



## Phylogenomic analysis of transferrin family from animals and plants



Lina Bai<sup>a</sup>, Mu Qiao<sup>b</sup>, Rong Zheng<sup>a</sup>, Changyan Deng<sup>a</sup>, Shuqi Mei<sup>b,\*</sup>, Wanping Chen<sup>c,\*</sup>

<sup>a</sup> College of Animal Sciences and Technology, Huazhong Agricultural University, Wuhan, Hubei Province 430070, China

<sup>b</sup> Institute of Animal Husbandry and Veterinary, Hubei Academy of Agricultural Sciences, Wuhan, Hubei Province 430064, China

<sup>c</sup> College of Food Science and Technology, Huazhong Agricultural University, Wuhan, Hubei Province 430070, China

### ARTICLE INFO

#### Article history:

Received 29 June 2015

Received in revised form 21 October 2015

Accepted 21 November 2015

Available online 25 November 2015

#### Keywords:

Iron-binding

Lactoferrin

Melanotransferrin

Plant transferrin

Serotransferrin

Transferrin

### ABSTRACT

Transferrins have been identified in animals and green algae, and they consist of a family of evolutionarily related proteins that play a central role in iron transport, immunity, growth and differentiation. This study assessed the transferrin genes among 100 genomes from a wide range of animal and plant kingdoms. The results showed that putative transferrins were widespread in animals, but their gene quantity and type differ greatly between animal groups. Generally, Mammalia possess abundant transferrin genes, whereas Trematoda contain few ones. Melanotransferrin and serotransferrin are widely distributed in vertebrates, while melanotransferrin-like and transferrin-like 1 are frequent in invertebrates. However, only a few plant species detected putative transferrins, and a novel transferrin member was first uncovered in Angiospermae and Pteridophyta. The structural comparison among transferrin family members revealed seven very well-repeated and conserved characteristic motifs, despite a considerable variation in the overall sequences. The phylogenetic analysis suggested that gene duplication, gene loss and horizontal transfer contributed to the diversification of transferrin family members, and their inferred evolutionary scenario was proposed. These findings help to the understanding of transferrin distribution, characteristic motifs and residues, and evolutionary process.

© 2015 Elsevier Inc. All rights reserved.

### 1. Introduction

The transferrin (Tf) family is a group of single-chain, glycosylated proteins defined by conserved residue motifs and iron-binding functions (Lambert et al., 2005). The Tf family spans both animal and lower plant groups, which is widely distributed in vertebrates and plays a central role in iron homeostasis, immunity, cell growth and differentiation (Gomme et al., 2005). The current study indicated that, the Tf family includes 13 members, with six major members in vertebrates (Hughes and Friedman, 2014), four members in the Insecta (Huebers et al., 1988; Bartfeld and Law, 1990; Thompson et al., 2003), and one member in marine algae (Fisher et al., 1997, 1998; Schwarz et al., 2003). The best known Tf representative is serotransferrin (STf), which performs essential iron transportation and delivery functions in plasma (Giblett et al., 1959). Closely related family members include lactoferrin (LTF) found in milk (Pakdaman et al., 1998), ovotransferrin (OTf) found in egg yolk (Graham and Williams, 1975; Lee et al., 1980), and melanotransferrin (MTf) anchored to the membrane as a membrane-bound Tf (Nakamasu et al., 1999; Geiser and Winzerling, 2012). Additional evolutionarily distant Tf family members include the mammalian inhibitor of carbonic anhydrase (ICA), amphibian

saxiphilin, and the salmonid otolith matrix protein (OMP) (Morabito and Moczydlowski, 1994; Eckenroth et al., 2010). Tf proteins also have been found in invertebrates and lower plants. In the Insecta species, at least four Tf family members (Tsf 1–4) have been characterized. Recently, three-lobed Tf-like proteins have been identified in marine algae, although these may represent a horizontal gene transfer event (Lambert et al., 2005).

Almost all members of the Tf family consist of a single polypeptide chain (MW ~ 80 kDa) that folds into two globular homologous halves referred to as the amino- and carboxy-lobes (Theil, 1987; Baker et al., 2001). The original Tf ancestor had only a single lobe, and the first duplication event resulted in tandem genes, which later fused to form an ancestral, double-lobed Tf (Williams et al., 1982). Over time, the lobes have been deleted (single lobe protein) and further duplicated (the triple-lobed molecules in algae), suggesting that Tfs have adapted over the years in a variety of ways based on the unique physiological demands of different organisms (Lambert et al., 2005); thus, the tracing of original events can be challenging. Both LTF and ICA originated from relatively recent but separate duplications of the STf gene, and multiple copies of the MTf gene may show lethal dominance; therefore, duplications cannot survive (Nakamasu et al., 1999).

In the current study, 100 representative genome sequences from a wide range of animals and plants were used to investigate the distribution of all possible Tf family members. Furthermore, the comprehensive phylogeny, characteristic motif and iron-binding domains of these proteins were analyzed. These analyses enabled the division of Tf family

\* Corresponding authors.

E-mail addresses: [msqlfe@163.com](mailto:msqlfe@163.com) (S. Mei), [chenwanping@mail.hzau.edu.cn](mailto:chenwanping@mail.hzau.edu.cn) (W. Chen).

members into different phylogenetic groups. Based on the phylogenetic analyses, reconstructed scenarios for the evolution of Tf s were proposed.

## 2. Materials and methods

### 2.1. Sequence data of animal and plant species

Based on the genome information of NCBI database (<http://www.ncbi.nlm.nih.gov/genome/browse/>), 100 representative species from a wide range of animal and plant kingdoms were chosen in this study, and their overall protein sequences extracted from their corresponding genomes were downloaded for surveying the genome-wide distribution of Tf s. The information of these species is shown in Table 1.

### 2.2. Annotation of putative Tf family

The Tf family shares a conserved domain, which has been characterized on Pfam database with the profile hidden Markov model (PF00405, <http://pfam.xfam.org/family/PF00405>) (Finn et al., 2013). In this study, the overall protein sequences of 100 representative animal and plant species were subjected to identify whether they have the Tf domains, and those with Tf characteristic domains were recognized as putative Tf s. The annotation pipeline of putative Tf s in the selected 100 animal and plant species was as follows: First, the Tf profile hidden Markov model was applied to search the homologs among the overall protein sequences with HMMER 3.1 (Eddy, 2011). In addition, the hits that passed Bias, Fwd, MSV and Vit filters (please refer to HMMER User's Guide, <http://hmmer.org/>) were recognized as putative Tf s. Second, the putative Tf s were annotated by BlastP comparisons against the NCBI database of non-redundant protein sequences (<http://blast.ncbi.nlm.nih.gov>). These predicted Tf proteins were annotated as MTf, STf, Lf, OTf, ICA, OMP, saxiphilin and melanotransferrin-like (M-like), based on their high levels of similarity of the hits.

### 2.3. Construction of Tf phylogenetic tree

In this study, the phylogenetic tree of Tf s was constructed as follows: First, the annotated Tf protein sequences were aligned to the Tf profile hidden Markov model with HMMER 3.1. Second, the phylogenetic tree of Tf alignments was inferred by FastTree 2.1.8 with maximum-likelihood method (Price et al., 2009). Third, the generated tree data was submitted to iTOL (<http://itol.embl.de/>) for making phylogenetic tree figures (Letunic and Bork, 2007, 2011).

### 2.4. Structural feature analysis of Tf phylogenetic groups

Sequence logos were generated among different Tf homologous groups for revealing their primary structural features. First, the homologous Tf s were aligned to the Tf profile hidden Markov model by HMMER 3.1. In addition, the mismatched residues were removed. Second, the consensus logos of alignments which showed the conservation of primary structure were generated by WebLogo 3 (<http://weblogo.threeplusone.com/create.cgi>) with plotting a stack of residues for each position (Schneider and Stephens, 1990; Crooks et al., 2004).

## 3. Results and discussion

### 3.1. Genome-wide distribution of putative Tf s in animals and plants

In this study, all tested animal genomes from classes Trematoda (flatworms), Cestoda (flatworms), Insecta (insects), Aves (birds), Teleostei (fishes), Reptilia (reptiles), and Mammalia (mammals) contained Tf genes (Table 1). It suggests that the Tf gene is conserved and plays an important role in animals. However, the search among

the completed fungal and botanic genomes from the NCBI database detected no putative Tf genes in the fungal genomes and a limited number of Tf genes within phyla Chlorophyta (algae), Pteridophyta (Fern) and Angiospermae (flowering plants), which implies that fungi and plants use specific iron uptake and intracellular iron storage systems that have been identified as siderophores for fungi (Lesuisse and Labbe, 1989; Johnson, 2008) and phytoferritin for plants (Seckback, 1982; Briat et al., 2010). Surprisingly, in the scope of authors' knowledge, putative Tf genes were first detected in the genomes from Pteridophyta and Angiospermae (Table 1).

The count of Tf genes varies greatly in the tested animal species (Table 1). For example, Mammalia species, such as those from the order Cetartiodactyla, Primates, and Rodentia, possess high numbers of Tf genes, which are approximately one to ten times more than certain species from the Reptilia and Aves classes. The Teleostei species show considerable numbers of Tf genes. Some of them, including *Tetraodon nigroviridis*, *Poecilia reticulata* and *Danio rerio*, possess a large number of Tf genes. The Insecta species have an abundance of Tf genes, whereas the Trematoda and Cestoda species possess a limited number of Tf genes. *Dendroctonus ponderosae*, a beetle Insecta that parasitizes on mountain pine, contained a high abundance of Tf genes.

There are 13 Tf family members among the putative Tf s from 100 animals and plants; those are MTf, STf, Lf, ICA, OTf, saxiphilin, OMP, M-like, Tf-like, Tsf 1, Tsf 2, Tsf 3 and Tsf 4 (Table 1). Based on the occurrence of Tf family members in these tested species, MTf and STf are widespread in vertebrates, Tsf 1 is presented in all Insecta species, and M-like is a specific member found only in Trematoda and Cestoda. The results indicate that Tsf 1 plays an essential role in Insecta. Moreover, the numbers of MTf and STf genes are conserved. Generally, most vertebrates have one copy each of MTf and STf; however, some species have two genes. There are three MTf genes in *Chlorocephalus sabaues* and three STf genes in *Bos taurus*. In addition, these genes are absent in certain species (*Microtus ochrogaster* does not have STf and *Cariama cristata* does not have MTf). The conserved distribution of the MTf and STf genes implies their important roles in vertebrates.

### 3.2. Phylogenetic relationship of putative animal and plant Tf s

In this analysis, the phylogenetic tree of putative Tf s from animals and plants was constructed and their clade features were also studied (Fig. 1). In general, an abundance of Tf family branches in the phylogenetic tree suggests their high divergence during long time evolution. However, the phylogenetic distribution of Tf s varied between taxonomic groups. Particularly, Tf s from the groups Mammalia and Insecta are highly evolved and widespread across many branches. However, Tf s from some taxonomic groups are conserved. Notably, Tf s from Trematoda and Cestoda are concentrated together in one branch, suggesting their evolutionary conservation over time.

Based on the phylogenetic relationships, putative Tf s were classified into 15 clades (Fig. 1), and the distribution of Tf family members and organism taxonomy was investigated in these clades (Table 2). Among them, Clade 8, specifically MTf, has the biggest branch and includes species members of all major vertebrate groups. This result is compatible with an ancient vertebrate Tf origin of MTf (Nakamasu et al., 1999; Lambert et al., 2005). Clades 11 (Tsf 1) and 12 (M-like) include members from all major invertebrate species. Notably, these two clades gathered in a big branch but could be divided by species specificity: Clade 11 is Insecta-specific, and Clade 12 is flatworm-specific. These results suggest that Tf s was probably in ancestral invertebrates, and M-like might be the ancestor of Tf s in all animal kingdoms. STf s are not gathered in a big branch but rather scattered in Clade 2 (Mammalia specific), Clade 5 (Teleostei specific) and Clade 7 (includes Aves, Reptilia and Insecta species). This scenario implies that STf might be resulted from three parallel lineage-specific duplications followed by convergent events of secretion into the blood serum. The position of *Anolis*

Download English Version:

<https://daneshyari.com/en/article/1978529>

Download Persian Version:

<https://daneshyari.com/article/1978529>

[Daneshyari.com](https://daneshyari.com)