

RESEARCH ARTICLE

Open Access

Identification and analysis of candidate fungal tRNA 3'-end processing endonucleases tRNase Zs, homologs of the putative prostate cancer susceptibility protein ELAC2

Wei Zhao, Haiyan Yu, Shuzhen Li, Ying Huang*

Abstract

Background: tRNase Z is the endonuclease that is responsible for the 3'-end processing of tRNA precursors, a process essential for tRNA 3'-CCA addition and subsequent tRNA aminoacylation. Based on their sizes, tRNase Zs can be divided into the long (tRNase Z^L) and short (tRNase Z^S) forms. tRNase Z^L is thought to have arisen from a tandem gene duplication of tRNase Z^S with further sequence divergence. The species distribution of tRNase Z is complex. Fungi represent an evolutionarily diverse group of eukaryotes. The recent proliferation of fungal genome sequences provides an opportunity to explore the structural and functional diversity of eukaryotic tRNase Zs.

Results: We report a survey and analysis of candidate tRNase Zs in 84 completed fungal genomes, spanning a broad diversity of fungi. We find that tRNase Z^L is present in all fungi we have examined, whereas tRNase Z^S exists only in the fungal phyla Basidiomycota, Chytridiomycota and Zygomycota. Furthermore, we find that unlike the Pezizomycotina and Saccharomycotina, which contain a single tRNase Z^L, *Schizosaccharomyces* fission yeasts (Taphrinomycotina) contain two tRNase Z^Ls encoded by two different tRNase Z^L genes. These two tRNase Z^Ls are most likely localized to the nucleus and mitochondria, respectively, suggesting partitioning of tRNase Z function between two different tRNase Z^Ls in fission yeasts. The fungal tRNase Z phylogeny suggests that tRNase Z^Ss are ancestral to tRNase Z^Ls. Additionally, the evolutionary relationship of fungal tRNase Z^Ls is generally consistent with known phylogenetic relationships among the fungal species and supports tRNase Z^L gene duplication in certain fungal taxa, including *Schizosaccharomyces* fission yeasts. Analysis of tRNase Z protein sequences reveals putative atypical substrate binding domains in most fungal tRNase Z^Ss and in a subset of fungal tRNase Z^Ls. Finally, we demonstrate the presence of pseudo-substrate recognition and catalytic motifs at the N-terminal halves of tRNase Z^Ls.

Conclusions: This study describes the first comprehensive identification and sequence analysis of candidate fungal tRNase Zs. Our results support the proposal that tRNase Z^L has evolved as a result of duplication and diversification of the tRNase Z^S gene.

Background

The endonuclease tRNase Z (also called RNase Z or 3'-tRNase) participates in maturation of tRNA 3'-end by removing the 3'-trailer sequence from tRNA precursors (pre-tRNAs, for reviews, see [1-4]). It belongs to the metallo- β -lactamase (MBL) superfamily, the members of

which have diverse functions from hydrolysis and inactivation of β -lactam antibiotics to processing of RNA precursors [5-9]. Other nucleases in the MBL superfamily that act on nucleic acids include members of the β -CASP (MBL-associated CISF Artemis SNM1/PSO2) family: the cleavage and polyadenylation specificity factor 73 kDa subunit (CPSF-73) [10] and the Integrator complex subunit 11 (Int11) [11], which are involved in eukaryotic mRNA and small nuclear RNA (snRNA) 3'-end formation, respectively; RNase J, which functions in rRNA

* Correspondence: yhuang@njnu.edu.cn
Nanjing Engineering and Technology Research Center for Microbiology,
Jiangsu Key Laboratory for Biodiversity and Biotechnology, School of Life
Sciences, Nanjing Normal University, Nanjing 210046, China

maturation and mRNA stability in bacteria [12] and the eukaryotic Pso2/Snm1/Artemis proteins, which play a role in DNA repair [7]. Although displaying distinct substrate specificity defined by their specific domains, these proteins appear to have a similar catalytic mechanism since their active sites are composed of highly conserved motifs including the histidine motif (HxHxDH, where x is any hydrophobic residues).

There are two forms of tRNase Z. The long form (tRNase Z^L) with 800-900 aa (amino acids) is about twice the size of the short form (tRNase Z^S) with 300-400 aa [9]. Sequence analysis suggests that tRNase Z^L has evolved by gene duplication from tRNase Z^S followed by sequence divergence [9]. The species distribution of tRNase Z is not homogenous. tRNase Z^S exists in all three domains of life (i.e. Bacteria, Archaea, and Eukarya) whereas tRNase Z^L has been found only in eukaryotes so far. The number of tRNase Zs varies among different organisms. The largest number of tRNase Zs was detected in the plant *Arabidopsis thaliana* (two tRNase Z^Ss and two tRNase Z^Ls) [13]. The fission yeast *Schizosaccharomyces pombe* contains two tRNase Z^Ls. Unexpectedly, the human genome encodes one tRNase Z^S and one tRNase Z^L. Human tRNase Z^L gene (also termed *ELAC2*) was originally identified as the first prostate cancer susceptibility gene by positional cloning [9]. However, the mechanism by which specific mutations in human tRNase Z^L lead to an increased risk of prostate cancer remains unknown. In contrast, the budding yeast *Saccharomyces cerevisiae*, the fruit fly *Drosophila melanogaster* and the nematode worm *Caenorhabditis elegans* have just one tRNase Z^L.

An intriguing question is why species have evolved to have more than one tRNase Z. One explanation is that additional tRNase Zs are targeted to organelles such as mitochondria and chloroplasts in which organelle-encoded pre-tRNAs must also be processed. Indeed, one of two *S. pombe* tRNase Z^Ls is targeted to the mitochondria, and has been suggested to play a role in mitochondrial-encoded pre-tRNA processing [14]. In *A. thaliana*, three of four tRNase Z^Ls are targeted to organelles [13]. However, it appears that the majority of tRNase Z^Ls identified thus far are imported both into the nucleus and mitochondria. Another explanation is that additional tRNase Zs may provide a back-up mechanism for nuclear tRNA 3'-end processing. The third explanation is that additional tRNase Zs may have different functions.

Recently, tRNase Z^L itself has been either demonstrated or suggested to have additional functions other than tRNA 3'-end processing. For example, human tRNase Z^L has been shown to play a role in generation of non-tRNA noncoding RNAs [15,16] and viral microRNAs (miRNAs) [17]. Moreover, human tRNase Z^L has

been proposed to cleave a subset of miRNAs in the cytoplasm [18]. In *S. cerevisiae*, tRNase Z^L has been suggested to have additional functions including rRNA biogenesis, mRNA splicing and mitochondrial maintenance [19]. Similarly, our previous study also suggested that the nuclear-localized tRNase Z^L in *S. pombe* may play a role beyond tRNA 3'-end processing [14].

Our current understanding of tRNase Z evolution is limited since there has been only one comprehensive survey on tRNase Z^Ss from prokaryotes [3]. Of eukaryotes, the Fungi is a large and diverse kingdom encompassing roughly 1.5 million species and spanning one billion years of evolution [20]. Sequence-based phylogenies show that the Chytridiomycota is the most basal phylum (group) among the Fungi, followed by the Zygomycota, with the Ascomycota and Basidiomycota as two largest phyla that together comprise the subkingdom Dikarya (also referred to as the "Higher Fungi") [21-24]. The Ascomycota (also known as sac fungi, yeasts or ascomycetes) is the largest and most diverse phylum in the Fungi, accounting for approximately 75% of all known fungi. Many popular model organisms such as *S. cerevisiae*, *S. pombe*, *Neurospora crassa*, *Aspergillus nidulans* and *Candida albicans* are classified in this phylum. This phylum is further divided into three major monophyletic subphyla (subgroups): Pezizomycotina, Saccharomycotina and Taphrinomycotina [25]. The Pezizomycotina (also known as euascomycetes) is the largest subphyla and contains over 90% of total Ascomycota species. They are multicellular filamentous fungi and grow by hyphal extension and branching. In contrast, the Saccharomycotina (also known as true yeasts) comprises the majority of unicellular species. The Taphrinomycotina is thought to be the earliest diverging group sister to the Saccharomycotina and Pezizomycotina. It constitutes a diverse group of organisms including unicellular yeast (for example, *Schizosaccharomyces*), multicellular filamentous fungi, and dimorphic fungi that can switch between yeast and hyphal growth forms. Like the Pezizomycotina, the Basidiomycota consists of primarily filamentous fungi.

Currently, most of eukaryotic species with sequenced genomes belong to the kingdom Fungi. The public fungal genome databases cover a broad range of fungal taxonomic groups with the majority coming from the Ascomycota and Basidiomycota. The availability of a large number of fungal genome sequences, together with the vast diversity of fungal morphology and lifestyle, provides an opportunity to identify tRNase Zs in the kingdom Fungi and to study evolution of eukaryotic tRNase Z.

In the present study, we performed a comprehensive survey of candidate tRNase Zs from 84 publicly available fungal genomes. To explore the evolutionary

relationship among fungal tRNase Zs, we conducted a phylogenetic analysis of predicted fungal tRNase Zs. Finally, we examined their domain architectures. Our results support the view that tRNase Z^L comes from tRNase Z^S.

Results

Identification of putative fungal tRNase Zs

As part of our efforts to better understand functional and structural diversity of tRNase Zs, we conducted extensive BLAST and PSI-BLAST homology searches against the publicly available fungal genome databases. Currently, the majority of sequenced species belong to the Dikarya, with a much higher proportion of Ascomycota species. Since other fungal phyla are poorly represented in public databases (three Zygomycota, three Microsporida, and three Chytridiomycota species), it is difficult to assess the true diversity of tRNase Z in these basal groups of fungi.

The initial candidates from the BLAST and PSI-BLAST were verified by multiple sequence alignment and reciprocal searches against the GenBank. Protein sequence alignment revealed a number of incorrectly predicted candidates, most likely due to misprediction of exon/intron boundaries or existence of gaps in the genome sequence. For example, the sequence (Broad accession no. CC1G_14814.2) annotated as the candidate Basidiomycete *Coprinopsis cinerea* tRNase Z in the fungal genome database at the Broad Institute has mispredicted exon/intron junctions. This 946-aa-long candidate is devoid of a histidine motif, which is a signature motif for the MBL superfamily, indicating that the exon encoding the histidine motif was likely mispredicted. After re-evaluating intron splicing pattern of the gene sequence, we were able to predict the exon encoding the histidine motif. The correctly predicted protein has 967 aa, and has the histidine motif. The sequence annotated as the candidate Pezizomycotina *Botrytis cinerea* (also named *Botryotinia fuckeliana*) tRNase Z (Broad accession no. BC1G_03733.1) is an example of misprediction due to the presence of sequence gaps in the genome. This sequence has 444 aa. However, examination of the genomic sequence revealed that its 5'-coding sequence contains gaps. Thus, this sequence was excluded. In back-searches, no candidate that shows homology to metallo- β -lactamase was found, but a limited number of candidates were found to show homology to the yeast homolog of CPSF-73 (Ysh1). Such candidates were also excluded from our final list. However, we cannot rule out the possibility that certain candidates may not be correctly predicted despite our efforts devoted to verification of these candidates.

We identified a total of 90 candidate tRNase Z^Ls and 19 candidate tRNase Z^Ss proteins from 84 fungal species

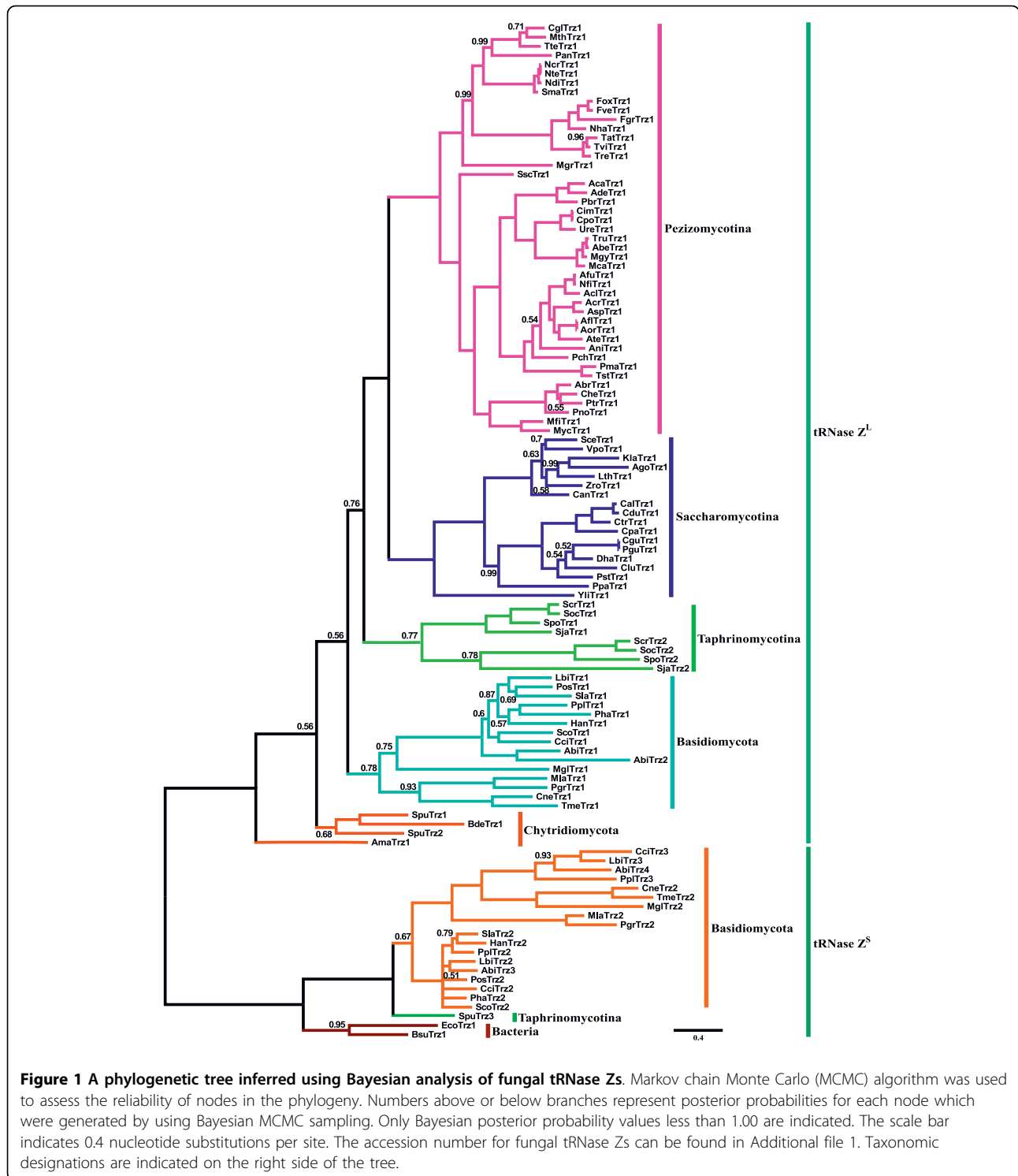
including 67 Ascomycota, 14 Basidiomycota and 3 Chytridiomycota (Additional file 1). Candidate tRNase Zs from two taxonomic groups, the Zygomycota and Microsporida, were not listed since their full-length protein sequences could not be correctly predicted. Of the proteins identified here, only tRNase Z^Ls from *S. cerevisiae* and *S. pombe* have been experimentally characterized [14,19,25-27].

All species of the Ascomycota we have examined lack tRNase Z^S. The Pezizomycotina and Saccharomycotina species have a single tRNase Z^L. Surprisingly, in contrast to the Pezizomycotina and Saccharomycotina species, all four sequenced *Schizosaccharomyces* species (*S. pombe*, *Schizosaccharomyces octosporus*, *Schizosaccharomyces japonicus* and a recently described *Schizosaccharomyces cryophobus*) in the Taphrinomycotina have two tRNase Z^Ls, which we term tRNase Z^{L1} and tRNase Z^{L2}, respectively. tRNase Z^{L1}s and tRNase Z^{L2}s have been either shown or predicted to localize to the nucleus and mitochondria, respectively (Additional file 2 and data not shown) [28]. Since in the current databases, all sequenced Taphrinomycotina species come from only *Schizosaccharomyces*, it would be interesting to see whether species in other genera also contain two tRNase Z^Ls.

Like Ascomycota species, all sequenced Basidiomycota species (except for *Agaricus bisporus*) have a single tRNase Z^L. However, unlike the situation in the Ascomycota, tRNase Z^S was found in all sequenced Basidiomycota species. While the majority of Basidiomycota species have a single tRNase Z^S, four Basidiomycota species (*A. bisporus*, *C. cinerea*, *Laccaria bicolor* and *Postia placenta*) have two tRNase Z^Ss. Among the Basidiomycota species examined, *A. bisporus* has the largest number of tRNase Zs (two tRNase Z^Ls and two tRNase Z^Ss). The number of tRNase Z seems to be variable in the three sequenced chytrid species. *Allomyces macrogynus* and *Spizellomyces punctatus* have two tRNase Z^Ls whereas *Batrachochytrium dendrobatidis* appears to have one tRNase Z^L. Moreover, tRNase Z^S was only identified in *S. punctatus*. Although we could not correctly predict the full-length tRNase Zs in three sequenced Zygomycota species (*Rhizopus oryzae*, *Mucor circinellodes* and *Phycomyces blakesleeanus*) and in three sequenced Microspordian species (*Encephalitozoon cuniculi*, *Enterocytozoon bieneusi* and *Nosema ceranae*), it is important to note that tRNase Z^S appears to exist in all sequenced Zygomycota fungi but not in sequenced Microspordian fungi known for extreme genome reduction and compaction [29].

Phylogenetic analysis

To explore evolutionary relationships among fungal tRNase Zs, we performed a phylogenetic analysis of the



amino acid sequences of all tRNase Zs predicted from the fungal genome databases. Figure 1 shows the phylogenetic tree for 109 fungal tRNase Zs. In addition to the fungi species, tRNase Z^Ss from *B. subtilis* and *E. coli* were included as reference. It is seen that tRNase Zs are

clearly separated into a small cluster containing both fungal and bacterial tRNase Z^Ss and a large cluster containing fungal tRNase Z^Ls. Moreover, within the tRNase Z^L cluster, tRNase Z^Ls can be grouped according to their taxonomic classification (Figure 1), although some

Bayesian posterior probability values for grouping are not strong. It appears that the phylogenetic relationships among fungal tRNase Z^Ls is basically congruent with the currently accepted fungi phylogenies based on cladistic analyses of RNA and/or protein sequences [21-24]. Notably, tRNase Z^{L2}s from four fission yeasts together form a group sister to a group formed by tRNase Z^{L1}s from the same fission yeasts, albeit with a posterior probability of only 0.77. Likewise, the two tRNase Z^Ls in the Basidiomycete *Agaricus bisporus* (AbiTrz1 and AbiTrz2) are sister to each other with a posterior probability value of 1.

Analysis of fungal tRNase Z^Ls

The sizes of predicted tRNase Z^Ls vary considerably among fungal species, ranging from 648 to 1140 aa with an average size of ~924 aa. The variation in protein size is due to a high degree of length and sequence variation of N-terminal and C-terminal extensions and many insertions and/or deletions. Remarkably, tRNase Z^Ls in *Sordaria macrospora* and three *Neurospora* species (*N. crassa*, *Neurospora discreta* and *Neurospora discreta*) have a very long N-terminal extension (~200 residues). This feature appears to be family-specific since all these species belong to the family *Sordariaceae*.

A number of fungal tRNase Z^Ls have a variable length N-terminal extension predicted to contain a canonical MTS (Additional file 2). In addition, tRNase Z^{L2}s from four *Schizosaccharomyces* species we have examined also contain a putative MTS in their N-terminal extensions. The N-terminal extensions found in fungal tRNase Z^Ls may serve as transit sequences for directing the proteins to the mitochondria. It is interesting to note that tRNase Z^Ls from *D. melanogaster* and humans also contain a canonical MTS.

To assess the extent of sequence and structural conservation among fungal tRNase Z^Ls, we aligned sixteen tRNase Z^L protein sequences from fifteen taxonomically diverse fungal species including eleven species of the Ascomycota, three species of the Basidiomycota, and one species of the Chytridiomycota (Table 1). Since the amino acid sequence of tRNase Z^L can be divided into an N-terminal half, which contains a substrate binding site, and a C-terminal half, which contains a catalytic center and most of the conserved motifs, we aligned the N- and C-terminal halves of tRNase Z^Ls separately, and first examined the C-terminal half. For comparison, we also included non-fungal eukaryotic tRNase Z^Ls from *D. melanogaster*, *A. thaliana* and humans. Figures 2 and 3 show sequence comparison of representative fungal and non-fungal eukaryotic tRNase Z^Ls (For a full list of all aligned fungal tRNase Z^Ls, see Additional file 3).

Although tRNase Z^Ls from closely related species share the high degree of sequence similarity, the

sequence similarity among fungal tRNase Z^Ls is low. Overall, sequence conservation of fungal tRNase Z^Ls is largely confined to highly conserved Motifs I-V (Motif II is also called the histidine motif) and the PxKxRN, HEAT and HST loop motifs at the C-terminal halves of the proteins (Figure 2). Except for the PxKxRN loop and Motif I, which were found to play a role in pre-tRNA acceptor stem binding and CCA anti-determination [30], all other motifs are involved in zinc binding and catalysis [31-33]. Motifs I-V contain invariant histidine and/or aspartate residues essential for the tRNase Z activity. In particular, the histidine and aspartate residues in the histidine motif, together with the histidine residues in Motifs III and V and the aspartate residue in Motif IV are involved in the coordination of the two zinc ions at the catalytic center [1,3,4].

All characterized tRNase Zs contain a characteristic domain of 30~50 aa residues, termed a flexible arm (or an exosite), which is important for substrate binding [34,35]. Based on sequence comparison, three types of flexible arms, termed the zinc-dependent phosphodiesterase (ZiPD)-, ELAC2- and *Thermotoga maritima* (TM)-type flexible arms, have been found in tRNase Zs. The ZiPD- and ELAC2-type flexible arms were typical for bacterial tRNase Z^Ss and eukaryotic tRNase Z^Ls, respectively, whereas the TM-type flexible arm is an atypical one found only in tRNase Z^Ss from *T. maritima* and *A. thaliana*. Interestingly, tRNase Z^S in *T. maritima* itself is an atypical enzyme as it cleaves CCA-containing pre-tRNAs after CCA. Although having sequence and length variations, both ZiPD- and ELAC2-type flexible arms comprise a GP motif rich in glycine and proline residues [35], followed by a Walker A-like motif [9]. However, unlike the ZiPD- and ELAC2-type flexible arms, the TM-type flexible arm is short and lacks the GP motif. Instead it contains a short stretch of mainly basic amino acids [35].

As anticipated, ELAC2-type flexible arm containing both the GP and Walker A-like motifs was found in the majority of the N-terminal halves of fungal tRNase Z^Ls (Figure 3 and Additional file 3). Unexpectedly, a subset of fungal tRNase Z^Ls appear to have an atypical ELAC2-type flexible arm which either lacks or contains an incomplete GP-motif (Figure 4). Moreover, unlike the TM-type flexible arm, this atypical flexible arm does not encompass a short cluster of basic amino acids.

Besides the flexible arm, the conserved domain search against the NCBI Conserved Domain Database (CDD) <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml> combined with manual evaluation revealed regions of sequences that match the Motifs I-IV and the PxKxRN loop in the N-terminal halves of tRNase Z^Ls. However, they appear to be nonfunctional as they differ from their original patterns in many positions including the key

Table 1 Representatives of candidate fungal tRNase Zs used in sequence alignment

Species [#]	Taxonomy	Protein name	Form	Accession Number	Database	No. aa ⁺
Ashbya gossypii (Ago)	Ascomycota	AgoTrz1	tRNase Z ^L	NP_984308	NCBI	821
Aspergillus nidulans (Ani)	Ascomycota	AniTrz1	tRNase Z ^L	ANID_11892.1	Broad	1083
Candida albicans (Cal)	Ascomycota	CalTrz1	tRNase Z ^L	XP_717703	NCBI	857
Coccidioides immitis (Cim)	Ascomycota	CimTrz1	tRNase Z ^L	CIHG_00067.1	Broad	962*
Fusarium graminearum (Fgr)	Ascomycota	FgrTrz1	tRNase Z ^L	FGSG_06635.3	Broad	840
Neurospora crassa (Ncr)	Ascomycota	NcrTrz1	tRNase Z ^L	NCU00232.4	Broad	1099
Pyrenophora tritici-repentis (Ptr)	Ascomycota	PtrTrz1	tRNase Z ^L	XP_001940780	NCBI	988*
Saccharomyces cerevisiae (Sce)	Ascomycota	SceTrz1	tRNase Z ^L	NP_013005	NCBI	838
Schizosaccharomyces pombe (Spo)	Ascomycota	SpoTrz1	tRNase Z ^{L1}	SPAC1D4.10	Broad	809
Schizosaccharomyces pombe (Spo)	Ascomycota	SpoTrz2	tRNase Z ^{L2}	SPBC3D6.03C	Broad	678
Sclerotinia sclerotiorum (Ssc)	Ascomycota	SscTrz1	tRNase Z ^L	XP_001586541	NCBI	832
Yarrowia lipolytica (Yli)	Ascomycota	YliTrz1	tRNase Z ^L	XP_500027	NCBI	815
Cryptococcus neoformans (Cne)	Basidiomycota	CneTrz1	tRNase Z ^L	CNBG_1589.2	Broad	1035
Malassezia globosa (Mgl)	Basidiomycota	MglTrz1	tRNase Z ^L	XP_001729151	NCBI	1109
Puccinia graminis (Pgr)	Basidiomycota	PgrTrz1	tRNase Z ^L	PGTG_11198.2	Broad	854
Spizellomyces punctatus (Spu)	Chytridiomycota	SpuTrz2	tRNase Z ^L	SPPG_00513.2	Broad	799*
Agaricus bisporus (Abi)	Basidiomycota	AbiTrz3	tRNase Z ^S	106812	JGI	376
Agaricus bisporus (Abi)	Basidiomycota	AbiTrz4	tRNase Z ^S	123256	JGI	469
Coprinopsis cinerea (Cci)	Basidiomycota	CciTrz2	tRNase Z ^S	CC1G_14603.2	Broad	390
Coprinopsis cinerea (Cci)	Basidiomycota	CciTrz3	tRNase Z ^S	CC1G_03806.2	Broad	544
Laccaria bicolor (Lbi)	Basidiomycota	LbiTrz2	tRNase Z ^S	XP_001876619	NCBI	406
Laccaria bicolor (Lbi)	Basidiomycota	LbiTrz3	tRNase Z ^S	XP_001874963	NCBI	476
Malassezia globosa (Mgl)	Basidiomycota	MglTrz2	tRNase Z ^S	EDP42443	NCBI	484*
Melampsora laricis-populina (Mla)	Basidiomycota	MlaTrz2	tRNase Z ^S	111950	JGI	473*
Postia placenta (Ppl)	Basidiomycota	PplTrz2	tRNase Z ^S	94043	JGI	386
Postia placenta (Ppl)	Basidiomycota	PplTrz3	tRNase Z ^S	92595	JGI	483*
Puccinia graminis (Pgr)	Basidiomycota	PgrTrz2	tRNase Z ^S	PGTG_13150.2	Broad	498
Spizellomyces punctatus (Spu)	Chytridiomycota	SpuTrz3	tRNase Z ^S	SPPG_06028.2	Broad	395

[#]Species abbreviations used in this study are shown in parentheses.

⁺ The number of amino acids in fungal tRNase Zs

* Indicates that mispredicted sequences obtained from the databases have been corrected.

residues critical for tRNase Z functions. We collectively termed these sequences pseudo-motifs. Figure 5 shows pseudo-motifs in representative candidate fungal tRNase Z^Ls. For comparison, we also include four metazoan tRNase Z^Ls from *C. elegans*, *D. melanogaster*, *A. thaliana* and humans. Since the pseudo-PxKxRN loops of *C. elegans* and human tRNase Z^Ls are indiscernible from their protein sequences, they are not included. Except for the pseudo-histidine motif [36], these pseudo-motifs have not been reported previously, which may reflect the difficulty in identifying these sequences. It appears that only the pseudo-histidine motif is widespread; the distributions of other pseudo-motifs are highly variable among fungal tRNase Z^Ls. Moreover, pseudo-Motif V could not be identified. It is likely that some of pseudo-motifs may have diverged too far and thus are no longer similar enough to the conserved motifs for homology to be detected by the NCBI conserved domain search. It is also notable that these pseudo-motifs were in the same

relative order as their original motifs in tRNase Z^Ss and in the C-terminal halves of tRNase Z^Ls.

Analysis of fungal tRNase Z^Ss

Like fungal tRNase Z^Ls, the lengths of predicted fungal tRNase Z^Ss are variable, ranging from 376 to 554 aa with an average size of ~442 aa. Sequence alignment of 15 selected representatives of fungal and non-fungal tRNase Z^Ss is presented in Figure 6. A list of all aligned fungal tRNase Z^Ss is provided in Additional file 4. Alignment revealed that like *B. subtilis* and human tRNase Z^Ss, fungal tRNase Z^Ss contain Motifs I-V and the PxKxRN, HEAT and HST loops (Figure 6). However, they display flexible arm diversity. A candidate flexible arm was also found in most of fungal tRNase Z^Ss. However, they exhibit considerable variation in amino acid sequence. Fungal tRNase Z^Ss can be grouped according to the presence or absence and the nature of sequences of the flexible arm. One group containing

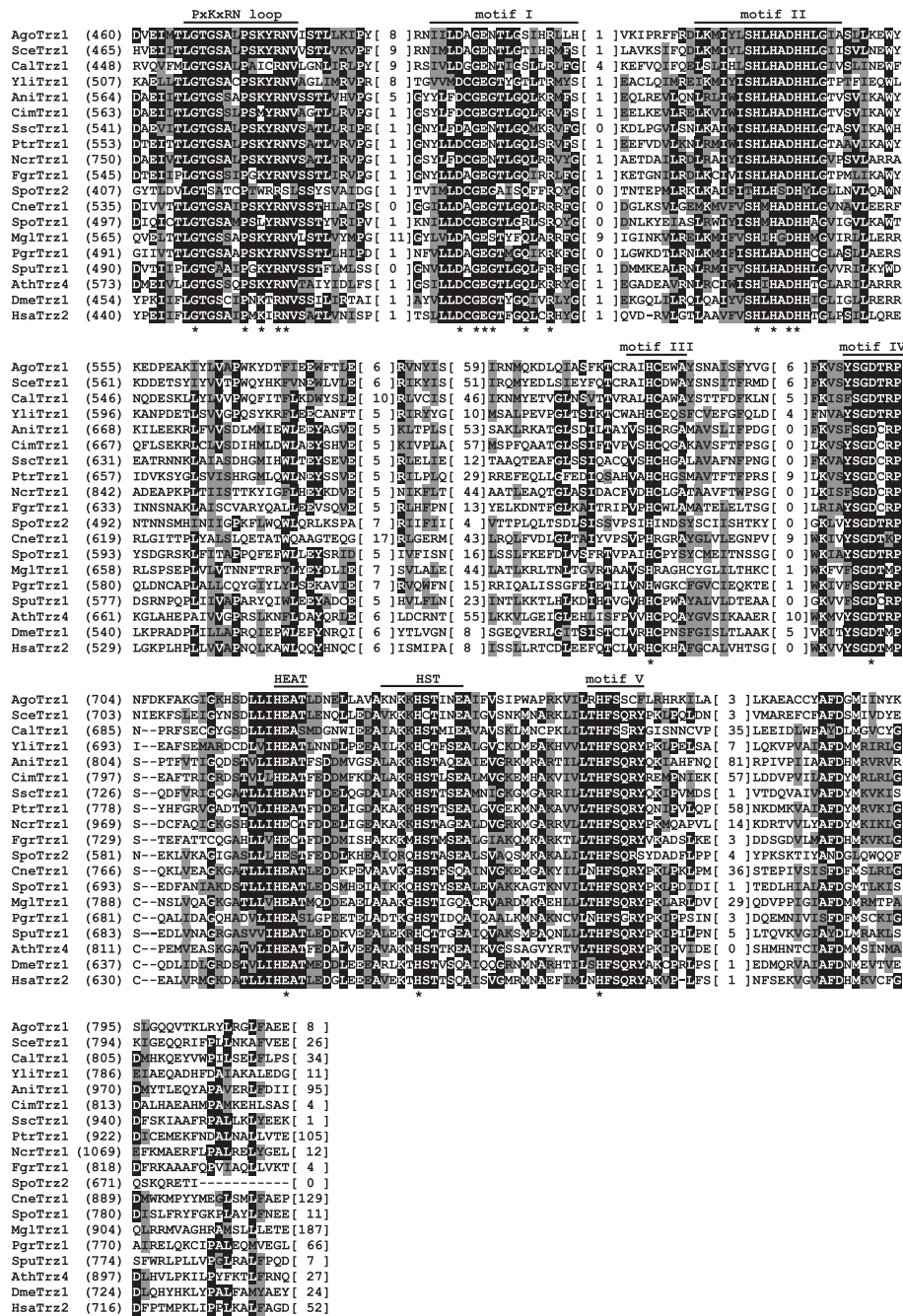
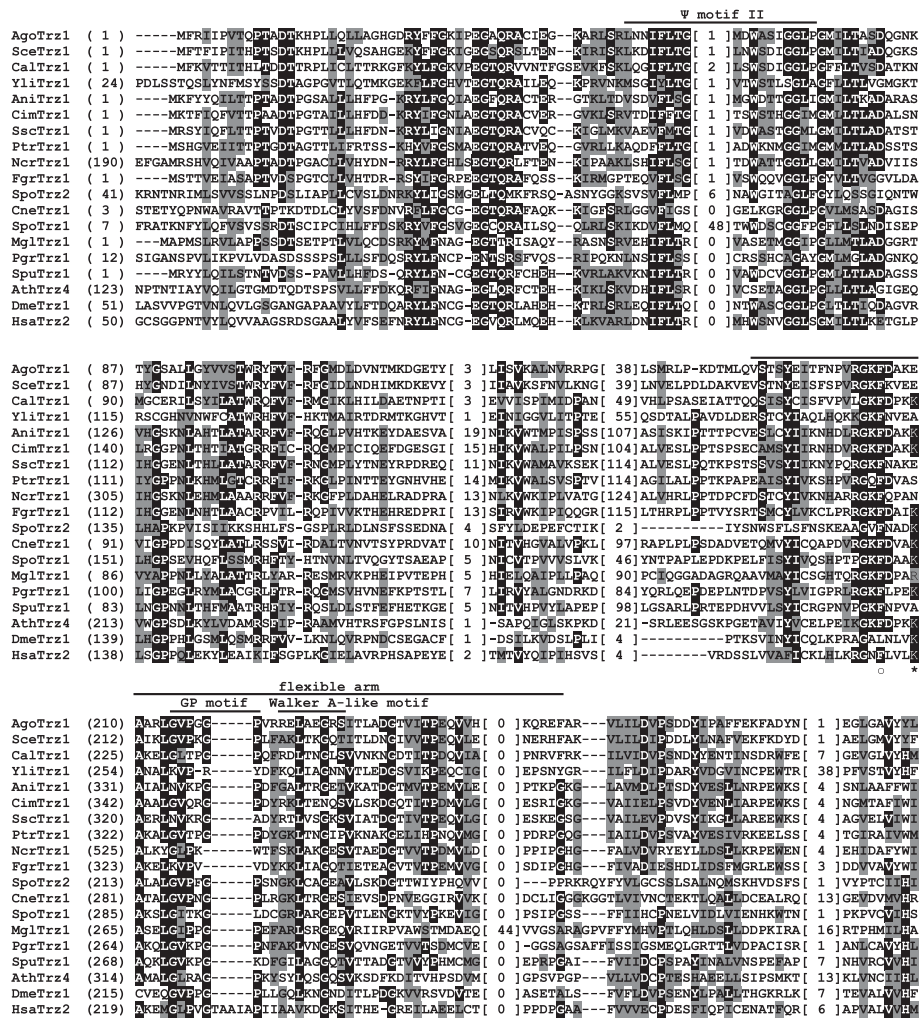


Figure 2 Sequence conservation between the C-terminal halves of fungal and non-fungal eukaryotic tRNase Z^Ls. Multiple sequence alignment of the C-terminal halves of representative fungal and non-fungal eukaryotic tRNase Z^Ls. Fungal tRNase Z^Ls are from *A. gossypii* (AgoTrz1), *S. cerevisiae* (SceTrz1) [19], *C. albicans* (CalTrz1), *Yarrowia lipolytica* (YliTrz1), *A. nidulans* (AniTrz1), *Coccidioides immitis* (CimTrz1), *Sclerotinia sclerotiorum* (SscTrz1), *Pyrenophora tritici-repentis* (PtrTrz1), *N. crassa* (NcrTrz1), *Fusarium graminearum* (FgrTrz1), *S. pombe* (SpoTrz1 and SpoTrz2) [14], *Cryptococcus neoformans* (CneTrz1), *M. globosa* (MglTrz1), *Puccinia graminis* (PgrTrz1) and *S. punctatus* (SpuTrz1). Non-fungal tRNase Z^Ls are from *A. thaliana* (AthTrz4) [13], *D. melanogaster* (DmeTrz1) [54] and *Homo sapiens* (HsaTrz2) [9]. Protein accession numbers are described in Table 1. The alignment was constructed using the program Clustal W [49]. Identical residues are on a black background and conserved residues shaded in gray. Also indicated above the sequence alignment are the conserved motifs involved in substrate binding and catalysis. The conserved motifs are labeled according to references [30,31,44]. Numbers in parentheses are the positions of the amino acid sequences. The numbers in brackets indicate the length of the region in tRNase Z, which are species-specific and could not be correctly aligned. Hyphens represent gaps introduced into sequences for maximum alignment. The positions of amino acid residues of the *D. melanogaster* tRNase Z^L critical for catalytic efficiency [30,31] are indicated by asterisk.



flexible arm

Walker A-like motif

GP motif

Figure 3 Sequence conservation between the N-terminal halves of fungal and non-fungal eukaryotic tRNase Z^Ls. Multiple sequence alignment of the N-terminal halves of representative fungal and non-fungal eukaryotic tRNase Z^Ls according to the same legend as in Figure 2. The positions of amino acid residues of the *D. melanogaster* tRNase Z^L that contribute to substrate binding are marked with an asterisk [44]. Open circle denotes the residue in the *D. melanogaster* tRNase Z^L that makes the largest contribution to substrate binding [44].

candidates from Basidiomycota species *A. bisporus* (AbiTrz4), *C. cinerea* (CciTrz3), *L. bicolor* (LbiTrz3) and *P. placenta* (PplTrz3) does not contain the flexible arm, which is located between Motif III and Motif IV. Interestingly, all these species have two tRNase Z^Ss.

The second group including candidates from Basidiomycota species *Agaricus bisporus* (AbiTrz3), *Coprinopsis cinerea* (CciTrz2), *Postia placenta* (PplTrz2) and *Laccaria bicolor* (LbiTrz2) lacks a recognizable GP motif but retain the Walker A-like motif. The third group

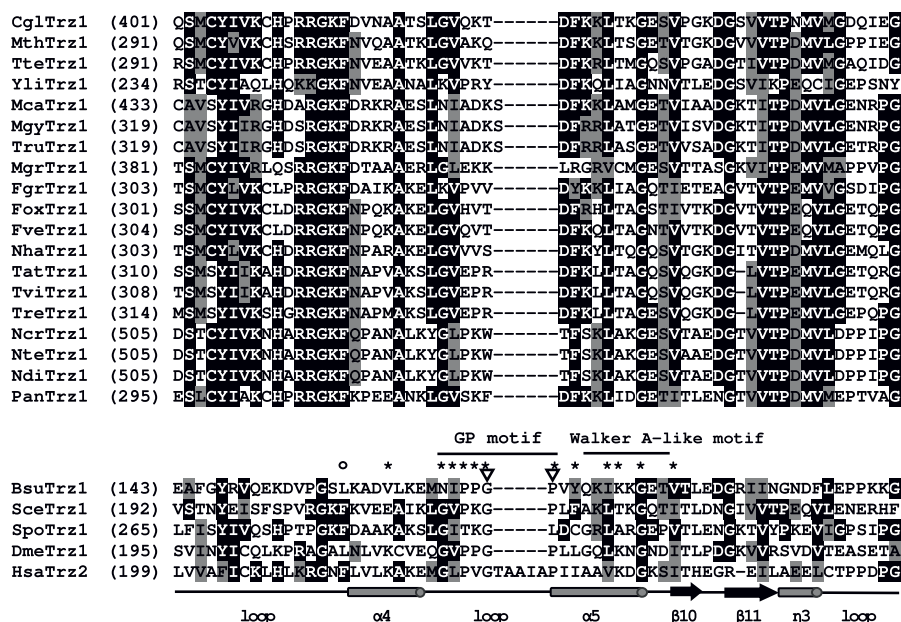


Figure 4 Amino acid alignments of putative atypical flexible arms in fungal tRNase Z^s. The top panel alignment shows putative atypical flexible arms found in certain fungal tRNase Z^s from *Chaetomium globosum* (CglTrz1), *Myceliophthora thermophila* (MthTrz1), *Thielavia terrestris* (TteTrz1), *Y. lipolytica* (YliTrz1), *Microsporium canis* (McaTrz1), *Microsporium gypseum* (MgyTrz1), *Trichophyton rubrum* (TruTrz1), *Magnaporthe grisea* (MgrTrz1), *F. graminearum* (FgrTrz1), *Fusarium oxysporum* (FoxTrz1), *Fusarium verticillioides* (FveTrz1), *Nectria haematococca* (NhaTrz1), *Trichoderma atroviride* (TatTrz1), *Trichoderma virens* (TviTrz1), *Trichoderma reesei* (TreTrz1), *N. crassa* (NcrTrz1), *N. tetrasperma* (NteTrz1), *N. discreta* (NdiTrz1), *Podospora anserina* (PanTrz1). The bottom panel alignment shows typical flexible arms found in tRNase Zs from *B. subtilis* (BsuTrz1) [55], *S. cerevisiae* (SceTrz1) [19], *S. pombe* (SpoTrz1) [14], *D. melanogaster* (DmeTrz1) [54] and *H. sapiens* (HsaTrz2) [9]. The predicted tRNase Z^s are defined in Additional file 1. The GP and Walker A-like motifs are indicated. Invariant and conserved amino acids are highlighted in Black and gray, respectively. The Asterisk indicates key residues in the *D. Melanogaster* tRNase Z^L that contribute to substrate binding [44]. The open circle denotes the residue in the *D. Melanogaster* tRNase Z^L that makes the largest contribution to substrate binding. Inverted triangles indicate the conserved glycine and proline residues in the GP motif, which are no longer conserved in certain fungal tRNase Z^s. Dark shading indicates identical residues, and gray shading designates conserved residues. Secondary structure assignment is based on the structure of the *B. subtilis* tRNase Z^S [34]. The secondary structures for α -helices, β -strands and 3_{10} -helix are indicated with Greek letters. Asterisks indicate key residues in the *D. Melanogaster* tRNase Z^L that contribute to substrate binding [44].

including Basidiomycota species *Malassezia globosa* (MglTrz2), *Melampsora laricis-populina* (MlaTrz2) and *Puccinia graminis* (PgrTrz2), and chytrid species *Spizellomyces punctatus* (SpuTrz3) lacks both recognizable GP and Walker A-like motifs and is considerably longer than those of the second group. Moreover, the sequence similarity between the flexible arms of fungal tRNase Z^s and the ZiPD-type flexible arm is mostly confined to their C-terminal sequences.

Discussion

The phylogenetic distribution of tRNase Z in fungi

tRNase Z^L is widespread in fungi and the majority of fungal species appear to have a single tRNase Z^L. This latter finding is somewhat unexpected given striking differences in the genome size, life cycle and morphology of species of fungi. The Pezizomycotina and Saccharomycotina belong to later diverging fungi. Their genomes vary considerably in size due to gene gain and loss events including tandem gene duplication, whole-

genome duplication and extensive gene loss [37,38]. The Saccharomycotina genome sizes vary from ~9 (*Pichia pastoris*) to ~24 Mb (*Candida parapsilosis*), whereas the genome sizes of the Pezizomycotina fungi range from 23 (*Microsporium canis*) to 43 Mb (*N. crassa*). Furthermore, Pezizomycotina and Saccharomycotina fungi range in complexity from unicellular yeasts to filamentous molds. However, despite their remarkable differences in genome size, life cycle and morphology, the Pezizomycotina and Saccharomycotina fungi tend to contain only one tRNase Z^L. These results indicate that the diversity of tRNase Z in fungi is not directly proportional to either the difference in genome size or the complexity of the life cycle and the morphology.

In contrast to tRNase Z^L, tRNase Z^S has a limited phylogenetic distribution. The apparent lack of the tRNase Z^S gene in the genomes of Ascomycota species suggests that it has been deleted from the genomes of Ascomycota fungi. It is possible that tRNase Z^S existed before the divergence of the Ascomycota from the

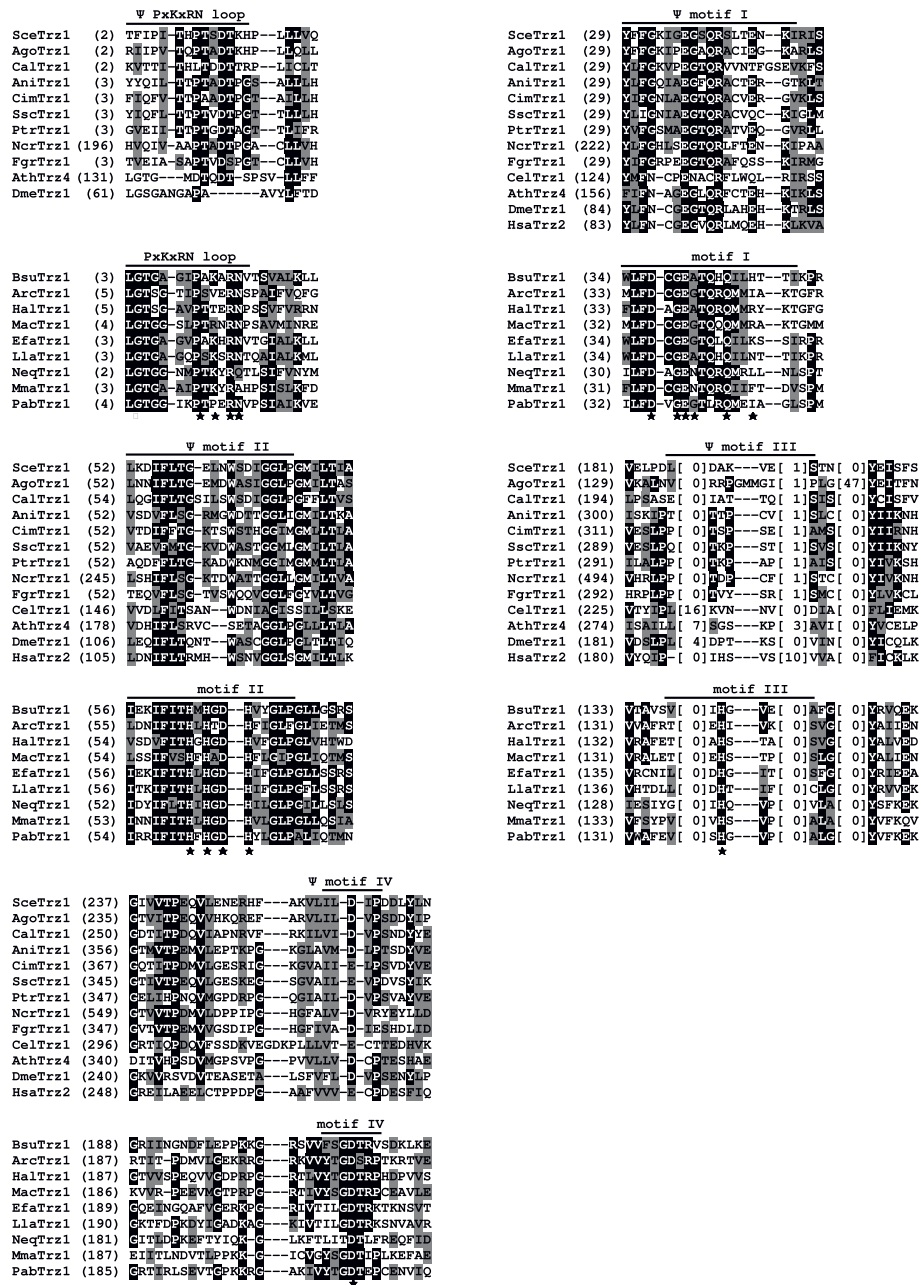


Figure 5 Pseudo-motifs in representative candidate fungal tRNase Z^Ls. The top panel shows alignment for pseudo substrate recognition and catalytic motifs found in tRNase Z^Ls from *S. cerevisiae* (SceTrz1) [19], *A. gossypii* (AgoTrz1), *C. albicans* (CalTrz1), *A. nidulans* (AniTrz1), *C. immitis* (CimTrz1), *S. sclerotiorum* (SscTrz1), *P. tritici-repentis* (PtrTrz1), *N. crassa* (NcrTrz1), *F. graminearum* (FgrTrz1), *C. elegans* (CelTrz1)[56], *D. melanogaster* (DmeTrz1) [54], *A. thaliana* (AthTrz4) [13] and humans (HsaTrz2) [9]. The bottom panel shows alignment for the corresponding conserved motifs in tRNase Zs from *B. subtilis* (BsuTrz1) [55], *Archaeoglobus fulgidus* (ArcTrz1; accession no. NP_069772), *Halobacterium salinarum* (HalTrz1; accession no. NP_280881), *Methanosarcina acetivorans* (MacTrz1; accession no. NP_617924), *Enterococcus faecalis* (EfaTrz1; accession no. NP_815399), *Lactococcus lactis subsp* (LlaTrz1; accession no. NP_266786), *Nanoarchaeum equitans* (NeqTrz1; accession no. NP_963358), *Methanococcus maripaludis* (MmaTrz1; accession no. NP_988026) and *Pyrococcus abyssi* (PabTrz1; accession no. NP_126781). The pseudo-PkKxRN loops of *C. elegans* and human tRNase Z^Ls are not included since they are indiscernible. The candidate fungal tRNase Z^Ls are defined in Additional file 1. GenBank accession numbers that are not listed in Additional file 1 are indicated. The alignment was extended to include surrounding regions to illustrate the further homology revealed by searching the Conserved Domain database (CDD), which could aid in the identification of pseudo-motifs. Similar or identical amino acid residues are shaded as described in the legend to Figure 2. The positions of amino acid residues of the *D. melanogaster* tRNase Z^L critical for catalytic efficiency [30,31] are indicated by asterisk.

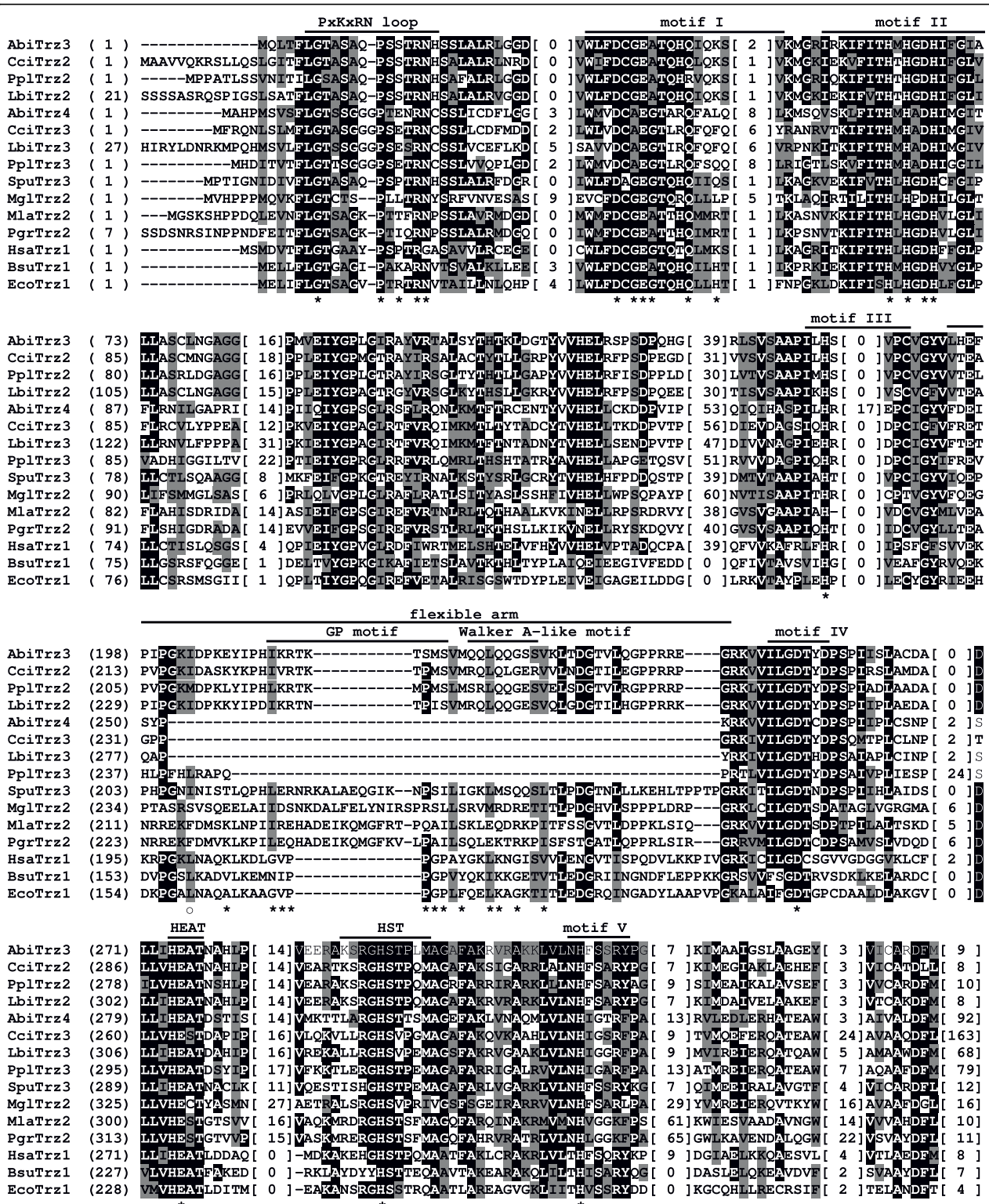


Figure 6 Multiple alignment fungal and non-fungal eukaryotic tRNase Z⁵s. tRNase Z⁵s are from *A. bisporus* (AbiTrz3 and AbiTrz4), *C. cinerea* (CciTrz2 and CciTrz3), *P. placenta* (PplTrz2 and PplTrz3), *L. bicolor* (LbiTrz2 and LbiTrz3), *S. punctatus* (SpuTrz3), *M. globosa* (MglTrz2), *Melampsora laricis-populina* (MlaTrz2) and *P. graminis* (PgrTrz2), as well as in *H. sapiens* (HsaTrz1) [9], *B. subtilis* (BsuTrz1) [55] and *E. coli* (EcoTrz1) [57]. See Table 1 for protein description. Alignment of eukaryotic tRNase Z⁵s as described in the legend to Figure 2.

Basidiomycota, and was subsequently lost after the appearance of a novel structure (tRNase Z^L). This is supported by the finding that tRNase Z^S is retained in the genomes of Basidiomycota, Chytridiomycota and Zygomycota fungi.

Most of Basidiomycota species examined contain one tRNase Z^S. The fungal tRNase Z^Ss appear to be unique among all known tRNase Z^Ss in either lacking the flexible arm or having an atypical flexible arm (see discussion below). One explanation of existence of tRNase Z^S genes in Basidiomycota species is that these genes may represent pseudo-tRNase Z^S genes. Another explanation, which we favor, is that fungal tRNase Z^S, at least some of them, may play a back-up or different role. Support for this hypothesis comes from recent studies of tRNase Z^S in *A. thaliana* and humans. In *A. thaliana*, one tRNase Z^S may represent a back-up for the nuclear tRNA 3'-end processing in case of dysfunction of nuclear-localized tRNase Z^L, whereas the other plays a role in chloroplasts [13]. In human cells, tRNase Z^S is located in the cytosol and likely have substrates other than pre-tRNA [39].

The flexible arms of candidate fungi tRNase Zs display remarkable diversity

An unexpected and striking result of this analysis is the diversity of the flexible arms within candidate fungal tRNase Zs, particularly tRNase Z^Ss. The most characteristic features of the typical flexible arm found in tRNase Zs are the GP- and Walker-A like motifs. A subset of fungal tRNase Z^Ls and all fungal tRNase Z^Ss appear to lack the GP-motif, and some fungal tRNase Z^Ss do not seem to have the Walker A-like motif. In the most extreme case, the flexible arm is missing in fungal tRNase Z^Ss. It is not yet understood why the flexible arms of fungal tRNase Zs display diversity in the primary sequence.

It is interesting to note that four Basidiomycota species contain two candidate tRNase Z^Ss, one of which lacks the flexible arm. It is likely that these two tRNase Z^Ss form heterodimers that would look like tRNase Z^L, where only the N-terminal half has a flexible arm.

The apparent lack of the GP-motif in some fungal tRNase Z^Ls and all fungal tRNase Z^Ss that we have examined raises the question of whether this motif is absolutely required for substrate binding. Structural and biochemical evidence suggests that the GP-motif may not be essential for pre-tRNA binding. To date, the three-dimensional structures of tRNase Z^Ss from *B. subtilis*, *E. coli* and *T. maritima* have been solved by X-ray crystallography [34,40-43]. Remarkably, the flexible arms of *T. maritima* tRNase Z^S lacking the GP motif and the other two tRNase Z^Ss harboring the GP motif have very similar structures, composed of a compact globular

domain and an extended two-stranded stalk, which extrude from the tRNase Z^S core. However, they have different lengths and globular domains. The globular domains at the end of the flexible arms of *B. subtilis* and *E. coli* tRNase Z^Ss are composed of two α -helices, two β -strands and one 3_{10} -helix, whereas the counterpart in *T. maritima* tRNase Z^S consists of one very short α -helix, one long helix and one 3_{10} -helix. The conserved GP-motif, particularly the proline residues, appears to add rigidity to two flexible arm helices since it is localized between them [42]. It would be interesting to know how the flexible arm lacking the GP-motif participates in substrate binding.

Recent biochemical studies have also suggested that the GP-motif may not be essential for substrate binding. Single alanine substitutions across the GP motif in *D. melanogaster* tRNase Z^L only moderately affect substrate binding. In contrast, substitution of a conserved leucine residue at the boundary of the globular domain and stalk with alanine almost completely abolishes substrate binding as the globular domain deletion [44]. Similarly, deletion of the GP motif in *B. subtilis* tRNase Z^S does not eliminate pre-tRNA binding but alters the cleavage specificity of the enzyme [45]. These results suggest that the GP motif may be important but not essential for substrate binding.

The evolutionary relationship between tRNase Z^S and tRNase Z^L

In eukaryotes, tRNase Z^L appears to take over tRNase Z^S in endonucleolytic 3'-end processing of pre-tRNAs, which raises the question of how it evolves. The protein sequence of tRNase Z^S is much more similar to the C-terminal half of tRNase Z^L than to the N-terminal half of tRNase Z^L. Furthermore, the C-terminal half of tRNase Z^L retains all conserved motifs for proper catalytic function but has lost the flexible arm involved in substrate binding, whereas the N-terminal half has lost all active motifs but contains the flexible arm. These observations led to the proposal that tRNase Z^L has evolved from tRNase Z^S by gene duplication and subsequent sequence divergence [9]. To assess whether phylogenetic evidence exists that is consistent with this notion, we estimated phylogenetic relationships among fungal tRNase Zs by using a Bayesian phylogenetic method. The clustering of all fungal tRNase Z^Ss with representative bacterial tRNase Z^Ss support the notion that tRNase Z^L comes to eukaryotes through duplication of tRNase Z^S gene. Further evidence that the N-terminal half of tRNase Z^L is derived from an ancient tRNase Z^S comes from our findings that the N-terminal half of fungal tRNase Z^L contains candidate pseudo-motifs and that these pseudo-motifs are present in the same relative order as their original motifs appeared in tRNase Z^S.

These pseudo-motifs likely represent relics of original tRNase Z^S motifs that were inactivated during diversification of the eukaryotic tRNase Z^L gene.

The reason for the adoption of tRNase Z^L over tRNase Z^S in eukaryotes is unknown. One possibility is that eukaryotic cells may require more efficient tRNase Z enzymes. Support for this proposal comes from biochemical characterization of human tRNase Zs. In vitro characterization of recombinant human tRNase Zs have shown that tRNase Z^L cleaves pre-tRNA significantly more efficiently compared to tRNase Z^S [46]. Although strong structural evidence to support that tRNase Z^L evolved into a more efficient enzyme than tRNase Z^S is still lacking, it is interesting to note that tRNase Z^S and tRNase Z^L may have different processing center numbers which would make much difference in the efficiency of pre-tRNA 3'-end processing. Three-dimensional structures of three bacterial tRNase Z^Ss have revealed that the proteins form homodimers [34,40-43]. In particular, the crystal structure of *B. subtilis* tRNase Z^S in complex with tRNA shows that the dimer has two identical processing centers with two substrate binding and catalytic sites. In contrast, a molecular modeling study has suggested that both the N-terminal and C-terminal halves of human tRNase Z^L can fold into two distinct MBL domains with one domain containing a fully functionally catalytic site and the other containing a candidate substrate binding domain [3].

Schizosaccharomyces fission yeasts have two tRNase Z^Ls, most likely targeted to the nucleus and mitochondria, respectively

Schizosaccharomyces fission yeasts including *S. pombe* appear to be unique among the Ascomycota in having two tRNase Z^Ls (tRNase Z^{L1} and tRNase Z^{L2}) that appear to be localized to the nucleus and mitochondria, respectively, as suggested in our previous study of tRNase Z^L in *S. pombe* [28]. Our fungal tRNase Z phylogeny shows that the two tRNase Z^Ls in fission yeasts may have arisen through gene duplication (Figure 1). Although the whole genome duplication found in the Saccharomycotina yeasts does not seem to occur in fission yeasts such as *S. pombe* [47], the tRNase Z^L gene could be duplicated by other mechanisms such as tandem and segmental gene duplication. The two tRNase Z^Ls in each *Schizosaccharomyces* species all have very limited homology with each other (20-23% identity and 31-33% similarity, see Additional file 5), indicating that these two proteins have diverged considerably from each other since their duplication. It is interesting to note that the tRNase Z^L gene could also be duplicated in non-fungal eukaryotic species *A. thaliana*. However, the two plant tRNase Z^Ls are highly related to each other (69% identity and 72% similarity).

Why do *Schizosaccharomyces* fission yeasts have two tRNase Z^Ls? In our previous study, it was found that the

nuclear-targeted tRNase Z^{L1} (SpoTrz1) is involved in nuclear pre-tRNA 3'-end processing in *S. pombe* [14]. Furthermore, its function can be compensated by either *S. cerevisiae* or human tRNase Z^L. Although the role of mitochondrial-targeted tRNase Z^{L2} (SpoTrz2) remains to be determined, it is likely that this protein plays an essential role in mitochondrial RNA processing [14]. Based on these results, it is possibly that the presence of two tRNase Z^Ls reflects that the nuclear and mitochondrial tRNA processing activities are associated with two different tRNase Z^Ls in *Schizosaccharomyces* fission yeasts. This may also hold true for wheat and potato since in these plants, enzymes involved in nuclear and mitochondrial tRNA 3'-end processing appear to be different [48]. However, it is important to note that the nuclear and organelle tRNase Z activities in the majority of organisms described to date seem to reside in the same enzyme.

Conclusions

A survey of fungal databases shows that tRNase Z^L appears to be universally present in fungi, whereas the presence of tRNase Z^S is restricted to certain fungal phyla, indicative of the fundamental role of tRNase Z^L in eukaryotic tRNA biogenesis. The apparent lack of tRNase Z^S in the Ascomycota suggests that tRNase Z^S may have lost before divergence of the Ascomycota and the Basidiomycota. A striking aspect of the tRNase Z^L distribution is that there are two different tRNase Z^Ls in *Schizosaccharomyces* fission yeasts. These two tRNase Z^Ls are likely present in different cellular compartments, suggesting functional partitioning between these two proteins. Phylogenetic analysis suggests that tRNase Z^S is ancestral to tRNase Z^L and that tRNase Z^L gene duplications may have occurred in certain fungal taxa, including *Schizosaccharomyces* fission yeasts. Sequence analysis reveals that the domain architecture of tRNase Z^Ls is highly conserved among fungi and metazoa. A surprising result of sequence analysis is the sequence diversity in the putative flexible arm of candidate fungal tRNase Zs. Our analysis also reveals pseudo-motifs at the N-terminal halves of tRNase Z^Ls. These findings support the view that tRNase Z^L evolved through duplication and divergence of the tRNase Z^S gene.

Methods

Fungal genome database search and protein sequence analysis

To identify candidate tRNase Zs, we conducted BLAST and PSI-BLAST searches using the known tRNase Z protein sequences as queries against fungal genomes databases including the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism=fungi), the Broad

Institute <http://www.broadinstitute.org/science/data>, the Joint Genome Institute <http://genome.jgi-psf.org/pages/fungi/home.jsf>, the Genome Center at Washington University <http://genome.wustl.edu/> and the Universal Protein Resource <http://www.uniprot.org>. All candidate sequences were obtained by using the cut-off E-value of 0.01. All candidate proteins were subjected to validation, which was carried out by using a variety of tests that evaluate the likelihood of annotation errors and the amino acid sequence conservation within and among taxonomic groups. First, confirmation of true candidate tRNase Zs was done by back-searching individual candidate protein sequence against the GenBank database. Second, the gene sequences for the predicted tRNase Zs were manually checked for possible sequence gaps. Third, multiple protein sequence alignment was used to identify candidate proteins that were discordant due to possible genomic sequencing errors and/or intron misprediction. In most cases, we changed the splicing pattern of candidate tRNase Z either using gene prediction programs Fgenesh <http://linux1.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind> and Geneid <http://genome.crg.es/geneid.html> or manually to restore the high degree of sequence conservation. Multiple sequence alignment was performed by using Clustal W [49], and the resulting alignment was further manually examined and adjusted to improve the detection of conserved regions. The putative subcellular localization signals of tRNase Zs were predicted by using the programs MitoProt <http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html> and PSORT II <http://psort.hgc.jp/>.

Phylogenetic analysis of fungal tRNase Zs

Full-length amino acid sequences of tRNase Zs from fungi and two bacteria, *B. subtilis* and *E. coli* were aligned by using Clustal W implemented in Mega 4.0 [50]. Conserved regions were selected and ambiguous aligned regions were removed by using the program Gblocks 0.91b [51]. tRNase Z^Ss from *B. subtilis* and *E. coli* were chosen as reference. The phylogenies were estimated by Bayesian inference with MrBayes 3.1.2 [52] using a mixture of the fixed amino acid models and the gamma distribution. Statistical confidence was assessed by using Markov Chain Monte Carlo (MCMC) sampling approaches. Four simultaneous Markov chains were run for one million generations sampling every 1,000 generation in two replicate runs. The first 250 trees were discarded as burn-in and the convergence of the chains was evaluated using AWTY implemented in MrBayes 3.1.2 [53].

List of abbreviations

pre-tRNA: tRNA precursor; tRNase Z: tRNA 3' endonuclease; tRNase Z^S: the short form of tRNase Z; tRNase

Z^L: the long form of tRNase Z; aa: amino acid; MBL: metallo- β -lactamase; NLS: nuclear localization signal; MTS: mitochondrial targeting signal; no: number; kDa: kiloDaltons

Additional material

Additional file 1: Distribution of candidate fungal tRNase Zs.

^aAbbreviations for species names are indicated in the parentheses. ^bThe number of amino acids in fungal tRNase Zs. ^calso known as *Histoplasma capsulatum*. ^dalso known as *Blastomyces dermatitidis*. ^ealso known as *Gibberella zeae* ^falso known as *Sporotrichum thermophile* ^galso known as *Fusarium solani* ^halso known as *Stagonospora nodorum* ⁱalso known as *Filobasidiella neoformans* ND denotes the sequence could not be predicted correctly likely due to sequencing errors. *Indicates that mispredicted sequences obtained from the databases have been corrected.

Additional file 2: Putative N-terminal mitochondrial targeting signals in candidate fungal tRNase Zs. The accession numbers for the proteins are listed in Additional file 1. The numbers refer to amino acid position starting from the N-terminus. ^aSpoTrz2 (SPBC3D6.03C) is localized to the mitochondria [28].

Additional file 3: Alignment of candidate fungal tRNase Z^Ls. Similar or identical amino acid residues are shaded as described in the legend to Figure 2. The conserved motifs are labeled according to references [30,31,44].

Additional file 4: Alignment of candidate fungal tRNase Z^Ss. Similar or identical amino acid residues are shaded as described in the legend to Figure 2. The conserved motifs are labeled according to references [30,31,44].

Additional file 5: Pairwise sequence comparisons of tRNase Z^Ss from *Schizosaccharomyces* species. The accession numbers for proteins are listed in Additional file 1. The pairwise percent identity (I) and percent similarity (S) between tRNase Z^Ss from *Schizosaccharomyces* species were calculated using the Clustal W program [49].

Acknowledgements

We thank Dr. Jie Yan and two anonymous reviewers for improving the quality of the manuscript. We are grateful to Dr. Guang Yang for the use of his laboratory facilities to construct the phylogenetic tree. This work was supported by grants from the National Science Foundation of China (30771178) and Nanjing Normal University (2007104XGQ0148).

Authors' contributions

WZ conducted online database searches and carried out the sequence alignment. SL constructed the phylogenetic tree. WZ, HY and YH analyzed the data. YH wrote the manuscript. All authors read and approved the final manuscript.

Received: 26 April 2010 Accepted: 6 September 2010

Published: 6 September 2010

References

1. Ceballos M, Vioque A: tRNase Z. *Protein Pept Lett* 2007, **14**:137-145.
2. Hartmann RK, Gossringer M, Spath B, Fischer S, Marchfelder A: Chapter 8 The Making of tRNAs and More - RNase P and tRNase Z. *Prog Nucleic Acid Res Mol Biol* 2009, **85C**:319-368.
3. Redko Y, Li de Lasierra-Gallay I, Condon C: When all's zed and done: the structure and function of RNase Z in prokaryotes. *Nat Rev Microbiol* 2007, **5**:278-286.
4. Spath B, Canino G, Marchfelder A: tRNase Z: the end is not in sight. *Cell Mol Life Sci* 2007, **64**:2404-2412.
5. Aravind L: An evolutionary classification of the metallo-beta-lactamase fold proteins. *In Silico Biol* 1999, **1**:69-91.

6. Daiyasu H, Osaka K, Ishino Y, Toh H: **Expansion of the zinc metallo-hydrolase family of the beta-lactamase fold.** *FEBS Lett* 2001, **503**:1-6.
7. Dominski Z: **Nucleases of the metallo-beta-lactamase family and their role in DNA and RNA metabolism.** *Crit Rev Biochem Mol Biol* 2007, **42**:67-93.
8. Schiffer S, Rosch S, Marchfelder A: **Assigning a function to a conserved group of proteins: the tRNA 3'-processing enzymes.** *EMBO J* 2002, **21**:2769-2777.
9. Tavtigian SV, Simard J, Teng DH, Abtin V, Baumgard M, Beck A, Camp NJ, Carillo AR, Chen Y, Dayananth P, Desrochers M, Dumont M, Farnham JM, Frank D, Frye C, Ghaffari S, Gupte JS, Hu R, Iliev D, Janecki T, Kort EN, Laity KE, Leavitt A, Leblanc G, McArthur-Morrison J, Pederson A, Penn B, Peterson KT, Reid JE, Richards S, et al: **A candidate prostate cancer susceptibility gene at chromosome 17p.** *Nat Genet* 2001, **27**:172-180.
10. Sullivan KD, Steiniger M, Marzluff WF: **A core complex of CPSF73, CPSF100, and Symplekin may form two different cleavage factors for processing of poly(A) and histone mRNAs.** *Mol Cell* 2009, **34**:322-332.
11. Baillat D, Hakimi MA, Naar AM, Shilatfard A, Cooch N, Shiekhattar R: **Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II.** *Cell* 2005, **123**:265-276.
12. Mathy N, Benard L, Pellegrini O, Daou R, Wen T, Condon C: **5'-to-3' exonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA.** *Cell* 2007, **129**:681-692.
13. Canino G, Bocian E, Barbezier N, Echeverria M, Forner J, Binder S, Marchfelder A: ***Arabidopsis* encodes four tRNase Z enzymes.** *Plant Physiol* 2009, **150**:1491-1502.
14. Zhao Z, Su W, Yuan S, Huang Y: **Functional conservation of tRNase Z^L among *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and humans.** *Biochem J* 2009, **422**:483-492.
15. Lee YS, Shibata Y, Malhotra A, Dutta A: **A novel class of small RNAs: tRNA-derived RNA fragments (tRFs).** *Genes Dev* 2009, **23**:2639-2649.
16. Wilusz JE, Freier SM, Spector DL: **3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA.** *Cell* 2008, **135**:919-932.
17. Bogerd HP, Karnowski HW, Cai X, Shin J, Pohlars M, Cullen BR: **A mammalian herpesvirus uses noncanonical expression and processing mechanisms to generate viral MicroRNAs.** *Mol Cell* 2010, **37**(1):135-142.
18. Elbarbary RA, Takaku H, Uchiumi N, Tamiya H, Abe M, Nishida H, Nashimoto M: **Human cytosolic tRNase Z^L can downregulate gene expression through miRNA.** *FEBS Lett* 2009, **583**:3241-3246.
19. Chen Y, Beck A, Davenport C, Chen Y, Shattuck D, Tavtigian SV: **Characterization of TRZ1, a yeast homolog of the human candidate prostate cancer susceptibility gene ELAC2 encoding tRNase Z.** *BMC Mol Biol* 2005, **6**:12.
20. Hawksworth DL: **Fungal diversity and its implication for genetic resource collections.** *Stud Mycol* 2004, **50**:9-18.
21. Fitzpatrick DA, Logue ME, Stajich JE, Butler G: **A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis.** *BMC Evol Biol* 2006, **6**:99.
22. James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung GH, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schüssler A, Longcore JE, O'Donnell K, Mozley-Standridge S, et al: **Reconstructing the early evolution of Fungi using a six-gene phylogeny.** *Nature* 2006, **443**:818-822.
23. Liu Y, Leigh JW, Brinkmann H, Cushion MT, Rodriguez-Ezpeleta N, Philippe H, Lang BF: **Phylogenomic analyses support the monophyly of Taphrinomycotina, including *Schizosaccharomyces fission* yeasts.** *Mol Biol Evol* 2009, **26**:27-34.
24. Robertse B, Reeves JB, Schoch CL, Spatafora JW: **A phylogenomic analysis of the Ascomycota.** *Fungal Genet Biol* 2006, **43**:715-725.
25. Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, Huhndorf S, James T, Kirk PM, Lücking R, Thorsten Lumbsch H, Lutzoni F, Matheny PB, McLaughlin DJ, Powell MJ, Redhead S, Schoch CL, Spatafora JW, Stalpers JA, Vilgalys R, Aime MC, Aptroot A, Bauer R, Begerow D, Benny GL, Castlebury LA, Crous PW, Dai YC, Gams W, Geiser DM, et al: **A higher-level phylogenetic classification of the Fungi.** *Mycol Res* 2007, **111**:509-547.
26. Takaku H, Minagawa A, Takagi M, Nashimoto M: **A candidate prostate cancer susceptibility gene encodes tRNA 3' processing endoribonuclease.** *Nucleic Acids Res* 2003, **31**:2272-2278.
27. Takaku H, Minagawa A, Takagi M, Nashimoto M: **The N-terminal half-domain of the long form of tRNase Z is required for the RNase 65 activity.** *Nucleic Acids Res* 2004, **32**:4429-4438.
28. Matsuyama A, Arai R, Yashiroda Y, Shirai A, Kamata A, Sekido S, Kobayashi Y, Hashimoto A, Hamamoto M, Hiraoka Y, Horinouchi S, Yoshida M: **ORFeome cloning and global analysis of protein localization in the fission yeast *Schizosaccharomyces pombe*.** *Nat Biotechnol* 2006, **24**:841-847.
29. Keeling PJ, Slamovits CH: **Simplicity and complexity of microsporidian genomes.** *Eukaryot Cell* 2004, **3**:1363-1369.
30. Zareen N, Hopkinson A, Levinger L: **Residues in two homology blocks on the amino side of the tRNase Z His domain contribute unexpectedly to pre-tRNA 3' end processing.** *RNA* 2006, **12**:1104-1115.
31. Karkashon S, Hopkinson A, Levinger L: **tRNase Z catalysis and conserved residues on the carboxy side of the His cluster.** *Biochemistry* 2007, **46**:9380-9387.
32. Minagawa A, Takaku H, Takagi M, Nashimoto M: **A novel endonucleolytic mechanism to generate the CCA 3' termini of tRNA molecules in *Thermotoga maritima*.** *J Biol Chem* 2004, **279**:15688-15697.
33. Spath B, Kirchner S, Vogel A, Schubert S, Meinschmidt P, Aymanns S, Nezzar J, Marchfelder A: **Analysis of the functional modules of the tRNA 3' endonuclease (tRNase Z).** *J Biol Chem* 2005, **280**:35440-35447.
34. de la Sierra-Gallay IL, Pellegrini O, Condon C: **Structural basis for substrate binding, cleavage and allostery in the tRNA maturase RNase Z.** *Nature* 2005, **433**:657-661.
35. Schilling O, Spath B, Kostecky B, Marchfelder A, Meyer-Klaucke W, Vogel A: **Exosite modules guide substrate recognition in the ZIPD/ElaC protein family.** *J Biol Chem* 2005, **280**:17857-17862.
36. Vogel A, Schilling O, Spath B, Marchfelder A: **The tRNase Z family of proteins: physiological functions, substrate specificity and structural properties.** *Biol Chem* 2005, **386**:1253-1264.
37. Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, De Montigny J, Marck C, Neuvéglise C, Talla E, Goffard N, Frangeul L, Aigle M, Anthouard V, Babour A, Barbe V, Barnay S, Blanchin S, Beckerich JM, Beyne E, Bleykasten C, Boisramé A, Boyer J, Cattolico L, Confanioleri F, De Daruvar A, Despons L, Fabre E, Fairhead C, Ferry-Dumazet H, et al: **Genome evolution in yeasts.** *Nature* 2004, **430**:35-44.
38. Kellis M, Birren BW, Lander ES: **Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*.** *Nature* 2004, **428**:617-624.
39. Takahashi M, Takaku H, Nashimoto M: **Regulation of the human tRNase Z^S gene expression.** *FEBS Lett* 2008, **582**:2532-2536.
40. Ishii R, Minagawa A, Takaku H, Takagi M, Nashimoto M, Yokoyama S: **Crystal structure of the tRNA 3' processing endoribonuclease tRNase Z from *Thermotoga maritima*.** *J Biol Chem* 2005, **280**:14138-14144.
41. Ishii R, Minagawa A, Takaku H, Takagi M, Nashimoto M, Yokoyama S: **The structure of the flexible arm of *Thermotoga maritima* tRNase Z differs from those of homologous enzymes.** *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2007, **63**:637-641.
42. Kostecky B, Pohl E, Vogel A, Schilling O, Meyer-Klaucke W: **The crystal structure of the zinc phosphodiesterase from *Escherichia coli* provides insight into function and cooperativity of tRNase Z-family proteins.** *J Bacteriol* 2006, **188**:1607-1614.
43. Li de la Sierra-Gallay I, Mathy N, Pellegrini O, Condon C: **Structure of the ubiquitous 3' processing enzyme RNase Z bound to transfer RNA.** *Nat Struct Mol Biol* 2006, **13**:376-377.
44. Levinger L, Hopkinson A, Desetty R, Wilson C: **Effect of changes in the flexible arm on tRNase Z processing kinetics.** *J Biol Chem* 2009, **284**:15685-15691.
45. Minagawa A, Ishii R, Takaku H, Yokoyama S, Nashimoto M: **The flexible arm of tRNase Z is not essential for pre-tRNA binding but affects cleavage site selection.** *J Mol Biol* 2008, **381**:289-299.
46. Yan H, Zareen N, Levinger L: **Naturally occurring mutations in human mitochondrial pre-tRNA^{Ser}(UCN) can affect the transfer ribonuclease Z cleavage site, processing kinetics, and substrate secondary structure.** *J Biol Chem* 2006, **281**:3926-3935.
47. Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouras J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A,

- Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, *et al.*: **The genome sequence of *Schizosaccharomyces pombe***. *Nature* 2002, **415**:871-880.
48. Mayer M, Schiffer S, Marchfelder A: **tRNA 3' processing in plants: nuclear and mitochondrial activities differ**. *Biochemistry* 2000, **39**(8):2096-2105.
 49. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: **Clustal W and Clustal X version 2.0**. *Bioinformatics* 2007, **23**:2947-2948.
 50. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0**. *Mol Biol Evol* 2007, **24**:1596-1599.
 51. Castresana J: **Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis**. *Mol Biol Evol* 2000, **17**:540-552.
 52. Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models**. *Bioinformatics* 2003, **19**:1572-1574.
 53. Nylander JA, Wilgenbusch JC, Warren DL, Swofford DL: **AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics**. *Bioinformatics* 2008, **24**:581-583.
 54. Dubrovsky EB, Dubrovskaya VA, Levinger L, Schiffer S, Marchfelder A: **Drosophila RNase Z processes mitochondrial and nuclear pre-tRNA 3' ends in vivo**. *Nucleic Acids Res* 2004, **32**(1):255-262.
 55. Pellegrini O, Nezzar J, Marchfelder A, Putzer H, Condon C: **Endonucleolytic processing of CCA-less tRNA precursors by RNase Z in *Bacillus subtilis***. *EMBO J* 2003, **22**(17):4534-4543.
 56. Smith MM, Levitan DJ: **The *Caenorhabditis elegans* homolog of the putative prostate cancer susceptibility gene ELAC2, *hoe-1*, plays a role in germline proliferation**. *Dev Biol* 2004, **266**(1):151-160.
 57. Ezraty B, Dahlgren B, Deutscher MP: **The RNase Z homologue encoded by *Escherichia coli* *elaC* gene is RNase BN**. *J Biol Chem* 2005, **280**(17):16542-16545.

doi:10.1186/1471-2148-10-272

Cite this article as: Zhao *et al.*: Identification and analysis of candidate fungal tRNA 3'-end processing endonucleases tRNase Zs, homologs of the putative prostate cancer susceptibility protein ELAC2. *BMC Evolutionary Biology* 2010 **10**:272.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

