Identification and Sequence Analysis of Metazoan tRNA 3’-End Processing Enzymes tRNase Zs

Zhikang Wang, Jia Zheng, Xiaojie Zhang, Jingjing Peng*, Jinyu Liu, Ying Huang*
Jiangsu Key Laboratory for Microbes and Genomics, School of Life Sciences, Nanjing Normal University, Nanjing, China

Abstract
tRNase Z is the endonuclease responsible for removing the 3’-trailer sequences from precursor tRNAs, a prerequisite for the addition of the CCA sequence. It occurs in the short (tRNase ZS) and long (tRNase ZL) forms. Here we report the identification and sequence analysis of candidate tRNase Zs from 81 metazoan species. We found that the vast majority of deuterostomes, lophotrochozoans and lower metazoans have one tRNase ZS and one tRNase ZL genes, whereas ecdysozoans possess only a single tRNase ZL gene. Sequence analysis revealed that in metazoans, a single nuclear tRNase ZL gene is likely to encode both the nuclear and mitochondrial forms of tRNA 3’-end processing enzyme through mechanisms that include alternative translation initiation from two in-frame start codons and alternative splicing. Sequence conservation analysis revealed a variant PxKxRN motif, PxPxRG, which is located in the N-terminal region of tRNase ZSs. We also identified a previously unappreciated motif, AxDx, present in the C-terminal region of both tRNase ZSs and tRNase ZLs. The AxDx motif consisting mainly of a very short loop is potentially close enough to form hydrogen bonds with the loop containing the PxKxRN or PxPxRG motif. Through complementation analysis, we demonstrated the likely functional importance of the AxDx motif. In conclusion, our analysis supports the notion that in metazoans a single tRNase ZL has evolved to participate in both nuclear and mitochondrial tRNA 3’-end processing, whereas tRNase ZS may have evolved new functions. Our analysis also unveils new evolutionarily conserved motifs in tRNase Zs, including the C-terminal AxDx motif, which may have functional significance.
the protein core, and is involved in the substrate binding primarily located on the opposite side of the active site and protrudes from domain and an extended two-stranded stalk. This flexible arm is flexible arm (also called exosite) consisting of a compact globular protein, which lacks the GP-motif. The number of tRNase Z genes in basal metazoans examined so far seems highly variable. The freshwater hydra *Hydra magnipapillata*, a representative cnidarian, and the marine demosponge *Amphimedon queenslandica* (phylum Porifera) each contain one tRNase Z S and one tRNase Z L genes. In the starlet sea anemone *Nematostella vectensis*, another representative cnidarian, one tRNase Z S and two tRNase Z L genes were identified. *Trichoplax adhaerens*, the sole representative of the phylum Placozoa, represents the simplest known animal with the smallest known animal genome. In this species, only one tRNase Z L gene was found, but no tRNase Z S genes were identified. In the choanoflagellate *Monosiga brevisolit*, which is the closest known unicellular relative to metazoans, two tRNase Z genes (one tRNase Z S gene and one tRNase Z L gene) were identified.

### Results

#### Identification of Candidate Metazoan tRNase Z Genes

As part of our effort to explore genomic distribution and sequence conservation of metazoan tRNase Zs, we performed BLAST similarity searches of metazoan genomic and EST databases. We obtained a total of 45 candidate tRNase Z S genes and 71 tRNase Z L genes from 82 species, including 77 bilaterians, 4 nonbilaterian metazoans, and 1 choanoflagellate (Table S1). It should be noted that tRNase Z L S was identified in all metazoan species examined, although the protein sequences of some tRNase Z L S could not be accurately predicted due to gaps in their nucleotide sequences. Although the number of basal metazoan and protozoan species that have had their complete genome sequences reported is very limited, the inclusion of these species in our analysis would enable us to have a better understanding of the origin and evolution of metazoan tRNase Zs. Among these candidate metazoan tRNase Z genes identified, only those from *D. melanogaster* [27,28], *C. elegans* [29] and humans have been experimentally characterized. To present the data in a concise manner, we have selected a list of tRNase Zs of representative species spanning a diverse range of taxonomic groups (Table 1).

It appears that the tRNase Z gene is variably distributed within the three major divisions of Bilateria (the Deuterostomia, the Ecdysozoa, and the Lophotrochozoa). All of the vertebrate deuterostomes examined to date have one tRNase Z S and one tRNase Z L genes. Similarly, four invertebrate deuterostomes, the amphioxus *Branchiostoma floridae* (cephalochordate), two sea squirts *Ciona intestinalis* and *Ciona savignyi* (urochordate), and the sea urchins *Strongylocentrotus purpuratus* (echinoderm) have one tRNase Z S and one tRNase Z L genes.

Unlike deuterostomes, protostomes contain variable numbers of tRNase Zs. Ecdysozoans, one of two major groups of protostomes and including animals such as arthropods and nematodes, appear to possess a single tRNase Z L gene but lack the tRNase Z S gene. Lophotrochozoa, which is another major group of protostomes and comprises annelids and mollusks, is significantly underrepresented in terms of whole genome data. Among sequenced lophotrochozoans, the marine mollusk *Lottia gigantea* appears to have the largest number of tRNase Zs: one tRNase Z S and two tRNase Z L. Like deuterostomes, the marine mollusk *Aplysia californica* and the two parasitic flatworms (*Schistosoma mansoni* and *Schistosoma japonicum*) harbor one tRNase Z S and one tRNase Z L genes. Like ecdysozoans, the two marine annelids *Helobdella robusta* and *Capitella teleta* appear to have one tRNase Z L gene, but have no tRNase Z S gene. In addition, *C. teleta* has one tRNase Z L-like protein, which lacks the GP-motif.

The prediction of the number of introns in metazoan tRNase Z S and tRNase Z L genes

To understand the structure of metazoan tRNase Z genes, we predicted their introns. Metazoan tRNase Z L genes appear to have few introns. For example, all mammalian tRNase Z S genes examined have just two introns (Table S2). Mammals and invertebrate tRNase Z S’s seem to have different numbers of introns. All mammalian tRNase Z L S genes examined contain 23 introns. In contrast, tRNase Z L genes of 3 Cnidaria species including *C. elegant* and 9 Drosophilida species including *D. melanogaster* contain 6 and 1 introns, respectively. The difference in intron numbers between mammalian and invertebrate tRNase Z L genes can be ascribed to massive, lineage-specific intron loss in worms and insects [30]. Interestingly, the intron found in *Drosophila* tRNase Z L genes splits the gene into two fragments that encode roughly the N- and C-terminal halves of the protein. This is consistent with the notion that proteins are often made from independently-folded functional domains encoded by separate exons, which can be separated by introns [31,32]. Unexpectedly, we found that the ascidian *C. intestinalis* tRNase Z L S gene encoding a protein of 727-aa apparently does not have any intron by comparing the gene sequence with corresponding EST sequences in NCBI databases.

### Alternative Splicing of tRNase Z L Transcripts

Alternative splicing can play a major role in modulating gene structure and function. To identify potential alternative spliced isoforms of tRNase Z L S, we performed BLAST searches against the NCBI EST database. Candidate splice variants of tRNase Z L S were largely found in vertebrates, consistent with the observation that vertebrates have substantially higher rates of alternative splicing compared to other species. Figure 1 shows a schematic representation of predicted alternatively spliced variants of metazoan tRNase Z L S. The common mRNA set (*Callithrix jaccus*) appears to have three alternatively spliced isoforms of tRNase Z L S, designated CjaTRZ2i1, CjaTRZ2i2 and CjaTRZ2i3. CjaTRZ2i1 [mRNAs] is created by using an
alternative splice donor within exon 13. The use of this alternative acceptor site (within tRNase Z L exon) causes a deletion of V703. CjaTRZ2i2 uses an alternative donor site within exon 1 and an acceptor site within exon 2 that results in a deletion of 57 nt. CjaTRZ2i3 is generated by skipping of exon 7 that causes an internal 120-nt deletion.

Table 1. Distribution of candidate tRNase Zs from representative metazoans.

| Species                  | Protein name | Form      | Accession number | Database       | No. aa | EST |
|--------------------------|--------------|-----------|------------------|----------------|--------|-----|
| **Deuterostomes**        |              |           |                  |                |        |     |
| Branchiostoma floridae  | BflTRZ1      | tRNase Z  | XP_002598072.1   | NCBI           | 367*   | +   |
| Branchiostoma floridae  | BflTRZ2      | tRNase Z  | XP_002603583.1   | NCBI           | 785*   |     |
| Canis familiaris         | CfaTRZ1      | tRNase Z  | XP_849493.1      | NCBI           | 374    |     |
| Canis familiaris         | CfaTRZ2      | tRNase Z  | XP_546630.2      | NCBI           | 821    | +   |
| Ciona intestinalis       | CinTRZ1      | tRNase Z  | 299192           | Metazome       | 383    |     |
| Ciona intestinalis       | CinTRZ2      | tRNase Z  | 295609           | Metazome       | 727    |     |
| Danio rerio              | DreTRZ1      | tRNase Z  | NP_001003503.1   | NCBI           | 372    |     |
| Danio rerio              | DreTRZ2      | tRNase Z  | XP_003198113.1   | NCBI           | 895    | +   |
| Homo sapiens             | HsaTRZ1      | tRNase Z  | NP_061166.1      | NCBI           | 363    |     |
| Homo sapiens             | HsaTRZ2      | tRNase Z  | NP_060597.4      | NCBI           | 826    |     |
| Loxodonta africana       | LafTRZ1      | tRNase Z  | ENSLAFP0000001283| Ensembl       | 364    |     |
| Loxodonta africana       | LafTRZ2      | tRNase Z  | ENSLAFP00000025345| Ensembl       | 819    |     |
| Mus musculus             | MmuTRZ1      | tRNase Z  | NP_444485.2      | NCBI           | 362    |     |
| Mus musculus             | MmuTRZ2      | tRNase Z  | CAI24609         | NCBI           | 824    |     |
| Salmo salar              | SsaTRZ1      | tRNase Z  |                  | NCBI           | 370    |     |
| Salmo salar              | SsaTRZ2      | tRNase Z  | COH8T6           | Uniprot       | 875    |     |
| Strongylocentrus purpuratus| SpuTRZ1  | tRNase Z  | XP_001189573.1   | NCBI           | 362    |     |
| Strongylocentrus purpuratus| SpuTRZ2 | tRNase Z  | XP_783904.2      | NCBI           | 931*   | +   |
| Xenopus tropicalis       | XtrTRZ1      | tRNase Z  | 476074           | Metazome       | 363    |     |
| Xenopus tropicalis       | XtrTRZ2      | tRNase Z  | 157533           | Metazome       | 835*   | +   |
| **Protostomes**          |              |           |                  |                |        |     |
| Bombyx mori              | BmoTRZ1      | tRNase Z  | Bmb008309        | Metazome       | 889    |     |
| Caenorhabditis elegans   | CelTRZ1      | tRNase Z  | NP_001023109.1   | NCBI           | 833    |     |
| Capitella teleta         | CteTRZ1      | tRNase Z  | 228344           | JGI            | 758    |     |
| Capitella teleta         | CteTRZ2      | tRNase Z  | 218896           | JGI            | 738    |     |
| Drosophila melanogaster  | DmeTRZ1      | tRNase Z  | NP_724916.1      | NCBI           | 766    |     |
| Lottia gigantea          | LgiTRZ1      | tRNase Z  | 218281           | JGI            | 356    |     |
| Lottia gigantea          | LgiTRZ2      | tRNase Z  | 156489           | JGI            | 738*   | +   |
| Lottia gigantea          | LgiTRZ3      | tRNase Z  | 230203           | JGI            | 780    |     |
| Tribolium castaneum      | TcaTRZ1      | tRNase Z  | XP_968692.2      | NCBI           | 827    |     |
| **Basal Metazoans**      |              |           |                  |                |        |     |
| Amphimedon queenslandica | AquTRZ1      | tRNase Z  | Aqu1.217211       | JGI            | 373*   |     |
| Amphimedon queenslandica | AquTRZ2      | tRNase Z  | XP_003385508.1   | NCBI           | 712*   |     |
| Hydra magnipapillata     | HmaTRZ1      | tRNase Z  | Hma2.228140      | JGI            | 359    |     |
| Hydra magnipapillata     | HmaTRZ2      | tRNase Z  | Hma2.206546      | JGI            | 723*   |     |
| Nemastostella vectensis  | NveTRZ1      | tRNase Z  | 186754           | JGI            | 342    |     |
| Nemastostella vectensis  | NveTRZ2      | tRNase Z  | 11980           | JGI            | 827*   |     |
| Nemastostella vectensis  | NveTRZ3      | tRNase Z  | 108544           | JGI            | 769*   | +   |
| Trichoplax adhaerens     | TadTRZ1      | tRNase Z  | B3RMG1           | Uniprot       | 728*   |     |
| **Protozoans**           |              |           |                  |                |        |     |
| Monosiga brevicollis     | MbrTRZ1      | tRNase Z  | 25108            | JGI            | 380*   |     |
| Monosiga brevicollis     | MbrTRZ2      | tRNase Z  | 23750            | JGI            | 812*   |     |

*The number of amino acids in metazoan tRNase Z proteins.
*Indicates that mispredicted sequences obtained from the databases have been corrected.
A plus sign indicates that gene expression is confirmed by the presence of ESTs.
doi:10.1371/journal.pone.0044264.t001
Schematic representation of the predicted alternative splice variants of mammalian and nematode tRNase Z genes (drawn to scale). (A) Mammalian tRNase Z genes are from H. sapiens (Hsa), C. jacchus (Cja), C. familiaris (Cfa), O. cuniculus (Ocu), B. taurus (Bta), M. musculus (Mmu), M. brevicollis (Mbr), H. sapiens (Hs), Caenorhabditis briggsae (Cbr), D. melanogaster (Dme), and D. rerio (Dre). Constitutive exons and introns are represented by the filled boxes and intervening horizontal lines respectively. Alternative exons are colored as follows: exon skipping in red, exon with alternative donors in blue and exon with alternative acceptors in green. The locations of the alternative exons in the tRNase Z genes and its splice variants are shown on the top of the exon. (B) Nematode tRNase Zs are from C. elegans (Cel), C. remanei (Cre) and C. briggsae (Cbr). CelTRZ2i1, CelTRZ2i2, CreTRZ2i1, CreTRZ2i2, CbrTRZ2i1, CbrTRZ2i2, MmuTRZ2i2, and RnoTRZ2i2 correspond to alternatively spliced variants. Constitutive exons and introns are represented by the filled boxes and intervening horizontal lines respectively. Alternative exons are colored as follows: exon skipping in red, exon with alternative donors in blue and exon with alternative acceptors in green. The locations of the alternative exons in the tRNase Z genes and its splice variants are shown on the top of the exon. (B) Nematode tRNase Zs are from C. elegans (Cel), C. remanei (Cre) and C. briggsae (Cbr). CelTRZ2i1, CelTRZ2i2, CreTRZ2i1, CreTRZ2i2, CbrTRZ2i1, CbrTRZ2i2, MmuTRZ2i2, and RnoTRZ2i2 correspond to alternatively spliced variants. Constitutive exons and introns are represented by the filled boxes and intervening horizontal lines respectively. Alternative exons are colored as follows: exon skipping in red, exon with alternative donors in blue and exon with alternative acceptors in green. The locations of the alternative exons in the tRNase Z genes and its splice variants are shown on the top of the exon.
Figure 2. Bayesian phylogenetic tree of predicted metazoan tRNase Zs. This analysis is based on the sequence alignment of the full-length metazoan tRNase Z’s with the C-terminal half of metazoan tRNase Z’s. The accession number for each candidate tRNase Z is identified in Table S1. Bayesian posterior probabilities are indicated at the nodes. The scale bar indicates 0.1 nucleotide substitutions per site. Taxonomic designations are indicated on the right side of the tree.

doi:10.1371/journal.pone.0044264.g002
Figure 3. Multiple sequence alignment of representatives of N-terminal halves of the metazoan and non-metazoan tRNase Z’s.

Representative metazoan tRNase Z’s are from ten deuteromes including C. familiaris (Cfa), H. sapiens (Hsa), L. Africana (Laf), M. musculus (Mmu), X. tropicalis (Xtr), D. rerio (Dre), S. salar (Ssa), C. intestinalis (Cin), B. floridae (Bfl), and S. purpuratus (Spy), six protostomes including B. mori (Bmo), D. melanogaster (Dme), T. castaneum (Tca), C. elegans (Cel), L. gigantean (Lgi), and C. teleta (Cte), and four basal metazoans including N. vectensis (Nve), H. magnipapillata (Hma), A. queenslandica (Agu), and T. adhaerens (Tad). Non-metazoan tRNase Z’s are from the unicellular chaoanoflagellate M. brevicollis (Mbr), green plant A. thaliana (Ath) and fungal S. pombe (Spn). See Table 1 for more information. The alignment was constructed using the computer program Clustal W [52]. Identical residues are on a black ground and conserved residues shaded in gray. The conserved motifs of tRNase Zs indicated above the alignment are labeled according to references [17,19,58]. The numbers in brackets indicate the length of the region in the protein, which are species-specific and could not be correctly aligned. Hyphens represent gaps introduced into sequences for maximum alignment.

Figure 4. Amino acid residues predicted to be critical for activity are indicated by a star. Protein, which are species-specific and could not be correctly aligned. Hyphens represent gaps introduced into sequences for maximum alignment.

Figure 5 shows alignments of representative metazoan tRNase Z’s (for a complete list, see Figure S2). Sequence alignment revealed that similar to tRNase Z^6^, tRNase Z^6^ is highly conserved throughout vertebrate evolution (see Figure 5 and Table S5). For example, the human tRNase Z^6^ shares approximately 92% and 57% sequence identity with the mouse and zebrafish tRNase Z^6^, respectively. In contrast, vertebrate tRNase Z^6^ show much lower overall sequence identity to tRNase Z^6^ from protostome invertebrates and basal metazoans. The amino acid sequence identity of the human tRNase Z^6^ with the tRNase Z^6^ from the human parasitic Schistosoma mansoni and starlet sea anemone Nematostella vectensis is 35% and 45%, respectively. The amino acid sequences of vertebrate tRNase Z^6^ are also markedly diverged from those of bacterial tRNase Z^6^’s. The sequence of the human tRNase Z^6^ shares 34% and 47% sequence identity to those from B. subtulis and E. coli, respectively.

Unlike bacterial tRNase Z^6^ and eukaryotic tRNase Z^6^, metazoan tRNase Z^6^’s appear to contain a variant PxKxRN motif (i.e., the PxPxRG motif) that contains Pro and Gly in place of Lys and Asn respectively (Figure 6). The Lys to Pro substitution is expected to increase the rigidity of the PxKxRN loop, whereas the Asn to Gly substitution may decrease the rigidity of the loop.

Careful examination of the sequences of metazoan tRNase Z^6^ revealed an AxDx motif located near the C-terminal region of the proteins (Figure 5). A reexamination of all previously identified tRNase Z^6^ sequences revealed that this motif is also present in the C-terminal regions of tRNase Z^6^’s and bacterial-type tRNase Z^6^’s but not in Thermotoga maritima (TM)-type tRNase Z^6^’s [5,7]. The AxDx motif contains an absolutely conserved Asp residue and a highly conserved Ala residue (Figure 6).

The Effect of the Point Mutation in the AxDx Motif on the Complementation Ability of the S. pombe tRNase Z^-C^-.

To establish the functional significance of the AxDx motif, we employed a Schizosaccharomyces pombe-based functional complementation assay to assess the functional consequence of the point mutation in the conserved AxDx motif of the S. pombe Trz1. As such, we introduced the D772A mutation into the AxDx motif of the S. pombe Trz1 and tested if these mutants could rescue a S. pombe trz1^-ts^ temperature-sensitive (ts) mutant strain trz1^-1^ as described before [37]. As shown in Figure 7, the wild type S. pombe Trz1 could complement the temperature sensitive growth phenotype of trz1^-1^ when expressed from the pREP4X plasmid. On the other hand, the D772A mutant protein showed an impaired ability to complement the temperature-sensitive growth defect of trz1^-1^ based on these results, we concluded that the highly conserved AxDx motif and particularly Asp772 is vital for the S. pombe Trz1 activity and function.

Discussion

The Origin and Evolution of tRNase Z in Metazoans

Our survey of available metazoan genomes revealed that the tRNase Z^6^ gene is ubiquitous across the metazoan kingdom, including nonbilateral metazoans. On the other hand, the tRNase Z^6^ gene does not exist in any of the sequenced ecdysozoan genomes, which is the largest clade of metazoans. The presence of the tRNase Z^6^ and tRNase Z^6^ genes in many higher metazoans such as lophophorozoans and deuteronomes, and in lower metazoans including Porifera and cnidarians indicates that the tRNase Z^6^ and tRNase Z^6^ genes were already present very early in the evolution of metazoans, and that the absence of tRNase Z^6^s in ecdysozoans is likely due to gene loss. This hypothesis is further reinforced by our finding that both tRNase Z^6^ and tRNase Z^6^ genes already existed in the unicellular choanoflagellate M. brevicollis, which is the closest known relative of metazoans. Our findings are consistent with the observations that ecdysozoans have high rates of gene loss [38,39].

There are two possible explanations for the loss of the tRNase Z^6^ gene in ecdysozoans. First, the function of the tRNase Z^6^ gene might largely represent the vertebrate-specific feature underlying the biological complexity of vertebrates. However, the role of the vertebrate tRNase Z^6^ gene has not been fully addressed (see below discussion). Second, the tRNase Z^6^ gene could be an animal ancestral gene that has been lost during Ecdysozoan evolution.

The Role of Metazoan tRNase Zs

In our early bioinformatics studies of fungi and plant tRNase Zs, we concluded that tRNase Z^6^ takes over the essential role of
tRNase Z's in the nuclear and mitochondrial tRNA 3'-end processing. The present bioinformatics analysis for metazoan tRNase Zs extends and reinforces this conclusion.

The conclusion that tRNase Z's is responsible for the 3'-end processing of nuclear and mitochondrial tRNA precursors in metazoa is further supported by the results of both in vitro and in vivo functional studies. In humans, tRNase ZL exhibits tRNA processing efficiency ~2000 fold higher than tRNase ZS in vitro [40]. Recent in vivo studies of tRNase ZL from Drosophila and humans provide direct evidence that tRNase ZL is responsible for endonucleolytic processing of the 3'-end of pre-tRNAs in metazoan nucleus and mitochondria [28]. The Drosophila tRNase ZL is encoded by the juvenile hormone [JH] inducible gene, jhl-1 [41]. RNA interference (RNAi)-mediated silencing of the jhl-1 gene expression in D. melanogaster results in the accumulation of 3'-end unprocessed tRNAs in both the nucleus and mitochondria [28], indicating that tRNase ZL is responsible for generating the mature 3'-end of pre-tRNAs in the nucleus and mitochondria of D. melanogaster. Likewise, in humans, tRNase ZL is the primary enzyme responsible for mitochondrial tRNA 3'-end since silencing of the human tRNase ZL gene inhibits the 3'-end processing of mitochondrial tRNAs [9,11]. However, among some of the mitochondrial tRNAs analyzed, siRNA against the human tRNase ZL gene only affects the mature level of a few mitochondrial tRNAs in humans [9,11].

In Drosophila, knockdown of the tRNase ZL gene significantly decreases the mature levels of the mitochondrial tRNAs but not the mature levels of all five tested nuclear tRNAs [28]. It is not known why the steady-state levels of many tested mature forms of tRNAs are not visibly affected. One possibility is that additional enzymes may also be involved in 3'-end processing of tRNAs.

Besides its primary role in tRNA maturation, the human tRNase ZL functions in other cellular pathways. For example, the human tRNase ZL has been demonstrated to be involved in the biogenesis of non-coding RNAs (ncRNAs), other than tRNAs, such as metastasis associated lung adenocarcinoma transcript 1 (MALAT1) [42]. tRNA-derived small RNAs [43,44] and viral microRNAs (miRNAs) [45,46]. It would be interesting to see if other metazoan tRNase ZLs also have additional functions apart from tRNA 3'-end processing.

So far, the metazoan tRNase ZLs have only been studied in humans and its function is still poorly understood. The human tRNase ZL gene appears to be nonessential since this gene was found to be deleted in a lung cancer cell line, Ma29 [47]. The human tRNase ZL is localized in the cytoplasm and has been proposed to play a role in fine-tuning of gene expression by processing and degradation of unstructured tRNAs [11]. It will be important to determine the RNA substrates of the metazoan tRNase ZLs by various methods including coimmunoprecipitation with an antibody against tRNase ZL, followed by sequencing of associated RNAs.

Mechanisms for Dual Localization of Metazoan tRNase ZLs

How does the metazoan tRNase ZL participate in both the mitochondrial and nuclear tRNA 3'-end processing? A recent study on the subcellular localization of tRNase ZL suggests that mammalian tRNase ZL genes encode both nuclear and mitochondrial forms of tRNase ZL by using alternative translation initiation sites [36]. Our results extend this conclusion to include other metazoan tRNase ZLs. We found that like mammalian tRNase ZL, most non-mammalian metazoan tRNase ZLs also contain a putative downstream AUG start codon in-frame with the first one (Figure S3). In additional, the nucleotide context around the first start codon is apparently suboptimal for translational initiation as defined by the consensus sequences [48,49]. We also found that most metazoan tRNase ZLs analyzed so far are characterized by the presence of the N-terminal MTS. Thus, it is likely that a single metazoan tRNase ZL mRNA is translated via leaky scanning to make proteins with or without an MTS by initiating translation at two in-frame AUG codons that should produce essentially the same proteins, one mitochondrial and the other nuclear. Indeed, changing the nucleotide sequences surrounding the first AUG start codon of the human tRNase ZL to conform to the eukaryotic consensus sequence results in exclusive mitochondrial localization of the protein, whereas removing the N-terminal region of the human tRNase ZL-encoding most of the predicted MTS causes the protein to be targeted exclusively to the nucleus [36].

Besides leaky scanning, other mechanisms including alternative splicing [50] may also be used to produce the nuclear and mitochondrial forms of the metazoan tRNase ZL from the same gene. Examination of the tRNase ZL genes from three Caenorhabditis nematodes revealed that all of the candidate genes contain a second putative in-frame start codon located in exon 3 (Figure 1). The two AUG start codons are separated by 406 nt (C. remanei tRNase ZL), 1248 nt (C. elegans) and 413 nt. (C. briggsae), respectively. These distances seem too long for efficient leaky scanning [51]. In C. elegans, RT-PCR analysis showed that the full length tRNase ZL and its splice variants that differ by the presence or absence of a N-terminal putative MTS are likely generated by alternative splicing of 5' exons [29]. Thus, it is possible that both the nuclear and mitochondrial forms of the Caenorhabditis tRNase ZL genes could be translated from a single mRNA species through an alternative splicing mechanism.

The Role of the AxDx Motif

Our sequence analysis revealed an AxDx sequence motif, which is conserved in the eukaryotic tRNase ZLs and bacterial tRNase ZLs, but not in the TM-type tRNase ZLs. In the crystal structures of human and B. subtilis tRNase ZLs, the two AxDx motifs have entirely the same folding topology (http://www.rcsb.org/pdb/explore.do?structureId = 3ZWF and [18]). The first residue of the motif (Ala) is located at the end of β15 while other residues are found in a short loop connecting β15 and β16. We noted that the AxDx motifs of human and Bacillus tRNase ZLs and the PksxRN motif or its variant PxPxRG are in close enough proximity to make hydrogen bonding interactions (Figure 8). The PksxRN motif or its variant PxPxRG forms part of a loop which acts like a flap covering the entrance to the active site. Strikingly, the hydrogen bonding networks formed by the AxDx motif and the PxPxRG or PksxRN motif are very similar with one notable difference: the OD2 atom of Asp353 and the main-chain amino groups of Thr10 in the human tRNase ZL make hydrogen bonds with the main-chain amino and carbonyl groups of Gly20 of the PxPxRG motif, respectively, whereas the corresponding residues in the Bacillus tRNase ZL form hydrogen bonds with the main-chain amino and carbonyl groups of Asn18 in the PksxRN motifs (Figure 8). Interestingly, our homology modeling of the human tRNase ZL based on the crystal structure of the human tRNase ZL suggests that the potential hydrogen bonding network involving...
Figure 5. Multiple sequence alignment of representatives of the metazoan and non-metazoan tRNase Zs.

Metazoan tRNase Zs are from ten deuteromes including *C. familiaris* (Cfa), *L. africana* (Laf), *H. sapiens* (Hsa), *M. musculus* (Mmu), *X. tropicalis* (Xtr), *D. rerio* (Dre), *S. salar* (Ssa), *C. intestinalis* (Cin), *B. floridae* (Bfl), and *S. purpuratus* (Spu), one protostome, *L. gigantean* (Lgi), and three basal metazoans, *N. vectensis* (Nve), *H. magnipapillata* (Hma), and *A. queenslandica* (Aqu). Non-metazoan tRNase Zs are from the unicellular choanoflagellate *M. brevicollis* (Mbr), *B. subtilis* (Bsu) and *E. coli* (Eco). The annotation of the alignment is as described in the legend to Figure 3.

doi:10.1371/journal.pone.0044264.g005
the AxDx motif is likely to be conserved in the structure of the human tRNase Z L (Figure S4).

The PxKxRN motif, in particular, the two basic residues in the motif, has been proposed to may play a role in cleavage site specificity and, by inference, CCA anti-determination, which prevents removal of mature 3′-end of tRNAs by tRNase Z, and thus avoids the repeated addition and removal of the CCA sequence [17,18]. In D. melanogaster, alanine substitutions of conserved residues in the PxKxRN motif significantly increased the Km value [17]. Considering its location and the possible role of the PxKxRN motif, we speculate that the AxDx motif may be indirectly involved in cleavage specificity through stabilization of an appropriate conformation of the PxKxRN/PxPxRG loop for tRNA processing.

Using a S. pombe-based functional complementation assay we have demonstrated the importance of the conserved Asp in the AxDx motif for the function of the S. pombe nuclear tRNase Z L. However, its exact role in tRNA processing is unclear and awaits further biochemical characterization.

Alternative Splicing

Our database searches revealed candidate alternatively spliced isoforms of tRNase Z L in mammalian and nematode species. Of the types of alternative splicing of the mammalian tRNase Z L, skipping of exon 7 appears to be the most common type of alternative splicing. This exon is skipped due a downstream splice site of exon 8. The sizes of exon 7 of mammalian tRNase Z L’s range from 108 (in dog) to 135 bp (in rat). The region encoded by exon 7 is located upstream from the exon 8 encoding the flexible arm. Further, the region encoded by exon 7 is not found in the fly and nematode tRNase Z L’s. Thus, exon 7 apparently does not required for the tRNA 3′-end processing activity of the metazoan tRNase Z L.

So far, alternative splicing of the mammalian tRNase Z L mRNA has only been experimentally verified in mouse and rat tissues [33]. The mouse and rat tRNase Z L’s have two mRNA splice forms. Real-time quantitative PCR analysis revealed that both splice forms of tRNase Z L are widely expressed in most mouse and rat tissues. In the mouse, the highest levels of expression of the two splice forms (MmuTRZ2 and MmuTRZ2i) were detected in the testis and spleen, respectively, whereas the lowest levels of MmuTRZ2 and MmuTRZ2i were observed in both the prostate and muscle, and the muscle, respectively. In the prostate and muscle, the MmuTRZ2i mRNA is the predominant transcript, whereas in other tissues, the MmuTRZ2 transcript is predominant. Similarly, in the prostate, the rat tRNase Z L expression is dramatically reduced and the alternative form of tRNase Z L is more abundant than the normally spliced form.

Figure 6. Sequence logos for the predicted PxRxRN, PxPxRG and AxDx motifs. The sequence logo of the PxRxRN motif was derived from 174 tRNase Z L’s from 75 fungi, 21 plants and 68 metazoans. The sequence log of the PxPxRG motif is derived from alignment of 44 metazoan tRNase Z L’s. The sequence logo of the AxDx motif is derived from 240 alignments including 178 tRNase Z L’s from 71 metazoan, 21 green plants, 77 fungi, and 62 tRNase Z L’s from 30 metazoans, 9 fungi, 2 bacteria and M. brevicollis. The sequence logos were created using WebLogo (http://weblogo.berkeley.edu). The height of each amino acid indicates the level of conservation at that position. Amino acids are colored as follows: red, basic; blue, acidic; orange, polar; and green, hydrophobic.

doi:10.1371/journal.pone.0044264.g006

Figure 7. Growth complementation analysis of S. pombe Trz1 with a mutation in the AxDx motif. A S. pombe strain bearing a temperature-sensitive allele of trz1 (trz1-1) was transformed with empty vector (4X) or vector expressing the wild-type Trz1 or the Trz1D772A mutant, and equal amounts of cells were spotted on EMM medium supplemented with leucine and adenine and grown at 26 C or 37 C. doi:10.1371/journal.pone.0044264.g007

PLOS ONE | www.plosone.org 12 September 2012 | Volume 7 | Issue 9 | e44264
AxDx motif is important to the function of the conserved structural feature of tRNase Zs. We showed that the tRNase Zs are highly examined. We predicted that potential hydrogen bonding interactions involving the AxDx and PxKxRN/PxPxRG motif are identified a conserved AxDx motif in all metazoan tRNase Zs. Using sequence conservation analysis, we predicted that potential hydrogen bonding interactions involving Ala351 (O) and Arg19 (NH1), Asp353 (OD2) and Gly20 (HN), Ala351 (HN) and Leu311 (O), Asp353 (OD1) and Thr10 (HN and OG1), and Gly20 (O) and Thr10 (HN). (B) In B. subtilus, the AxDx strand/loop-centered hydrogen bond network involves Ala296 (O) and Arg17 (NH1), Asp298 (OD2) and Asn18 (HN), Ala296 (HN) and Leu267 (O), Asp298 (OD2) and Thr8 (HN and OG1), and Asn18 (O) and Thr8 (HN).

Figure 8. Depiction of the predicted hydrogen bonding network in the AxDx strand/loop region of human and B. subtilus tRNase Zs. The atomic coordinates of human and B. subtilus tRNase Zs were obtained from the Protein Data Bank, www.rcsb.org (PDB accession codes 3ZWF and 1Y44, respectively). Images were made with Swiss-PdbViewer [59]. Potential hydrogen bonds were determined using the Swiss-PdbViewer and Insight II. Secondary structure elements are colored green. The AxDx strand/loop is colored yellow and the PxPxRG or PxKxRN loop is colored red. Hydrogen bonds are represented by a dashed green line. (A) In the human tRNase Zs, the AxDx strand/loop-centered hydrogen bond network involves Ala351 (O) and Arg19 (NH1), Asp353 (OD2) and Gly20 (HN), Ala351 (HN) and Leu311 (O), Asp353 (OD1) and Thr10 (HN and OG1), and Gly20 (O) and Thr10 (HN). (B) In B. subtilus, the AxDx strand/loop-centered hydrogen bond network involves Ala296 (O) and Arg17 (NH1), Asp298 (OD2) and Asn18 (HN), Ala296 (HN) and Leu267 (O), Asp298 (OD2) and Thr8 (HN and OG1), and Asn18 (O) and Thr8 (HN).

However, the spliced variants of mouse and rat tRNase ZL’s have been experimentally validated on the mRNA level, but not on the protein level. Furthermore, the function of these spliced variants of tRNase ZL remains to be characterized.

Conclusions

The present study complements our previous work by identifying and analyzing metazoan tRNase Zs. We found that although both tRNase ZSs and tRNase ZLs arose very early in metazoan evolution, only tRNase ZL has been retained throughout metazoan evolution. Similar to the situation in fungi and plants, in metazoas, a single tRNase ZL seems to be responsible for both nuclear and mitochondrial tRNA 3′-end processing. In contrast, tRNase ZSs, which is retained only in deuterostomes, lophotrochozoans and lower metazoas, may have a function other than tRNA 3′-end processing. We also found the existence of the predicted alternative splice variants of the mammalian and nematode tRNase ZL genes in processing. We also found the existence of the predicted alternative splice variants of the mammalian and nematode tRNase ZL genes in the GenBank database. Using sequence conservation analysis, we identified a conserved AxDx motif in all metazoan tRNase Zs examined. We predicted that potential hydrogen bonding interactions involving the AxDx and PxKxRN/PxPxRG motif are highly conserved structural feature of tRNase Zs. We showed that the AxDx motif is important to the function of the S. pombe tRNase ZL in a functional complementation assay. Our results should open up interesting new perspectives for future studies of tRNase Z.

Methods

Metazoan Genome Database Searches and Protein Sequence Analyses

Candidate tRNase Zs were identified by blastp and tblastn searches of genomic, protein and EST databases, which include the National Center for Biotechnology Information nonredundant protein sequence database (NCBI; http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism = euk), Metazome (http://www.metazome.net/search.php?show = blast), Ensemble (http://www.ensembl.org/index.html), ExPASy (http://www.expasy.ch/), the Joint Genome Institute (JGI; http://www.jgi.doe.gov/), the Broad Institute (http://www.broadinstitute.org/scientific-community/data), Wormbase (http://www.wormbase.org/), the Universal Protein Resource (UniProt; http://www.uniprot.org/), and SpongeBase (http://spongebase.unimainz.de/), using default parameters.

The resulting sequences were subject to validation as described [5,7]. Briefly, each prediction was first subjected to reciprocal BLAST searches against the GenBank database. In back-searches, a candidate was confirmed if the best hits from reciprocal BLAST are tRNase Zs. Accuracy of prediction was further evaluated by multiple sequence alignment. All discordant candidate sequences were checked manually for possible errors including sequencing errors, intron mispredictions and existence of gaps in the genome sequences. When possible, we also used mRNA/EST sequences to check the predicted intron positions. We kept the candidate if it could be accurately predicted, otherwise we discarded it. However, despite our best efforts, we could not conclusively rule out the possibility that some of candidates might contain errors. The splicing pattern was verified using the Fgenesh and Fgenesh_GC programs provided at the Softberry website (http://linux1.softberry.com/berry.phtml?topic = fgenesh). Prediction of subcellular localization of proteins was made using web-based prediction programs such as MITOPROT (http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html) and PSORT (http://psort.hgc.jp/form.html). Multiple sequence alignments were done by Clustal W [32].

Phylogenetic Analysis

The phylogenetic tree for candidate metazoan tRNase Zs was constructed as described [5,7]. Accession numbers for the sequences used in the analysis are summarized in Table S1. Briefly, a multiple amino acid sequence alignment of metazoan tRNase ZSs and C-terminal halves of metazoan tRNase ZLs was generated with Clustal W implemented in MEGA5.0 [33] were manually adjusted. Model selection was performed using ProtTest 2.4 [54]. The Bayesian phylogenetic tree was inferred by using MrBayes version 3.1.2 [55] and a mixture of the fixed amino acid + G distribution. Statistical confidence was assessed by using Markov Chain Monte Carlo (MCMC) sampling approaches. Two separate runs including a total of four independent tree searches were conducted. All searches consisted of one ‘cold’ and three ‘heated’ Markov chains estimated for 10^7 generations, and every 100 generations were sampled. The burn-in parameter was estimated by plotting lnL against the generation number using Tracer (v1.4.1, http://beast.bio.ed.ac.uk/Tracer), and the retained trees were used to estimate the consensus tree and the Bayesian posterior probabilities.
Preparation of candidate tRNase Zs

The aspartate to alanine mutation in the 770AxDx773 motif of SpTrz1 was generated by the overlapping PCR method ([56]). Briefly, the plasmid expressing wild type Sprz1 ([57]) was used as template for PCR. The upstream region of the mutant was generated using the following primers: 5'GACCCTGGTTTCCTGCTGGAAGAACCTGCTTAAAATITCGG (designated sprz1-D-re). The downstream region of the mutant was created using the following primer pairs: 5'CCGATATCTTATGTCCAGCTGAAGCTGCAAGTAATATGTAGTGC (designated sprz1-D-re) and 5'AGCTGTCGTCGCCGGTTAGACAGTAAGTC (designated spTrz1-re). The up and downstream fragments of the mutant from the first round of PCR were mixed in a 1:1 molar ratio, diluted and used as template for the second round of PCR. The pair of primer SalI and sprz1-re was used in the second round of PCR. The resulting PCR product containing mutated sprz1 (sprz1-D772) was digested and cloned into the Sal I/Sma I sites of pREP4X. The point mutation was verified by DNA sequencing.

Supporting Information

Figure S1 Alignment of candidate tRNase Zs identified in metazoans. The accession numbers for the candidates are listed in Table S1. The annotation of the alignment is described in the legend to Figure 3.

Figure S2 Alignment of candidate tRNase Zs identified in metazoans. The accession numbers for the candidates are listed in Table S1. The annotation of the alignment is described in the legend to Figure 3.

Figure S3 Comparison of the sequences flanking the two in-frame AUG start codons in 58 metazoan tRNase Z+ mRNAs. Sequences were aligned by the start codon. The AUG start codons are indicated in red. In each sequence, nucleotides at positions +3 and +4, which are the most important determinants of context are shown in blue and green, respectively. Although the consensus sequences surrounding AUG start codons vary considerably between eukaryotic groups, they share a strong preference for purines at the position +3 and G at the position +4 for translational initiation.

Figure S4 The predicted hydrogen bonding network in the AxDx strand/loop region of human tRNase Z+. The published structure of human tRNase Z+ (PDB code 3ZW6) was used as a template to build a modeled structure of C-terminal region of human tRNase Z+ (residues 481–754) using the homology modeling server SWISSMODEL (http://swissmodel.expasy.org/). The picture is labeled as described in the legend to Figure 8. The potential hydrogen bond network is formed by the O atom of Ala746 and the NH1 atom of Arg497 (3.07 Å), the OD2 atom of Asp748 and the OG1 atom of Thr488 (2.71 Å), the OG1 atom of Thr488 and the NH1 atom of Arg497 (2.81 Å) and the O atom of Asn496 and the HN atom of Thr488 (3.01 Å). The figure was prepared using Swiss-PdbViewer ([59]).

Table S1 Distribution of candidate tRNase Zs identified in metazoans. Abbreviations for species names are indicated in the parentheses. *The number of amino acids in metazoan tRNase Z and tRNase Z-like proteins. *Indicates that mispredicted sequences obtained from the databases have been corrected. **Indicates the sequence could not be correctly predicted.

Table S2 Prediction of the number of introns in metazoan tRNase Z genes.

Table S3 Subcellular localization prediction of metazoan tRNase Z+’s. The putative NLs were predicted using PSORT (http://psort.hgc.jp/form.html), while the putative MTs were predicted using MITOPROT (http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html). Nuc: stand for the nucleus; Mit: mitochondria. N indicates nuclear localization and M denotes mitochondrial localization. “-”, the localization or targeting sequence could not be predicted. The numbers refer to amino acid positions starting from the N-terminus.

Table S4 Percentage amino acid identity among tRNase Z+’s from selected metazoans. The pairwise percent identity scores were generated with Clustal W ([52]). tRNase Z+’s from H. sapiens (Hsa), C. familiaris (Cfa), M. musculus (Mmu), R. norvegicus (Rno), O. cuniculus (Ocu), S. scrofa (Ssc), A. carolinensis (Aca), X. tropicalis (Xtr), D. rerio (Dre), G. aculeatus (Gam), C. intestinalis (Cin), C. savignyi (Cas), B. mori (Bmo), M. melangaster (Dme), C. elegans (Cel), H. robusta (Hro), S. mansoni (Sma), T. adhaerens (Tad), M. brevicolis, (Mbr), A. thaliana (Ath), and S. pome (Spo).

Table S5 Percentage amino acid identity among tRNase Z+’s from selected metazoans. The pairwise percent identity scores were generated with Clustal W ([52]). H. sapiens, Hsa; M. musculus, Mmu; R. norvegicus, Rno; O. cuniculus, Ocu; S. scrofa, Ssc; A. carolinensis, Aca; X. tropicalis, Xtr; D. rerio, Dre; G. aculeatus, Gam; C. intestinalis, Cin; C. savignyi, Cas; B. mori, Bmo; M. melangaster, Dme; C. elegans, Cel; H. robusta, Hro; S. mansoni, Sma; T. adhaerens, Tad; M. brevicolis, Mbr; B. subtilis, Bsu; E. coli, Eco.

Acknowledgments

We thank Lijuan Fan for initial efforts in this work, Houchane Dai and Haiyan Yu for technical help, and Jie Yan for help in construction of phylogenetic trees and for valuable discussion.

Author Contributions

Conceived and designed the experiments: YH. Performed the experiments: ZW, JZ, JP, XZ, JL. Analyzed the data: YH. Wrote the paper: YH.

References

1. Cechalos M, Vioque A (2007) tRNase Z. Protein Pept Lett 14: 137–145.
2. Hartmann RK, Gossringer M, Spath B, Fischer S, Marchfelder A (2009) Chapter 6 The Making of tRNAs and More - tRNase P and tRNase Z. Prog Nucleic Acid Res Mol Biol 85C: 319–368.
3. Redko Y, Li de Lasierra-Gallay I, Condon C (2007) When all’s zed and done: the structure and function of RNase Z in prokaryotes. Nat Rev Microbiol 5: 271–286.
4. Spath B, Canino G, Marchfelder A (2007) tRNase Z: the end is not in sight. Cell Mol Life Sci 64: 2404–2412.
32. Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV (2005) Analysis of 30. Roy SW, Gilbert W (2005) Complex early genes. Proc Natl Acad Sci U S A 102:
29. Smith MM, Levitan DJ (2004) The Caenorhabditis elegans homolog of the 27. Dubrovsky EB, Dubrovskaya VA, Levinger L, Schiffer S, Marchfelder A (2004)
23. Kostelecky B, Pohl E, Vogel A, Schilling O, Meyer-Klaucke W (2006) The 18. de la Sierra-Gallay IL, Pellegrini O, Condon C (2005) Structural basis for 17. Zareen N, Hopkinson A, Levinger L (2006) Residues in two homology blocks on 16. Minagawa A, Ishii R, Takaku H, Yokoyama S, Nashimoto M (2008) The 15. Schiffer S, Rosch S, Marchfelder A (2002) Assigning a function to a conserved 12. Aravind L (1999) An evolutionary classification of the metallo-beta-lactamase 11. Takahashi M, Takaku H, Nashimoto M (2008) Regulation of the human tRNase 10. Rossmanith W (2011) Of P and Z: Mitochondrial tRNA processing enzymes. 9. Brzezniak LK, Bijata M, Szczesny RJ, Stepien PP (2011) Involvement of human 8. Ishii R, Minagawa A, Takaku H, Takagi M, Nashimoto M, et al. (2005) Crystal 7. Gutierrez-Camps JM, Aragona E, Aravind L (2003) Comparative analysis of the 6. Tavtigian SV, Simard J, Teng DH, Abtin V, Baumgard M, et al. (2001) A 5. Fan L, Wang Z, Liu J, Guo W, Yan J, et al. (2011) A survey of green plant tRNA 4. Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of 3. Dumont M, Frank D, Moisan AM, Tranchant M, Soory P, et al. (2004) Structure of 2. Aravind L (1999) An evolutionary classification of the metallo-beta-lactamase 1. Takahashi M, Takaku H, Nashimoto M (2006) Regulation of the human tRNase ZS gene expression. FEBS Lett 582: 2392–2396.

Identification of Metazoan tRNase Zs