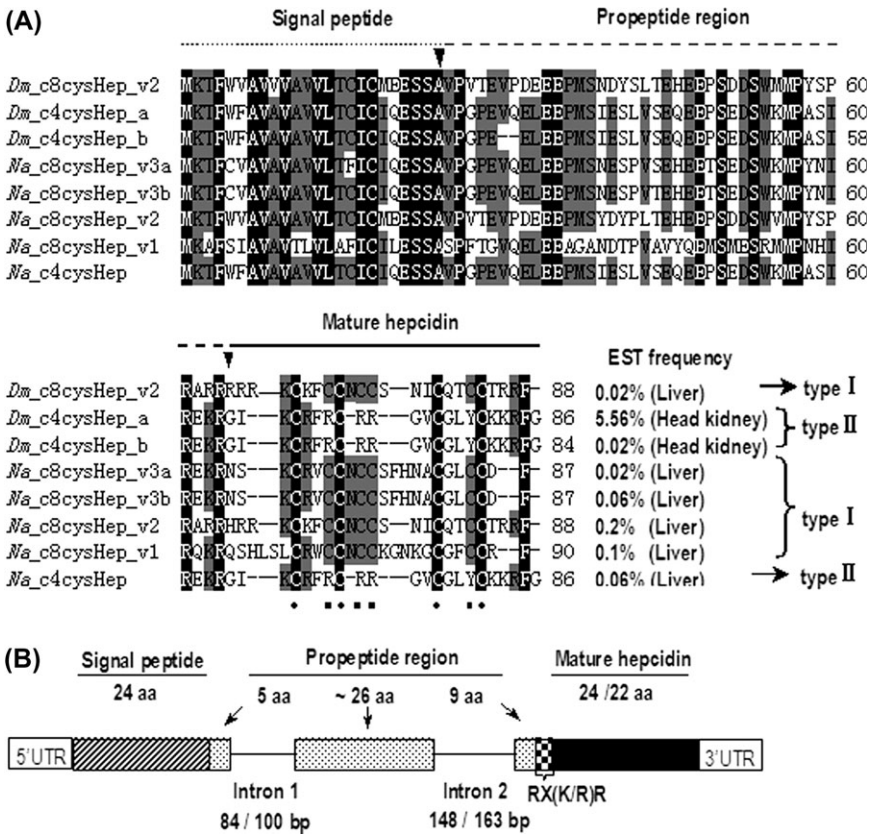

FIG. 1.—Notothenioid hepcidin variants and gene structure. (*A*) Alignment of the predicted amino acid sequences from 3 *Dissostichus mawsoni* and 5 *Notothenia angustata* hepcidin cDNAs using ClustalW. The 3 functional regions of the prepropeptide are denoted on the top of each region with different lines. The number of amino acids and the EST frequency of each variant in the respective transcriptomes are shown to the end of each sequence. The putative cleavage sites of the signal peptide and the mature hepcidin are marked with inverted triangles. Levels of amino acid conservation are indicated by black (identical), gray (2 variations), and white (>2 variations) highlights. Hepcidins with 8 and 4 cysteines in the mature hepcidins are designated as type I and type II, respectively. The character "c" before "8cys" and "4cys" in the gene names denotes cDNA sequence. Small solid circles denote cysteines conserved among all variants and those absent from the type II hepcidins are denoted by small squares, underneath the alignment. (*B*) The intron/exon organization of type I (*Dm*_g8cysHep_v2) and type II (*Dm*_g4cysHep_a) hepcidin variants from *D. mawsoni*. The character "g" before 8cys and 4cys in the gene names denotes genomic sequence. The corresponding functional domains of the prepropeptide are presented by differentially filled rectangles. The number of amino acids in each domain is given under each region. Type I and type II hepcidin variants differ in the mature peptide region with 24 and 22 amino acids, respectively. The intron lengths in base pairs are listed in the order of type I/type II below each intron. The region between the prodomain and the mature hepcidin is the putative proteolytic cleavage site RX(K/R)R of the mature hepcidin.

Dissostichus mawsoni head kidney (fig. 2*A*; lower panel, lane 4) showed highly intense hybridization to the type II hepcidin probe, in keeping with the very high EST frequency of 4cys hepcidins observed in the transcriptome (fig. 1*A*). Another *D. mawsoni* hematopoietic tissue, the spleen, also showed significant 4cys hepcidin expression (fig. 2*A*; lower panel). In contrast, little or no expression of the 8cys type I hepcidin was observed in *D. mawsoni* head kidney or spleen (fig. 2*A*; upper panel). Except for *D. mawsoni* liver, abdominal viscera (pyloric ceca, liver, mesentery which contains adipose and pancreatic tissue), and gonads (testis and ovary) showed little or undetectable expression of either type of hepcidin. The significant hybridization to the 8cys type I hepcidin probe in *D. mawsoni* liver RNA (fig. 2*A*; upper panel, lane 2), contradicted the low (0.02%) of the EST frequency observed in the *D. mawsoni* liver transcriptome (fig. 1*A*, *Dm*_c8cysHep_v2). The cause of the discrepancy is at present unclear but may stem from variation in expression levels among individuals due to variations in physiological status. For *N. angustata* liver, 8cys hepcidin is clearly expressed, whereas the 4cys hepcidin expression was undetectable, in keeping with their EST frequencies of 0.2% and 0.006%, respectively (fig. 1*A*). No hybridization to either probe was observed in liver RNA of the unrelated Antarctic eelpout *L. dearborni*, although it has the homologs of each hepcidin type (see later section in Results).

Distribution of Type I and Type II Hepcidin Genes in Notothenioidei

The large divergence in the mature hepcidin sequences prompted a survey of the distribution of type I and type II hepcidin genes in members of Notothenioidei families. We examined select species from 7 of the 8 notothenioid families. Figure 2*B* shows the hybridization of products from PCR amplification of genomic DNA that targeted the mature hepcidin cds, to a probe derived from the same region

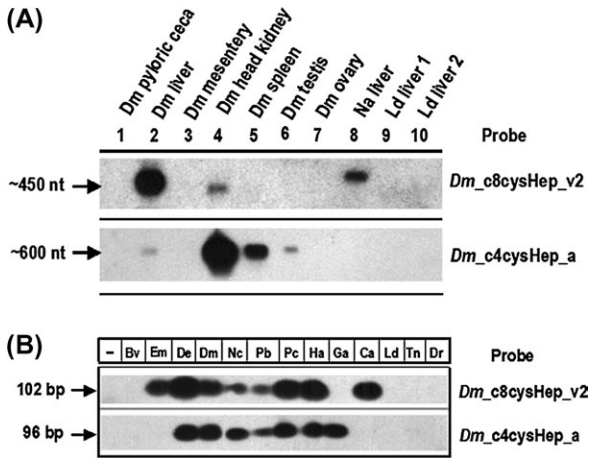

FIG. 2.—(*A*) Tissue-specific expression of notothenioid hepcidin variants. Northern blot hybridization showing the differential expression patterns of type I (top panel) and type II (bottom panel) hepcidin in tissues of *Dissostichus mawsoni* (Dm), *Notothenia angustata* (Na), and *Lycodichthys dearborni* (Ld). The distinct mature hepcidin–coding region the pertinent hepcidin cDNA from *D. mawsoni* (*Dm*_c8cysHep_v2, type I and *Dm*_c4cysHep_a, type II) were used as probes. The sizes of the respective messenger ribonucleic acids were shown to the left of each panel. (*B*) Lineage-confined distribution of type I and type II notothenioid hepcidin variants. Southern blot of products from PCR amplification of the mature peptide–coding sequence of type I (8cys variant 2) and type II (4cys) hepcidin from genomic DNA hybridized to probes derived from the same region from *D. mawsoni* hepcidin (*Dm*_c8cysHep_v2, upper panel and *Dm*_c4cysHep_a, lower panel). The sizes of positive PCR products were shown to the left of each panel. The species names are shown on top of each lane: –, negative control; Bv, *Bovichtus variegatus*; Em, *Eleginops maclovinus*; De, *Dissostichus eleginoides*; Dm, *D. mawsoni*; Nc, *Notothenia coriiceps*; Pb, *Pagothenia borchgrevinki*; Pc, *Pogonophryne cerebropogon*; Ha, *Harpagifer antarcticus*; Ga, *Gymnodraco acuticeps*; Ca, *Chaenocephalus aceratus*; Ld, *L. dearborni*; Tn, *Tetraodon nigroviridis*; and Dr, *Danio rerio*.

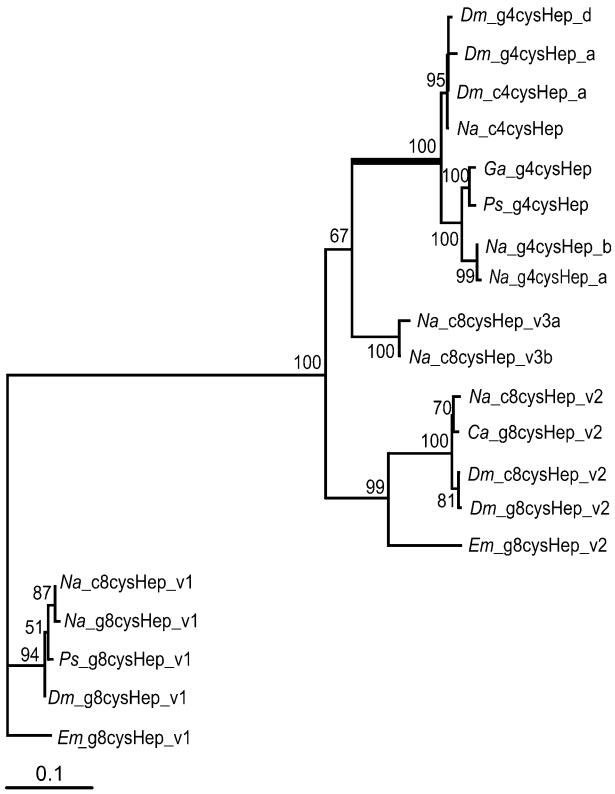

FIG. 3.—Evolutionary relationship of notothenioid hepcidin variants. The NJ reconstruction of the evolutionary relationship of 20 notothenioid hepcidin–coding sequences using HKY + I model of sequence evolution. Node supports were evaluated with 1,000 bootstrap pseudoreplicates. The branch to the 4cys hepcidin clade (foreground branch) for lineage-specific selection analysis was indicated in bold. The characters "c" and "g" before "8cys" and "4cys" in the gene names denote cDNA and genomic sequence, respectively.

of type I (*Dm*_c8cysHep_v2; fig. 2*B* upper panel) or type II (*Dm*_c4cysHep; fig. 2*B* lower panel) hepcidin. Type I 8cys hepcidin_v2 gene was absent in *B. variegatus* belonging to the most basal non-Antarctic notothenioid family Bovichtidae but present in *E. maclovinus*, the monotypic species of the basal non-Antarctic family Eleginopidae and closest sister taxon to the 5 Antarctic families, and in members of 4 of the 5 Antarctic families—Nototheniidae (*D. mawsoni*, *D. eleginoides*, *N. coriiceps*, and *P. borchgrevinki)*, Artedidraconidae (*P. cerebropogon*), Harpagiferidae (*H. antarcticus*), and the most derived family Channichthyidae (*C. aceratus*) but not detected in the Bathydraconidae species *G. acuticeps* (fig. 2*B* upper panel). In contrast, 4cys type II hepcidin genes were found exclusively in the Antarctic species, except in the channichthyid *C. aceratus* (fig. 2*B*, lower panel). No PCR product and thus no hybridization to either hepcidin probe was obtained for the unrelated Antarctic eelpout *L. dearborni* and the tropical green spotted pufferfish *T. nigroviridis* and zebrafish *D. rerio*, indicating that their hepcidin gene sequences are too divergent from notothenioid sequences for annealing of notothenioid primers. (*L. dearborni* has hepcidin genes [see later section in Results], and both zebrafish and the green spotted pufferfish both have known hepcidin sequences in the databases).

Evolutionary Relationship of Notothenioid Hepcidin Variants

Twenty distinct preprohepcidin-coding sequences from genomic and cDNA clones representing distinct hepcidin genes that we obtained from 6 notothenioid species (supplementary fig. S2, Supplementary Material online) were selected for phylogenetic analysis to assess their evolutionary relationship. Figure 3 shows NJ reconstruction of the evolutionary relationship of the 20 hepcidins implementing HKY + I. One well-supported clade of 4cys hepcidins was recovered, as well as 3 distinct clades of 8cys hepcidins (fig. 3), which lead to and validated our initial designation of the 3 distinct type I hepcidin variants—v1, v2, and v3 (fig. 1*A*). The 8cys hepcidin_v1 clade is the longest branching clade, and the gene is present in the basal non-Antarctic notothenioid *E. maclovinus* (fig. 3). The 4cys type II hepcidins are found only in the species of the more derived Antarctic families (fig. 2*B*). The temporal evolution of the hepcidin variants in notothenioid fishes can be further inferred from sequence comparison to hepcidin genes of more basal teleostean lineages such as zebrafish (order Cypriniformes, supplementary fig. S1 [Supplementary Material online]). Zebrafish has 2 hepcidin genes and they are closely related to those of amphibians and

Table 1
Parameter Estimates and Log Likelihoods in Various Models of d_N/d_S Ratio Test on Regions of Notothenioid Hepcidin Genes

Test Region	Models	Parameter Estimates	*l*
Mature hepcidin	M0	ω = 2.44	−398.83
	M1a	ω0 = 0.00, p0 = 0.16, (ω1 = 1) p1 = 0.84	−393.45
	M2a	ω0 = 0.00, p0 = 0.16, (ω1 = 1) p1 = 0.08, ω2 = 3.50, p2 = 0.76	−389.64
	M3	ω0 = 0.00, p0 = 0.16, ω1 = 2.09, p1 = 0.52, ω2 = 7.94, p2 = 0.32	−388.7
	M7	p = 3.32, q = 0.005	−401.17
	M8	p0 = 0.17, p = 0.005, q = 99.00, (p1 = 0.83) ω = 3.11	−389.7
Propeptide region	M0	ω = 0.51	−647.63
	M1a	ω0 = 0.28, p0 = 0.53, (ω1 = 1) p1 = 0.47	−645.84
	M2a	ω0 = 0.28, p0 = 0.53, (ω1 = 1) p1 = 0.35, ω2 = 1.00, p2 = 0.12	−645.8
	M3	ω0 = 0.00, p0 = 0.09, ω1 = 0.37, p1 = 0.45, ω2 = 0.84, p2 = 0.46	−645.46
	M7	p = 1.01, q = 0.78	−645.5
	M8	p0 = 1.00, p = 1.01, q = 0.78, (p1 = 0.00) ω = 1.00	−645.5

mammals (Shi and Camus 2006), indicating that the zebrafish ortholog might have appeared in the common ancestor of the vertebrates. Among the hepcidin variants of notothenioids, 8cys hepcidin_v1 shows the highest sequence similarity (77% amino acid identity in mature hepcidin region) to the zebrafish hepcidin, suggesting that the 8cys hepcidin_v1 type is the oldest hepcidin lineage in the notothenioids and the other variants are from more recent diversifications. The NJ tree of teleost hepcidins rooted with zebrafish hepcidins (fig. 5; see subsequent Results) corroborates this evolutionary inference.

Positive Selection on Amino Acid Sites in Mature Hepcidin Peptides of Notothenioids

The diversity of notothenioid hepcidin paralogs including the novel 4cys type II hepcidins is suggestive of diversifying positive selection operating on this gene family, which we assessed with d_N/d_S rate ratio tests using the NJ notothenioid hepcidin tree topology (fig. 3) as input. Table 1 shows the parameter estimates and likelihood values from separate analyses of the propeptide and the mature peptide under 6 site-specific models. Three of these, the selection model (M2a), the discrete model (M3), and the beta plus ω model (M8) allow a fraction of amino acid sites to be under positive selection, that is, $\omega > 1$ (Yang and Nielsen 2000). Estimates from these 3 models suggest that a large proportion of mature hepcidin sites ($p_2 = 76\%$, $p_1 + p_2 = 84\%$, and $p_2 = 83\%$, for M2a, M3, and M8, respectively) are under diversifying positive selection (ω between 2.09 and 7.94). Even the 1-ratio model suggested positive selection, with an average ω of 2.44 over all sites. In contrast, the propeptide sequences are substantially more conserved, with no sites having ω ratio >1, suggesting absence of positive selection. For the mature hepcidin peptide sequences, the LRT statistic (2Δl) for comparing the 3 nested model pairs is highly significant for the M0 versus M3 and M7 versus M8 comparisons and significant for the M1 versus M2 comparison (table 2), rejecting the null model in favor of presence of positive selective pressure.

Whether the branch (foreground) leading to the 4cys hepcidin clade is under positive selection was tested using branch-site models (Yang and Nielsen 2002; Zhang et al. 2005), which let ω ratio vary both among sites and among lineages to estimate ω_2 of the foreground branch, which may be >1. For both model A (test 1 based on Zhang et al. 2005) and model B (Yang and Nielsen 2002), the foreground ω_2 was infinite with $p_2 = 1$ (i.e., 100% probability) (table 3), indicating high certainty of positive selection on the marked branch (fig. 3). LRT statistic (2Δl) comparing model A with the neutral model M1a and model B with M3 ($k = 2$), rejected the null hypothesis by significant *P* values (table 3).

Independent Evolution of Type II Hepcidin in Unrelated Antarctic Fish

To assess if the evolution of 4cys type II hepcidin in Antarctic notothenioids may be related to inhabitation in freezing marine environment, we cloned the hepcidin genes from an Antarctic zoarcid (eelpout) fish *L. dearborni* that is unrelated to notothenioids. Both type I (*Ld*-g8cysHep) and type II (*Ld*-g4cysHep) hepcidin genes (fig. 4*A*) were isolated from recombinant clones derived from *L. dearborni* genomic DNA. Like the notothenioid hepcidins, the prepropeptides of eelpout type I and type II hepcidins share substantial sequence similarities but are highly divergent in the mature peptide including the reduction to 4cys in the type II molecules (fig. 4*A*). However, the eelpout mature hepcidin peptide sequences differ significantly from those of the notothenioids, which would account for the absence of hybridization to notothenioid hepcidin probes in Northern blot of eelpout liver RNA (fig. 2*A*) and unproductive PCR amplification of hepcidin genes from eelpout DNA using notothenioid hepcidin primers (fig. 2*B*). Additionally, in eelpout 4cys hepcidin, the positions of cysteine loss correspond to the third, fourth, fifth, and seventh cysteines of eelpout 8cys hepcidin, differing from the notothenioid in 1 position (loss of second, fourth, fifth, and seventh cysteines) (fig. 4*A* and *B*). This results in a different putative intramolecular disulfide bond, between second and sixth cysteines for

Table 2
LRT for Site Models

Test Region	Models	2Δ*l*	Degrees of Freedom	*P* Value
Mature hepcidin	M0 versus M3	21.08	2	<0.0001
	M1a versus M2a	7.62	2	0.0221
	M7 versus M8	22.78	2	<0.0001

Table 3
Parameter Estimates and LRTs for the Branch-Site Models

Models	Parameter Estimates	*l*	2Δ*l*	Degrees of Freedom	*P* Value
Model A versus M1a (test 1)	ω0 = 0.00, p0 = 0, (ω1 = 1) p1 = 0, ω2 = infinity, p2 = 1	−390.60	5.7	2	0.0578
Model B versus M3[a]	ω0 = 0.00, p0 = 0, ω1 = 3.46, p1 = 0, ω2 = infinity, p2 = 1	−386.36	6.84	2	0.0372

[a] Indicates the parameters ($k = 2$, $l = -389.78$).

eelpout versus between third and sixth cysteines for notothenioid (fig. 4*B*) based on the 8cys hepcidin numbering system and known disulfide linkages of vertebrate hepcidins (Hunter et al. 2002; Lauth et al. 2005). Thus, Antarctic eelpout 4cys hepcidin was derived from a different hepcidin homolog independent of the notothenioid version.

In NJ phylogenetic analyses of translated preprohepcidin sequences from Antarctic and non-Antarctic teleost lineages (fig. 5), the eelpout 8cys hepcidin and notothenioid 8cys hepcidin_v1 occur in a well-supported clade with 8cys hepcidin from other members of Perciformes (medaka and 3-spine stickleback) and from the more derived olive and winter flounders (Pleuronectiformes). This clade is phylogenetically closer to the two 8cys hepcidin genes of zebrafish belonging to the basal lineage Cypriniformes (fig. 5), indicating that notothenioid 8cys hepcidin_v1 and its teleost orthologs represent an older hepcidin gene prevalent across teleost lineages. The eelpout 4cys hepcidin was recovered as a distinct branch more closely related to the 8cys hepcidin_v1 clade than the notothenioid 4cys hepcidin clade (fig. 5), consistent with its independent evolution from a distinct precursor. The NJ analysis additionally illustrates an overall teleostean diversification of hepcidin paralogs in lineage- or species-specific manner. The 13 hepcidin genes from the related black and red sea breams and Japanese seabass (Percoidei) clustered in a large group, whereas hepcidins of the flounders (Pleuronectoidei) and notothenioids (Notothenioidei) clustered among both related species and unrelated species.

Discussion

Adaptive Evolution of Hepcidins in Notothenioid Fishes

Vertebrate hepcidins have been well characterized in mammals, particularly human (Ganz and Nemeth 2006), and in a number of fishes (Shi and Camus 2006). They are encoded as prepropeptides of generally ~80–90 amino acids in length, encompassing a conserved 24-residue signal peptide and propeptide (~40 residues), and a more variable mature hepcidin (Pigeon et al. 2001; Chen et al. 2005). Despite sequence variations, almost all known vertebrate mature hepcidins have 8 highly conserved cysteines, forming 4 positionally specific intramolecular disulfide bonds, 3 across the stem of the hairpin structure of the mature peptide, and 1 unusual vicinal bond between the adjacent fourth

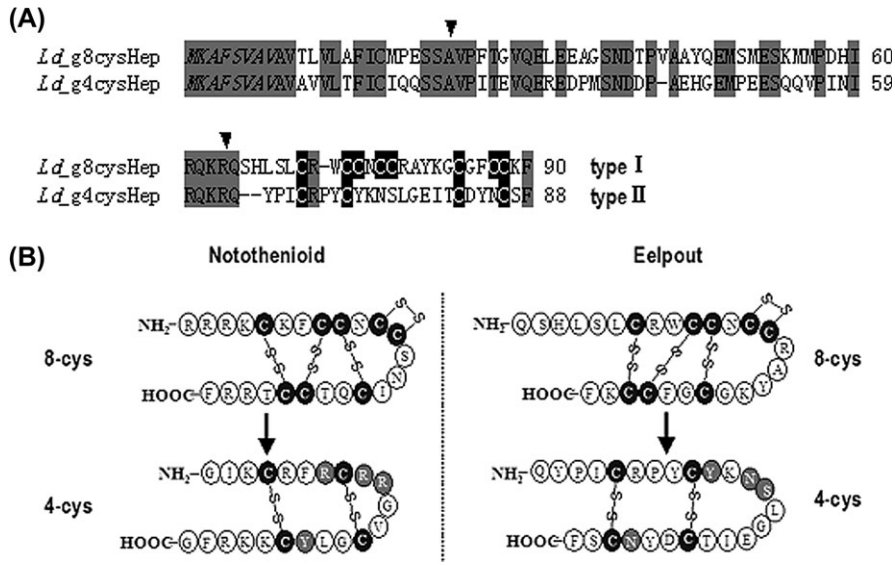

FIG. 4.—Hepcidin genes from an Antarctic eelpout and the independent origins of the notothenioid and eelpout 4cys hepcidins. (*A*) Alignment of the deduced amino acid sequences of 8cys (type I) and 4cys (type II) hepcidin genes from the Antarctic eelpout, *Lycodichthys dearborni*. Identical amino acids between the encoded by the 2 genes are highlighted in gray and the cysteine residues of the mature peptide region in black. The putative cleavage sites for the signal peptide and propeptide (denoted by inverted triangles) were predicted using the same rules as in the notothenioid preprohepcidins. The sequences covered by the 5′ PCR primer were in italic. (*B*) Diagrammatic illustration of the independent evolutionarily reduction of the cysteine residues from their respective 8cys containing hepcidin precursors in notothenioids and the unrelated Antarctic eelpout. Black circles indicate the locations of the cysteine residues, and the gray circles indicate the positions where the cysteines were replaced by other amino acids in the 4cys hepcidins. The notothenioid illustrations were based on the sequences of *Dm*_c8cysHep_v2 (8cys) and *Dm*_c4cysHep_a (4cys) and the eelpout illustrations based on *Ld*_g8cysHep and *Ld*_g4cysHep.

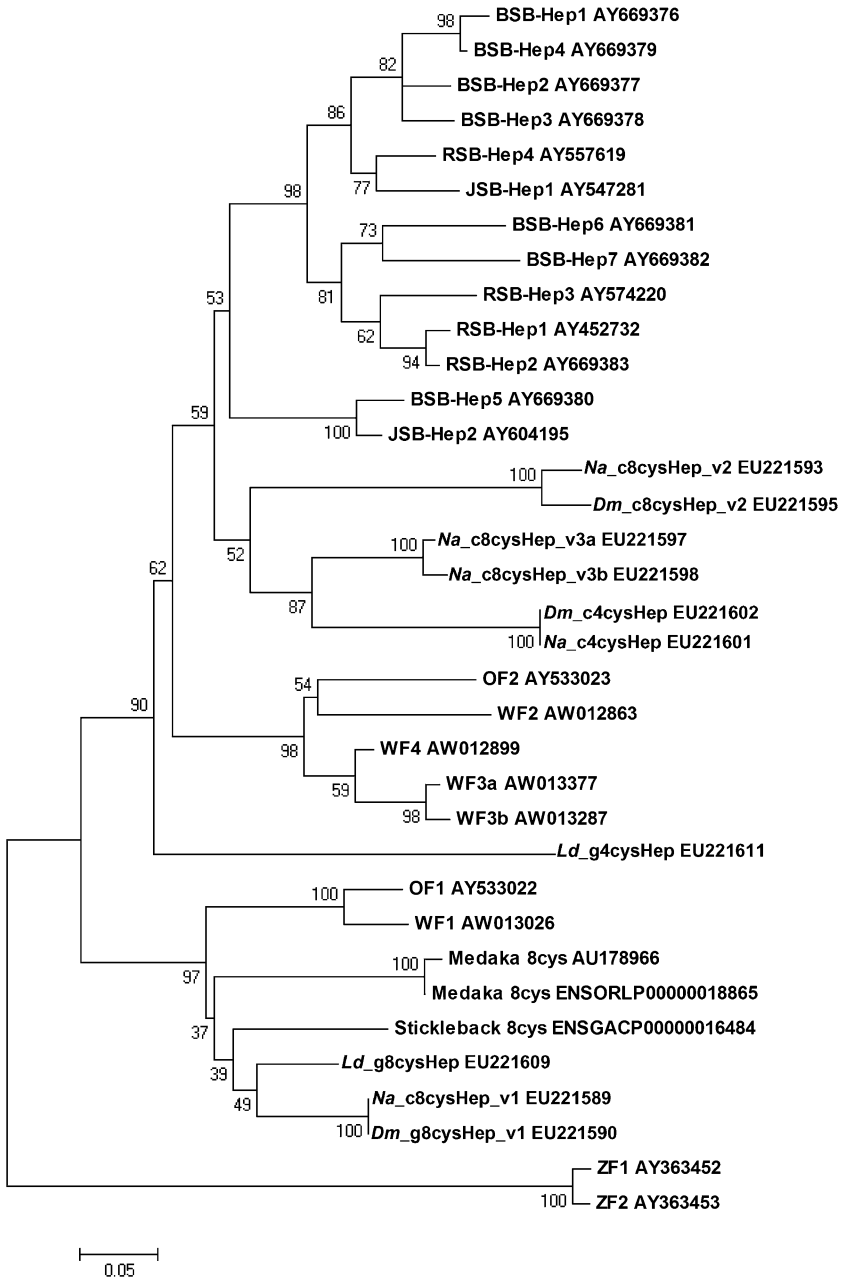

FIG. 5.—Evolutionary relationships of teleost hepcidins. The NJ reconstruction of the evolutionary relationship based on the deduced amino acid sequences of 35 hepcidin genes from 11 species representing Teleostei lineages of Cypriniformes (zebrafish), Beloniformes (medaka), Perciformes (black and red seabreams, Japanese black seabass, 3-spined stickleback, notothenioids, and eelpout), and Pleuronectiformes (olive and winter flounders). Sequences are aligned using ClustalW and manually checked to ensure the proper alignment of the conserved cysteine residues. The evolutionary distances were computed using *P* distance model. The phylogenetic relationship was inferred using MEGA4.0 with 1,000 bootstrap replicates. Because zebrafish is a species of the most basal teleost lineage (Cypriniformes) used in this study, the tree was rooted using the zebrafish hepcidin genes as the outgroup.

and fifth cysteines at the loop region (Hunter et al. 2002; fig. 4*B*). In our transcriptome analyses of 2 notothenioid fish, *D. mawsoni* and *N. angustata*, we identified 3 and 5 distinct hepcidin cDNAs, respectively, in these 2 species. We cloned the hepcidin genomic sequences from these 2 species, as well as a number of species from other families of Notothenioidei, and all have similar gene structures to known vertebrate hepcidin genes (fig. 1*B*). However, peculiar to the notothenioids, besides the homologs of the widespread 8cys hepcidins (designated as type I), we found a second, distinctive type of hepcidin (designated as type II) in which only 4 cysteines are present in the mature peptide. The other 4, corresponding to the second, fourth, fifth, and seventh cysteines of the 8cys homologs have been lost or replaced by other amino acids, leading to a hepcidin molecule with only 2 of the 4 possible intramolecular disulfide bonds (fig. 4*B*). One of the 4cys type II hepcidin gene was expressed at extraordinarily high frequency (5.56% of all

transcripts) in *D. mawsoni* head kidney (fig. 1*A*) and at high levels in the spleen but not in gonads or several abdominal visceral components based on Northern blot (fig. 2*A*), suggesting an important functional role of this reduced cysteine form in hematopoietic tissues.

The 8cys hepcidins of *D. mawsoni* and *N. angustata* show sequence divergence in the mature peptide based on which 3 variants are distinguishable, which we designated as 8cysHep_v1, v2, and v3 (fig. 1*A*). Phylogenetic analyses of the collection of 20 distinct hepcidin genes from species across Notothenioidei validated the existence of these 3 variants in their recovery as 3 distinct clades in NJ reconstruction (fig. 3). The NJ tree infers that 8cys hepcidin_v1 is the oldest hepcidin lineage in notothenioids, and successive gene duplication events lead to the other variants, with the 4cys type II hepcidins being the most recent members of the hepcidin diversification. This is corroborated by the results of genomic PCR and Southern hybridization survey of hepcidin variants across Notothenioidei, showing the 8cys hepcidin genes to be present in both basal non-Antarctic (*E. maclovinus)* and Antarctic notothenioids, whereas 4cys hepcidin genes are present in species of the Antarctic notothenioid radiation (Eastman 2005) predominantly endemic to the freezing Antarctic environment or later diverged to cool temperate waters (*D. elegnioides* and *N. angustata*) (fig. 2*B*). Secondary loss appeared to have occurred in some species—8cys hepcidin_v2 in *G. aculiceps* and 4cys hepcidin in *C. aceratus* (fig. 2*B*). The cause of the loss is unknown, but for 4cys hepcidin, which is highly expressed in hematopoietic tissues (head kidney and spleen) in *D. mawsoni* (figs. 1*A* and 2*A* upper panel), its loss in *C. aceratus*, a member of the derived notothenioid family of icefishes, might be related to loss of hemoglobin and absence of hematopoiesis (Cocca et al. 1995; Near et al. 2006).

What is clear is hepcidin diversification had occurred, leading to the emergence of the novel 4cys type II hepcidin exclusively in species of the Antarctic notothenioid radiation, suggesting that freezing selection might have driven the adaptive hepcidin evolution in Antarctic notothenioids. Our estimates of nonsynonymous/synonymous codon substitution rate ratios ($\omega = d_N/d_S$) under 6 site-specific models (Yang and Nielsen 2000) showed a substantial proportion of sites in the mature hepcidin peptides with large ω (>1) but not on the propeptides (table 1), suggesting presence of positive diversifying selection on the mature hepcidin molecule. These 6 models constitute 3 nested pairs of a null and a selection model (M0 and M3, M1a and M2a, and M7 and M8), and LRTs comparing each pair are consistently significant for positive selection in the mature hepcidin–coding sequences (table 2). We also evaluated if functional divergence after gene duplication might have caused the adaptive evolution of the novel 4cys type II hepcidin lineage by testing with branch-site models (Yang and Nielsen 2002; Zhang et al. 2005). Results of LRT of model B against M3 (Yang and Nielsen 2002) and model A against M1a (test 1 of Zhang et al. 2005) are significant for positive selection on the branch leading to the 4cys hepcidin clade (table 3). However, test 2 (Zhang et al. 2005) did not reject the null hypothesis of no lineage-specific selection (results not shown). Test 2 (Zhang et al. 2005) is a more direct test for identifying positive selections in the foreground branch because significant LRT from test 1 can result from either positive selection or relaxed selective constraint in the foreground branch (Zhang et al. 2005). However, the extraordinarily high 4cys hepcidin transcript frequency (fig. 1*A*) and strong Northern hybridization with 4cys hepcidin probe (fig. 2*A*) in *D. mawsoni* hematopoietic tissues are indicative of an important functional role for this new hepcidin type in the cold-adapted notothenioid and argue against the possibility of relaxed constraints.

The exclusive presence of the 4cys type II hepcidin genes in species of the Antarctic notothenioid radiation but not their basal non-Antarctic relatives are consistent with the adaptive evolution of 4cys hepcidin genes on selection by freezing Antarctic water temperatures. The independent evolution of a structurally distinct 4cys hepcidin homolog (*Ld*-g4cysHep), we found in the Antarctic eelpout *L. dearborni* (figs. 4 and 5) that is unrelated to Antarctic notothenioids but occurs in the same freezing water, lends support to this hypothesis of cold-temperature driven adaptive evolution of the reduced cysteine form of hepcidin. Hepcidins with fewer number of cysteine reductions have also been found in other teleosts. Black sea bream has a 7cys hepcidin (BSB7 in fig. 5) (Yang et al. 2007), and several flat fishes (Pleuronectiformes) including winter flounder *Pleuronectes americanus* have a 6cys hepcidin (WF2 in fig. 5) (Douglas et al. 2003). Some of these flat fishes have a northern high–latitude distribution, particularly winter flounder, which is a north subpolar species and which, like the Antarctic notothenioids, synthesizes antifreeze protein to avoid freezing (Graham et al. 1985), suggesting that cold marine temperatures may be a contributing selection on the evolution of its 6cys variant. In the Antarctic notothenioids, among the amino acid sites determined to be under positive selection (supplementary fig. S3, Supplementary Material online), 2 sites correspond to the fourth and seventh cysteine residues of the 8cys hepcidins, both of which were lost in the 4cys variant and in the winter flounder, the fourth cysteine is lost in the 6cys hepcidins. This suggests that cold selection might have exerted influence in the cysteine content of the hepcidins in cold-water fishes.

The functional consequence of reduced intramolecular disulfide bonds in the 4cys hepcidins of the Antarctic fishes remains to be determined. Currently, structure–function studies of teleost hepcidins are lacking. Studies of human hepcidin showed that synthetic mutants with 1 disulfide bond removed by pairwise substitution of cysteines with alanine exhibited only minor decrease in ferroportin-binding activity in vitro and over half (54%) of the activity persisted even when 3 of the 4 bonds were removed leaving only the most terminal bond (between first and eighth cysteines) intact (Nemeth et al. 2006). If the structure–function relationships of human hepcidins are applicable to Antarctic fish hepcidins, the 4cys type II hepcidins with 2 of the 3 possible cross hairpin-stem disulfide bonds (fig. 4*B*) intact may have full ferroportin-binding activity and therefore retain the iron regulation function. Transgenic overexpression of the 8cys or 4cys notothenioid hepcidins in fertilized zebrafish eggs both lead to high embryo mortality, indicated that both

hepcidin variants possess the iron regulation function (Chen L, unpublished data).

The reduction to 4 cysteines in the Antarctic notothenioid hepcidin is by far the most dramatic and likely would increase the flexibility of the hairpin structure. Structure–function comparisons of lactate dehydrogenase A_4 orthologs (A_4-LDHs) from Antarctic and non-Antarctic notothenioid species showed that amino acid substitution leads to increase in molecular flexibility in small areas of the cold-adapted A_4-LDHs that affect the mobility of adjacent active sites resulting in modulation of the enzyme's catalytic rate (*K*cat) and substrate-binding affinity (*K*m) (Fields and Somero 1998). One general hypothesis of protein cold adaptation from this line of investigation is the acquisition of small sequence changes leads to increased molecular flexibility in the areas near the sites of conformational change during catalysis and thus reduce activation energy in the cold (Fields 2001; Somero 2004). In the 4cys hepcidin of Antarctic fish, the most obvious substitutions are the cysteine residues that lead to the reduction of intramolecular bonds, and this would likely increase structural flexibility. Functionally, the N-terminal 6 amino acids before the most terminal disulfide bond are critical ones for ferroportin-binding activity in human (Nemeth et al. 2006). It is possible that the increased flexibility by the reduced disulfide bonds in the 4cys hepcidin may facilitate the conformational change in the nearby N-terminal active site for proper binding of the notothenioid hepcidin to ferroportin in the cold. Thus, the adaptive evolution of the notothenioid hepcidins appears to follow the similar strategy as the enzymatic adaptation in the Antarctic fishes. Characterizing the ferroportin gene in this group of fish and to see whether structural variations have occurred in the hepcidin receptor to accommodate the structural diversities of the ligands would be informative to reveal the mechanisms underlying positive selection of the notothenioid hepcidins.

Hepcidin was initially discovered as antimicrobial peptide synthesized by human liver (Park et al. 2001) before its dual function as a hormonal regulator of iron metabolism in human and mice was determined. In vitro study using synthetic bass hepcidin showed active inhibition of Gram-negative bacteria and fungi at 10–100 nM concentrations, similar to those required for other antimicrobial peptides (Lauth et al. 2005). The iron regulation function and the antimicrobial function are not independent, at least in mammals, where the antibacterial function of hepcidin could be achieved by depleting the plasma iron essential for the proliferation of invading bacteria (Ganz and Nemeth 2006). In teleost fishes, however, due to the existence of multiple hepcidin genes, functional separation between the hepcidin types was suggested (Shi and Camus 2006). Thus, some or all the notothenioid hepcidin variants may participate in innate immunity, a function particularly in demand in extreme cold temperatures known to depress adaptive immunity (O'Neill 1989; Bly and Clem 1991). Given this potential function, the adaptive evolution of the notothenioid hepcidins could also be driven by the changing pathogenic challenges, which is widely attributed as the cause for adaptive evolution of vertebrate immunoglobulin variable gene segments (Ota et al. 2000).

Lineage-Specific Expansion of Hepcidin Genes in Marine Fishes

Human has a single hepcidin gene (Krause et al. 2000; Park et al. 2001; Hunter et al. 2002), whereas in mouse, zebrafish, and olive flounder, 2 copies were reported (Douglas et al. 2003; Ilyin et al. 2003; Shike et al. 2004; Kim et al. 2005). Characterization of the hepcidin genes in a number of marine fishes reveals multiple copies in winter flounder (Douglas et al. 2003), Japanese flounder (Hirono et al. 2005), Japanese sea bass (Ren et al. 2006), red sea bream (Chen et al. 2005), black sea bream (Yang et al. 2007), and others. Many hepcidin genes showed lineage- or species-specific clustering as shown in figure 5, indicating frequent species or lineage-specific expansion of the hepcidin genes in the teleost fishes. In the case of notothenioids, we have identified at least 4 hepcidin variants (8cys hepcidin_v1, v2, v3, and 4cys hepcidin), and each variant could be encoded by multiple distinct genes in the genome. From distinct hepcidin genomic clones we obtained from *D. mawsoni* and *N. angustata* by PCR amplification with variant-specific primers (primer pairs 1–3 in supplementary table 1 [Supplementary Material online]) and distinct hepcidin cDNAs from the transcriptomes of the 2 species (fig. 1*A*), we identified three 4cysHep, two 8cysHep_v2 and one 8cysHp_v1 genes in *D. mawsoni*, and three 4cys and five 8cys genes in *N. angustata*, (fig. 3). Thus, in total, at least 6 and 8 distinctive hepcidin genes are present in *D. mawsoni* and *N. angustata*, respectively. The gene number may even be larger on more exhaustive genomic cloning and sequencing.

An interesting question is why diversification of hepcidin genes occurred in many marine teleosts. A recent analysis showed that the hepcidins from teleost species of Perciformes and Pleuronectiformes are under positive Darwinian selection (Padhi and Verghese 2007), as we have detected the same in notothenioid hepcidins, but what factors drive the adaptive evolution remain as a matter of conjecture. We posit that the evolution of teleost hepcidin diversity relates to the dual environmental responding roles hepcidins play as iron regulator and antibacterial agent. In both mammals and teleost fishes, hepcidin gene expression is altered by inflammation (Frazer et al. 2004; Nemeth, Rivera, et al. 2004), anemia (Latunde-Dada et al. 2004), hypoxia (Nicolas et al. 2002; Leung et al. 2005), and iron concentration (Nemeth, Tuttle, et al. 2004). Fluctuating water temperatures, O_2 saturation and iron concentrations as well as pathogen diversity in different regions and depths of vast seas, in which marine fishes inhabit and may migrate between, conceivably could be driving forces for adaptive diversification of hepcidin genes. The diversity of hepcidins in a given teleost species may be a reflection of the complexity and challenges of environmental conditions the fish encounters and copes with during its life history.

The oceanographically and thermally isolated Antarctic marine environment since mid-Miocene time (Kennett and Thunell 1977; Shevenell et al. 2004) created the harshest body of water for ectothermic fish, characterized by chronic frigid temperatures (−1.9 °C to +2 °C) and abundant ice in the water column (DeVries and Cheng 2005). Antarctic notothenioid underwent an adaptive radiation

and now dominates the fish fauna in this challenging environment in terms of species and biomass (Eastman 2005). The diversification was accompanied by genetic changes, most visibly the gain of antifreeze proteins that prevent freezing and the paradoxical loss of respiratory hemoproteins and red blood cells in the derived channichthyid family (Cheng and Detrich 2007). Living in extreme cold and therefore oxygen-rich marine environment, a significant challenge is elevated reactive oxygen species (ROS) that may lead to unwarranted apoptosis when uncontrolled (Zager et al. 2000; Abele and Puntarulo 2004), and free cellular Fe^{2+} is a potent catalyst for ROS generation via the lipid peroxidation chain reactions (Barrier et al. 1998). In addition, hematocrits in Antarctic notothenioids are low (Wells et al. 1980) contributing to a reduction of blood viscosity in the cold, and the availability of Fe^{2+} is a critical regulator of the erythropoiesis. Therefore, the need for tight control of iron absorption and storage is of particular importance in the notothenioids. In fact, the complement of the hepcidin variants among the notothenioid fishes could vary, for example, the hemoglobinless icefish *C. aceratus* is missing the 4cys hepcidin gene, whereas *G. acuticeps* might not possess the 8cys hepcidin_v2. Such secondary loss may be genetically random events, or they may confer physiological adaptive values. In this regard, the dynamic evolution of hepcidin genes in the Antarctic notothenioid fishes provides a system to delineate the relationships between specific hepcidin variants and the particular biotic or abiotic factors to which the variants evolved to respond.

Supplementary Material

Supplementary table 1 and figures S1–S3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Dr Arthur DeVries for assistance in collecting the fish species used in this study. We thank Drs Yinqiu Wang, Yan Li, and Xin Li of Kunming Institute of Zoology, Chinese Academy of Sciences for assistance in the selection analyses of hepcidin genes. This work is supported in part by grants from the Natural Science Foundation of China (33030080) and the 973 Program of the Ministry of Science and Technology of China (2004CB117404) to L.C. and from the US National Science Foundation (grants OPP 0231006 and OPP 0636696) to C.-H.C.C. and A.D.

Literature Cited

Abele D, Puntarulo S. 2004. Formation of reactive species and induction of antioxidant defense systems in polar and temperate marine invertebrates and fish. Comp Biochem Physiol A Mol Integr Physiol. 138:405–415.

Aisen P, Enns C, Wessling-Resnick M. 2001. Chemistry and biology of eukaryotic iron metabolism. Int J Biochem Cell Biol. 33:940–959.

Barrier L, Page G, Fauconneau B, Juin F, Tallineau C. 1998. Autoxidation of rat brain homogenate: evidence for spontaneous lipid peroxidation. Comparison with the characteristics of Fe2+- and ascorbic acid-stimulated lipid peroxidation. Free Radic Res. 28:411–422.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: signalP 3.0. J Mol Biol. 340:783–795.

Bly JE, Clem LW. 1991. Temperature-mediated processes in teleost immunity: in vitro immunosuppression induced by in vivo low temperature in channel catfish. Vet Immunol Immunopathol. 28:365–377.

Chen L, DeVries AL, Cheng CH. 1997. Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. Proc Natl Acad Sci USA. 94:3811–3816.

Chen SL, Xu MY, Ji XS, Yu GC, Liu Y. 2005. Cloning, characterization, and expression analysis of hepcidin gene from red sea bream (*Chrysophrys major*). Antimicrob Agents Chemother. 49:1608–1612.

Chen ZZ, Xue CH, Zhu S, Zhou FF, Liu BL, Chen LB. 2005. GoPipe: streamlined gene ontology annotation for batch anonymous sequences with statistics. Prog Biochem Biophys. 32(2):187–190.

Cheng CH, Detrich HW. 2007. Molecular ecophysiology of Antarctic notothenioid fishes. Philos Trans R Soc Lond B Biol Sci. 362:2215–2232.

Cocca E, Ratnayake-Lecamwasam M, Parker SK, Camardella L, Ciaramella M, di Prisco G, Detrich HW 3rd. 1995. Genomic remnants of alpha-globin genes in the hemoglobinless antarctic icefishes. Proc Natl Acad Sci USA. 92:1817–1821.

Detrich HW 3rd, Parker SK, Williams RC Jr, Nogales E, Downing KH. 2000. Cold adaptation of microtubule assembly and dynamics. Structural interpretation of primary sequence changes present in the alpha- and beta-tubulins of Antarctic fishes. J Biol Chem. 275:37038–37047.

DeVries AL. 1971. Glycoproteins as biological antifreeze agents in antarctic fishes. Science. 172:1152–1155.

DeVries AL, Chen C-HC. 2005. Antifreeze proteins and organismal freezing avoidance in polar fishes. In: Farrell AP, Steffenson JF, editors. Fish physiology series, the physiology of polar fishes. Vol. 22. San Diego (CA): Elsevier Academic Press. p. 155–201.

Douglas SE, Gallant JW, Liebscher RS, Dacanay A, Tsoi SC. 2003. Identification and expression analysis of hepcidin-like antimicrobial peptides in bony fish. Dev Comp Immunol. 27:589–601.

Eastman JT. 2005. The nature of the diversity of Antarctic fishes. Polar Biol. 93–107.

Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8:175–185.

Fields PA. 2001. Review: protein function at thermal extremes: balancing stability and flexibility. Comp Biochem Physiol A Mol Integr Physiol. 129:417–431.

Fields PA, Somero GN. 1998. Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. Proc Natl Acad Sci USA. 95:11476–11481.

Fleming RE, Sly WS. 2001. Hepcidin: a putative iron-regulatory hormone relevant to hereditary hemochromatosis and the anemia of chronic disease. Proc Natl Acad Sci USA. 98:8160–8162.

Frazer DM, Wilkins SJ, Millard KN, McKie AT, Vulpe CD, Anderson GJ. 2004. Increased hepcidin expression and hypoferraemia associated with an acute phase response are not affected by inactivation of HFE. Br J Haematol. 126: 434–436.

Ganz T, Nemeth E. 2006. Regulation of iron acquisition and iron distribution in mammals. Biochim Biophys Acta. 1763: 690–699.

Graham MS, Fletcher GL, Haedrich RL. 1985. Blood viscosity in arctic fishes. J Exp Zool. 234:157–160.
Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98NT. Nucleic Acids Symp Ser. 41:95–98.
Hirono I, Hwang JY, Ono Y, Kurobe T, Ohira T, Nozaki R, Aoki T. 2005. Two different types of hepcidins from the Japanese flounder *Paralichthys olivaceus*. FEBS J. 272:5257–5264.
Huang X, Madan A. 1999. CAP3: a DNA sequence assembly program. Genome Res. 9:868–877.
Hunt BM, Hoefling K, Cheng C-HC. 2003. Annual warming episodes in seawater temperatures in McMurdo Sound in relationship to endogenous ice in notothenioid fish. Antarct sci. 15:333–338.
Hunter HN, Fulton DB, Ganz T, Vogel HJ. 2002. The solution structure of human hepcidin, a peptide hormone with antimicrobial activity that is involved in iron uptake and hereditary hemochromatosis. J Biol Chem. 277:37597–37603.
Ilyin G, Courselaud B, Troadec MB, Pigeon C, Alizadeh M, Leroyer P, Brissot P, Loreal O. 2003. Comparative analysis of mouse hepcidin 1 and 2 genes: evidence for different patterns of expression and co-inducibility during iron overload. FEBS Lett. 542:22–26.
Kennett JP, Thunell RC. 1977. On explosive cenozoic volcanism and climatic implications. Science. 196:1231–1234.
Kim YO, Hong S, Nam BH, Lee JH, Kim KK, Lee SJ. 2005. Molecular cloning and expression analysis of two hepcidin genes from olive flounder Paralichthys olivaceus. Biosci Biotechnol Biochem. 69:1411–1414.
Krause A, Neitz S, Magert HJ, Schulz A, Forssmann WG, Schulz-Knappe P, Adermann K. 2000. LEAP-1, a novel highly disulfide-bonded human peptide, exhibits antimicrobial activity. FEBS Lett. 480:147–150.
Latunde-Dada GO, Vulpe CD, Anderson GJ, Simpson RJ, McKie AT. 2004. Tissue-specific changes in iron metabolism genes in mice following phenylhydrazine-induced haemolysis. Biochim Biophys Acta. 1690:169–176.
Lauth X, Babon JJ, Stannard JA, Singh S, Nizet V, Carlberg JM, Ostland VE, Pennington MW, Norton RS, Westerman ME. 2005. Bass hepcidin synthesis, solution structure, antimicrobial activities and synergism, and in vivo hepatic response to bacterial infections. J Biol Chem. 280:9272–9282.
Leung PS, Srai SK, Mascarenhas M, Churchill LJ, Debnam ES. 2005. Increased duodenal iron uptake and transfer in a rat model of chronic hypoxia is accompanied by reduced hepcidin expression. Gut. 54:1391–1395.
Near TJ, Parker SK, Detrich HW 3rd. 2006. A genomic fossil reveals key steps in hemoglobin loss by the antarctic icefishes. Mol Biol Evol. 23(11):2008–2016.
Nemeth E, Preza GC, Jung CL, Kaplan J, Waring AJ, Ganz T. 2006. The N-terminus of hepcidin is essential for its interaction with ferroportin: structure-function study. Blood. 107:328–333.
Nemeth E, Rivera S, Gabayan V, Keller C, Taudorf S, Pedersen BK, Ganz T. 2004. IL-6 mediates hypoferremia of inflammation by inducing the synthesis of the iron regulatory hormone hepcidin. J Clin Invest. 113:1271–1276.
Nemeth E, Tuttle MS, Powelson J, Vaughn MB, Donovan A, Ward DM, Ganz T, Kaplan J. 2004. Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. Science. 306:2090–2093.
Nicolas G, Bennoun M, Devaux I, Beaumont C, Grandchamp B, Kahn A, Vaulont S. 2001. Lack of hepcidin gene expression and severe tissue iron overload in upstream stimulatory factor 2 (USF2) knockout mice. Proc Natl Acad Sci USA. 98:8780–8785.
Nicolas G, Chauvet C, Viatte L, Danan JL, Bigard X, Devaux I, Beaumont C, Kahn A, Vaulont S. 2002. The gene encoding the iron regulatory peptide hepcidin is regulated by anemia, hypoxia, and inflammation. J Clin Invest. 110:1037–1044.
O'Neill JG. 1989. Ontogeny of the lymphoid organs in an Antarctic teleost, *Harpagifer antarcticus* (Notothenioidei: Perciformes). Dev Comp Immunol. 13:25–33.
Ota T, Sitnikova T, Nei M. 2000. Evolution of vertebrate immunoglobulin variable gene segments. Curr Top Microbiol Immunol. 248:221–245.
Padhi A, Verghese B. 2007. Evidence for positive Darwinian selection on the hepcidin gene of Perciform and Pleuronectiform fishes. Mol Divers. doi:10.1007/s11030-007-9066-4.
Park CH, Valore EV, Waring AJ, 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, Turlin B, Brissot P, Loreal O. 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.
Posada D. 2006. ModelTest Server: a web-based tool for the statistical selection of models of nucleotide substitution online. Nucleic Acids Res. 34:W700–W703.
Ren HL, Wang KJ, Zhou HL, Yang M. 2006. Cloning and organisation analysis of a hepcidin-like gene and cDNA from Japan sea bass, *Lateolabrax japonicus*. Fish Shellfish Immunol. 21:221–227.
Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406–425.
Shevenell AE, Kennett JP, Lea DW. 2004. Middle Miocene Southern Ocean cooling and Antarctic cryosphere expansion. Science. 305:1766–1770.
Shi J, Camus AC. 2006. Hepcidins in amphibians and fishes: antimicrobial peptides or iron-regulatory hormones? Dev Comp Immunol. 30:746–755.
Shike H, Shimizu C, Lauth X, 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.
Somero GN. 2004. Adaptation of enzymes to temperature: searching for basic "strategies. Comp Biochem Physiol B Biochem Mol Biol. 139:321–333.
Swofford DL. 2002. PAUP* 4.0: phylogenetic analysis using parsimony (* and other methods). Sunderland (MA): Sinauer Associates.
Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876–4882.
Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.
Weinstein DA, Roy CN, Fleming MD, Loda MF, Wolfsdorf JI, Andrews NC. 2002. Inappropriate expression of hepcidin is associated with iron refractory anemia: implications for the anemia of chronic disease. Blood. 100:3776–3781.
Wells RMG, Ashby MD, Duncan SJ, McDonald JA. 1980. Comparative study of the erythrocytes and haemoglobins in nototheniid fishes from Antarctica. J Fish Biol. 17(5):517–527.

Wong WS, Nielsen R. 2004. Detecting selection in noncoding regions of nucleotide sequences. Genetics. 167:949–958.
Yang M, Wang KJ, Chen JH, Qu HD, Li SJ. 2007. Genomic organization and tissue-specific expression analysis of hepcidin-like genes from black porgy (*Acanthopagrus schlegelii* B). Fish Shellfish Immunol. 23:1060–1071.
Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci. 13:555–556.
Yang Z. 2002. Likelihood and Bayes estimation of ancestral population sizes in hominoids using data from multiple loci. Genetics. 162:1811–1823.
Yang Z, Nielsen R. 2000. Estimating synonymous and non-synonymous substitution rates under realistic evolutionary models. Mol Biol Evol. 17:32–43.
Yang Z, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol. 19:908–917.
Zager RA, Burkhart KM, Johnson A. 2000. Sphingomyelinase and membrane sphingomyelin content: determinants of proximal tubule cell susceptibility to injury. J Am Soc Nephrol. 11:894–902.
Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol. 22:2472–2479.

Jianzhi Zhang, Associate Editor

Accepted February 21, 2008