Molecular Cloning and Copy Number Variation of a Ferritin Subunit (Fth1) and Its Association with Growth in Freshwater Pearl Mussel *Hyriopsis cumingii*

Zhiyi Bai[1], Yiming Yuan[1], Genhua Yue[3], Jiale Li[1,2]*

1 Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources, Shanghai Ocean University, Ministry of Education, Shanghai Ocean University, Shanghai, China, 2 E-Institute of Shanghai Universities, Shanghai Ocean University, Shanghai, China, 3 Molecular Population Genetics Group, Temasek Life Sciences Laboratory, National University of Singapore, Singapore, Singapore

Abstract

Iron is one of the most important minor elements in the shells of bivalves. This study was designed to investigate the involvement of ferritin, the principal protein for iron storage, in shell growth. A novel ferritin subunit (Fth1) cDNA from the freshwater pearl mussel (*Hyriopsis cumingii*) was isolated and characterized. The complete cDNA contained 822 bp, with an open reading frame (ORF) of 525 bp, a 153 bp 5′ untranslated region (UTR) and a 144 bp 3′ UTR. The complete genomic DNA was 4125 bp, containing four exons and three introns. The ORF encoded a protein of 174 amino acids without a signal sequence. The deduced ferritin contained a highly conserved motif for the ferroxidase center comprising seven residues of a typical vertebrate heavy-chain ferritin. It contained one conserved iron associated residue (Try27) and iron-binding region signature 1 residues. The mRNA contained a 27 bp iron-responsive element with a typical stem-loop structure in the 5′-UTR position. Copy number variants (CNVs) of Fth1 in two populations (PY and JH) were detected using quantitative real-time PCR. Associations between CNVs and growth were also analyzed. The results showed that the copy number of the ferritin gene of in the diploid genome ranged from two to 12 in PY, and from two to six in JH. The copy number variation in PY was higher than that in JH. In terms of shell length, mussels with four copies of the ferritin gene grew faster than those with three copies ($P<0.05$), suggesting that CNVs in the ferritin gene are associated with growth in shell length and might be a useful molecular marker in selective breeding of *H. cumingii*.

Introduction

As a cofactor in many biochemical reactions, the trace element iron is important to all living organisms through its role in regulating metabolism, electron transport, oxidative phosphorylation and DNA biosynthesis [1]. However, excess labile iron is detrimental due to its participation in oxidation-reduction reactions that generate harmful free radicals. Therefore, iron homeostasis is important in living cells [1,2]. Ferritin, a major iron storage protein in living cells, plays a crucial role in intracellular iron homeostasis. Ferritin has 24 subunits, which form a hollow shell and can store up to 4,500 iron (III) atoms [3]. There are two types of ferritin subunit: the heavy chain (H) and the light chain (L). H-ferritin subunits occur in multiple forms in animals, plants and bacteria, while L-ferritin subunits are usually restricted to vertebrates [4,5]. Ferritin has been shown to be involved in many physiological activities such as development [6,7,8], immunity [9,10] and other cellular mechanisms [11,12,13].

Among bivalves, ferritin genes from the pacific oyster (*Crassostrea gigas*), pearl oyster (*Pinctada fucata*), bay scallop (*Argopecten irradians*) and clam (*Meretrix meretrix*) have been characterized [14,15,16,17].

In *M. meretrix*, ferritin is reported to play a role in larval shell development [16]. In *P. fucata*, studies suggest that ferritin is involved in adult shell formation [17]. However, understanding of ferritin gene sequence variation and the biological function of ferritins is limited to freshwater bivalve species. *Hyriopsis cumingii* is a freshwater bivalve that is widely distributed in China and is commercially the most important mussel species exploited for freshwater pearl production in this country [18]. In this study, the full length cDNA and genomic DNA of one ferritin H subunit (Fth1) of *H. cumingii* were isolated and characterized.

Genetic variation takes many forms, ranging from large, microscopically visible chromosome anomalies to single nucleotide changes. Copy number variants (CNVs) represent segments of DNA larger than 1 kb present at in variable copy numbers in comparison with a reference genome [19]. They can be responsible for altered gene expression [20] leading to striking phenotypic variance, including disease associated traits [21,22]. CNVs have been reported to be present genome-wide not only in humans [23], but also in chimpanzees [24], mice [25], nematodes [26], fruit flies [27] and pigs [28]. However, no study of the extent and impact of CNVs in the molluscan genome has been made

until now. We investigated whether the copy number of the ferritin gene differed between a wild population and a cultured population, and whether a functional relationship existed between copy number and growth in *H. cumingii*. To our knowledge, this is the first study of gene copy number variation in freshwater bivalves.

Results

cDNA and genomic organization of ferritin gene in *H. cumingii*

An 822 bp cDNA of a ferritin subunit (Fth1), containing a 525 bp open reading frame (ORF), a 153 bp 5′ UTR and a 144 bp 3′ UTR, was isolated from *H. cumingii* (Fig. 1). A putative polyadenylation signal (ATTAAA) was recognized at the nucleotide position 708–713 (Fig. 1). The translated amino acid sequence comprised 174 residues with a calculated molecular mass of 20.19 kDa and an isoelectric point of 5.08. SignalP software analysis found no signal peptide. The genomic structure of the ferritin gene was determined by sequencing the genomic fragment amplified by PCR. From the first transcription initiation site, the *H. cumingii* g-ferritin gene extended 4125 bp to the end of the 3′ UTR and contained four exons and three introns. All of the 5′ and 3′ ends of the introns showed canonical splicing motifs (GT/intron/AG).

An iron-responsive element (IRE) and a stem-loop structure (SLS) in the ferritin subunit of *H. cumingii*

An SLS (5′-ctttgctgcgtcagtgaacgtacgagc-3′) in the 5′ UTR of ferritin subunit cDNA of *H. cumingii* was recognized as an iron-responsive element (or iron regulatory element, IRE) (Fig. 1). IRE

alignment showed a high identity between *H. cumingii*, other species of mollusk and chicken, with only three or four bases different in their flank sequences (Fig. 2A). The most conserved motif in the IRE, 5′-CAGTGA-3′, is in good agreement between species, with the exception of humans (Fig. 2A). This putative IRE could be folded into a typical secondary SLS (Fig. 2B) matching all characteristics of IREs, including the six nucleotide loop 5′-CAGUGA-3″, the proximal stem of five paired bases followed by a bulged cysteine, and the six nucleotide bottom stem.

Homology comparison and phylogenetic analysis of ferritin in mollusks

Homologous sequence alignment using ClustalW revealed uniformly high similarities between Fth1 of *H. cumingii* and ferritin of other species, ranging from 64% (*H.sapiens*) to 81% (*M.gallopro-vincialis*) (Table 1). Homologous sequence alignment also indicated seven amino acid residues, identified as metal ligands at the H-specific ferroxidase center in mammalian ferritins, that were completely conserved in the ferritin subunit. They were Glu25, Tyr32, Glu59, Glu60, His63, Glu105 and Gln139. A potential biomineralization residue, Tyr27, was also conserved (Fig. 3). Phylogenetic analysis based on ferritin subunit amino acid sequences showed *H. cumingii* to be distant from seawater bivalves (Fig. 4).

Copy number variation of ferritin in wild and cultured populations

Extensive variations in the copy number of the ferritin gene were observed in both wild (PY) and cultured (JH) populations. There were more copy number variations in the wild population than in the cultured population, ranging from two to 12 and two to

```
  1                                 ggacttgaaacggtataaaagcagcgtctggac
 34   attttcatcatttctggtgttatttgctttgctgcgtcagtgaacgtacgagcaaataaa
 94   ccttccaacttaaaacgttttgataaaaacaaatcaggttgattaattaccagtgaaaga
154   atggcacagacaagaccacgtcagaactaccatgccgaaaatgaagccggtatcaacaag
       M   A   Q   T   R   P   R   Q   N   Y   H   A   E   N   E   A   G   I   N   K        20
214   cagataaacatggagctttatgccagctatgtctaccagtccatgtcttactactttgac
       Q   I   N   M   E   L   Y   A   S   Y   V   Y   Q   S   M   S   Y   Y   F   D        40
274   cgggatgatgtggccctgaaaggatttgccaaattttttcaagaagtcttctgaggaggaa
       R   D   D   V   A   L   K   G   F   A   K   F   F   K   K   S   S   E   E   E        60
334   agagagcatgcagagaagttcatgaagtatcagaacaagcgtggtggccgcatcgtcctc
       R   E   H   A   E   K   F   M   K   Y   Q   N   K   R   G   G   R   I   V   L        80
394   caggacatcaagaagcctgaccgggatgagtggggcactgggctggatgcaatggaagtg
       Q   D   I   K   K   P   D   R   D   E   W   G   T   G   L   D   A   M   E   V       100
454   gctcttgatcttgagaagagtgtgaatcaggcccttctcgaccttcacaaggtggcagat
       A   L   D   L   E   K   S   V   N   Q   A   L   L   D   L   H   K   V   A   D       120
514   tctcatggtgacccccagatgatggacttcatcgagcacgagttcctagaagaacaggtg
       S   H   G   D   P   Q   M   M   D   F   I   E   H   E   F   L   E   E   Q   V       140
574   agttccattaaggagctctgtaaccacattactaacctcaagcgtgttggacctggactt
       S   S   I   K   E   L   C   N   H   I   T   N   L   K   R   V   G   P   G   L       160
634   ggggaatacatgtttgacaaggagaccctggatggtgacagctaaactggggcattgata
       G   E   Y   M   F   D   K   E   T   L   D   G   D   S   *                          174
694   caagatccagcccctaataaattggactttctgtctatctttagaggctcttgttcaaag
754   tcatagggctgaagaaaactgtttttaacagtattcaaagaaaaaattaataaaaaaatgg
814   aaacttgagaaaaaaaaaaaaaaaaaaaaaaaaaa
```

Figure 1. Complete cDNA sequence and deduced amino acid sequence of *H. cumingii* ferritin. The IRE in the 5′ UTR is shaded. Seven residues in bold and italics represent a putative active site of ferroxidase. The iron associated residue Tyr27 is in bold and underlined. Iron-binding region signature 1 is in bold. The star represents the stop codon. The poly(A) signal is underlined.
doi:10.1371/journal.pone.0022886.g001

A

```
H.cumingii_fern     1  CTTTGCTGCGTCAGTGAACGTACGAGC
Pinctada_fern       1  GCTTGCTGCGTCAGTGAACGTACGGGC
Abalone_fern_2__Abf2_ 1  TCTTGCTGCGTCAGTGAACGTACGGGC
Oyster_GF2          1  TTTTGCTGCGTCAGTGAACGTACGGAC
Oyster_GF1          1  TCTTGCTGCGTCAGTGAACGTACGGAC
Lymneas_fern        1  TCTTGCTGCGTCAGTGAACGTACGGAC
Rana_fern           1  TCTTGCTGCGTCAGTGAACGTACGGAC
Chicken_fern        1  TCTTGCTGCGTCAGTGAACGTACGGAC
Human_HF            1  TCCTGCTTCAACAGTGCTTGGACGGAA
consensus           1  ...****.*..*****...*.***...
```

B

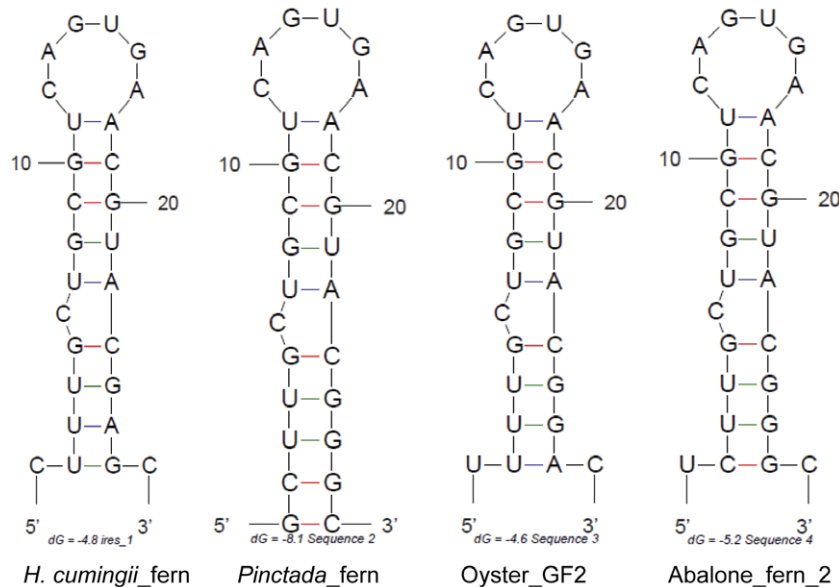H. cumingii_fern   Pinctada_fern   Oyster_GF2   Abalone_fern_2

Figure 2. Alignments of IRE and SLS of *H. cumingii* with those of other species. (A) Alignment of the IRE with selected ferritin IREs. Consensus residues are shaded and identical residues are in bold. Freshwater pearl mussel (*H. cumingii*), Fth1; pearl oyster (*P. fucata*), AAQ12076.1; abalone (*Haliotis discus hannai*), DQ821494; pacific oyster (*C. gigas*), AY321299 and AY321300; bullfrog (*Rana catesbeiana*), M12120; great pond snail (*Lymnaea stagnalis*), X56778; chicken (*Gallus gallus*), NM205086; and human (*Homo sapiens*), NM002032. (B) A comparison of the IRE SLS of *H. cumingii* and known ferritins of *P. fucata* (AAQ12076.1), *C. gigas* (AY321300) and *H. discus* (DQ821494). Typical CAGUGA loop structure residues and bulged cysteines are seen in all four species.
doi:10.1371/journal.pone.0022886.g002

six respectively. In the wild population, seven, eight, nine, and 10 CNVs were the principal genotypes (representing 23.3%, 16.7%, 20.0% and 15.0% respectively); other CNVs were less common. In the cultured population, three and four CNVs were the principal genotypes (40.0% and 49.0% respectively); again, other CNVs were less common (Fig. 5).

Analysis of association between CNVs and growth

The growth of individuals from the JH population was recorded before genotyping of the ferritin gene. Differences between individuals with three and those with four CNVs were analyzed. The t-test revealed that, in terms of shell length, mussels with four copies of ferritin grew faster than those with three copies ($P<0.05$), but no significant differences were observed in shell width and body weight (Table 2).

Discussion

In this study, we cloned one ferritin gene in an important freshwater pearl mussel species, *H. cumingii*. According to the bioinformatics analysis, this ferritin is likely to be a homolog of vertebrate H-ferritin. This is further supported by the finding of a putative ferroxidase site in the deduced peptide, where seven amino acid residues were located in the same positions as in vertebrate H-ferritin [3]. The Tyr27 in ferritin is important for iron binding [29], enabling iron biomineralization. An IRE was identified, which is important for the regulation of ferritin translation [5]. When cellular iron is low, the iron regulatory protein (IRP) can repress the expression of ferritin by binding to the SLS of IRE. When cellular iron is high, IRP unbinds from the IRE and ferritin mRNA translates ferritin to store extra iron [3,30]. The IRE and IRP have been greatly conserved during evolution [31]. The IRE of *H. cumingii* ferritin cDNA is homologous with those of other species (Fig. 2). In this case, the ferritin might be regulated on a mainly translational level.

Iron is an important trace element in the prismatic and nacreous layers of bivalve shells [32,33]. The mantle epithelial cells of bivalves have been shown to accumulate iron and other elements in lysosomes, which participate actively in the incorporation of metals into the shell [34,35]. Calcified concretions containing calcium, iron and other metals have been described in the mantles of *Hyridella depressa* and *Margaritifera margaritifera*, which were proposed to accumulate metals and contribute to shell formation [36,37,38]. To further understand the role of ferritin in

Table 1. Pairwise similarities between Fth1 of *H. cumingii* and ferritin genes of other species as specified.

Ferritin genes of other Species	Similarities to Fth1 of *H. cumingii*
C. gigas, AAP83794.1;	77
C. ariakensis, ABE99842.1	77
P. fucata, AAQ12076.1	77
P. maxima, ACS72281.1	78
C. farreri, AAV66906.1	77
A. irradians, AAV66907.1	76
M. galloprovincialis, ACM86786.1	81
S. constricta, ACZ65230.1	77
V. decussatus, ACB38006.1	78
M. meretrix, AAZ20754.1	74
H. rufescens, ACZ73270.1	75
H. discus hannai, ABH10672.1	74
H. diversicolor, ABY87353.1	75
L. stagnalis, P42577.2	80
A. californica, ABF21074.1	76
E. dofleini, AAD29639.1	70
H. sapiens, AAA52437.1	64

shell formation by *H. cumingii*, the relationship between genotype and phenotype is an effective investigation. CNVs are present genome-wide and usually associated with phenotypic variance [39]. In humans, the ferritin gene copy number varies widely. In this study, ferritin CNVs and their association with the growth of *H. cumingii* were determined.

Recently, real-time quantitative PCR has become an attractive method for gene copy number measurements, being fast, precise and reproducible and having a high throughput [40,41,42,43]. In contrast, traditional methods for measuring DNA copy numbers, such as fluorescence in situ hybridization, are difficult to perform with a high throughput and fail to detect small deletions/duplications [44,45]. Southern blotting and chromosomal comparative genomal hybridization require relatively large amounts of genomic DNA, and Southern blotting is time consuming and labor intensive [44,45]. Multiplex amplification and probe hybridization (MAPH) and multiplex ligation dependent probe amplification are promising new methods to detect dosage changes. However, neither is able to detect translocations in balanced form [45]. Furthermore, MAPH can detect a large relative difference in copy number, such as one copy instead of two, but it is not sufficiently sensitive to discriminate between high copy numbers [46]. In this study, a method of real-time quantitative PCR to detect copy number polymorphisms in *H. cumingii* ferritin was established.

Compared with other real-time quantitative PCR methods reported previously, the method introduced here was improved by a minor modification. To estimate copy number using a comparative cycle threshold (CT) method, it is necessary to identify a single-copy gene as the control. However, only a small number of genomic and cDNA sequences are known for *H. cumingii*, and no single-copy gene has yet been identified. Justen et al. developed an efficient method for amplifying single-copy nuclear loci by designing a primer anchored in the 3′ UTR [47]. In this study, we designed primers in the 3′ UTR to amplify single-copy nuclear loci as control genes, and in the conserved region of exon2 for target genes. This method will provide a powerful tool for screening copy number polymorphisms in non-model species.

Using this method, the CNVs of ferritin in wild and cultured populations were determined. To our knowledge, these are the first genotype data obtained for CNVs in freshwater bivalves. Our finding that there were more copy number variations in the wild population than in the cultured population is in agreement with other results obtained by simple sequence repeats and other markers in our laboratory [18,48]. In the wild population, seven, eight, nine, and 10 copy number variations were the principal genotypes, which suggested the mussels with higher copy number in ferritin would be easier to survive in lake, and implied that ferritin dosage may convey a growth or resistance advantage. In the cultured population, three and four copy number variations were the principal genotypes. It is most possible that the cultured population constructed on several parent mussels with lower copy number in ferritin. One female mussel can release at least 2–4 hundred thousand of offsprings, so offsprings reproduced by several mussels are sufficient for the production demand of one pearl mussel farm. In this study, preliminary analysis of the association between CNVs and growth suggested that ferritin may be involved shell formation, which may be improved by a greater ferritin copy number. However, this is only the first step towards an understanding of the functions of the ferritin gene in shell formation. Further studies will help to discover the role of the ferritin copy number polymorphisms.

In conclusion, we cloned a ferritin gene (Fth1) in an important freshwater pearl mussel species *H. cumingii*, developed an approach to estimate ferritin gene CNVs and detected a potential association between CNVs and growth, for the first time in a freshwater bivalve. Further studies on the mechanisms underlying the association between growth and ferritin CNVs are required.

Materials and Methods

Ethics Statement

Collecting mussels of wild stock from Duchang County of Jiangxi Province was permitted by Department of Fishery of Poyang Lake (DFPL2008WAC11). All handling of mussels was conducted in accordance with the guidelines on the care and use of animals for scientific purposes set up by the Institutional Animal Care and Use Committee (IACUC) of Shanghai Ocean University, Shanghai, China.

Samples and DNA extraction

Adult *H. cumingii* were randomly collected from wild and cultured stocks in 2008. The wild stock was collected from Duchang County of Jiangxi Province, around Poyang Lake (29.2°N 116.2°E). The cultured stock was Jinhua (JH) cultured stock from the Weiwang pearl farm of Jinhua City, Zhejiang Province. Sixty adults were sampled randomly from the wild stock. From the cultured stock, 100 3-year-old individuals that had been produced at the same time and cultured in the same pond were sampled. Shell length, shell width and body weight were recorded for each individual from the Jinhua stock. Mantle samples from each individual were stored in 95% ethanol until DNA extraction. Genomic DNA was isolated from the tissues using a Genomic DNA Extraction Kit (Innogent, Shenzhen, China) according to the manufacturer's instructions. Compared with other DNA extraction kits, this kit was recommended for the acquisition of a high quality DNA template, which was important for the reliability of this experiment. The concentration of the extracted DNA was estimated spectrophotometrically. The DNA was diluted in
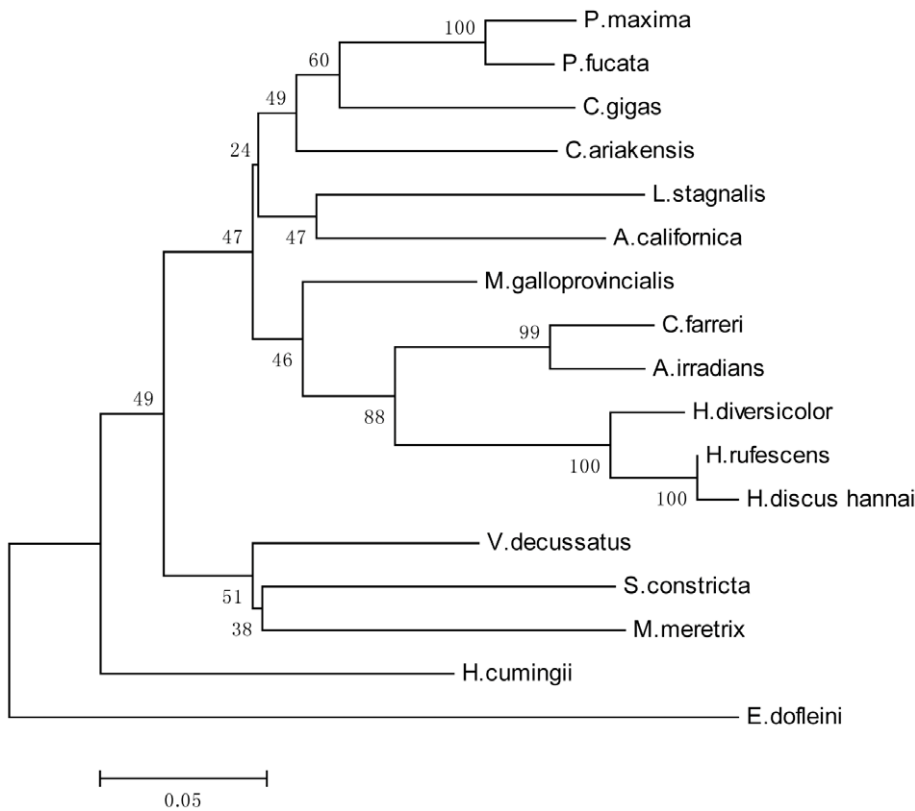
Figure 3. Multiple alignments of *H. cumingii* ferritin with homolog from other mollusk species and human. Identical amino acid residues are darkly shaded, similar amino acids are lightly shaded, unrelated residues have a white background. *C. gigas*, AAP83794.1; *Crassostrea ariakensis*, ABE99842.1; *P. fucata*, AAQ12076.1; *Pinctada maxima*, ACS72281.1; *Chlamys farreri*, AAV66906.1; *A. irradians*, AAV66907.1; *Mytilus galloprovincialis*, ACM86786.1; *Sinonovacula constricta*, ACZ65230.1; *Venerupis decussatus*, ACB38006.1; *M. meretrix*, AAZ20754.1; *Haliotis rufescens*, ACZ73270.1; *H. discus hannai*, ABH10672.1; *H. diversicolor*, ABY87353.1; *Lymnaea stagnalis*, P42577.2; *Aplysia californica*, ABF21074.1; *Enteroctopus dofleini*, AAD29639.1; *Homo sapiens*, AAA52437.1.
doi:10.1371/journal.pone.0022886.g003

10 mM pH 8.5 Tris buffer to a concentration of approximately 50 ng/µL and stored at −20°C until use.

Isolation of full length cDNA and genomic DNA

One expressed sequence tag (EST) of ferritin (GW694187) was originally identified from the mantle cDNA library constructed in our laboratory [49]. The full length cDNA of the ferritin was identified by further sequencing of the original EST clone using a primer walking approach. The complete genomic DNA sequence of the ferritin was amplified using PCR with primers (gFER-F1 and gFER-R1, Table 3) by long distance PCR, and sequenced using Bigdye chemicals (Applied Biosystems). Exons and introns were identified by aligning the cDNA and genomic DNA sequences of ferritin using Sequencher (Gene Codes). The full

Figure 4. Neighbor-joining phylogenetic tree of ferritins from mollusk species. *C. gigas*, AAP83794.1; *C. ariakensis*, ABE99842.1; *P. fucata*, AAQ12076.1; *P. maxima*, ACS72281.1; *C. farreri*, AAV66906.1; *A. irradians*, AAV66907.1; *M. galloprovincialis*, ACM86786.1; *S. constricta*, ACZ65230.1; *V.s decussatus*, ACB38006.1; *M. meretrix*, AAZ20754.1; *H. rufescens*, ACZ73270.1; *H. discus hannai*, ABH10672.1; *H. diversicolor*, ABY87353.1; *L. stagnalis*, P42577.2; *A. californica*, ABF21074.1; *E. dofleini*, AAD29639.1; *E. dofleini*, AAD29639.1.
doi:10.1371/journal.pone.0022886.g004

length cDNA sequence and genomic DNA of ferritin were submitted to GenBank with accession numbers HQ896721 and HQ896722, respectively.

Sequence analysis of full length cDNA

The alignment search tool BLASTn, from the National Center for Biotechnology Information was used to search homologous sequences in GenBank. The cDNA of ferritin was translated into its potential ORF by the ORF Finder algorithm (http://www.ncbi.nlm.nih.gov/gorf/). The putative amino acid sequences were

analyzed for the presence of signal peptides using SignalP software (http://www.cbs.dtu.dk/services/signalP/) [50]. Domain analyses were undertaken using several resources, including a Simple Modular Architecture Research Tool (http://smart.gembl-heidelberg.de/) [51], Pfam 20.0 (http://pfam.wustl.edu/) [52] and ScanProsite (http://www.expasy.org/tools/scanprosite/) [53]. Multiple sequence alignments were performed using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html) and BOXSHADE (http://www.ch.embnet.org/software/BOX_form.html). SLS prediction was performed using the DINAMelt Web Server (Prediction of Melting
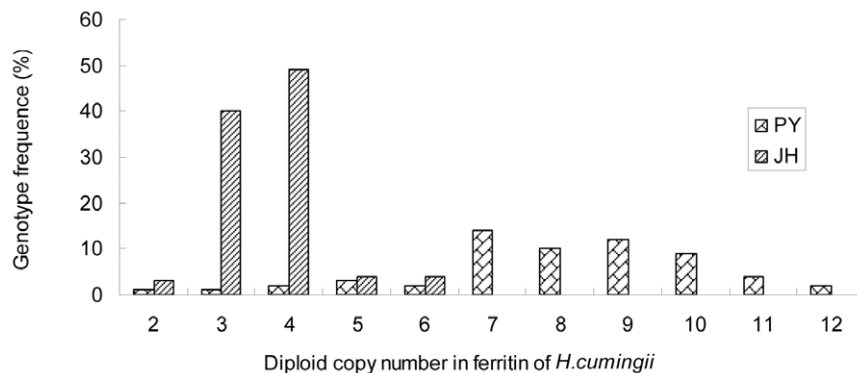

Figure 5. Genotype frequencies of diploid copy number in ferritin gene in wild (PY) and cultured (JH) populations of *H. cumingii*.
doi:10.1371/journal.pone.0022886.g005

Table 2. Growth *H. cumingii* with different ferritin copy numbers (mean ± SE).

Copy Number	Individuals	Shell Length	Shell Width	Weight
3	40	133.37±1.29[a]	33.23±1.35[a]	256.85±15.66[a]
4	49	136.57±1.30[b]	32.98±0.74[a]	267.39±9.97[a]

Same superscript represents no obvious difference; different superscript represents obvious difference ($P<0.05$).
doi:10.1371/journal.pone.0022886.t002

Table 3. PCR primers used for amplification of genomic DNA and estimation of Fth1 copy numbers in *H. cumingii*.

Primers	(Primer sequence) (5′–3′)
gFER-F	GCTTTGCTGCGTCAGTGAAC
gFER-R	TTAGGGCTGGATCTTGTATCAATG
CNV-F	ATGATGTGGCCCTGAAAGGATTTG
CNV-R	CGGCCACCACGCTTGTTCTG
3′ UTR F	CTCAAGCGTGTTGGACCTGGACT
3′ UTR R	CAGCCCTATGACTTTTGAACAAGA

doi:10.1371/journal.pone.0022886.t003

Profiles for Nucleic Acids, http://mfold.rna.albany.edu/?q = DINA-Melt/Quickfold) using default RNA3.0 parameters [54].

The phylogenetic tree was constructed based on the deduced full length amino acid sequences using the neighbor-joining algorithm in MEGA 4.0 [55], and the reliability of the analysis was assessed with 1000 bootstrap replicates.

Copy number estimation

Two pairs of primers were designed according to the ferritin cDNA sequence using the EditSeq program in DNA Star 7 (DNASTAR, Inc., Madison, WI). One pair of primers was anchored in the 3′ UTR as a control (diploid copy numbers of the locus = 2) and the other pair was anchored in the conserved coding region of exon 2 (Table 3). Optimized primer pairs were selected based on their amplification efficiencies and specificities [56]. Amplification efficiencies were calculated based on the generation of standard curves using genomic DNA dilution series according to the equation: $E = (10^{-1/\text{slope}} - 1)100$ [57].

The amplification efficiencies were 97% and 98% for the 3′ UTR and CNV primer pairs respectively. A correction for different amplification efficiencies was introduced in the sample quantification process [58]. Melting curve analysis was used to check the specificities of PCR reactions.

Real-time PCR was performed with Bio-Rad IQ$^{\text{TM}}$5 Multicolor Real-Time PCR Detection System in a 25 μL reaction system containing the following components: 4 μL DNA sample, 12.5 μL 2×SYBR Green Supermix (Bio-Rad), 0.5 μL of each primer and 7.5 μL ddH$_2$O. Briefly, following denaturation at 95°C for 10 min, real-time PCR was performed, with 40 cycles at a melting temperature of 95°C for 15 s, an annealing temperature of 60°C for 30 s and an extension temperature of 72°C for 30 s. Finally, a melting curve analysis was undertaken. All samples were amplified in triplicate from the same DNA preparation and the mean value was determined. The ferritin copy numbers were calculated using the comparative CT method according to the equation: diploid copy number = $2*2^{(-(\text{Ct}_{\text{Target gene}} - \text{Ct}_{\text{Control gene}}))}$. Cut-off values of 0.3 and 0.7 were used to define the copy numbers.

Analysis of association between CNVs and growth

The statistical analysis was performed using SPSS 17.0 (SPSS, Inc., Chicago, IL). In the cultured population, three and four CNVs were the principal genotypes; other CNVs were less common (Fig. 5). Differences in growth (shell length, shell height and body weight) between individuals with three and four CNVs were assessed using the t-test. P-values were considered statistically significant if <0.05.

Author Contributions

Conceived and designed the experiments: ZB YY GY JL. Performed the experiments: ZB YY GY JL. Analyzed the data: ZB YY GY JL. Contributed reagents/materials/analysis tools: ZB YY GY JL. Wrote the paper: ZB YY GY JL.

References

1. Andrews NC (2008) Forging a field: the golden age of iron biology. Blood 112: 219–230.
2. Sargent PJ, Farnaud S, Evans RW (2005) Structure/function overview of proteins involved in iron storage and transport. Curr Med Chem 12: 2683–2693.
3. Harrison PM, Arosio P (1996) The ferritins: molecular properties, iron storage function and cellular regulation. Biochim Biophys Acta 1275: 161–203.
4. Kim SR, Lee KS, Yoon HJ, Park NS, Lee SM, et al. (2004) Molecular cloning, expression and characterization of cDNAs encoding the ferritin subunits from the beetle, *Apriona germari*. Comp Biochem Physiol B Biochem Mol Biol 138: 423–433.
5. Theil EC (2003) Ferritin: at the crossroads of iron and oxygen metabolism. J Nutr 133: 1549S–1553S.
6. Chen T, Amons R, Clegg JS, Warner AH, MacRae TH (2003) Molecular characterization of artemin and ferritin from *Artemia franciscana*. Eur J Biochem 270: 137–145.
7. Levenson CW, Fitch CA (2000) Effect of altered thyroid hormone status on rat brain ferritin H and ferritin L mRNA during postnatal development. Brain Res Dev Brain Res 119: 105–109.
8. Missirlis F, Kosmidis S, Brody T, Mavrakis M, Holmberg S, et al. (2007) Homeostatic mechanisms for iron storage revealed by genetic manipulations and live imaging of Drosophila ferritin. Genetics 177: 89–100.
9. Li M, Saren G, Zhang S (2008) Identification and expression of a ferritin homolog in amphioxus *Branchiostoma belcheri*: evidence for its dual role in immune response and iron metabolism. Comp Biochem Physiol B Biochem Mol Biol 150: 263–270.
10. Zhang J, Li F, Wang Z, Zhang X, Zhou Q, et al. (2006) Cloning, expression and identification of ferritin from Chinese shrimp, *Fenneropenaeus chinensis*. J Biotechnol 125: 173–184.
11. Chevion M, Leibowitz S, Aye NN, Novogrodsky O, Singer A, et al. (2008) Heart protection by ischemic preconditioning: a novel pathway initiated by iron and mediated by ferritin. J Mol Cell Cardiol 45: 839–845.
12. Larade K, Storey KB (2004) Accumulation and translation of ferritin heavy chain transcripts following anoxia exposure in a marine invertebrate. J Exp Biol 207: 1353–1360.
13. Yamashita M, Ojima N, Sakamoto T (1996) Molecular cloning and cold-inducible gene expression of ferritin H subunit isoforms in rainbow trout cells. J Biol Chem 271: 26908–26913.
14. Durand JP, Goudard F, Pieri J, Escoubas JM, Schreiber N, et al. (2004) *Crassostrea gigas* ferritin: cDNA sequence analysis for two heavy chain type subunits and protein purification. Gene 338: 187–195.
15. He X, Zhang Y, Wu X, Xiao S, Yu Z (2010) Cloning and characterization of two ferritin subunit genes from bay scallop, *Argopecten irradians* (Lamarck 1819). Mol Biol Rep: (doi:10.1007/s11033-11010-10339-11030).
16. Wang X, Liu B, Xiang J (2009) Cloning, characterization and expression of ferritin subunit from clam *Meretrix meretrix* in different larval stages. Comp Biochem Physiol B Biochem Mol Biol 154: 12–16.
17. Zhang Y, Meng Q, Jiang T, Wang H, Xie L, et al. (2003) A novel ferritin subunit involved in shell formation from the pearl oyster (*Pinctada fucata*). Comp Biochem Physiol B Biochem Mol Biol 135: 43–54.
18. Li JL, Wang GL, Bai ZY (2009) Genetic diversity of freshwater pearl mussel (*Hyriopsis cumingii*) in populations from the five largest lakes in China revealed by inter-simple sequence repeat (ISSR). Aquaculture International 17: 323–330.
19. Feuk L, Carson AR, Scherer SW (2006) Structural variation in the human genome. Nat Rev Genet 7: 85–97.

20. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, et al. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315: 848–853.

21. Lupski JR (2007) Genomic rearrangements and sporadic disease. Nat Genet 39: S43–47.

22. Lupski JR, Stankiewicz P (2005) Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet 1: e49.

23. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, et al. (2004) Large-scale copy number polymorphism in the human genome. Science 305: 525–528.

24. Perry GH, Yang F, Marques-Bonet T, Murphy C, Fitzgerald T, et al. (2008) Copy number variation and evolution in humans and chimpanzees. Genome Res 18: 1698–1710.

25. Li J, Jiang T, Mao JH, Balmain A, Peterson L, et al. (2004) Genomic segmental polymorphisms in inbred mouse strains. Nat Genet 36: 952–954.

26. Maydan JS, Flibotte S, Edgley ML, Lau J, Selzer RR, et al. (2007) Efficient high-resolution deletion discovery in *Caenorhabditis elegans* by array comparative genomic hybridization. Genome Res 17: 337–347.

27. Dopman EB, Hartl DL (2007) A portrait of copy-number polymorphism in *Drosophila melanogaster*. Proc Natl Acad Sci U S A 104: 19920–19925.

28. Fadista J, Nygaard M, Holm LE, Thomsen B, Bendixen C (2008) A snapshot of CNVs in the pig genome. PLoS One 3: e3916.

29. Waldo GS, Theil EC (1993) Formation of iron(III)-tyrosinate is the fastest reaction observed in ferritin. Biochemistry 32: 13262–13269.

30. Bomford AB, Munro HN (1992) Ferritin gene expression in health and malignancy. Pathobiology 60: 10–18.

31. Rothenberger S, Mullner EW, Kuhn LC (1990) The mRNA-binding protein which controls ferritin and transferrin receptor expression is conserved during evolution. Nucleic Acids Res 18: 1175–1179.

32. Almeida MJ, Machado J, Moura G, Azevedo M, Coimbra J (1998) Temporal and local variations in biochemical composition of *Crassostrea gigas* shells. Journal of Sea Research 40: 233–249.

33. Narasimhulu KV, Rao JL (2000) EPR and IR spectral studies of the sea water mussel *Mytilus conradinus* shells. Spectrochim Acta A Mol Biomol Spectrosc 56A: 1345–1353.

34. Fowler BA, Wolte DA, Hettler WF (1975) Mercury and iron uptake by cytosomes in mantle epithelial cells of quahog clams (*Mercenaria mercenaria*) exposed to mercury. J Fish Res Bd Can 32: 1767–1775.

35. Thompson JD, Pirie BJS, George SG (1985) Cellular metal distribution in the pacific oyster *Crassostrea gigas* determined by quantitative X-ray microprobe analysis. J Exp Mar Biol Ecol 85: 37–46.

36. Adams SM, Shorey CD, Byrne M (1997) An ultrastructural and microanalytical study of metal-ion content in granular concretions of the freshwater mussel *Hyridella depressa*. Micron 28: 1–11.

37. Marigomez I, Soto M, Cajaraville MP, Angulo E, Giamberini L (2002) Cellular and subcellular distribution of metals in molluscs. Microsc Res Tech 56: 358–392.

38. Pekkarinen M, Valovirta I (1997) Histochemical and X-ray studies on tissue concretions and shells of *Margaritifera margaritifera* (Linnaeus). J Shellfish Res 16: 169–177.

39. Itsara A, Cooper GM, Baker C, Girirajan S, Li J, et al. (2009) Population analysis of large copy number variants and hotspots of human genetic disease. Am J Hum Genet 84: 148–161.

40. Aarskog NK, Vedeler CA (2000) Real-time quantitative polymerase chain reaction. A new method that detects both the peripheral myelin protein 22 duplication in Charcot-Marie-Tooth type 1A disease and the peripheral myelin protein 22 deletion in hereditary neuropathy with liability to pressure palsies. Hum Genet 107: 494–498.

41. Kindich R, Florl AR, Jung V, Engers R, Muller M, et al. (2005) Application of a modified real-time PCR technique for relative gene copy number quantification to the determination of the relationship between NKX3.1 loss and MYC gain in prostate cancer. Clin Chem 51: 649–652.

42. Laccone F, Junemann I, Whatley S, Morgan R, Butler R, et al. (2004) Large deletions of the MECP2 gene detected by gene dosage analysis in patients with Rett syndrome. Hum Mutat 23: 234–244.

43. Suo Z, Daehli KU, Lindboe CF, Borgen E, Bassarova A, et al. (2004) Real-time PCR quantification of c-erbB-2 gene is an alternative for FISH in the clinical management of breast carcinoma patients. Int J Surg Pathol 12: 311–318.

44. Ginzinger DG (2002) Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. Exp Hematol 30: 503–512.

45. Rooms L, Reyniers E, Kooy RF (2005) Subtelomeric rearrangements in the mentally retarded: a comparison of detection methods. Hum Mutat 25: 513–524.

46. Hollox EJ, Armour JA, Barber JC (2003) Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. Am J Hum Genet 73: 591–600.

47. Whittall JB, Medina-Marino A, Zimmer EA, Hodges SA (2006) Generating single-copy nuclear gene data for a recent adaptive radiation. Molecular phylogenetics and evolution 39: 124–134.

48. Li JL, Wang GL, Bai ZY (2009) Genetic variability in four wild and two farmed stocks of the Chinese freshwater pearl mussel (*Hyriopsis cumingii*) estimated by microsatellite DNA markers. Aquaculture 287: 286–291.

49. Bai ZY, Yin YX, Hu SN, Wang GL, Zhang XW, et al. Identification of Genes Potentially Involved in Pearl Formation by Expressed Sequence Tag Analysis of Mantle from Freshwater Pearl Mussel (*Hyriopsis Cumingii* Lea). Journal of Shellfish Research 29: 527–534.

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

51. Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A 95: 5857–5864.

52. Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, et al. (2006) Pfam: clans, web tools and services. Nucleic Acids Res 34: D247–251.

53. de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, et al. (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34: W362–365.

54. Markham NR, Zuker M (2005) DINAMelt web server for nucleic acid melting prediction. Nucleic Acids Res 33: W577–581.

55. 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.

56. D'Haene B, Vandesompele J, Hellemans J Accurate and objective copy number profiling using real-time quantitative PCR. Methods 50: 262–270.

57. Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, et al. (2004) Guideline to reference gene selection for quantitative real-time PCR. Biochem Biophys Res Commun 313: 856–862.

58. Marino JH, Cook P, Miller KS (2003) Accurate and statistically verified quantification of relative mRNA abundances using SYBR Green I and real-time RT-PCR. Journal of immunological methods 283: 291–306.