Pak. J. Agri. Sci., Vol. 52(4), 1055-1063; 2015
ISSN (Print) 0552-9034, ISSN (Online) 2076-0906
http://www.pakjas.com.pk

DECIPHERING THE EVOLUTION OF FERRITIN GENE FAMILY IN VARIOUS LIVING ORGANISMS

Shabana Memon[1,2], Rongchao Gao[1], Ruizhu Jiang[3], Weina Si[1] and Xiaohui Zhang[1,*]

[1]School of life Sciences, Nanjing University, Nanjing, 210093, China; [2]Department of Plant Breeding and Genetics, Sindh Agriculture University, Tando Jam, Hyderabad-70060, Pakistan; [3]Department of Forests and Ecosystem Science, the University of Melbourne, Victoria 3010, Australia.
[*]Corresponding author's e-mail: xiaohuizhang@nju.edu.cn

Ferritin, an iron binding protein, is an essential element in biological systems. Study of the two main subunits of ferritin; i.e., the heavy chain (FtH) subunit and ferritin light chain (FtL) subunit, has provided new insights into living organisms and improved the efficacy of ferroxidase activity. In this study, ferritin genes were identified in 24 organisms, including microbes, plants and animals. Mammalian ferritin genes exhibited the highest numbers and were found in clusters on X chromosomes considering duplication might have occurred. In addition, ferritin genes contained FtH subunits have different distribution compared with FtL subunits in mammalian. A phylogenetic tree was constructed according to the nucleotide coding sequences and 7 groups were classified. Fast gene loss and duplication were discovered. Investigation of conserved motif in different groups implies diverse evolution. Remarkably, the copy numbers of ferritin heavy polypeptide-like 17 homologs in mammalian were strongly related to the offspring number.
Keywords: Ferritin genes, mammals, mitochondrial ferritins, sex chromosomes, *Fthl17* genes

INTRODUCTION

Ferritin is an essential multi-meric protein found in the cells of almost all living organisms. It primarily functions in iron storage and is responsible for regulating levels of intracellular iron (Gaymard *et al.*, 1996; Petit *et al.*, 2001; Vanarsa *et al.*, 2012). It is utilized as a buffer either against iron deficiency or when iron is in excess, thus helping to control iron homeostasis in vertebrates and invertebrates (Gaymard *et al.*, 1996; Zhang, 2011). During metabolic processes, these iron-binding proteins also prevent iron toxicity of Fe^{2+} and toxic hydroxyl radicals ensuing from redox reactions and they maintain iron in a soluble bioavailable form (e.g., apoferritin, which is capable of sequestering large amounts of iron) (Hamburger *et al.*, 2005; Vanarsa *et al.*, 2012; Theil, 2013; Galay *et al.*, 2014). In biomedicine, its external protein shell is used as a precursor for some metallic nanoparticles that are applied as multimodal probes (Domínguez-Vera *et al.*, 2010; Theil, 2013). Ferritins, which are inside the cells of every type of tissue, are composed of 24 polypeptides and they have about 4,500 iron atoms within its hollow spherical shell. Profoundly, mammalian ferritins are ubiquitous cytoplasmic proteins consisting of two types of subunit chains; i.e., the ferritin heavy chain (FtH), subunit, with ferroxidase activity, and the ferritin light chain (FtL) subunit, that assists in iron nucleation, mineralization, and long-term storage. Variation in ferritin subunit composition may affect the rates of iron uptake and release in different tissues (Stevens *et al.*, 1987; Briat, 1996; Theil, 2013; Wen and Paine, 2013). In addition to cytosolic ferritins (CFts), mitochondrial ferritins (MtF) have been characterized in humans, plants, fruit flies, and mice, which are encoded by a nuclear gene with ferroxidase activity (Missirlis *et al.*, 2006; Zhang, 2011). MtFs are not ubiquitous and are different from CFts. CFt is highly expressed in tissues, especially in the liver and spleen, whereas the expression of MtF is restricted (Nichol and Locke, 1999; Bartnikas *et al.*, 2010; Vanarsa *et al.*, 2012). Furthermore, another kind of ferritin, known as a 'ferritoid', is present in avians, existing in the corneal epithelial (CE) cells of birds (a nuclear protein that protects DNA from UV damage). Ferritoids are cytoplasmic, suggesting that the nuclear transport function requires an interaction with ferritin (Nichol and Locke, 1999; Beazley, 2008; Beazley *et al.*, 2009; Bartnikas *et al.*, 2010; Vanarsa *et al.*, 2012). Microbes have two types of ferritins; i.e., heme containing bacterioferritins (BFRs) and non-heme containing bacterial ferritins (Gaymard *et al.*, 1996; Chiancone *et al.*, 2004). Beyond iron metabolism, ferritin genes also play a vital role in the cellular defense system against stress and inflammation and produce an immune response to the organism functioning as a regulator, which is crucial in autoimmune diseases (Hamburger *et al.*, 2005; Vanarsa *et al.*, 2012).

It has been postulated that ferritin was first isolated and discovered from the spleens of horses (Granick, 1942; Zhang, 2011). Plant ferritins are different from animal ferritins and are mostly localized in chloroplasts and visible in plastids but not in the cytoplasm (Petit *et al.*, 2001; Zancani *et al.*, 2004). Insect ferritin is considered to be extracellular or

within a vacuolar system of tissues including hydrophobic peptides (Nichol and Locke, 1999; Hamburger *et al.*, 2005). Recent studies have discovered that ferritin binding proteins vary by sex and age in mammals. Sex chromosomes, which are highly active in both males and females in early embryonic stages, have been recognized as having a highly specialized role in gene expression patterns in mammals. There is a report that the ferritin heavy polypeptide-like 17 (*Fthl17*) subfamily of genes in mice are imprinted and expressed from the paternal X chromosome as early as the two-cell stage. This suggests that by the time zygotic genome activation starts, there are already differences in gene expression between male and female mouse embryos (Wang *et al.*, 2001). Such exclusive early sex-linked differences may have some effects at later stages (Santambrogio *et al.*, 2007; Kobayashi *et al.*, 2010). In addition to mice, MtF has been highly expressed in humans and *Drosophila melanogaster* (*Fer3HCH*), mainly in the testes. The third ferritin gene; i.e., *CG4349* (*Fer3HCH*) of *D. melanogaster,* located on the X chromosome, has been suggested to be a gene coding for CFt (Missirlis *et al.*, 2006)
The studies of ferritins have been focused on their gene structure, function, and importance in different organisms. Few studies have been performed on the evolution of the ferritin genes among living organisms. To demonstrate the presence of ferritin genes in different living organisms and survey the evolutionary patterns of ferritins, we used 30 different species to identify these genes. The complete genomic sequences of these living species provided an opportunity to focus on the importance of ferritin genes and to investigate the evolutionary patterns of this gene family. Totally, 166 ferritin genes were identified in 24 species. However, ferritin genes have not been found in the remaining six species; *i.e. Anopheles gambiae, Chlamydomonasreinhardtii, ciona, Encephalitozooncuniculi, Ornithorhynchusanatinus,* and*Physcomitrella patens*. A phylogenetic tree was constructed to evaluate the evolutionary relationships of ferritin genes in these species. Chromosomal distribution of the genes was explored to view the potential mechanisms leading to their species-specific expansion in organisms. A motif search by MEME was carried out among paralogs to demonstrate the evolutionary pattern in ferritin genes. This systematic ferritin gene family study will lead to an understanding of the importance of functional genes in living organisms.

MATERIALS AND METHODS

***Acquisition of ferritin gene sequences for different species*:**In this study, we selected a total of 30 whole-genome sequenced species, including microbes, plants, invertebrates, non-mammalian vertebrates, and mammals. They were: *Acyrthosiphonpisum, Anopheles gambiae, Arabidopsis thaliana, Bos Taurus, Caenorhabditis elegans, Canis lupus, Chlamydomonasreinhardtii, Ciona intestinalis, Cyanidioschyzonmerolae, Danio rerio, Drosophila melanogaster, Drosophila pseudoobscura, Encephalitozooncuniculi, Gallus gallus, Homo sapiens, Hydra magnipapillata, Macacamulatta, Monodelphisdomestica, Mus musculus, Nostocverrucosum, Ornithorhynchusanatinus, Oryctolaguscuniculus, Oryza sativa, Ostreococcustauri, Pan troglodytes, Physcomitrella patens, Rattusnorvegicus, Strongylocentrotuspurpuratus, Taeniopygiaguttata, and Xenopus (Silurana) tropicalis*.
Most of the sequences were obtained from the National Center for Biotechnology Information GenBank (http://www.ncbi.nlm.nih.gov) and the other sources were specified as http://merolae.biol.s.u-tokyo.ac.jp/ and http://genome.jgi-psf.org/.
For each species, protein entries matching the ferritin-like domain (Pfam: PF00210) in the Pfam database V23.0 were identified as ferritin genes using HMMER searches with an E-value cut-off of 10-4.
***Nomenclature and identification of ferritin subunit genes*:**Ferritin subunit identification was based on the identified ferritin gene information from NCBI Genbank. We used BLAST for each ferritin gene in the NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to search out gene descriptions and general protein information. Most that had already been identified previously contained a gene description and subunit identification. Consequently, we symbolized these data to name each sequence as H for an FtH gene, L for an FtL gene, M for an MtF gene, and D for a ferritoid gene, and Un was used for a gene we identified from a genome that was not predicted previously in NCBI Genbank.
The particular name of each gene we codified was classified into four main categories; i.e., species, chromosome, number, and subunit. In the first category, we used five alphabetic letters to represent each species name. The first three letters were taken from the first three letters of the generic name and the last two letters were taken from the first two letters of the species name. The second category designation came from the gene location on a certain chromosome. Those genes that were identified from GenBank and did not contain any chromosomal location information were not labelled in this category. If there were more than two genes located on the same chromosome or belonging to the same species that did not have any chromosomal location information, we have utilized the third category to give each gene a different and distinctive number. Subunit information was the fourth category included in the gene name.
***Phylogenetic analysis and motif identification*:**Multiple sequence alignments of amino acid sequences of ferritin proteins were performed using ClustalW with default options (Thompson *et al.*, 1994). The resulting alignments were then used to guide the alignment of nucleotide coding sequences using MEGA Version 5.0 (Tamura *et al.*, 2007).

Phylogenetic trees were then generated based on the neighbor-joining (NJ) method with the Kimura two-parameter model (Kimura 1980). The confidence for each branching node was assessed by bootstrap analysis with 1,000 replicates.

Conserved motifs of the paralogous ferritin proteins were identified statistically using MEME (Bailey and Elkan, 1995) with the maximum number of motifs set at 3 and the motif length at 2 to 140.

RESULTS

***Identification and characterization of ferritin genes from different species*:**From the genomic analysis of 30 living organisms, we identified a total of 166 ferritin genes in only 24 genome species. Living kingdom species were specified into five groups; i.e. microbes (bacteria), plants, invertebrates, non-mammalian vertebrates, and mammals. The ferritin gene number for each species is presented in Table 1. The ferritin genes could be found in prokaryotes, plants, and animals, indicating that these genes are ancient and have very important roles in the life of the organisms. The results revealed that the mammalian ferritin gene number was higher than the gene number in other species. More than 10 ferritin genes were identified in most mammals, in comparison to other species, which had less than five members. It is suggested that an expansion of ferritin genes has taken place in these mammal species. However, two invertebrate species; i.e., *Hydra magnipapillata* and *Danio rerio,* also had a large number of ferritin genes (18 and 12, respectively).

Additionally, the distribution of ferritin genes in chromosomes was different between mammals and other species. Most of the ferritin genes of mammalian species were dispersed on both autosomes and the X chromosome, and not on the Y chromosome, whereas those genes in other species were not present on the X chromosome, except in fruit flies. In mammals, many ferritin genes (2–8) were distributed on the X chromosome, whereas only one or two ferritin genes were found on each autosome. It can be inferred that these ferritin genes on the X chromosome might be related to sex. Exceptionally, 8 of 12 ferritin genes of *D. rerio* (an invertebrate species) resided on chromosome 3.

The iron-binding protein is an essential element of living organisms, which comprises of both cytoplasmic and mitochondrial ferritin protein molecules. CFts is composed

Table 1. Ferritin gene Distribution in species and Chromosome location.

	Species	Total number	Chromosome location
Bacteria	*Nostoc*	3	
Plant	*Arabidopsis thaliana*	4	
	Cyanidioschyzonmerolae	1	
	Oryza sativa	4	
	Ostreococcustauri	1	
Invertebrates	*Acyrthosiphonpisum*	2	
	Caenorhabditis elegans	2	
	Drosophila melanogaster	3	3R(2)、X
	Drosophila pseudoobscura	3	
	Hydra magnipapillata	18	
	Strongylocentrotuspurpuratus	1	
Other Vertebrates	*Danio rerio*	12	3(8)、4、24、25(2)
	Gallus gallus	2	5(2)
	Taeniopygiaguttata	3	5(2)、Un
	Xenopus (Silurana) tropicalis	3	
Mammal	*Bostaurus*	15	1(2)、6(2)、7、X(3)、11、15、18、21、23、25、29
	Canis lupus	17	1(2)、4、6、8、11(4)、15、28、31、32、X(4)
	Homo sapiens	6	3、5、11、19、X(2)
	Macacamulatta	12	2(2)、6、13(2)、14、19、X(5)
	Monodelphisdomestica	5	2、4(2)、5(2)
	Mus musculus	15	3、4、7、13、16、18、19、X(8)
	Oryctolaguscuniculus	7	3、13、14、X(4)
	Pan troglodytes	11	1、5、8、11(2)、13、19(2)、20、X(2)
	Rattusnorvegicus	16	1(4)、2、7、11、18、X(8)

*The number in the () represent the gene number located on this chromosome.

** The blank in last column which do not fill the chromosome number means the genome sequences have not been

of heavy- (H) and light- (L) chain subunits, enabling ferrioxidase activity and provide better nucleation in the cells (Santambrogio *et al.*, 2007; Finazzi and Arosio, 2014). In mammals, FtH and FtL subunits were found to be in various proportions located on the X chromosomes and autosomes (Table 2). For 36 of these ferritin genes residing on the X chromosome, 26 (72.2%) genes were FtH and 7 (19.4%) were FtL genes. However, on the autosomes, 21 FtH and 32 FtL genes were elucidated, respectively. About 55.3% of the FtH genes existed on the X chromosome, whereas most of the FtL genes (82.1%) were distributed on the autosomes and only 17.9% of the FtL genes were on the X chromosome. Ferritin gene expression in the subunits varied from species to species and the amount of the heteropolymers situated within the species also diversified. Such complication in the composition of ferritin protein is generally due to the regulation of gene expression. Consequently, we inferred that the FtL gene has a locational bias, with preference to locate on an autosome. MtF is a nuclear-encoded iron-sequestering protein that specifically localizes in mitochondria. The ancient mammalian species *Monodelphisdomestica*did not contain the MtF gene, while other mammalian species each exhibited only one MtF gene and all were located on autosomes.

Table 2. Distribution of different ferritin genes on mammalian X chromosome and autosome.

Ferrtin	Number of ferritin on X chromosome (%)	Number of ferritin on Autosome (%)
H (FtH)	26 (72.2%)	21 (31.3%)
L (FtL)	7 (19.4%)	32 (47.8%)
M (FtMt)	0	8 (11.9%)
Un	3 (8.3%)	6 (9.0%)
Total	36	67

FtH: Ferritin heavy chain subunit, FtL: Ferritin light chain subunit, FtMt: Mitochondrial ferritin, Un: identified ferritins genes with unknown subunit or type.

***Phylogenetic analysis and classification of ferritin genes*:**To investigate the phylogenetic relationships and evolutionary patterns among ferritin genes in 24 species, a phylogenetic tree was constructed. According to the topology structure of the tree and the types of ferritin genes, we divided the phylogenetic tree into seven groups (Groups 1–7, Figs. 1 and S1).

Group 1 consisted of 23 ferritin gene members from vertebrates and they were all FtH located on autosome. One to five ferritin genes of each species were clustered in group 1. In this group, only one human ferritin gene was identified while five and four genes from chimpanzees and rhesus monkeys, respectively, were found, indicating a rapid gain or loss of ferritin genes in Homininae species. Group 2 is an MtF-specific group with nine members in eight mammals. Group 3 represented the mammalian family including 28 members and all were found to be FtH-subunit genes located on the X chromosome, searched with a bootstrap support value of 85%. Genes in group 3 were mostly clustered in a species-specific manner, suggesting that these genes were produced under rapid duplication after these species split from each other. Recent studies reported that the X-chromosome-linked mouse gene; e.g., ferritin heavy polypeptide-like-17 (*Fthl17*) gene has a predominant role in the pre-meiotic stages of mammalian spermatogenesis (Wang *et al.*, 2001).

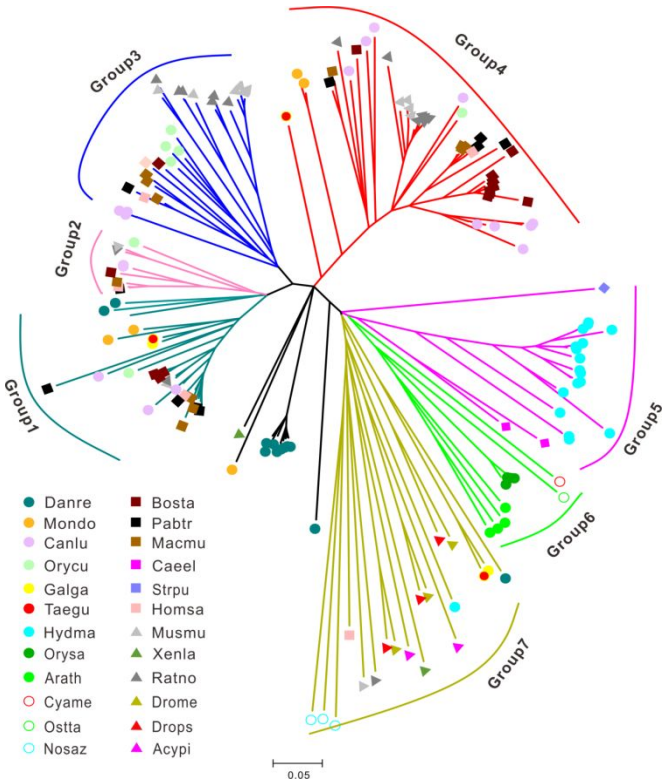


Figure 1. Phylogenetic tree of ferritin genes from different species.

The phylogenetic tree was built according to the nucleotide coding sequences of ferritin genes in 24 species and divided into seven groups. Ferritin genes in different species were marked with different colors or shapes. Species names were simplified to five words by using the first three letters of the generic name and the first two letters of the species name (Table 1).

The homologs of the *Fthl17* gene from other species in group 3 might have similar functions to the *Fthl17* gene in mice. Obviously, the gene copy number among different species in this group varied (Table 3). Notably, the species that have a higher growth rate and produce larger numbers of offspring in one birth may have more *Fthl17* genes in this group, such as mice and rats. Figure 2 shows the number

of*FtH* genes present on the X chromosome which are highly related to the offspring number in mammals. When plotting the number of *Fthl17* genes vs. the average offspring number in one birth, a significant positive relationship was found (R = 0.97276, P<0.0001). Almost all of the FtH genes were divided into group 2 or group 3. These results demonstrate that FtH genes have specific structures or motifs that could be distinguished from other types of ferritin genes. The MtF-specific group; i.e., group 2, resided between two FtH gene groups; i.e., group 1 and group 3, indicating that FtH genes and MtF genes are more similar than other ferritin genes and they might have some relationship in evolution.

Table 3. Number of copies and offsprings during birth of species in phylogenetic group 3.

Species	Copy number in group 3	Offspring number in one birth	
		Average	Range
Homo sapiens	2	1.5	1~2
Pan troglodytes	1	1	1
Macacamulatta	3	2	1~3
Bostaurus	1	1	1
Canis lupus	3	4.5	1~10
Oryctolaguscuniculus	4	5.5	5~12
Rattusnorvegicus	6	8	5~16
Mus musculus	8	10	8~15

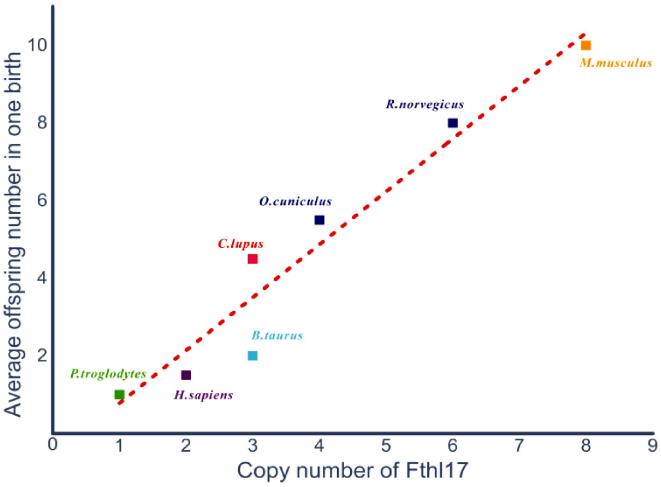


Figure 2. Correlationship between the copy number of Fthl17 and average offspring number in one birth of mammals.

The largest group in the phylogenetic tree was group 4, comprised of 44 members with a bootstrap support value of 68%. In this group, all the members were FtL genes from mammals except for one gene from *Taeniopygiaguttata,* which could be considered as an outgroup. Large expansion took place in *C. lupus* and *B. Taurus*, because 9 and 11 FtL genes were identified in these two species, respectively, within this group. The small clade containing nine FtL genes of *B. Taurus* in the tree denote a very rapid recent duplication. The FtL genes in mice and rats also duplicated after they split from each other. More interestingly, humans only had one FtL gene whereas chimpanzees and rhesus monkeys each had four FtL genes, indicating that gene loss or duplication events are lacking in humans. More importantly, most of the higher animals exhibited FtL genes on an X chromosome, such as mice, rats, chimpanzees, and rhesus monkeys, while none of these genes were found on the X chromosome in humans. To illuminate the reason for the loss of these ferritin genes in humans, we used the chimpanzee FtL gene on the X chromosome to do a BLASTN searches on human X chromosome sequences. A pseudogene in humans that had a stop codon existing in the middle of the sequence was detected. Thus, the FtL gene on the human X chromosome revealed a non-functional protein that could not translate. These two sequences were highly similar, as the nucleotide divergence was 0.035. A high Ka/Ks ratio (=1.87) was detected among these FtL genes on X chromosomes, suggesting that these genes evolved under selective pressure. We presume that the FtL gene on the X chromosome might be functionally redundant, with the copy being mutated to a functionless sequence.

The ferritin genes in group 5 were invertebrate species-specific. Remarkably, 17 *H. magnipapillata*ferrin genes were identified in group 5, which was more than other invertebrate species. All plant ferritin genes were clustered in group 6, which emphasized the divergence of ferritin genes between plants and animals. In rice and Arabidopsis, four members were observed in each genome. Also, the ferritin genes of these two species, each formed special clades with opposite features, having an extremely short branch in rice and a long branch in Arabidopsis. Group 7 is the most complex group, including 19 members ranging from prokaryotes to eukaryotes, from lower animals to humans. Consequently, all the insect ferritins existed only in this group and were scattered. The average length of the branch in this group, which represents the nucleotide divergence among members, was the longest among all the groups, hence demonstrating the ancient origin of these genes. We suggest that this group represents a relatively ancient phylogenetic relationship of ferritin genes. Except for the plant-specific groups; i.e., groups 6 and 7, the phylogenetic trees of ferritin genes generally corresponded to a classical animal phylogeny. The genes from the same species gathered together and the branch lengths were extremely short, indicating a recent species-specific expansion.

***Sequence analysis of conserved motifs for each phylogenetic group*:**A functional motif-based recognition of ferritin proteins can contribute to the understanding of the

evolutionary history of the ferritin gene family. A motif search by MEME was identified among paralogs in the seven groups. In each group, two or three conserved motifs were identified (Fig. S2). Among these seven groups, distinct distribution patterns of conserved motifs were attributed. Each group revealed a quite different strategy compared to their positions. Every group appeared to have a long conserved motif consisting of more than 100 amino acids, except for group 4 (FtL) and the complex group 7. The lengths of the three motifs in group 4 were all short and nearly equal. In addition, the conserved region of FtH sequences was wider than that of FtL sequences. Although group 1 and group 3 were both FtH-specific groups, their conserved motif sequences and the distribution patterns of motifs varied, suggesting that FtH genes on autosomes and sex chromosomes were distinguished and might have some functional differentiation. The ancient group 7 showed only two short conserved motifs, indicating the ferritin genes in this group were more divergent.

For further interpretation, we analyzed the phylogenetic relationship of all the motifs by building a motif phylogenetic tree (Fig. 3). Only one conserved motif of the FtH groups; i.e., group 1 and group 3, was included in the tree, representing low similarity to other ferritin genes. The biggest clade, whose bootstrap values were higher than 80%, consisted of five groups, except for group 3 (FtH on the X chromosome) and group 5 (the plant-specific group). The topology structure of the clade resembled that of the ferritin gene tree, indicating that this motif in each group is the common conserved sequence of most ferritin genes,

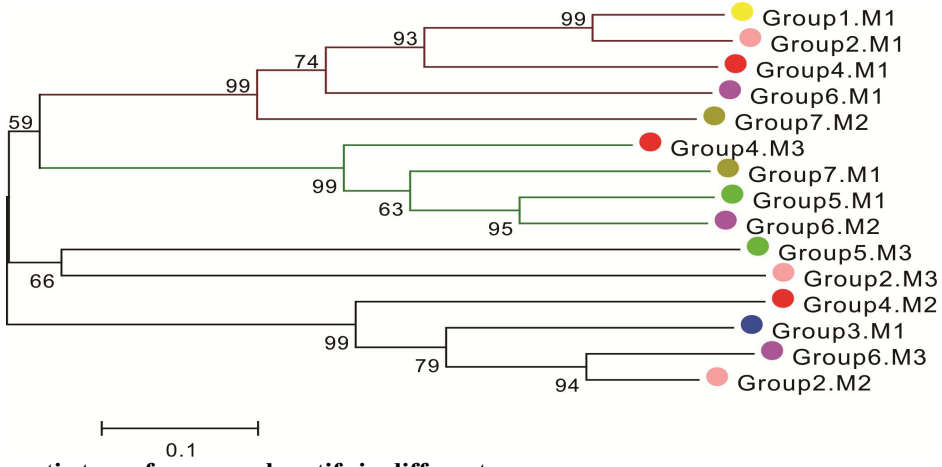
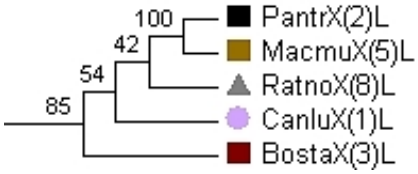


Figure 3. Phylogenetic tree of conserved motifs in different groups.
M1, M2, and M3 correspond to the respective motif sequences of the seven groups in Fig.1. Several motif sequences were excluded as they are too short and too divergent to build a phylogentic tree.

(A)



(B)

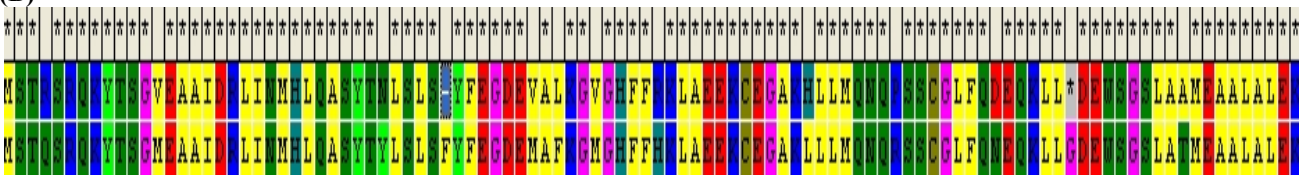


Figure 4. A FtL pseudogene found in humans.
(A) NJ phylogenetic tree of all FtL genes on X chromosome from different genomes. (B) This figure reveals the blast result between the human and the chimpanzee. The red box cites a stop codon in between the human sequence. The upper portion of the sequence is the blast result from human X chromosome, whereas, the lower FtL sequence is from chimpanzee X chromosome.

especially the animal ferritin genes. The corresponding sequences in group 3 and the plant group 5 might have evolved rapidly to produce new special functions. A clade containing conserved motif sequences identified from groups 4, 5, 6, and 7, seemed to be a primitive and ancient clade of ferritin genes.

DISCUSSION

***Role of sex-linked Fthl17 genes in mammalian species*:**Wang *et al.* (2001) identified 25 genes among which *Fthl17* genes were also expressed. Out of 25 genes, 3 genes of which were Y-linked and 10 were X-linked genes, indicating that the X chromosome has a predominant role in pre-meiotic stages of mammalian spermatogenesis. Also, the X-linked mouse gene; e.g, *Fthl17,* shows predominantly female expression prior to meiotic stages in mouse embryos. This discovery led to knowledge of the importance of early sex differentiation (Wang *et al.*, 2001; Kobayashi *et al.*, 2010). This gene was found to be orthologous to the full-length human *FTHL17* sequence (Wang *et al.*, 2001). Our analysis informed us that the species that contributes a larger number of offspring and has fast growth rate, such as mice and rats, harbor more copy numbers of FtH genes on the X chromosome (Fig. 2). It has been hypothesized that the paternally derived X chromosome and/or genes expressed from both X chromosomes retard development prior to gonadal sex differentiation. This means that the expression of *Fthl17* genes may be due to earlier sex differentiation. Furthermore, it has been pointed out that the paternally expressed *Fthl17* family and other X-linked genes that are not activated could serve as good candidates for the cause of sex differentiation in the rate of growth. Therefore, the large number of candidates can promote a high growth rate. Thus, species with a high growth rate will acquire more X-linked genes. Similar to the copy number of *Fthl17* genes in group 3, it is associated with the offspring number in each species.

***Duplication and functional diversification of ferritin genes*:**Ferritin provides a heterogeneous protein family that is useful for aerobic metabolism (Ekman *et al.*, 2014). The ferritin genes were detected in most living organisms, from prokaryotes to eukaryotes, from plants to animals, from lower animals to humans, suggesting great importance. (Grossman *et al.*, 1992; Briat *et al.*, 2010). Ferritin was found to be a superfamily with diverse functions within species (Chiancone *et al.*, 2004; Andrews, 2010; Theil, 2013). The plant ferritin proteins diverged from animal ferritin, but they both evolved from the same ancestor a long time ago (Briat *et al.*, 2010). Plant ferritin genes sharing 39%–49% identity with animal ferritin genes are localized in different cellular compartments. Plant ferritin is more similar to lower animals and the ferritin gene number in each species is small. In mammals, the ferritin gene number was found to be more than those in other species, inferring that they might have experienced duplication events after splitting from other animals.

Gene duplication constitutes an efficient role in multigene families (Ohta, 1987). The duplication of a gene results in an additional gene copy that is free from selective pressure. These gene copies formed clusters to promote a novel function maintained by selection (Ober, 2005). In mammals, the ferritin protein diverged to FtH subunits and FtL subunits. Each gains a new function that may benefit mammals with adaptations to varied iron conditions and improved efficiency of ferroxidase activity (Hamburger *et al.*, 2005; Finazzi and Arosio, 2014). The conserved motifs of FtH proteins and FtL proteins are both partly similar to ferritin proteins in lower animal motifs and also have their own conserved sequences that are not homologous to other groups. Genes splitting across different lineages, accumulating speciation events and comprising a common ancestor, resemble orthologs, whereas the genes produced by gene duplication in the same species are paralogous genes (Hurles, 2004; Organ and Edwards, 2011). Paralogous genes found in the same genome are probably created by gene duplication events and might have different functions. In our analysis, after the differentiation of FtHs and FtLs, the FtH proteins have undergone two duplication events independently. First, the duplication occurred on the X chromosome as genes of each species in group 3 were found in clusters. Gene clusters are probably created by gene duplication and the duplicated genes easily acquire novel functions (neofunctionalization), which evolves as a new copy with the old one still performing the original function (Fan *et al.*, 2008) or performing part of the original function (subfunctionalization) (Organ and Edwards, 2011). In group 3, tandem duplications were prominent and the function of the ferritin genes in this group seemed different from other ferritins, which may be related to sex differentiation. Secondly, the FtHs on the autosomes (group 1) revealed another duplication event. The gene number from each species in group 1 was less than that in group 3. In mammals, in addition to ferritin proteins having H- and L-cytosolic subunits, a ferritin protein with long N-terminal mitochondrial localization amino acids has been reported. The MtF consists of a distinctive conserved motif. According to the phylogenetic tree, the MtF genes are more similar to FtH genes, but the function of MtF is distinctive to FtHs. In humans, MtF has been found in specific organs and cells, such as in testes and spermatozoa. It has been reported that MtF may protect the mitochondria against iron toxicity and oxidative damage (Harrison and Arosio, 1996). Generally, there is only one MtF gene in each species, assuming that the function of MtF is important and conserved among various species.

Sometimes the duplicated genes have no effect on the genotype of the organism and they lose their function or they become silenced (nonfunctionalization) due to degenerative

mutation (Ohta, 1987). In humans, the FtL gene on the X chromosome was not functional and became pseudogenized. The pseudogene in humans exhibited a stop codon causing premature termination. It can be proposed that it did not gain any function and may have functional redundancy because the gene did not express in humans.

***The expansion patterns of ferritin genes in species*:**According to the phylogenetic tree, many ferritin genes have shown species expansion, which indicates a recent duplication (Fig. 1). In animals, most genes in groups 1, 3, 4, and 5 have expanded in a species-specific manner. The phenomenon of gene expansion is more obvious in mice, rats, and *B. taurus*. The short length of the clades containing ferritin genes from these species indicates very recent gene duplication events.

In *H. magnipapillata* and *D. rerio,* two different kinds of water living animals, 18 and 12 ferritin genes were found, respectively. The number of these genes is the same as in mammals. More interestingly, most of these genes in each species clustered together. In detail, 17 of 18 *H. magnipapillata* genes and 9 of 12 *D. rerio* genes were respectively divided into two clades, which are clearly separated from other ferritin genes. It can be assumed that ferritin genes may have undergone an intraspecific expansion. It has been reported that genome duplication occurred early during the evolution of ray-finned fishes. The large number of ferritin genes in zebrafish might be the result of whole genome duplication, together with less gene loss and other small gene duplications (Taylor *et al.*, 2003). The bursts of horizontal gene transfer in *H. magnipapillata*might offer an explanation for the expansion of ferritin genes (Chapman *et al.*, 2010). Based on the phylogenetic tree, recent duplication also contributed to the expansion of ferritin genes in *H. magnipapillata*.

A number of studies have focused on ferritin gene structure, function, and ferritin's importance in different organisms. Our study led to the importance of the ferritin gene family, which shows an evolutionary relationship among the living organisms. Mammalian ferritin proteins are made of two different subunit types, coded by FtH and FtL genes, respectively. The FtH genes that reside on autosomes and the X chromosome are clearly divided into two different groups, indicating conserved motifs and different functions for each type of FtH gene. The representative of FtH genes on the X chromosome; i.e., *Fthl17* gene, is found to be related to early sex differentiation. We found the species that contribute a larger number of offspring and have a fast growth rate, harbor a larger copy number of *Fthl17* homologs and in each species these homologs are clustered, indicating a species-specific expansion. This advantageous gene should be explored further for more efficient study of mammalian spermatogenesis.

***Conclusions*:** The ferritin genes in different organisms varied accordingly with FtH subunits compared with Ftl, in which mammals were found to have more FtH genes. Hence Fast duplication and gene loss of ferritin genes was found. Accordingly, the copy number of *Fthl17* genes resembled the offspring number in species. This gene could also be utilized for explanation for polyembryony.

***Acknowledgements*:** This work was supported by the National Natural Science Foundation of China (91331205, J1103512 and J1210026), Program for Changjiang Scholars and Innovative Research Team in University (IRT1020).

REFERENCES

Andrews, S.C. 2010. The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. Biochimica et BiophysicaActa (BBA)-General Subjects. 1800: 691-705.

Bailey, T.L. and C. Elkan. 1995. The value of prior knowledge in discovering motifs with MEME. Ismb. 3: 21-29.

Bartnikas, T.B., D.R. Campagna, B. Antiochos, H. Mulhern, C. Pondarré and M.D. Fleming. 2010. Characterization of mitochondrial ferritin deficient mice. Amer. J. Hematol. 85: 958-960.

Beazley, K.E. 2008. Regulation of ferritoid expression and function in the avian corneal epithelium. Ph. D diss., Dept. Cell Mol. and Development Biol., Sackler School of Graduate Biochemical Sciences, Tufts Univ.,United States.

Beazley, K.E., J.P. Canner and T.F. Linsenmayer. 2009. Developmental regulation of the nuclear ferritoid–ferritin complex of avian corneal epithelial cells: Roles of systemic factors and thyroxine. Exp. Eye Res. 89: 854-862.

Briat, J.F., C. Duc, K. Ravet and F. Gaymard. 2010. Ferritins and iron storage in plants. Biochimica et BiophysicaActa (BBA)-General Subjects 1800: 806-814.

Briat, J. 1996. Roles of ferritin in plants. J. Plant Nutr. 19:1331-1342.

Chapman, J.A., E.F. Kirkness, O. Simakov, S.E.Hampson, T. Mitros, T. Weinmaier, T. Rattei, P.G. Balasubramanian, J. Borman and D. Busam. 2010. The dynamic genome of Hydra. Nature. 464: 592-596.

Chiancone, E., P. Ceci, A. Ilari, F. Ribacchi and S. Stefanini. 2004. Iron and proteins for iron storage and detoxification. Biometals. 17: 197-202.

Domínguez-Vera, J.M., B.Fernández and N.Gálvez.2010. Native and synthetic ferritins for nanobiomedical applications: recent advances and new perspectives. Future Med. Chem. 2: 609-618.

Ekman, M., G. Sandh, A. Nenninger, P. Oliveira and K. Stensjö. 2014. Cellular and functional specificity among ferritin-like proteins in the multicellular cyanobacterium Nostocpunctiforme. Environ. Microbiol. 16: 829-844.

Fan, C., Y. Chen and M. Long. 2008. Recurrent tandem gene duplication gave rise to functionally divergent genes in Drosophila. Mol. Biol. Evol. 25: 1451-1458.

Finazzi, D., P. Arosio. 2014. Biology of ferritin in mammals: an update on iron storage, oxidative damage and neurodegeneration. Arch. Toxicol. 88: 1787-1802.

Galay, R.L., R. Umemiya-Shirafuji, E.T. Bacolod, , H. Maeda, K. Kusakisako, J. Koyama, N. Tsuji, M. Mochizuki, K .Fujisaki and T. Tanaka. 2014. Two Kinds of Ferritin Protect Ixodid Ticks from Iron Overload and Consequent Oxidative Stress. Plos One .9: e90661.

Gaymard, F., J. Boucherez and J. Briat. 1996. Characterization of a ferritin mRNA from Arabidopsis thaliana accumulated in response to iron through an oxidative pathway independent of abscisic acid. Biochem. J. 318: 67-73.

Granick, S. 1942. Ferritin I. Physical and chemical properties of horse spleen ferritin. J. Biol. Chem. 146: 451-461.

Grossman, M., S. Hinton, V. Minak-Bernero, C. Slaughter and E.I. Stiefel. 1992. Unification of the ferritin family of proteins. Proc. Natl. Acad. Sci. U.S.A. 89: 2419–2423

Hamburger, A.E., A.P. West Jr, Z.A. Hamburger, P.Hamburger and P.J.Bjorkman. 2005. Crystal structure of a secreted insect ferritin reveals a symmetrical arrangement of heavy and light chains. J. Mol. biol. 349:558-569.

Harrison, P.M and P. Arosio. 1996. Ferritins: Molecular properties, iron storage function and cellular regulation. Bba-Bioenergetics. 1275: 161-203.

Hurles, M. 2004. Gene duplication: the genomic trade in spare parts. Plos Biol. 2: 206.

Kobayashi, S., Y. Fujihara, N. Mise, K. Kaseda, K. Abe, F. Ishino and M. Okabe. 2010. The X-linked imprinted gene family *Fthl17* shows predominantly female expression following the two-cell stage in mouse embryos. Nucleic Acids Res. 38: 3672-81.

Missirlis, F., S. Holmberg, T. Georgieva, B.C. Dunkov, T.A. Rouault and J.H. Law. 2006. Characterization of mitochondrial ferritin in Drosophila. Proc. Natl. Acad. Sci. U.S.A. 103: 5893-5898..

Nichol, H. and M. Locke. 1999. Secreted ferritin subunits are of two kinds in insects: molecular cloning of cDNAs encoding two major subunits of secreted ferritin from Calpodesethlius. Insect Biochem. Mol. Biol. 29: 999-1013.

Ober, D. 2005. Seeing double: gene duplication and diversification in plant secondary metabolism. Trends Plant Sci. 10: 444-449.

Ohta, T. 1987. Simulating evolution by gene duplication. Genetics. 115: 207-213.

Organ, C.L. and S.V.Edwards. 2011. 13 Major Events in Avian Genome Evolution. pp. 325-337. In G. Dyke and G. Kaiser (eds). Living Dinosaurs: The Evolutionary History of Modern Birds..Wiley-Blackwell.

Petit, J., J. Briat and S. Lobreaux. 2001. Structure and differential expression of the four members of the Arabidopsis thaliana ferritin gene family. Biochem. J. 359: 575-582.

Santambrogio, P., G. Biasiotto, F. Sanvito, S. Olivieri, P. Arosio and S. Levi. 2007. Mitochondrial ferritin expression in adult mouse tissues. J. Histochem.Cytochem. 55: 1129-1137.

Stevens, P.W., J. Dodgson, J. Engel. 1987. Structure and expression of the chicken ferritin H-subunit gene. Mol. Cell. Biol. 7: 1751-1758.

Tamura, K., J. Dudley, M. Nei and S. Kumar. 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596-1599.

Taylor, J.S., I. Braasch, T. Frickey, A. Meyer and Y. Van de Peer. 2003. Genome duplication, a trait shared by 22,000 species of ray-finned fish. Genome Res. 13: 382-390.

Theil, E.C., 2013. Ferritin: the protein nanocage and iron biomineral in health and in disease. Inorg. Chem. 52: 12223-12233.

Thompson, J.D., D.G. Higgins and T.J. Gibson. 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.

Vanarsa, K., Y. Ye, J. Han, C. Xie, C. Mohan and T. Wu. 2012. Inflammation associated anemia and ferritin as disease markers in SLE. Arthritis Res. Ther. 14:R182.

Wang, P.J., J.R. McCarrey, F. Yang and D.C. Page. 2001. An abundance of X-linked genes expressed in spermatogonia. Nature Genet. 27: 422-426.

Wen, X. and M.L. Paine. 2013. Iron deposition and ferritin heavy chain (Fth) localization in rodent teeth. BMC Res. Notes 6: 1.doi:10.1186/1756-0500-*6-1*

Zancani, M., C. Peresson, A. Biroccio, G. Federici, A. Urbani, I. Murgia, C. Soave, F. Micali, A. Vianello and F. Macrì. 2004. Evidence for the presence of ferritin in plant mitochondria. Eur. J. Biochem. 271: 3657-3664.

Zhang, L., 2011. Interaction of ferritin with transition metal ions and chelates. Ph. D diss., Dept. Chemistry, University of Pennsylvania, United States.