Cloning, characterization, and gene expression analysis of a novel cadmium metallothionein gene in *Tetrahymena pigmentosa*

Lina Guo a,b,1, Chengjie Fu a,b,1, Wei Miao a,⁎

a State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, PR China
b Graduate School of Chinese Academy of Sciences, Beijing, 100039, PR China

**A R T I C L E  INFO**

Article history:
Received 1 February 2008
Received in revised form 25 April 2008
Accepted 25 April 2008
Available online 22 May 2008

Keywords:
Gene duplication
Genetic diversity
Heavy metals
Internal similarity
Real-time quantitative PCR
Oxidative stress

**A B S T R A C T**

A novel cadmium-inducible metallothionein (MT) gene (*Tpig-MT1*) was cloned and sequenced from the ciliate *Tetrahymena pigmentosa*. The number of deduced amino acids is 118. The polypeptide possesses CCC and CC clusters characteristic of typical *Tetrahymena* Cd-inducible MTs. The structure of *Tpig-MT1* is different from the reported Cd-MT in *T. pyriformis, T. thermophila* and *T. pigmentosa*. *Tpig-MT1* contains two intragenic tandem repeats with 72.9% identity described as *Tpig-MT1* (repeat A1) and *Tpig-MT1* (repeat A2). The transcriptional response of *Tpig-MT1* gene to different heavy metals (Cd, Cu, Zn, Hg, Pb) and oxidative stress (H2O2) was measured using real-time quantitative PCR. The results showed that the gene was quickly induced (1 h) by the five heavy metals and the order of expression level was Hg > Pb > Cd > Cu > Zn. The induction effect of H2O2 was 5-fold after about 15 min, but soon decreased to a non-significant level (30 min). The genetic diversity of *Tetrahymena* MT genes is discussed in relation to the unique structure of the *Tpig-MT1* gene and other reported Cd-MT isoforms.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Metallothioneins (MTs) are a family of low molecular weight (<7–10 kDa), cysteine-rich, aromatic residues-lacking, heavy metal-binding (such as Zn, Cd, Cu and Hg) proteins. They were first isolated from horse kidney as Cd-binding proteins (Margoshes and Vallee, 1957), and have been identified throughout the animal kingdom, higher plants, eukaryotic microorganisms, and also in many prokaryotes (Kagi, 1991). Although their precise cellular function is uncertain, by virtue of their high affinity of binding with heavy metal ions, they take part in protection against heavy metal toxicity via sequestration (Cd and Hg), and are also involved in essential-metal homeostasis (Cu and Zn). Functional diversity between MT isoforms was hypothesized to be due to their different evolutionary adaptation to specific (external) environment and metabolic (internal environment) requirements (Coyle et al., 2002). In animals, multiple MT genes with distinct biological properties have been demonstrated in different tissues, but it is not clear what important physiological functions they possess (Palmiter, 1998). Invertebrate species have also been shown to have multiple isoforms of MTs, which can exhibit differential expression and regulation (Dondero et al., 2005).

Different MT isoforms have been described in the unicellular ciliated protozoan genus *Tetrahymena* (*T. pyriformis, T. pigmentosa, T. thermophila* and *T. tropicalis*; Piccinni et al., 1999; Shang et al., 2002; Boldrin et al., 2003; Fu and Miao, 2006; Díaz et al., 2007, Shuja and Shakoori, 2002). According to their phylogenetic relationships, the pattern of clustering of Cys residues, and which metal inducer they preferentially respond to, they were grouped into Cd-MTs (subfamily 7a) or Cu-MTs (subfamily 7b), showing little sequence identity with one another. The *Tetrahymena* isoforms so far characterized show unusually longer peptide chains than almost all MTs reported in other organisms and showed only limited similarities with a unique internal homology (Boldrin et al., 2003; Díaz et al., 2007). Previous studies indicated that the two different *Tetrahymena*-MT polypeptides carry out different biological functions in the process of metal homeostasis and detoxification (Boldrin et al., 2002). However, the function of the specific MT isoforms in different *Tetrahymena* species is not fully clarified. Thus, identification and characterization of new *Tetrahymena* metallothionein isoforms will provide important information about the functional diversity and evolution within both unusual MT subfamilies.

In this study, we cloned a novel MT gene from *T. pigmentosa*, named *Tpig-MT1*. It possesses the CCC and CC clusters found in a typical *Tetrahymena* Cd-MT. The expression level of *Tpig-MT1*, as
determined by real time quantitative PCR, was elevated to different extents by exposure to various heavy metals (Cd, Cu, Zn, Hg, Pb) or a reactive oxygen species (H₂O₂). The genetic diversity of Tetrahymena MT genes is also discussed.

2. Materials and methods

2.1. Cell culture and genomic DNA isolation

*Tetrahymena pigmentosa* (strain HG2) was grown axenically at 27 °C in 2% proteose peptone (Difco) and 0.1% bacto-yeast extract. Extraction of genomic DNA was according to Gaertig et al. (1994). Cells grown in mid log phase (3-5 x 10⁷ cells per ml) were starved in 10 mM Tris–HCl (pH 7.5) for 18–24 h at 3 x 10⁷ cells per ml. 25 ml of cells were spun down at 3000 rpm for 5 min and resuspended in 3.5 ml lysate buffer (42% urea, 0.35 M NaCl, 0.01 M Tris pH 7.4, 0.01 M EDTA, 1% SDS). The cell pellet was gently shaken until it was homogeneous, the lysate was extracted twice with an equal volume of phenol: chloroform: isoamyl alcohol (24:1). The lysate was spun for 15 min between steps. The DNA concentration was determined using a DU 800 spectrophotometer (Beckman Instruments Inc., Fullerton, USA). Cell density was determined by using a counting chamber.

2.2. Standard PCR reactions and PCR product sequencing

A pair of primers (MT-1 Fw, MT-1 Re) (Table 1) was used to amplify the coding regions of *T. pyriformis*, *T. pigmentosa* and *T. thermophila* MTs. The total volume of PCR reactions was 25 μl containing 1X Taq buffer with KCl–MgCl₂, 2.0 mM MgCl₂, 5 mM dNTPs, 50 pmol of each primer, 0.75 U Taq DNA polymerase (Fermentas) and 100 ng of genomic DNA. PCR reaction conditions were: 10 min at 94 °C, 35 cycles of 20 s at 94 °C, 20 s at 46 °C and 40 s at 72 °C; and 72 °C for 10 min. The PCR products were analyzed by 1.5% agarose gel electrophoresis and one band was obtained. The amplified product was purified with a Glass Milk kit (Biostar, Canada) and sequenced.

2.3. RNA isolation, cDNA synthesis and 5’ 3’ RACE

For RNA extraction, 10 ml exponential phase cells grown in the same medium were treated with CdCl₂·2.5H₂O at a final non-toxic concentration of 36 μM for 50 min. Cells were harvested by centrifugation at 5000 × g for 10 min and total RNA was isolated using TRIzol reagent (Invitrogen, USA). RNA purity was determined by the A₂₆₀/₂₈₀ ratio (1.9–2.0) and its integrity was judged by visualization of rRNA in ethidium bromide-stained gels. Poly (A)⁺ RNA was purified from total RNA using PolyAtract mRNA Isolation Systems (Promega, Madison, WI), and the cDNA from about 0.5 μg of this Poly (A)⁺ RNA was synthesized by long-distance PCR (LD PCR), using a SMART cDNA Library Construction Kit (Clontech Laboratories, Inc.) following the manufacturer’s instructions.

For amplification of the 5’ cDNA ends, 3’ RACE (Rapid Amplification of cDNA Ends) was performed using a pair of primers (Smart L, SMART R; Table 1). Smart L was synthesized according to the Smart IV™ Oligonucleotide (Table 1) and SMART R was designed on the basis of the coding sequence of obtained from the 140-bp fragment described above. For 3’ cDNA ends analysis (3’ RACE), the pair of primers (3’UTR2, SMART R) were employed. Smart R was designed according to CDSIII/3PCR Primer (Table 1), and 3’UTR2 was designed according to the sequence of the 140-bp fragment. Both Smart L and Smart R primers were used for the first-strand full-length cDNA synthesis, while SMART R and 3’UTR2 were gene-specific primers. PCR reaction was performed as follows: 94 °C for 10 min, 35 cycles (94 °C for 15 sec, 58 °C for 15 s and 72 °C 15 s), and 72 °C 15 min in a 25 μl PCR reaction volumes.

2.4. DNA cloning

RACE PCR products were gel purified as described above, ligated into pGEM T-Vector (Promega, Madison, WI), and transformed into E. coli DH5α cells. At least 3 positively screened clones from each RACE reaction were sequenced with an ABI model 377 Stretch automated DNA sequencer (PE Applied Biosystems).

2.5. Gene expression analysis by real-time quantitative PCR

Exponential phase cells were treated for 1 h with each of the following five heavy metals at sublethal concentrations: Cd 44.5 μM (CdCl₂·2.5H₂O), Hg 10 μM (HgCl₂), Cu 630 μM (CuCl₂·2 H₂O), Zn 500 μM (ZnCl₂), Pb 603 μM (Pb(NO₃)₂). The heavy metals were dissolved in sterile double-distilled water. For H₂O₂ treatment (500 μM), the solution was prepared immediately before the brief exposure times (15, 30 min). Total RNA of *T. pigmentosa* in each treatment was extracted using TRIzol reagent (Invitrogen, USA), digested by DNase (Invitrogen) and then were reverse transcribed into double strand cDNA using M-MLV reverse transcriptase RNaseH− (TOYOBO) and Random Primers. EvaGreen™ (Biotium) was used as the fluorescent dye for real-time quantitative PCR on a Chromo 4 96-well reactor with optical caps (M Research). Primers of both MT target gene (HG2Q-MT L, HG2Q-MT R) and 18S rRNA gene (18S L, 18S R), which was used as the non-regulated reference in the normalization of expression data, are shown in Table 1. The pair of primers (HG2Q-MT L, HG2Q-MT R) was designed from the most variable region of the Tpig-MT1 gene sequence and produced a single 120 bp product. The *T. pyriformis* 18S ribosomal RNA gene was used as the non-regulated reference; primers 18S L and 18S R resulted in a single 185 bp product. Quantitative real-time PCR reactions were performed in a final 20 μl volume containing 1X EvaGreen™, 0.25 μM primers and 10 ng cDNA. The cycling conditions were 7 min at 94°C, 40 cycles×(1 s at 94 °C, 15 s at 56 °C, and 20 s at 72 °C, 1 s at 77 °C (plate read for Tpig-MT1) or 1 s at 80 °C (plate read for 18S)), and 10 min at 72 °C. Then a melting
curve of PCR products (65–95 °C, with a heating rate of 0.3 °C/s) was performed to ensure the absence of artifacts. The data from triplicate reactions performed for each cDNA sample were analyzed using the Opticon Monitor 2 software (MJ Research). To evaluate the amplification efficiency for validation purposes, four-fold serial dilutions of cDNA over a range of 80 ng to 0.32 ng were carried out and the efficiency requirement was met for the 18S rRNA gene and the target gene (differ <10%). The relative expression data was normalized against 18S RNA transcript in each Q-PCR experiment and calculated as relative induction value (n-fold) with respect to the untreated control using the Relative Expression Software Tool, REST (Pfaffl et al., 2002), based on mean threshold cycle (Ct) differences between the sample and the control group. The expression ratios results under the investigated conditions were tested for significance by a pair-wise fixed reallocation randomization test and were considered to be significant when p < 0.05 (Pfaffl et al., 2002).

2.6. Sequence alignment and data analysis

The nucleotide and deduced amino acid sequence alignments used the ClustalX program (Thompson et al., 1997). The program BLAST 2 on the NCBI web site (Tatusova and Madden, 1999) was used for calculating sequence identity over the entire lengths of the gene isoforms. Phylogenetic trees were constructed using MEGA 4.0 program with the neighbor-joining method (NJ) (Tamura et al., 2007). The numbers at the nodes represent the bootstrap values for 1000 replicates produced by the Poisson correction model with pairwise deletion gaps. The nucleotide sequences used in this paper are available from the GenBank under the following accession numbers: 18S rRNAs: X56167; T. pigmentosa MT1; X56171; T. thermophila MTT1; X56172; T. hegewischi; X56173; T. empidokyrea; X56174; T. hegewischi; X56175; Tetrahymena atlantis; X56167; Tetrahymena borealis; M98020; Tetrahymena caligata; X56176; Tetrahymena caligata; X56177; Tetrahymena empidokyrea U3622; Tetrahymena hyperangulata; X56173; Tetrahymena melanocentrus M26360; Tetrahymena patula; X56174; Tetrahymena pigmentosa M26358; Tetrahymena pyriformis; X56171; Tetrahymena thermophila; M10932; Tetrahymena tropicalis X56168; Lamberonnella sp; AF364043; Tetrahymena thermophila; M10932; Tetrahymena tropicalis X56168; Lamberonnella sp; AF364043; Tetrahymena thermophila; M10932; Tetrahymena tropicalis X56168; Lamberonnella sp; AF364043; Tetrahymena thermophila; M10932; Tetrahymena tropicalis X56168; Lamberonnella sp; AF364043; Tetrahymena thermophila; M10932; Tetrahymena tropicalis X56168; Lamberonnella sp; AF364043; Tetrahymena thermophila; M10932; Tetrahymena tropicalis X56168; Lamberonnella sp; AF364043; Tetrahymena ther

Fig. 1. PCR amplification of Cd-MT from the genomic DNA of different Tetrahymena strains using one pair of primers (MT-1 Fw, MT-1 Re). Products were amplified from the TpMT-1 and TpMT-2 genes of T. pyriformis and the MT1 gene of T. thermophila. No product could be amplified from T. pigmentosa DNA. Lanes: 1 = T. pyriformis GL; 2 = T. thermophila B5; 3 = T. pigmentosa HG2; M = molecular weight markers.

Surprisingly, this pair of primers, which was designed to amplify the almost identical MT genes from T. pyriformis (AF050800), T. thermophila (AF479587) and T. pigmentosa (AF509328), only gave the expected size products with T. pyriformis DNA. Another pair of primers (TpMt-1L, TpMt-1R) was constructed to amplify a fragment of the Cd-MT genes and a PCR product about 140 bp was found in each genomic DNA of the above three Tetrahymena species by standard PCR. The 140-bp DNA fragment isolated from T. pigmentosa was sequenced. This sequence of deduced amino acids is cysteine-rich, which revealed that it was MT-like new protein. Gene-specific primers (5UTR1 and 3UTR2) for 5’ RACE and 3’ RACE were designed based on this 140 bp sequence from T. pigmentosa. The full-length cDNA sequence of Tp-MT1 gene (GenBank accession No. EU420056) was then obtained using both 5′ and 3′ RACE (Rapid Amplification of cDNA Ends) by reverse-transcription PCR. The resulting sequence was 521 bases long including as 357-bp open reading frame (ORF), a 73-bp 5’-untranslated regions (UTR) and a 91-bp 3’-UTR (Fig. 2).

Both 5’-UTR and 3’-UTR are A-T rich regions. The 3’-UTR does not include the conserved consensus polyadenylation signal element AATAAA used by eukaryotic organisms. Instead, another putative signal AATAATA at nt 457-462 was identified which is located beyond the TGA stop codon, 59 nt before the site of poly A tail addition. The 357 bp ORF encodes a predicted protein of 118 amino acids (aa) containing 34 Cys residues, organized in 4 Cys-Cys-Cys, 6 Cys-Cys, 2 Cys-X-Cys, 2 Cys-X-Cys motifs (Fig. 2).

3.2. Sequence alignment

Cd-MTs in Tetrahymena can be divided into several regions according to internal similarity. The previous report showed that MT1 in T. pyriformis contained two regions of internal similarity, MT1 (a) and MT1 (b), which have 50.8% identity (Piccinni et al., 1999; Boldrin et al., 2003). According to this criterion all reported Cd-MTs can be divided up into several parts, and multiple sequence alignments among each part of all reported Cd-MTs with Tp-MT1 revealed that it also contains two similar regions. The two parts of Tp-MT1 have 72.9% identity and have more similarity to T. pyriformis MT (a), 69%, than MT (b) 53%; they have been labeled TpMT-1 and TpMT-2 (ORF), a 73-bp 5’-UTR and 3’-UTR (Fig. 3A). The two portions of Tp-MT1 share a common sequence pattern summarized as: x2KxTCCCGx2AKPCxxDPNSGCCxSKxMNKxCDKxDPNSGCC, where x indicates a non-conserved amino acid residue (Fig. 3B). Tp-MT1 (A1) and Tp-MT1 (A2) contain a slightly different
conserved sequence, C2XaCCXXXCXXCQ, but share significant sequence similarity with other *Tetrahymena* MTs (Fig. 4).

### 3.3. Gene expression of Tpig-MT1 gene determined by real-time quantitative PCR

*Tpig-MT1* gene expression changes induced by a short exposure (1 h) to five different heavy metals and H2O2 (15, 30 min) were determined in separate experiments by real-time quantitative PCR using Eva-Green chemistry. The specificity of the *Tpig-MT1* target gene and 18S rRNA reference gene were confirmed by agarose gel electrophoresis. The PCR amplification efficiencies were 1.93 for the 18S rRNA and 1.94 for the *Tpig-MT1* (Table 2), which validate the expression calculations using REST (Fig. 5).

The expression measurement is impacted mainly by the real-time PCR amplification efficiency (E). REST requires the same or similar E between target gene and reference gene. To generate the data basis for the determination of each transcript E, we use a serial dilution of 80, 20, 5, 1.25, 0.31 (ng/μl) each cDNA of *Tpig-MT1* and 18S (Table 2).

All inducers, including essential metals such as Cu (630 μM) and Zn (500 μM), the well known toxicants Cd (44.5 μM), Hg (10 μM) and Pb (603 μM), and the oxyl-radicals H2O2 (500 μM), were tested at sublethal concentrations and stimulated to various extents the expression of the *Tpig-MT1* gene (Fig. 5). Exposure to Hg elicited the most effective induction (193-fold) of *Tpig-MT1* mRNA levels, followed by Pb (176-fold). Cd elicited an intermediate level induction (55-fold), while Cu...
(20-fold) and Zn (5.9-fold) induced Tpig-MT1 to a lesser degree. The induction effect of H\textsubscript{2}O\textsubscript{2} was 5-fold after the very short time (15 min), but decreased to a non-significant level after 30 min (data not shown).

4. Discussion

4.1. The characterization of Tpig-MT1 gene

Cd-MTs have been previously reported in T. pyriformis, T. pigmentosa, T. thermophila and T. tropicalis (Piccinni et al., 1994, 1999; Shang et al., 2002; Boldrin et al., 2003; Fu and Miao, 2006; Diaz et al., 2007, Shuja and Shakoori, 2007). The Tpig-MT1 gene is unusually long (521 bp), uses TGA as the stop codon, and is intronless. The deduced amino acid sequence of the protein reveals a high degree of internal similarity, as seen for other Tetrahymena MTs (Boldrin et al., 2003; Diaz et al., 2007).

Tpig-MT1 gene 3′UTR does not contain the canonical poly (A) signal AATAAA (Fig. 2). The putative polyadenylation signals AATAAA and ATAAA were identified in the MT-1 of T. pyriformis (Boldrin et al., 2003). The AATAA at nt 457-462 in the 3′UTR of Tpig-MT1 gene, 56 nt before the poly A tail, differs by only one residue from the usual eukaryotic organisms polyadenylation signal and has not been shown in other MT genes. A "TTTT" motif reported in other Tetrahymena MT genes (Piccinni et al., 1999; Santovito et al., 2001; Dondero et al., 2004; Fu and Miao, 2006), which was suggested to function in rapid degradation of certain short-lived mRNAs as revealed in mammals, was also found in Tpig-MT1 gene 3′UTR. However, since the average GC content of the Tetrahymena genome is 22% and intergenic sequences tend to be even more AT-rich (Eisen et al., 2006), such AT-rich motifs are easily found throughout the genome. So far, none of those motifs has been experimentally demonstrated to have a functional role in any gene in any species of this genus, and thus remain hypothetical (Diaz et al., 2007).

Tpig-MT1 contains 34 cysteines, organized in 4 Cys-Cys-Cys, 6 Cys-Cys, 2 Cys-X-Cys, 2 Cys-X-Cys-Cys motifs. Thus, it contains the well conserved CCC and CC clusters found in the five previously reported members of the Tetrahymena MT subfamily 7a (Diaz et al., 2007). The common structural pattern GTXXXCKXCKC found in TpMT-1, MTT1, and MT-2 genes of the three Tetrahymena species was proposed to be a conserved sequence of Tetrahymena MTs (Boldrin et al., 2003). However, this sequence is not present in the deduced amino acid sequence for Tpig-MT1; instead a slightly longer sequence, GDXXXXCKXCKC, is found.

The length of Tpig-MT1 chain is 118 residues, and the ORF (open reading frame) could be divided into two parts according to internal similarity, showing 72.9% identity (Fig. 3A), from codons 1 to 59 and from 60 to 118. The TpMT-1 of T. pyriformis can be divided into two portions (MT A and MT B) on the basis of the internal similarity (Piccinni et al., 1994). According to this criterion, all of the reported Cd-MT isoforms (except TMcd1, which does not have a complete cDNA) contain both MT (A) and MT (B) (Fig. 4, 6). When compared to MT (A) and MT (B), the two parts of Tpig-MT1 are the same length as MT (A), 59 residues. They also share more similarities with MT (A) (69% identity) than MT (B) (53% identity), and will subsequently be described as Tpig-MT1 (A1) and Tpig-MT1 (A2). Multiple sequence alignments show that MT (B) is present in all the other Tetrahymena Cd-MTs, but is absent from Tpig-MT1 (Fig. 4).

Fig. 6. Phylogenetic distance tree for tetrahymenid ciliates inferred from small subunit ribosomal DNA sequences using MEGA 4.0 with the neighbor joining method (NJ). The nucleotide sequences used are available from the GenBank databases referred by Strüder-Kypke et al. (2001). Tetrahymena species with sequenced MTs are shown marked with red triangles. The MTs are diagrammed on the right using rectangles. White color portion represents module “A” and black color represents module B. The broken line represents the unsequenced portion of MT A for T. tropicalis.
4.2. Induction of Tpig-MT1 transcription by heavy metals and hydrogen peroxide

The intronless MT genes found in three *Tetrahymena* species may support the idea that intronless MT genes may allow a rapid transcriptional response upon exposure to heavy metals (Leignel et al., 2005). The biological functions of the ciliate MT can be inferred by gene expression analysis (Dondero et al., 2004), albeit indirectly. *Tetrahymena* MTs are proposed to function in both heavy metal homeostasis and detoxification, and different isoforms exhibit unique expression patterns in response to exposure to different metals (Boldrin et al., 2002). Previous studies indicated that *Tetrahymena* Cd-MTs are induced to the highest level of expression by Cd (Boldrin et al., 2002; Fu and Miao, 2006). In this study, the transcriptional activation of *Tpig-MT1* mRNA was determined after a short exposure (1 h) to different heavy metal. The time was chosen according to the reported plateau of expression for Cd-MT (Santovito et al., 2001; Shang et al., 2002; Fu and Miao, 2006). The concentration of the various heavy metals was in a sublethal range and did not affect normal cell growth rates (unpublished observations).

All heavy metals were able to stimulate *Tpig-MT1* mRNA in this experiment, but showed different expression levels. Interestingly, Hg elicited the most effective induction (Fig. 5), while previous reports had indicated that, after a short exposure to various heavy metals, Cd was the most effective inducer of Cd-MTs in *T. pyriformis* (Santovito et al., 2000; Fu and Miao, 2006, Diaz et al., 2007). Hg can stimulate MT mRNA (*MTT1*, *TgMT-1* and *TgMT-2*) synthesis but it was not as effective as Cd (Boldrin et al., 2002; Fu and Miao, 2006). However, in this study Hg and Pb resulted in higher *Tpig-MT1* mRNA levels than Cd. It is proposed that in mammals, mollusks and fish, MTs play a key role in Hg or Cd detoxification, but Hg or Cd stimulated different inductions of MT mRNA (Bourdinadea et al., 2006; Krishna et al., 2006). In the black scabbard fish, there are some relationships between MT and Hg concentrations, although this relationship is tissue dependent (Bebianno et al., 2007). MT or MT-like proteins were reported to be induced by Pb in rats, human, and fish (Ikebuchi et al., 1986; Church et al., 1993; Cheung et al., 2005). The *MTT5* gene in *T. thermophila* showed stronger induction to Pb than Cd (Diaz et al., 2007). *Tpig-MT1* mRNA was expressed at its highest level after a short exposure to either Hg or Pb, not Cd (Fig. 5). Thus, a similar biological function might be inferred between *Tpig-MT1* and *MTT5*, although their exact physiological roles are not yet clear.

MTs also play an important role in regulating cellular levels of essential metals as Zn and Cu which are considered as natural MT inducers, or after exposure to ROS species such as hydrogen peroxide (Kägi, 1993). It also has been proposed that different MT isoforms possibly carry out distinct functions during the processes of metal homeostasis and detoxification (Sadhu and Gedamu, 1988; Dondero et al., 2005). In mammals, MTs are speculated to be the predominant single liganding species for Zn in cells that have not been treated with other heavy metals, and they can release Zn in response to induction with Cd or Cu (Suzuki et al., 1990). Cu and Zn also stimulate *Tpig-MT1* mRNA although to lower expression level. The unique transcriptional response of *Tpig-MT1* is not yet understood; more information, including the characteristics of its MRES, will be needed to decipher its gene expression patterns.

4.3. Genetic diversity of MT genes of *Tetrahymena*

The *Tpig-MT1* gene described in this paper is substantially different from the *T. pigmentosa* MT-1 gene (AF509328) reported by Piccinni’s group (Piccinni et al., 1994; Boldrin et al., 2002, 2003). Surprisingly, the *T. pigmentosa* MT-1 gene (AF509328), the *T. thermophila* MT-1 gene (AF479587), and the *T. pyriformis* Cd-induced MT gene (AJ005080) are 99 to 100% identical in nucleotide sequence, including flanking sequences well outside the coding sequence. Based on the genetic distance separating these species (Fig. 6; Nanney et al., 1998), the near identity of these genes is extremely unlikely. Gene-specific primers failed to amplify MT1 from either *T. pigmentosa* DNA or from *T. thermophila* DNA, but did amplify an appropriate sized product from *T. pyriformis* DNA (Fig. 1). This suggests that both the previously described *T. pigmentosa* MT-1 gene (AF509328) and the *T. thermophila* MT-1 gene (AF479587) are actually *T. pyriformis* genes attributed to the other species.

The Cd-MTs in *Tetrahymena* were grouped into subfamily 7a according to the criteria proposed by Diaz et al. (2007). A notably regular and characteristically hierarchical modular organization can be inferred among those gene isoforms based on sequence alignment at the amino acid level. Those different Cd-MT isoforms could be divided up into modules, where every module carries a CXCC motif at its C-terminus. Modules result from the joining of two types of submodules; Type 1 submodule has the sequence C2X6CxX5, while Type 2 submodule can be represented as C2X6CXXCXXCX. The module 1 and module 2 correspond to MT (A) and MT (B), respectively, in this article. *Tpig-MT1* (A1) and *Tpig-MT1* (A2) both have module 1 and lack module 2 (Fig. 4). Both halves of the *Tpig-MT1* (A1 and A2), contain the C2X6CxX5 Type 1 submodule, but their type 2 submodule has a slightly modified sequence C2X6CXCCXXCXX. It is suggested that different combinations of the elementary motifs, submodules and modules would probably lead to the final successive subfamily 7a MT formation (Diaz et al., 2007).

In this study, the putative protein sequence of *Tpig-MT1* can be represented as having a repeated “A-A” modular organization and missing a B module. The latter feature makes it rather similar to *T. thermophila* MTT5 and distinguishes it from the other full-length subfamily 7a MT genes that have been identified in the three *Tetrahymena* species, which end with the “B” module. The major difference between module A and B is that B lacks an elementary motif (C2X6X) within the type 2 submodule. It has been proposed previously that the genetic mechanism of replication slippage during gene duplication events, at the submodule level or through the well-conserved motifs, may have given rise to the intragenic repeats and resulted in the formation of *Tetrahymena* subfamily 7a MT genes (Boldrin et al., 2003; Fu and Miao, 2006). However, the detailed evolutionary history of this category of MTs is difficult to reconstruct due to both the possible paralog recombination events and the force of convergent evolution on structural subunits of the protein under functional pressure (Diaz et al., 2007). The primary sequence of *Tpig-MT1* gene provides new evidence that the formation of *Tetrahymena* subfamily 7a MTs genes might be caused by “large” structure units, and module A and B may already exist respectively as the ancestral subfamily 7a MT component. We hypothesize that a *Tetrahymena* subfamily 7a MT gene composed of only module B could exist (now or in the past). Further investigation is needed to understand the genetic diversity of MT genes in the various *Tetrahymena* species.

Acknowledgements

The authors wish to thank Prof. Eduardo Orias (University of California at Santa Barbara) for kindly providing the cells of *Tetrahymena pigmentosa* (strain HG2) and his critical review of the manuscript, Judy Orias (University of California at Santa Barbara) for improving the English of the manuscript, Prof. Fukang Gu (East China Normal University, Shanghai, China) for *Tetrahymena thermophila* (strain Bf5), and two anonymous reviewers for their valuable comments and improvements of the manuscript. This work was supported by: National Natural Science Foundation of China (No.30670298) and Projects of International Cooperation and Exchanges Ministry of Science and Technology of China (No.2006DFA31960) to W. Miao.
References


