A survey of green plant tRNA 3′-end processing enzyme tRNase Zs, homologs of the candidate prostate cancer susceptibility protein ELAC2

Lijuan Fan1,2,3, Zhikang Wang1,2,3, Jinyu Liu1,2,3, Weili Guo1,2,3, Jie Yan4 and Ying Huang1,2,3*

Abstract

Background: tRNase Z removes the 3′-trailer sequences from precursor tRNAs, which is an essential step preceding the addition of the CCA sequence. tRNase Z exists in the short (tRNase ZS) and long (tRNase ZL) forms. Based on the sequence characteristics, they can be divided into two major types: bacterial-type tRNase ZS and eukaryotic-type tRNase ZL, and one minor type, Thermotoga maritima (TM)-type tRNase ZS. The number of tRNase Zs is highly variable, with the largest number being identified experimentally in the flowering plant Arabidopsis thaliana. It is unknown whether multiple tRNase Zs found in A. thaliana is common to the plant kingdom. Also unknown is the extent of sequence and structural conservation among tRNase Zs from the plant kingdom.

Results: We report the identification and analysis of candidate tRNase Zs in 27 fully sequenced genomes of green plants, the great majority of which are flowering plants. It appears that green plants contain multiple distinct tRNase Zs predicted to reside in different subcellular compartments. Furthermore, while the bacterial-type tRNase ZSs are present only in basal land plants and green algae, the TM-type tRNase ZSs are widespread in green plants. The protein sequences of the TM-type tRNase ZSs identified in green plants are similar to those of the bacterial-type tRNase ZSs but have distinct features, including the TM-type flexible arm, the variant catalytic HEAT and HST motifs, and a lack of the PxKxRN motif involved in CCA anti-determination (inhibition of tRNase Z activity by CCA), which prevents tRNase Z cleavage of mature tRNAs. Examination of flowering plant chloroplast tRNA genes reveals that many of these genes encode partial CCA sequences. Based on our results and previous studies, we predict that the plant TM-type tRNase ZSs may not recognize the CCA sequence as an anti-determinant.

Conclusions: Our findings substantially expand the current repertoire of the TM-type tRNase ZSs and hint at the possibility that these proteins may have been selected for their ability to process chloroplast pre-tRNAs with whole or partial CCA sequences. Our results also support the coevolution of tRNase Zs and tRNA 3′-trailer sequences in plants.

Keywords: tRNA precursor (pre-tRNA), endonuclease, tRNase Z, ELAC2, post-transcriptional processing, green plant, phylogenetics

Background
tRNA 3′-end maturation is a process through which the 3′-trailer sequence of precursor tRNAs (pre-tRNAs) is removed, and processed tRNAs acquire the CCA end which is absolutely essential for tRNA aminoacylation and protein synthesis (for reviews, see [1-3]). In prokaryotes, this process can be either exonucleolytic or endonucleolytic depending on whether the 3′-CCA sequence is genomically encoded. CCA-containing pre-tRNAs are generally processed by the exonucleases that tend to stop removing nucleotides from the 3′-end upon encountering the transcriptionally encoded CCA, whereas CCA-less pre-tRNAs are processed by a 3′-endonuclease termed tRNase Z (also termed RNase Z or 3′-tRNase; for reviews, see [4-7]) that cleaves immediately after the N73 discriminator nucleotide (the first
unpaired base after the acceptor stem) to allow subsequent addition of the CCA sequence.

Unlike prokaryotic pre-tRNAs, eukaryotic and organellar pre-tRNAs generally lack the 3′-CCA sequence (which is added post-transcriptionally) and their 3′-trailer sequences are removed by tRNase Z. Also unlike prokaryotic pre-tRNAs, eukaryotic nuclear pre-tRNAs contain oligo (U) at their 3′-ends, which are recognized and bound by the La protein (for reviews see [2,8]). In the budding yeast Saccharomyces cerevisiae and fission yeast Schizosaccharomyces pombe, the endonucleolytic cleavage of nuclear pre-tRNAs requires the presence of the yeast La protein [9,10]. In the absence of the yeast La protein, the 3′-trailer sequence of nuclear pre-tRNAs is trimmed by 3′-exoribonucleases including Rex1p [11]. However, organellar pre-tRNAs lack terminal oligo (U). Furthermore, unlike nuclear pre-tRNAs which are typically monostronic, most organellar pre-tRNAs are polystronic [12,13].

tRNase Z is present in all kingdoms of life. It exists in two forms: tRNase Z\(^S\) [300-400 amino acids (aa)] and tRNase Z\(^L\) (700-800 aa), which are encoded by different genes. It is believed that the tRNase Z\(^L\) gene has evolved from a tandem duplication of the tRNase Z\(^S\) gene, followed by divergence of the sequence [14]. In prokaryotes, only tRNase Z\(^S\) is identified. By contrast, all eukaryotes possess tRNase Z\(^L\), and some have both forms.

The species distribution of tRNase Z is complex. The majority of eukaryotic species analyzed to date, including S. cerevisiae, the fruit fly Drosophila melanogaster and the nematode worm Caenorhabditis elegans contain a single tRNase Z\(^L\) [15-17]. In contrast, S. pombe have two tRNase Z\(^L\)'s [18,19]. Interestingly, two tRNase Z\(^L\)'s and two tRNase Z\(^S\)'s have been experimentally identified in the flowering plant Arabidopsis thaliana [20]. In humans, one tRNase Z\(^S\) (also termed ELAC1) and one tRNase Z\(^L\) (also termed ELAC2) are found [14]. Our BLAST searches against public genomic and expressed sequence tag (EST) databases reveal that with few exceptions, vertebrates contain one tRNase Z\(^L\) and one tRNase Z\(^S\) (a detailed description of tRNase Z protein distribution in the animal kingdom will be provided elsewhere).

tRNase Z belongs to the metallo-\(\beta\)-lactamase (MBL) superfamily [14,21-24]. The typical MBL domain contains five conserved sequence motifs termed Motifs I-V. Motifs I and IV each harbor an invariant Asp, Motif II (HxHxDH), which is also called the His motif, is the signature motif of the superfamily, whereas Motifs III and V each contain a conserved His residue. Structural studies of tRNase Z\(^S\)'s from E. coli, T. maritima and B. subtilis [25-28] and mutation analyses of tRNase Z\(^S\) from a variety of species [29-35] reveal that the His and Asp residues of Motifs II-V form the active site for coordination of two catalytic zinc ions. In particular, the Asp residue of Motif II may participate in both zinc ion coordination and act as a general base to generate a hydroxide ion for nucleophilic attack on the scissile phosphodiester bond at the cleavage site [25,29]. The Asp residue of Motif I is also catalytically important and appears to stabilize the catalytic site [33].

Besides tRNase Z\(^S\), some nucleic acid processing enzymes are also members of the MBL superfamily. Most of these proteins belong to the \(\beta\)-CASP (MBL-associated CISF Artemis SNM1/PSO2) subfamily of the MBL [23]. This subfamily includes the 73-kD subunit of the cleavage and polyadenylation specificity factor (CPSF-73) and its yeast homolog Ysh1p, which are involved in endonucleolytic cleavage of pre-mRNA, the Integrator complex subunit 11 (Int11) involved in the 3′-end formation of small nuclear RNAs (snRNA) [36], bacterial RNase J, which participates in rRNA 5′-end maturation [37] and RNA decay [38], and the eukaryotic Pso/Snm1/Artemis proteins, which function in DNA repair and V(D)J recombination [23]. However, unlike tRNase Z\(^S\), \(\beta\)-CASP proteins contain conserved \(\beta\)-CASP sequence motifs in place of Motif V.

tRNase Z is distinguished from other MBL members by their unique substrate binding domain termed the flexible arm (also termed the exosite). Based on flexible arm type, there are two major types (bacterial- and eukaryotic-types) and one minor type [T. maritima (TM)-type] of tRNase Z\(^S\) [39]. The bacterial-type tRNase Z\(^S\), which are present predominantly in bacteria, possess the bacterial-type flexible arm. The bacterial-type flexible arm is \(\sim 55\) aa in length and contains the Gly- and Pro-rich GP motif (GxPxGP, sometimes GxPPGP) [39]. The eukaryotic-type tRNase Z\(^S\), which are found only in eukaryotes, contain the ELAC2-type flexible arm. This type of flexible arm harbors the GP motif and is \(\sim 62\) aa long, which is slightly longer than the bacterial-type flexible arm.

The TM-type tRNase Z was believed to be the minor type at the time of discovery since it was found only in T. maritima and A. thaliana [39]. The flexible arm found in TM-type tRNase Z\(^S\) appears to be shorter (\(\sim 30\) aa) and lacks the GP motif but instead contains one short basic residue-rich region [39]. In addition, both the bacterial- and eukaryotic-type tRNase Z\(^S\) contain the PxKxRN, HEAT and HST motifs, which form part of loop structures, whereas the TM-type tRNase Z\(^S\) lacks these motifs [33,40,41]. The PxKxRN motif has been suggested to function in CCA anti-determination (tRNase Z activity is inhibited by 3′-CCA) [25,33], whereas the HEAT and HST motifs have been suggested to play a role in facilitating proton transfer at the final stage of reaction [25,29,40].
tRNase Z has diverse functions besides its primary role in tRNA 3’-end processing. This is perhaps best exemplified by ELAC2, which serves a multitude of functions within cells. Recent studies have shown that ELAC2 is involved in the generation of MALAT1, a cancer-associated long noncoding RNA which participates in regulation of pre-mRNA splicing [42], tRNA-derived small RNAs [43,44], and viral microRNAs (miRNAs) [45,46]. Overexpression of ELAC2 delays cell cycle progression, suggesting that ELAC2 may be involved in cell cycle control either directly or indirectly via its role as tRNA processing enzyme [47]. ELAC2 also potentiates TGF-β(transforming growth factor-β)/Smad-induced transcription response, indicating a role for ELAC2 in TGF-β/Smad signaling mediated growth arrest [48]. Interestingly, a recent study has shown that destruction of human mitochondria through depletion of mitochondrial DNA results in down-regulation of ELAC2 and a delay in cell cycle progression [49]. Since ELAC2 may be involved in cell cycle regulation, it is likely that ELAC2 may link mitochondrial function and cell cycle control. It is important to note that ELAC2 is a candidate prostate cancer susceptibility gene as its mutations are associated with prostate cancer [14]. However, the underlying mechanisms are unknown. In S. cerevisiae, either inactivating mutations or overexpression of tRNase ZL causes a petite phenotype, suggesting that the action of tRNase ZS may be related to mitochondrial function [15]. In addition, the S. cerevisiae tRNase ZL has also been suggested to play a role in 35S rRNA processing [50].

The study of tRNase Z evolution has been facilitated by the increasing availability of genome sequences. A previous study showed that only tRNase ZS is found in bacteria and that its presence in bacteria is widespread [6]. We recently reported on a systematic survey of tRNase Zs in fungi [51]. Our analysis reveals that while the majority of fungal species contain one tRNase ZL, all four sequenced Schizosaccharomyces species contain two distinct tRNase ZL’s either demonstrated or predicted to be localized to the nucleus and mitochondria, respectively. In addition, the presence of tRNase ZS in fungi is restricted to the phylum Basidiomycota and the basal fungal phyla.

Green plants (Viridiplantae) represent a monophyletic group of land plants and green algae that evolved near the base of the tree of eukaryotic life. Flowering plants (angiosperms), which are typically polyplody, represent the largest, most diverse and most evolutionary advanced phylum of land plants making up 90% of the plant kingdom. It can be divided into two major groups: dicotyledons (dicots), which accounts for the majority of the angiosperm species, and monocotyledons (monocots). At present, there are at least 27 sequenced and annotated genomes representing the major taxonomic groups within green plants, although the majority of them are those of flowering plants. The public availability of these genome sequences enabled us to identify tRNase Zs in green plants and to study their evolution.

In this study, we undertook a comprehensive survey of candidate tRNase Zs from annotated green plant genomes. To understand the evolutionary relationships among green plant tRNase Zs, we further conducted a phylogenetic analysis of these newly identified candidates. Finally, we presented a detailed sequence analysis of tRNase Zs with the intent of further delineating the distinct features of green plant tRNase Zs.

**Results**

**Identification of candidate green plant tRNase Zs**

To extend our previous study of tRNase Z diversity and evolution, we searched public genome databases for putative green plant tRNase Zs with significant matches to known bacterial and eukaryotic tRNase Zs. Since most of candidate sequences identified from the databases are computationally generated without subsequent manual annotation, it is likely that many predictions may contain errors. Therefore, we verified each candidate. We first validated each prediction by reciprocal searches against the GenBank. In back-searches, a candidate was confirmed if reverse BLAST also gave tRNase Z hits in the top matches. 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. We found that many candidate sequences are apparently incomplete or contain annotation errors. For example, the predicted coding sequence of the monkeyflower (*Mimulus guttatus*) MguTRZ2 (Phytozome accession no. mgv1a024577 m.g) in the database was incomplete lacking the N-terminal region. We were able to predict this region from the genomic DNA sequence based on sequence similarity. The predicted full-length coding region of MguTRZ2 has 364 aa. As another example, the sequence annotated as the candidate castor bean (*Ricinus communis*) tRNase ZL (Phytozome accession no. 30146.t000117) appears to be mispredicted due to the presence of sequence gaps. Thus, this sequence was excluded from the list.

Several incorrect predictions are apparently caused by the presence of the non-canonic GC-AG splice site pairs. While the GT donor splice site is a canonical 5’-splice site for introns in eukaryotic genes, the GC donor splice sites account for the majority of the non-canonical donor splice site. Thus, for those that cannot be accurately predicted by the conventional FGENESH program, we carried out gene prediction using FGENESH.
_GC, which is a new version of the FGENESH program including noncanonical GC dinucleotide in donor splice sites. Indeed, by doing so, we could predict some exons encoding missing conserved motifs. For example, the HEAT motif was originally missing in several annotated candidate tRNase Z’s from flowering plants including papaya (Carica papaya CpaTRZ3), cassava (Manihot esculenta MesTRZ4), Medicago (Medicago truncatula MtTRZ3) and black cottonwood (Populus trichocarpa PtTRZ3). After re-evaluation of intron splice sites using FGENESH_GC, we were able to recover their HEAT motifs. Because the intron sequences of some candidate sequences appear to be extremely difficult to predict correctly, we could not conclusively rule out the possibility of errors in certain candidates.

In total, 54 candidate tRNase ZS and 32 candidate tRNase ZL were identified from 27 green plant species including 21 flowering plants, 1 moss, 1 lycophyte and 4 green algae (Additional file 1). The names of these candidates follow the A. thaliana tRNase Z nomenclature [52]. Of these, only tRNase Zs from A. thaliana have been experimentally characterized [20]. While most of these sequenced genomes examined belong to the flowering plants, the availability of the genome sequences from two basal land plants (the bryophyte Physcomitrella patens and the lycophyte Selaginella moellendorffii) and four green algae (Chlamydomonas reinhardtii, Volvox carteri, Micromonas pusilla and Ostreococcus lucimarinus) allows for evaluation of differences between flowering plant tRNase Zs and those from basal land plants and green algae.

The flowering plants examined to date appear to contain multiple tRNase Zs. Foxtail millet (Setaria italica) contains the largest number of tRNase Zs (5) so far identified in a flowering plant. Most flowering plant species have a single tRNase Z. In contrast, six flowering plants including two Arabidopsis species (A. thaliana and Arabidopsis lyrata) harbor two tRNase Zs. It should be noted that the presence of two tRNase Zs is not species-specific since the plant species containing two tRNase Zs come from diverse taxonomic groups.

Unlike most flowering plants that possess two tRNase Zs, two members of the Panicoideae subfamily of grasses, sorghum (Sorghum bicolor) and foxtail millet have three tRNase Zs and four tRNase Zs, respectively. The presence of multiple tRNase Zs appears not to be grass-specific, since three other grass plants including rice (Oryza sativa), Brachypodium (Brachypodium distachyon) and maize (Zea mays) contain only two tRNase Zs. Calculation of the percentage identity and similarity between candidate tRNase Zs from these two grass species shows strong conservation of the proteins at the amino acid level with the most identity (93%) and similarity (95%) between S. bicolor SbiTRZ2 and S. italica SitTRZ2 (Additional files 2 and 3). To our surprise, all candidate tRNase ZSs identified in flowering plants are highly similar and belong to the TM-type tRNase ZSs (see below for a detailed discussion).

The number and type of tRNase Zs appear to be highly variable in the two primitive plant species and four green algae (Table 1 and Additional file 1). The largest number of tRNase Zs is found in the moss P. patens, which contains two TM-type and one bacterial-type tRNase ZSs, and one tRNase ZL. In contrast, two green algae C. reinhardtii and V. carteri have the least number of tRNase Zs, comprising one TM-type tRNase ZS and one tRNase ZL. The two other green algae M. pusilla and O. lucimarinus contain one TM-type and one bacterial-type tRNase ZSs and one tRNase ZL. The lycophyte S. moellendorffii, which has the smallest genome size of any land plant reported, contains two tRNase ZSs in addition to one TM-type tRNase ZS.

Interestingly, tRNase Z-like proteins (TLP) which apparently lack one or more conserved motifs of tRNase Zs necessary for the enzymatic activity of the protein are widespread in basal land plant and green algal genomes that have been analyzed here (Table 1 and Additional file 1). Many tRNase Z-like proteins appear to lack the flexible arm but contain all other conserved motifs of tRNase Zs (Figure 1 and see below for a more detailed discussion of the motifs). At one extreme, one S. moellendorffii tRNase Z-like protein (SmoTLP2) appears to lack all of the conserved motifs.

Based on their sizes, tRNase Z-like proteins can be divided into tRNase ZS-like and tRNase ZL-like proteins which are comparable in size to tRNase ZSs and tRNase ZLs, respectively. The number and form of tRNase Z-like proteins vary among the species, being largest in C. reinhardtii (two tRNase ZS-like and two tRNase ZL-like proteins) and next largest in S. moellendorffii (two tRNase ZS-like and one tRNase ZL-like proteins). The basal land plant P. patens and the two green algae M. pusilla and O. lucimarinus contain one tRNase ZL-like protein, whereas the green alga V. carteri contains one tRNase ZS-like and one tRNase ZL-like proteins. In contrast, black cottonwood (P. trichocarpa) appears to be the only species among the flowering plant genomes examined that contains the tRNase ZL-like protein. This species contains one tRNase ZL-like protein in which the second His in the His motif is mutated to Gln (data not shown).

**Prediction of subcellular localization of candidate tRNase Zs from flowering plants**

To help understand the function of candidate tRNase Zs, we predicted *in silico* the subcellular localization of each of flowering plant tRNase Zs using different methods. The results show that tRNase Zs are distributed in various cellular compartments, with some tRNase Zs being primarily localized to the cytoplasm, plastids, or mitochondria. This suggests that tRNase Zs play specific roles in protein synthesis in different cellular compartments.
| Table 1 Distribution of candidate tRNase Zs from representative green plants |
|---------------------------------------------------------------------------|
| **Species** | **Protein name** | **Type** | **Accession number** | **Database** | **No. aa** |
|---------------|-----------------|----------|----------------------|--------------|------------|
| DICOTS        |                 |          |                      |              |            |
| Aquilegia coerulea | AcoTRZ1       | TM-type tRNase Z^2 | AcoGoldSmith_v1.019418 m.g | Phytozome    | 283        |
| Aquilegia coerulea | AcoTRZ2       | TM-type tRNase Z^3 | AcoGoldSmith_v1.007495 m.g | Phytozome    | 356        |
| Aquilegia coerulea | AcoTRZ3       | tRNase Z^3 | AcoGoldSmith_v1.000749 m.g | Phytozome    | 997*       |
| Aquilegia coerulea | AcoTRZ4       | tRNase Z^4 | AcoGoldSmith_v1.024294 m.g | Phytozome    | 848*       |
| Arabidopsis thaliana | AthTRZ1      | TM-type tRNase Z^2 | NP_1776082 | NCBI        | 280        |
| Arabidopsis thaliana | AthTRZ2      | TM-type tRNase Z^2 | NP_178532.2 | NCBI        | 354        |
| Arabidopsis thaliana | AthTRZ3      | tRNase Z^2 | AcoGoldSmith_v1.019418 m.g | Phytozome    | 354        |
| Arabidopsis thaliana | AthTRZ4      | tRNase Z^2 | AcoGoldSmith_v1.019418 m.g | Phytozome    | 354        |
| Canica papaya | CpaTRZ1       | TM-type tRNase Z^2 | evm.TU.supercontig_17.217 | Phytozome    | 306        |
| Canica papaya | CpaTRZ2       | TM-type tRNase Z^2 | evm.TU.supercontig_132.51 | Phytozome    | 359*       |
| Citrus clementina | CclTRZ1       | TM-type tRNase Z^2 | clementine09_017657 m.g | Phytozome    | 288        |
| Citrus clementina | CclTRZ2       | TM-type tRNase Z^2 | clementine09_014182 m.g | Phytozome    | 355        |
| Citrus clementina | CclTRZ3       | tRNase Z^2 | clementine09_014182 m.g | Phytozome    | 355        |
| Cucumis sativus | CsaTRZ1       | TM-type tRNase Z^2 | Cusca369470  | Phytozome    | 301*       |
| Cucumis sativus | CsaTRZ2       | TM-type tRNase Z^2 | Cusca185180  | Phytozome    | 347        |
| Eucalyptus grandis | EgrTRZ1       | TM-type tRNase Z^2 | Egrandis_v1_0.020509 m.g | Phytozome    | 309        |
| Eucalyptus grandis | EgrTRZ2       | TM-type tRNase Z^2 | Egrandis_v1_0.021229 m.g | Phytozome    | 357*       |
| Eucalyptus grandis | EgrTRZ3       | tRNase Z^2 | Egrandis_v1_0.001658 m.g | Phytozome    | 971        |
| Glycine max | GmaTRZ1       | TM-type tRNase Z^2 | Glycma20g01480 | Phytozome    | 279        |
| Glycine max | GmaTRZ2       | TM-type tRNase Z^2 | ACUI8735.1 | NCBI        | 353        |
| Glycine max | GmaTRZ3       | tRNase Z^2 | Glycma13g43270 | Phytozome    | 923*       |
| Glycine max | GmaTRZ4       | tRNase Z^2 | Glycma15g02070 | Phytozome    | 931*       |
| Mimulus guttatus | MguTRZ1       | TM-type tRNase Z^2 | mgv1a010642 m.g | Phytozome    | 306        |
| Mimulus guttatus | MguTRZ2       | TM-type tRNase Z^2 | mgv1a024577 m.g | Phytozome    | 364*       |
| Mimulus guttatus | MguTRZ3       | tRNase Z^2 | mgv1a009391 m.g | Phytozome    | 976        |
| Vitis vinifera | VviTRZ1       | TM-type tRNase Z^2 | XP_002730581 | NCBI        | 290        |
| Vitis vinifera | VviTRZ2       | TM-type tRNase Z^2 | XP_002272741.1 | NCBI        | 354        |
| Vitis vinifera | VviTRZ3       | tRNase Z^2 | XP_002278956.1 | NCBI        | 951        |
| MONOCOTS       |                 |          |                      |              |            |
| Oryza sativa japonica | OsaTRZ1     | TM-type tRNase Z^2 | NP_001046280.1 | NCBI        | 302        |
| Oryza sativa japonica | OsaTRZ2     | TM-type tRNase Z^2 | LOC_Os09g30466 | Phytozome    | 365        |
| Oryza sativa japonica | OsaTRZ3     | tRNase Z^2 | B9EU17 | UniProt       | 964        |
| Zea mays      | ZmaTRZ1       | TM-type tRNase Z^2 | GRMZM2G147727 | Phytozome    | 302        |
| Zea mays      | ZmaTRZ2       | TM-type tRNase Z^2 | ACGR7847.1 | NCBI        | 357        |
| Zea mays      | ZmaTRZ3       | tRNase Z^2 | GRMZM2G379286 | Phytozome    | 930        |
| MOSSES        |                 |          |                      |              |            |
| Physcomitrella patens | PpaTRZ1   | TM-type tRNase Z^2 | XP_001785449.1 | NCBI        | 295        |
| Physcomitrella patens | PpaTRZ2   | TM-type tRNase Z^2 | Pp1s17_61V6 | Phytozome    | 420        |
| Physcomitrella patens | PpaTRZ3   | Bacterial-type tRNase Z^2 | Pp1s1409_46V6 | Phytozome    | 336*       |
| Physcomitrella patens | PpaTRZ4   | tRNase Z^2 | Pp1s126_39V6 | Phytozome    | 984*       |
| Physcomitrella patens | PpaTLP1   | tRNase Z^2-like | Pp1s3_58V6 | Phytozome    | 923*       |
| LYCOPHYTES    |                 |          |                      |              |            |
| Selaginella moellendorfii | SmoTRZ1   | TM-type tRNase Z^2 | 16771 | Phytozome    | 299        |
| Selaginella moellendorfii | SmoTRZ2   | tRNase Z^2 | 111409 | Phytozome    | 771*       |
| Selaginella moellendorfii | SmoTRZ3   | tRNase Z^2 | 97243 | Phytozome    | 819*       |
| Selaginella moellendorfii | SmoTLP1   | tRNase Z^2-like | 404484 | Phytozome    | 303        |
| Selaginella moellendorfii | SmoTLP2   | tRNase-like | XP_002967989 | NCBI        | 346        |
| Selaginella moellendorfii | SmoTLP3   | tRNase Z^2-like | 416540 | Phytozome    | 731        |
bioinformatic prediction programs. Most flowering plants have two tRNase Zs (tRNase ZS1 and tRNase ZS2). All tRNase ZS1 proteins apparently lack any predictable signal sequences and therefore predicted to be cytoplasmic proteins, with the exception of *O. sativa* tRNase ZS1 (OsaTRZ1), which is predicted to be localized in the chloroplasts (Table 2). On the other hand, all tRNase ZS2 except *P. trichocarpa* tRNase ZS2 (PtrTRZ2), contain a predicted chloroplast targeting signal (Table 2). The predicted chloroplast targeting signal ofPtrTRZ2 is unusual in that it is only seven residues long since the majority of the known chloroplast targeting signals consist of N-terminal 20-80 amino acid residues [53]. In the two flowering plants (*S. bicolor* and *S. italica*) carrying multiple tRNase Zs, only one tRNase ZS (tRNase ZS2) from each species is predicted to be chloroplast-localized, whereas the rest are predicted to be cytoplasmic (Table 2).

Besides tRNase ZS, flowering plants also contain one (tRNase ZL1) or two (tRNase ZL1 and tRNase ZL2) tRNase ZLs. Most tRNase ZL1 proteins contain a predicted mitochondrial targeting signal between two putative translational initiation sites at the N-terminus and also have a nuclear localization signal (Table 3). Some tRNase ZL2 proteins have a putative mitochondrial targeting signal, while others are predicted to have both nuclear and mitochondrial targeting signals. The length of the putative mitochondrial targeting signals found in these candidates are within the expected size-range (20-80 aa) [54]. Consistent with the prediction, *A. thaliana* tRNase ZL1 (AthTRZ3) was found in both the nucleus and the mitochondria [20]. However, *A. thaliana* tRNase ZL2 (AthTRZ4) predicted to have both nuclear and mitochondrial targeting signals is actually found only in the mitochondria [20].

**Phylogenetic analysis**

To gain insights into the evolutionary relationship among plant tRNase Zs, we reconstructed the phylogeny of 86 candidate sequences using Bayesian phylogenetics. Although most of plant species are flowering plants, they are taxonomically diverse. In addition to tRNase Zs from plants, we also include tRNase Zs from *T. maritima* as outgroup. Phylogenetic analysis reveals the presence of two well-supported clades: one formed by including all TM-type tRNase Zs, and the other all tRNase ZLs and bacterial-type tRNase ZSs (Figure 2). The latter clade is further divided into two subclades, of which one contains all tRNase ZSs and bacterial-type tRNase ZSs (Figure 2). The former clade is composed of two subclades, of which one contains all tRNase ZSs and bacterial-type tRNase ZSs (Figure 2). The latter clade is further divided into two subclades, of which one contains all tRNase ZSs and bacterial-type tRNase ZSs (Figure 2). The former clade is composed of two subclades, of which one contains all tRNase ZSs and bacterial-type tRNase ZSs (Figure 2).
Alignment of bacterial-type and tRNase Z-like candidates in green plants

Bacterial-type tRNase Z’s are from *P. patens* (PpaTRZ3), *O. lucimarinus* (OtTRZ2), *M. pusilla* (MpuTRZ2), *Synechocystis* sp. PCC 6803 (slr0050), *E. coli* (EcoTRZ1) [72], *B. subtilis* (BsuTRZ1) [60] and humans (*HsaTRZ1*) [14]. Plant tRNase Z-like (TLP) candidates are from *S. moellendorffii* (SmoTLP1 and SmoTLP2), *V. carteri* (VcaTLP1), *C. reinhardtii* (CreTLP1 and CreTLP2), TM-type tRNase Z’s from *T. maritima* (TmaTrz1) [30] and *A. thaliana* (AthTRZ1) [20] are included for comparison. The alignment was constructed using Clustal W [60]. Identical residues are on a black ground and conserved residues shared in gray. Also indicated above the alignment are conserved motifs: Motif III, Motif IV, Flexible arm, HEAT and HST motifs. 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. Amino acid residues predicted to be critical for activity are indicated by a star.
| Protein name | Form | Chl | Chloroplast targeting signal |
|--------------|------|-----|-----------------------------|
| AcoTrz1      | AcoTrz1 | -      | -                           |
| AcoTrz2      | AcoTrz2 | C    | MKISLGISTSPILLPSKPTLHNRVICIP (31) |
| AlyTrz1      | AlyTrz1 | -      | -                           |
| AlyTrz2      | AlyTrz2 | C    | MQLSSFPIPSKIFFTFKYHAPPVIHQLAACLQISHRSNVPSVWSVYGSSISRAIEEVEYRKAR (67) |
| AthTrz1      | AthTrz1 | -      | -                           |
| AthTrz2      | AthTrz2 | C    | MQLSSFPIPSKPIFSPSTHIHKPVTHOQLAACLQISHRNRHFFPSVWSVYGSSISRAIEEVEYRKAR (68) |
| BdiTrz1      | BdiTrz1 | -      | -                           |
| BdiTrz2      | BdiTrz2 | C    | MATISLFLPSLRLTRGLSSGASRFQTL (32) |
| CpaTrz1      | CpaTrz1 | -      | -                           |
| CpaTrz2      | CpaTrz2 | C    | MLYSVPSLQSVPSFLHSHISHPTHIQFQLFYV (40) |
| CctTrz1      | CctTrz1 | -      | -                           |
| CctTrz2      | CctTrz2 | C    | MQLSPLTPSPLKTFPHPSISIPPSIQPHQTVRQHLSQHGVGLNAKLSAGFSSISRAIDEEVEYRKAR (68) |
| CsiTrz1      | CsiTrz1 | -      | -                           |
| CsiTrz2      | CsiTrz2 | C    | MQLSPLTPSPLKTFPHPSISIPPSIQPHQTVRQHLSQHGVGLNAKLSAGFSSISRAIDEEVEYRKAR (68) |
| CsaTrz1      | CsaTrz1 | -      | -                           |
| CsaTrz2      | CsaTrz2 | C    | MQSIPIPSLQPSVQPPFQHQLHQHPGVVQLSHNPSFSDGSLTGVE (55) |
| EgrTrz1      | EgrTrz1 | -      | -                           |
| EgrTrz2      | EgrTrz2 | C    | MRISLPISPSQSIPQFPHHIIHRHRHQICLDRRRLPALSAPAVNILKSGYLSTIHR (59) |
| GmaTrz1      | GmaTrz1 | -      | -                           |
| GmaTrz2      | GmaTrz2 | C    | MQISLSDACKTQLFPHFIFKPPPLNHQVST (35) |
| MesTrz1      | MesTrz1 | -      | -                           |
| MesTrz2      | MesTrz2 | C    | MQTSFVPISPSKIFPSFHPHPIHHNPTAPTENHICLQLQTYKRPINSTSLKSSGFAISGNAIEEVEYKAR (73) |
| MtrTrz1      | MtrTrz1 | -      | -                           |
| MtrTrz2      | MtrTrz2 | C    | MQIPLTNPFTKQFQPHHIIIFPSKQFQLSFPT (33) |
| MguTrz1      | MguTrz1 | -      | -                           |
| MguTrz2      | MguTrz2 | C    | MIFLPPTSNVNLHPPFHPTPHATIKIHLHPQTRRVSIAVCVSDAVV (45) |
| OsaTrz1      | OsaTrz1 | -      | -                           |
| OsaTrz2      | OsaTrz2 | C    | MANSGKSSPAATSSATAPPPGRPKAKAPPLTEGISIFNGITQCVF (52) |
| PteTrz1      | PteTrz1 | -      | -                           |
| PteTrz2      | PteTrz2 | C    | MQISIPL (7)                  |
| PtrTrz1      | PtrTrz1 | -      | -                           |
| PtrTrz2      | PtrTrz2 | C    | MQISIPL (7)                  |
| RcoTrz1      | RcoTrz1 | -      | -                           |
| RcoTrz2      | RcoTrz2 | C    | MQTSFVPISPSKIFPSFHPHPIHHNPTAPTENHICLQLQTYKRPINSTSLKSSGFAISGNAIEEVEYKAR (73) |
| SitTrz1      | SitTrz1 | -      | -                           |
| SitTrz2      | SitTrz2 | C    | MATASLFPISPSLRLSSRTATLRLSFQTLAARKPP (34) |
| SitTrz3      | SitTrz3 | -      | -                           |
| SbiTrz1      | SbiTrz1 | -      | -                           |
| SbiTrz2      | SbiTrz2 | C    | MATASLPSLRALSRTSARCSRFQTLAARKPVESSSTATS (43) |
| SbiTrz3      | SbiTrz3 | -      | -                           |
| VviTrz1      | VviTrz1 | -      | -                           |
| VviTrz2      | VviTrz2 | C    | MQISLPFSTSPVPSLPLNPPLTPQPLTPKPHKYTI (39) |
| ZmaTrz1      | ZmaTrz1 | -      | -                           |
| ZmaTrz2      | ZmaTrz2 | C    | MVTASLFLPSLRRLSRTSAHGPPQFLAAKRPVESSTTSGSRRGSGGKAGLSSLVLR (59) |

Identification of the putative chloroplast targeting signals was made by using ChloroP http://www.cbs.dtu.dk/services/ChloroP/. Chl stands for the chloroplast. C indicates chloroplast localization. "-", the localization could not be predicted. The numbers refer to amino acid positions starting from the N-terminus.
other with a posterior probability value of 1. These observations suggest that tRNase ZL gene duplication took place in certain species.

Conservation of candidate green plant tRNase ZSs
To assess the presence and conservation of motifs in candidate green plant tRNase Zs, we performed multiple sequence alignments of identified tRNase ZS sequences. Candidate tRNase ZSs are analyzed first. A list of aligned sequence conservation between the green plant tRNase ZSs and non-plant species is largely confined to the flexible arms show only weak protein sequence homology to the bacterial-type flexible arms, and have distinctive features including the absence of the GP motif and a harbothe TM-type tRNase ZS.

The putative nuclear localization signals were predicted using PSORT http://psort.hgc.jp/form.html, while the putative mitochondrial targeting signals were predicted using MITOPROT http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html. Nuc: stand for the nucleus; Mito: mitochondria. N indicates nuclear localization and M denotes mitochondrial localization. "-", the localization could not be predicted. The numbers refer to amino acid positions starting from the N-terminus.

From Nuc Mito Predicted mitochondrial targeting sequences

| Protein name | From Nuc Mito Predicted mitochondrial targeting sequences |
|--------------|--------------------------------------------------------|
| AcoTrz3      | N                                                     |
| AcoTrz4      | M                                                     |
| AtyTrz3      | N                                                     |
| AtyTrz4      | N                                                     |
| AthTrz3      | N                                                     |
| AthTrz4      | N                                                     |
| BdiTrz3      | N                                                     |
| CpaTrz3      | N                                                     |
| CctTrz3      | N                                                     |
| CsiTrz3      | N                                                     |
| CsaTrz3      | N                                                     |
| EgrTrz3      | N                                                     |
| GmaTrz3      | N                                                     |
| GmaTrz4      | N                                                     |
| MesTrz3      | N                                                     |
| MesTrz4      | N                                                     |
| MguTrz3      | N                                                     |
| OsaTrz3      | M                                                     |
| PfrTrz3      | N                                                     |
| PpeTrz3      | N                                                     |
| SbiTrz4      | N                                                     |
| VviTrz3      | N                                                     |
| ZmaTrz3      | N                                                     |

The putative nuclear localization signals were predicted using PSORT http://psort.hgc.jp/form.html, while the putative mitochondrial targeting signals were predicted using MITOPROT http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html. Nuc: stand for the nucleus; Mito: mitochondria. N indicates nuclear localization and M denotes mitochondrial localization. "-", the localization could not be predicted. The numbers refer to amino acid positions starting from the N-terminus.
Figure 2: Phylogram showing phylogenetic relationships among candidate green plant tRNase Zs. Posterior probabilities, which are indicated at the nodes, are generated by the Bayesian analysis. The scale bar indicates 0.1 nucleotide substitutions per site. For each protein, the species, the accession number and the database can be found in Additional file 1. Taxonomic designations are indicated on the right side of the tree.
This Lys- and Leu-rich motif was previously unappreciated in *A. thaliana*, likely due to the limited number of plant sequences available at the time analysis was performed [39]. As might be expected, there are variations in the consensus sequence (Figures 3 and 4).

Second, most green plant tRNase Z's lack the PxKxRN motif normally present on the amino side of Motif I. Instead, they share a Gly-rich consensus sequence EGxSxxG in an analogous position to the PxKxRN motif (Figure 3). In some cases, variants in the consensus sequence are found (Figures 3 and 4). Notably, there is very weak sequence similarity between this motif and the corresponding region in *T. maritima* tRNase Zs.

Third, most green plant tRNase Z's contain conserved xExT and HxH motifs in place of the HEAT and HST motifs, respectively. The xExT motif is related to the HEAT motif except that the invariant His is replaced by a Thr/Met/Val/Leu residue and the invariant Ala is often replaced by Ser, or occasionally Met, Thr and Cys (Figures 3 and 4). In contrast, the HxH motif is related to the HST motif, except that the conserved Ser is mostly replaced by Thr, and the conserved Thr is substituted by His (Figures 3 and 4). Based on the structural and mutagenesis studies, it has been suggested that the Glu of the HEAT motif and His of the HST motif play a role in facilitating proton transfer at the final stage of reaction [25,29,40]. In particular, the His residue is implicated as the proton donor. It is highly likely that the conserved Glu-His pair from the xExT and HxH motifs in the TM-type tRNase Z's may also participate in the terminal proton transfer reaction.

Multiple sequence alignment of bacterial-type tRNase Z's from the basal plant *P. patens* (*PpaTRZ3*) and green algae reveals that in addition to the well conserved Motifs I-V, these candidates possess a bacterial-type flexible arm domain containing the GP, HEAT and HST motifs, and the variant PxKxRN motif, in which Lys is replaced with Leu (Figure 1).

**Conservation of tRNase Z's in green plants**

Since tRNase Z's can be divided into the N-terminal and C-terminal halves, which are related by weak sequence similarity, we aligned these two halves separately. The alignment of the N-terminal and C-terminal halves of representative candidate tRNase Z's from diverse green plant species are shown in Figures 5 and 6, respectively. Two non-plant eukaryotic tRNase Z's from *D. melanogaster* and humans were included for comparison. A complete list of all aligned green plant tRNase Z's identified here is given in Additional file 5.

Except for the N-terminal regions, candidate tRNase Z's from land plants share a high degree of overall sequence similarity to each other. In contrast, candidate tRNase Z's from the four green algae show relatively low overall sequence similarity to each other and to the land plant tRNase Z's. In particular, their N-terminal halves are highly divergent compared to those from the land plants. Furthermore, sequence similarity between land plant tRNase Z's and their homologs in green algae, *D. melanogaster* and humans is limited to the conserved motifs identified in tRNase Zs.

The N-terminal halves of all green plant tRNase Z's contain a conserved eukaryotic-type flexible arm domain containing the GP motif. In addition, the N-terminal halves of the proteins also contain a pseudo-PxKxRN motif and pseudo-Motifs I and II, which differ from their original motifs in many positions including critical residues for tRNase Z functions (Figure 5). These pseudo-motifs were previously identified in candidate fungi tRNase Z's [51]. In contrast, the C-terminal halves of green plant tRNase Z's harbor conserved Motifs I-V and the PxKxRN, HEAT and HST motifs (Figure 6). These motifs appear in the same relative order in all the sequences in which they are present.

**Many chloroplast tRNA genes encode a portion of the 3'-CCA sequence**

It has been suggested that many chloroplast tRNA genes encode partial CCA sequences [55]. However, this conclusion was primarily based on the examination of the 3'-flanking sequences of tRNA genes from three chloroplast genomes. To determine if this conclusion holds true when more chloroplast genome sequences are now available for analysis, we examined the presence of tRNA genes encoding whole or partial CCA sequences in 15 chloroplast genomes. The results obtained were tabulated in Table 4.

Indeed, we found that many tRNA genes in the chloroplast genomes encode partial CCA sequences, which could serve as part of the CCA sequence (Table 4). For instance, in the *A. thaliana* chloroplast genome, 31% of tRNA genes encode the whole or partial CCA sequences in 15 chloroplast genomes. Moreover, many chloroplast tRNA genes encode the first base of the CCA sequence. For example, of the 37 *A. thaliana* chloroplast tRNA genes examined, 11 had C after the discriminator.

**Discussion**

**The presence of multiple tRNase Zs in green plants**

Unlike *C. elegans, D. melanogaster*, humans and most fungal species examined to date, green plants are unique in that they possess multiple tRNase Zs. Of the 27 complete green plant genomes analyzed, the majority of them encode two tRNase Z's and one or two tRNase Z's. It is reasonable to expect that the existence of multiple tRNase Zs would be common to green plants. This phenomenon appears to have arisen from genome doubling (polyploidy), which seems to be a driving force in...
Figure 3 Amino acid sequence alignment of the representative TM-type green plant and non-green plant tRNase Z’s. Plant TM-type tRNase Z’s are from C. reinhardtii (CreTRZ1), V. carteri (VcaTRZ1), M. pusilla (MpuTRZ1), O. lucimarinus (OuTRZ1), P. patens (PpaTRZ1 and PpaTRZ2), S. moellendorffii (SmoTRZ1), O. sativa japonica (OsaTRZ1 and OsaTRZ2), Z. mays (ZmaTRZ1 and ZmaTRZ2), A. thaliana (AthTRZ1 and AthTRZ2) [20], C. saliva (CasTRZ1 and CasTRZ2), G. max (GmaTRZ1 and GmaTRZ2), C. clementina (CclTRZ1 and CclTRZ2), M. guttatus (MguTRZ1 and MguTRZ2), A. coerulea (AcoTRZ1), V. vinifera (VvTRZ1 and VvTRZ2) and E. grandis (EgTRZ1 and EgTRZ2). Non-green plant tRNase Z’s are from T. maritima (TmaTRZ1) [30], B. subtillis (BsuTRZ1) [60], humans (HsaTRZ1) [14]. Protein accession numbers are shown in Table 1. Alignment presentations are as described in the legend to Figure 1.
Figure 4 Sequence logos of unique motifs found in candidate TM-type tRNase Z’s from green plants. The sequence logos of motifs specific to candidate TM-type tRNase Z’s from green plants were derived from the 51 alignments. 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.
Figure 5 Sequence alignment of N-terminal halves of representative green plant and non-green plant tRNase Z's. Green plant tRNase Z's are from *O. sativa japonica* (OsaTRZ3), *Z. mays* (ZmaTRZ3), *A. thaliana* (AthTRZ3 and AthTRZ4) [20], *M. guttatus* (MguTRZ3), *G. max* (GmaTRZ3), *C. sativus* (CsaTRZ3), *C. clementina* (CcTRZ3), *C. papaya* (CpaTRZ3), *E. grandis* (EgrTRZ3), *V. vinifera* (VviTRZ3), *A. coerulea* (AcoTRZ3 and AcoTRZ4), *S. moellendorfii* (SmoTRZ2 and SmoTRZ3), *P. patens* (PpaTRZ4), *O. lucimarinus* (OluTRZ3), *M. pusilla* (MpuTRZ3), *Chlamydomonas reinhardtii* (CreTRZ2) and *V. carpere* (VcaTRZ2). Non-green plant *O. melongaster* (OmeTRZ1) [16] and humans (HsaTRZ2) [14] are included for comparison. The annotation of the alignment is as described in the legend to Figure 1.
Figure 5. Sequence alignment of C-terminal halves of representative green plant and non-green plant tRNase ZLs.

Figure 6. Sequence alignment of C-terminal halves of representative green plant and non-green plant tRNase ZLs. Same legend as in Figure 5.
plant evolution and variation [56]. However, there seems to be no correlation between the number of tRNase Zs and genome size (Table 5).

An unexpected observation in this study is that most green plant tRNase Zs, which are clearly distinct from the bacterial-type tRNase Zs, represent the TM-type tRNase ZS. A comparison of motifs found in three different types of tRNase Zs from green plants is provided in Figure 7. Our phylogenetic analysis also supports the existence of both the TM- and bacterial-types of tRNase ZSs in green plants (Figure 2). The TM-type tRNase ZS was previously found only in the hyperthermophilic bacterium T. maritima and the flowering plant A. thaliana, and thus, it was originally thought to be a minor type. Our data greatly expand the repertoire of this type of tRNase Z. Although it remains to be determined if the TM-type tRNase ZS is also widespread in other taxonomic groups, the prevalence of the TM-type tRNase ZS in green plants suggests that this type might be plant-specific.

In contrast to the TM-type tRNase ZS, which is widespread in the green plants, the bacterial-type tRNase ZS is only present in the basal land plants and green algae. Sequence analysis reveals that the two types of tRNase ZSs are related but distinct from each other, suggesting that they likely arose from the same ancestral

---

**Table 4 The distribution of the trinucleotide sequences immediately after the discriminator nucleotide in chloroplast tRNA genes**

| Species                 | $^{74}$CCA | $^{74}$CCN | $^{74}$CNN | Total genes examined | %  |
|-------------------------|------------|------------|------------|----------------------|----|
| Arabidopsis thaliana    | 0          | 1          | 11         | 37                   | 32 |
| Brachypodium distachyon | 1          | 1          | 8          | 38                   | 26 |
| Canica papaya           | 0          | 1          | 8          | 37                   | 24 |
| Citrus sinensis         | 2          | 1          | 7          | 37                   | 27 |
| Cucumis sativus         | 0          | 1          | 7          | 37                   | 22 |
| Eucalyptus grandis      | 0          | 1          | 8          | 36                   | 25 |
| Glycine max             | 0          | 1          | 9          | 37                   | 27 |
| Manihot esculenta       | 0          | 1          | 7          | 37                   | 22 |
| Medicago truncatula     | 0          | 0          | 5          | 29                   | 17 |
| Oryza sativa japonica   | 1          | 2          | 7          | 38                   | 26 |
| Populus trichocarpa     | 0          | 0          | 10         | 37                   | 27 |
| Prunus persica          | 0          | 1          | 8          | 37                   | 24 |
| Sorghum bicolor         | 0          | 2          | 7          | 38                   | 24 |
| Vitis vinifera          | 0          | 1          | 8          | 37                   | 24 |
| Zea mays                | 1          | 2          | 7          | 37                   | 27 |

Nucleotide sequences of chloroplast tRNA genes and their 3’ flanking regions were obtained from the GenBank genome database http://www.ncbi.nlm.nih.gov/genomes/genlist.cgi?taxid=2759&type=4&name=Eukaryotae%20Organelles, the Chloroplast Genome Database (ChloroplastDB; http://chloroplast.bio.psu.edu/organism.cgi), and the Organelle Genome Database (GObASE; http://gobase.bcm.umontreal.ca/searches/gene.php). The discriminator nucleotide was predicted using tRNAscan-SE 1.21 http://lowelab.ucsc.edu/tRNAscan-SE/. Percentage of chloroplast tRNA genes encoding $^{74}$CNN, $^{74}$CCN and $^{74}$CCA.

---

**Table 5 The genome size and ploidy level of representative green plants and the number of tRNase Zs encoded by their genomes**

| Species                  | Common name         | No. tRNase Z* | Ploidy level | Genome size (Mbp) | Refs |
|--------------------------|---------------------|---------------|--------------|-------------------|------|
| Arabidopsis thaliana     | Mouse-ear cress     | 4             | Diploid      | 125               | [73] |
| Brachypodium distachyon  | Purple false brome  | 3             | Diploid      | 272               | [74] |
| Canica papaya            | Papaya              | 3             | Diploid      | 372               | [75] |
| Cucumis sativus          | Cucumber            | 3             | Diploid      | 367               | [76] |
| Glycine max              | Soybean             | 4             | Diploid      | 1100              | [77] |
| Oryza sativa japonica    | Rice                | 3             | Diploid      | 389               | [78] |
| Populus trichocarpa      | Black cottonwood or poplar | 3 | Diploid | 485 | [79] |
| Ricinus communis         | Castor bean         | 3             | Diploid      | 350               | [80] |
| Setaria italica          | Foxtail millet      | 5             | Diploid      | 490               | [81] |
| Sorghum bicolor          | Sorghum             | 4             | Diploid      | 730               | [82] |
| Vitis vinifera           | Grape               | 3             | Diploid      | 505               | [83] |
| Zea mays                 | Maize               | 3             | Diploid      | 2300              | [84] |

The genome size and ploidy level were obtained from the Plant DNA C-values database http://www.kew.org/cvalues/.

*The number of tRNase Zs
tRNase Z$^5$ gene, followed by sequence divergence. Since the flowering plants possess only the TM-type tRNase Z$^5$, the bacterial-type tRNase Z$^5$ has apparently been lost during the course of plant evolution whereas the TM-type tRNase Z$^5$ has been maintained. Interestingly, the bacterial-type tRNase ZS is also found in humans. Although the precise function of human tRNase ZS is unknown, it may play a role in degradation of a subset of microRNAs (miRNA) in the cytoplasm [57].

It is likely that multiple tRNase Zs found in the flowering plants are localized in different subcellular compartments. In most of the flowering plants examined, the two TM-type tRNase Z$^5$s (tRNase Z$^{51}$ and tRNase Z$^{52}$) are predicted to be either in the cytosol or in the chloroplast, whereas one of tRNase Z$^1$ (tRNase Z$^{1.1}$) is predicted to contain both nuclear and mitochondrial targeting signals. The second tRNase Z$^1$ (tRNase Z$^{1.2}$) found in some plants either contains or lacks a predicted mitochondrial targeting signal. These predictions are generally supported by recent subcellular localization studies of A. thaliana tRNase Zs [20]. A. thaliana tRNase Z$^{51}$ (AthTRZ1) and tRNase Z$^{52}$ (AthTRZ2) are localized in the cytoplasm and chloroplast, respectively, whereas A. thaliana tRNase Z$^{1.1}$ (AthTRZ3) and tRNase Z$^{1.2}$ (AthTRZ4) are targeted to both the nucleus and mitochondria, and the mitochondria, respectively [20].

Based on their predicted subcellular localization, it is most likely that tRNase Z$^{52}$ and tRNase Z$^{1.1}$ are involved in the 3’-ends processing of pre-tRNAs in the chloroplasts, and both the nucleus and mitochondria, respectively, whereas tRNase Z$^{1.2}$ either serves as a back-up for mitochondrial pre-tRNA 3’-end processing or plays a specialized, yet to be discovered function. The function of tRNase Z$^{51}$ is unknown. It may play a role in repair of incorrectly processed 3’-ends of tRNAs in the cytosol.
Alternatively, tRNase ZS could also participate in nuclear pre-tRNA 3'-end processing, as suggested by the observation that deletion of both A. thaliana tRNase ZS1 (AthTRZ1) and tRNase ZS1 (AthTRZ3) results in a lethal phenotype [20].

The plant TM-type tRNase ZSs may not recognize the CCA sequence as an anti-determinant

The CCA triplet following the discriminator generally inhibits tRNase Z activity, with the first C at nucleotide position 74 having the greatest effect [25,33,58]. This anti-determinant property of the CCA sequence is thought to prevent removal of the 3'-terminal CCA sequence from mature tRNAs, and thus avoids futile cycles of CCA addition and removal [33,58-60]. However, unlike other types of tRNase Zs, the TM-type tRNase Z does not appear to recognize the CCA sequence that is downstream the discriminator and present in the trailer sequence as an anti-determinant. The recombinant T. maritima tRNase ZS cleaves after the CCA sequence which is encoded by all but one tRNA gene, leaving the CCA sequence intact [30]. Similarly, the recombinant A. thaliana tRNase ZS can process pre-tRNAs with the 3'-terminal CCA sequences embedded in the trailer sequence, albeit at a different position [61]. Like the recombinant proteins, a partially purified spinach chloroplast tRNase Z fraction can also cleave pre-tRNAs containing complete or partial CCA sequences after the first C74 regardless of the sequence of the flanking region [55]. Importantly, the remainder of the CCA sequence can be added by chloroplast tRNA nucleotidyltransferase [55].

However, A. thaliana tRNase ZS appears to be unusual in that it can also cleave off the 3'-terminal CCA sequence from mature tRNAs in vitro [61]. Since mature tRNAs must be protected from counterproductive cleavage by tRNase ZS in vivo, it is likely that the plant chloroplast protein may acquire the ability to recognize the 3'-terminal CCA sequence as the mature tRNA 3'-end via cofactors [61].

Sequence determinants potentially involved in cleavage specificity

The mechanisms responsible for cleavage site selection and CCA inhibition are not fully understood, but appear to involve a combination of sequence features. First, the flexible arm has been suggested to contribute to cleavage site selection and the inhibitory effect of the CCA sequence [62]. This T. maritima tRNase ZS-specific element is located on the opposite side of the active site and binds primarily the D and T loops of the pre-tRNA [28,39,63]. Notably, the flexible arms of the T. maritima and plant tRNase ZSs lack the GP motif but contain the KL motif. Additionally, they are significantly smaller in length relative to those of other tRNase Zs possessing the anti-determinant function. Structural studies reveal that although TM- and bacterial-types of tRNase Zs possessing a similar overall structure which is composed of a compact globular domain and an extended two-stranded stalk and protrudes from the protein core, they have different globular domains [25-28,64]. In the TM-type flexible arm, the globular domain consists of one very short α-helix, one long helix and one 310-helix, whereas in the bacterial-type flexible arm, it is composed of two α-helices, two β-strands and one 310-helix. It would be interesting to know how the differences in the sequence feature, length and topology of the flexible arm may contribute to cleavage specificity by the enzymes.

Another possible motif involved in selection of the cleavage site has been suggested to be Motif I which appears to participate in binding the acceptor stem of pre-tRNA substrates [33,62]. In vitro studies using the recombinant T. maritima tRNase ZS has suggested that the Ser31 and Thr33 residues of Motif I are involved in the cleavage site selection, with the former residue being more critical [30]. However, only individual, but not simultaneous mutations of these two non-Gln residues to Gln (which are found at the corresponding positions in other tRNase Zs with the CCA anti-determinant) in T. maritima tRNase ZS affects the cleavage site selection [62].

In addition to the flexible arm and Motif I, the PxKxRN motif, in particular, the two basic residues in the motif, has also been suggested to be involved in the cleavage specificity and, by inference, CCA inhibition of tRNase Zs, since this motif is absent from T. maritima and A. thaliana [33]. Strikingly, tRNase ZSs found in all flowering plants examined so far also lack this motif (Figure 3). However, since the archaean Methanococcus jannaschii tRNase ZS harbors a PxKxRN motif but cannot recognize the CCA sequence as an anti-determinant in vitro, suggesting that other sequence elements may also be involved in the cleavage specificity of tRNase Z [61]. Alternatively, it has been suggested that a high enzyme concentration used in the assay may contribute to the lack of a CCA anti-determinant effect with the M. jannaschii tRNase ZS [33].

Why is the TM-type tRNase ZS developed for chloroplast pre-tRNA 3‘-end processing?

The discovery that the TM-type tRNase ZS is widespread in green plants raises a question as to why plants adopt the TM-type tRNase ZS over other types of tRNase Z for the 3‘-end processing of chloroplast pre-tRNAs. One possible explanation is that the TM-type tRNase ZS identified in plants has evolved to adapt to chloroplast pre-tRNA 3‘-end processing. It has previously been suggested that many plant chloroplast
tRNA genes encode C74 based on a limited number of available chloroplast tRNA genes [55]. To see if this conclusion could be extended to more flowering plants, we have examined the 3′-flanking region for tRNA genes in additional chloroplast genomes. Indeed, many chloroplast tRNA genes seem to encode partial CCA sequences (Table 4).

The development of the TM-type tRNase ZS may be particularly important for chloroplasts, which have a limited but sufficient number of self-encoded tRNA species and do not import tRNAs [65,66]. As all possible codons are used in the chloroplast protein-encoding genes, all chloroplast tRNA species appear to be used in protein synthesis [66]. Thus, efficient 3′-end processing of each pre-tRNA by tRNase Z may be critical in ensuring maximum efficiency in chloroplast protein synthesis.

The presence of candidate tRNase Z-like proteins in green plants

tRNase ZS-like proteins have previously been found in cyanobacteria including *Synechocystis* sp. PCC 6803 [67] and fungi [51], whereas tRNase ZL-like proteins have not been reported. In *Synechocystis*, one tRNase ZS-like protein (sll1036) has been identified. This protein does not exhibit any tRNase Z activity in vitro, consistent with the lack of some of the most conserved motifs of tRNase Zs in the protein sequence [67]. Since candidate tRNase Z-like proteins found in basal land plants and green algae either lack all or some of the conserved motifs of tRNase Zs, they most likely possess no tRNase Z activity.

It has been suggested that tRNase ZS has arisen from the fusion of duplicated tRNase ZS genes with further sequence diversification. It is possible that tRNase ZS-like proteins found in the basal land plants and green algae may represent relics of original tRNase ZS that were mutated during diversification of eukaryotic tRNase Z genes. Alternatively, these proteins may play species-specific functions as suggested for cyanobacterial tRNase ZS-like proteins [67]. The predominant presence in the basal land plants and green algae of tRNase ZS-like proteins also suggests that duplication of ancestral tRNase ZS genes may occur early in green plant evolution.

Conclusions

This study represents the first large-scale identification and analysis of green plant tRNase Zs. Our survey of current plant genome databases shows that green plants are represented by multiple tRNase Zs, which include one or two tRNase ZS and two tRNase ZSs. One tRNase ZS is predicted to participate in 3′-end processing of nuclear and mitochondrial pre-tRNA, whereas the other is likely to provide a backup for mitochondrial pre-tRNA processing. It appears that most tRNase ZSs, which is widespread throughout the green plants, belong to a minor but highly distinct type of tRNase ZS (TM-type). In contrast, the typical bacterial-type tRNase ZS is restricted to basal land plants and green algae. The apparent lack of the bacterial-type tRNase ZS in flowering plants suggest that while both types were present in the basal land plants, the bacterial-type tRNase ZS was discarded in favor of TM-type during plant evolution. Based on our results and previous studies, we propose that like *T. maritima* tRNase ZS, TM-type tRNase ZSs found in green plants seem not to recognize the CCA sequence as an anti-determinant and that the rise of this type of tRNase ZS appears to accommodate the 3′-end processing of chloroplast pre-tRNAs with partial or whole CCA sequences. This unusual property of green plant tRNase ZSs is likely due to multiple sequence determinants including the TM-type-specific flexible arm comprising the KL motif, Motif I and lack of the PxKxRN motif.

However, it should be noted that bioinformatics analysis alone cannot resolve possible differences in cleavage specificity among TM-type tRNase ZSs. The complete understanding of the mechanisms of the cleavage specificities of the TM-type tRNase ZSs awaits the determination of the structures of these enzymes bound to pre-tRNA and the accumulation of more precise biochemical data.

Methods

Plant genome database search and protein sequence analysis

Candidate tRNase Zs were identified by BLAST and PSI-BLAST searches against the genome databases using known tRNase Zs as query sequences. The databases used include Phytozome http://www.phytozome.net/, the NCBI nonredundant protein sequence database http://blast.ncbi.nlm.nih.gov/Blast.cgi, Joint Genome Institute (JGI; http://www.jgi.doe.gov/) and Universal Protein Resource (Uniprot; http://www.uniprot.org/). An E-value cutoff of 0.001 was used in all searches. With this value, no β-CASP protein or other MBL protein was found. The resulting sequences were subject to validation as described [51]. 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, PSORT http://psort.hgc.jp/form.html and ChloroP http://www.cbs.dtu.dk/services/ChloroP/. Multiple sequence alignments were done by Clustal W [68].
**Phylogenetic analysis**

Phylogenetic analysis was performed using the Bayesian approach, with *T. maritima* tRNase Z\(^3\) (TmaTrz1) as an outgroup as described [51]. Briefly, full-length amino acid sequences of candidate plant tRNase Zs and TmaTrz1 were aligned by using Clustal W implemented in MEGA 5.0 [69]. After excluding gaps and the ambiguous sites, we used ProtTest 2.4 [70] to choose the most appropriate evolutionary model for our data set. The phylogenies were estimated by Bayesian inference with MrBayes 3.1.2 [71] using a mixture of the fixed amino acid models and I + 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 search were conducted. All searches consisted of one ‘cold’ and three ‘heated’ Markov chains estimated for 10\(^7\) generations, and every 1000 generations were sampled. The burn-in parameter was estimated by plotting -lnL against the generation number using Tracer 1.4.1 http://beast.bio.ed.ac.uk/Tracer, and the retained trees were used to estimate the consensus tree and the Bayesian posterior probabilities.

**Additional material**

Additional file 1: Distribution of candidate tRNase Zs identified in green plants. Abbreviations for species names are indicated in the parentheses. *The number of amino acids in plant 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.

Additional file 2: Pairwise comparisons of S. bicolor and S. italica tRNase Zs\(^2\). The accession numbers for proteins are listed in Additional file 1. The pairwise percent identity (I) and percent similarity (S) between tRNase Z\(^2\)'s from *S. bicolor* and *S. italica* were calculated using the Clustal W program [88].

Additional file 3: Alignment of candidate tRNase Z\(^2\)'s from *S. bicolor* and *S. italic*. The accession numbers for the candidates are listed in Additional file 1. The annotation of the alignment is described in the legend to Figure 1.

Additional file 4: Alignment of candidate TM-type tRNase Z\(^2\)s identified in green plants. The accession numbers for the candidates are listed in Additional file 1. Annotation is as described in the legend to Figure 1.

Additional file 5: Alignment of candidate tRNase Z\(^2\)s identified in green plants. The accession numbers for the candidates are listed in Additional file 1. The annotation of the alignment is as described in the legend for Figure 1.

**List of abbreviations**

pre-tRNAs: tRNA precursors; miRNA: microRNAs; tRNase Z: tRNA 3′ endonuclease; tRNase Z\(^2\): the short form of tRNase Z; tRNase Z\(^3\): the long form of tRNase Z; β-CASP: MBL-associated CSF. Artemis SNM1/PDS2; CPSF-73: the cleavage and polyadenylation specificity factor; MBL: the metallo-β-lactamase superfamily; TGFB: transforming growth factor-β; TM: T. maritima.

**Acknowledgements**

We thank three anonymous reviewers for improving the quality of the manuscript, Jingjing Peng and Wei Zhao for valuable discussion. This work was supported by grants from the National Science Foundation of China (31070703), the Priority Academic Program Development of Jiangsu Higher Education Institutions and Nanjing Normal University (20071040XGQ0148).

**Author details**

1Laboratory of Yeast Genetics and Molecular Biology, School of Life Sciences, Nanjing Normal University, 1 Wenyuan Road, Nanjing 210046, China.
2Jiangsu Key Laboratory for Microbes and Genomics, School of Life Sciences, Nanjing Normal University, 1 Wenyuan Road, Nanjing 210046, China.
3Jiangsu Engineering and Technology Research Center for Industrialization of Microbial Resources, School of Life Sciences, Nanjing Normal University, 1 Wenyuan Road, Nanjing 210046, China.
4Jiangsu Key Laboratory for Biodiversity and Biotechnology, School of Life Sciences, Nanjing Normal University, 1 Wenyuan Road, Nanjing 210046, China.

**Authors’ contributions**

LF, ZW and J. performed online database searches and sequence analysis. JY carried out a phylogenetic analysis. YH conceived this study, analyzed the data and drafted the manuscript. All authors have read and approved the final version of the manuscript.

**Competing interests**

The authors declare that they have no competing interests.

**Received:** 8 May 2011 **Accepted:** 23 July 2011 **Published:** 23 July 2011

**References**

1. Phizicky EM, Hopper AK: tRNA biology charges to the front. Genes Dev 2010, 24:1892-1891.
2. Bayfield MA, Yang R, Maraia RJ: Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). Biochim Biophys Acta 2010, 1799:365-378.
3. Maraa RJ, Lamichhane TN: 3′ processing of eukaryotic precursor tRNAs. WIREs RNA 2011, 2:362-375.
4. Geballez M, Vaque A: tRNase Z. Protein Pept Lett 2007, 14:137-145.
5. Hartmann RK, Gossringer M, Spath B, Fischer S, Marshfelder A, Chapter & The Making of tRNAs and More - RNase P and tRNase Z. Prog Nucleic Acid Res Mol Biol 2009, 85:319-368.
6. Redko V, Li de Laserreria-Galay, I Condron C: When all’s zed and done: the structure and function of RNase Z in prokaryotes. Nat Rev Microbiol 2007, 5:278-286.
7. Spath B, Canino G, Marchfelder A: tRNase Z: the end is not in sight. Cell Mol Life Sci 2007, 64:2404-2412.
8. Wolin SL, Cedervall T: The La protein. Annu Rev Biochem 2002, 71:375-403.
9. Van Horn DJ, Yoo CJ, Xue D, Shi H, Wolin SL: The La protein in Schizosaccharomyces pombe: a conserved yet dispensable phosphoprotein that functions in tRNA maturation. RNA 1997, 3:1434-1445.
10. Yoo CJ, Wolin SL: The yeast La protein is required for the 3′ endonucleolytic cleavage that matures tRNA precursors. Cell 1989, 59:399-402.
11. Copela LA, Fernandez CF, Sherrer I1L, Wolin SL: Competition between the NCl1 exonuclease and the La protein affects both Trf4p-mediated RNA quality control and pre-tRNA maturation. RNA 2008, 14:1214-1227.
12. Schafer B: RNA maturation in mitochondria of S. cerevisiae and S. pombe. Gene 2005, 354:80-85.
13. Sugita M, Sugiura M: Regulation of gene expression in chloroplasts of higher plants. Plant Mol Biol 1996, 32:315-326.
14. Tavtigian SV, Simard J, Teng DH, Abtin V, Baumgard M, Beck A, Camp NJ, Carillo AR, Chen Y, Dayananth P, Desochers M, Dumont M, Femham JM, Frank D, Frye C, Ghaffari S, Gupta J, Hu R, Iliev D, Janedki T, Kort EN, Lally 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.
15. Chen Y, Beck A, Davenport C, Chen Y, Shattuck D, Tavtigian SV: Characterization of TRZ\(^1\), a yeast homolog of the human candidate prostate cancer susceptibility gene ELAC2 encoding tRNase Z. BMC Mol Biol 2005, 6:12.
16. 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, 32:255-262.
17. Smith MM, Levitan DJ: *Caenorhabditis elegans* eleomorph of the putative prostate cancer susceptibility gene ELAC2, hoz-1, plays a role in germline proliferation. *Dev Biol* 2004, 266:151-160.

18. Zhao Z, Su W, Yuan S, Huang Y: Functional conservation of tRNase Z among Saccharomyces cerevisiae, Schizosaccharomyces pombe and *humans*. *Biochem J* 2009, 422:483-492.

19. Gan X, Yang J, Li J, Yu H, Dai H, Liu J, Huang Y: The fission yeast Schizosaccharomyces pombe has two distinct tRNase Zs encoded by two different genes and differentially targeted to the nucleus and mitochondria. *Biochem J* 2011, 435:103-111.

20. Canino G, Roclan E, Barberis N, Echeverria M, Fomer J, Binder S, Marchfelder A: Arabidopsis encodes four tRNase Z enzymes. *Plant Physiol* 2009, 150:1491-1502.

21. Aravind L: An evolutionary classification of the metallo-beta-lactamase fold proteins. *In Silico Biol* 1999, 1:69-91.

22. Dayasu 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.

23. 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.

24. 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.

25. de la Sierra-Gallay II, Pellegrini O, Condron C: Structural basis for substrate binding, cleavage and allottery in the tRNA maturase RNase Z. *Nature* 2005, 433:657-661.

26. Iishi R, Minagawa A, Takaku H, Takagi M, Nashimoto M, Yokoyama S: Crystal structure of the tRNA 3' processing endonucleobonuclease tRNase Z from Thermotoga maritima. *J Biol Chem* 2005, 280:14138-14144.

27. Kostedek 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.

28. Li de la Sierra-Gallay I, Mathy N, Pellegrini O, Condron C: Structure of the ubiquitous 3' processing enzyme RNase Z bound to transfer RNA. *Nat Struct Mol Biol* 2006, 13:376-377.

29. Minagawa A, Takaku H, Iishi R, Takagi M, Yokoyama S, Nashimoto M: Identification by Mn²⁺ rescue of two residues essential for the proton transfer of tRNase Z catalysis. *Nucleic Acids Res* 2006, 34:3811-3818.

30. Minagawa A, Takaku H, Iishi R, 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.

31. Schilling O, Vogel A, Kosteleycky B, Natal da Luz H, Spermann D, Spah B, Marchfelder A, Troger W, Meyer-Klaucke W: Zinc- and iron-dependent cytosolic metallo-beta-lactamase domain proteins exhibit similar zinc-binding affinities, independent of an anlytical glutamate at the metal-binding site. *Biochem J* 2005, 383:145-153.

32. Vogel A, Schilling O, Meyer-Klaucke W: Identification of metal binding residues for the binuclear zinc phosphodiesterase reveals identical coordination as glyoxylase. *Biochemistry* 2004, 43:10379-10386.

33. Zareen N, Hopkinson A, Levinger L: Residues in two homology blocks on the amino side of the tRNase Z domain contribute unexpectedly to pre-tRNA 3' processing. *RNA* 2004, 10:148-155.

34. Zareen N, Yan H, Hopkinson A, Desetty R, Wilson C: Identification and analysis of candidate pre-tRNA binding site. *Genes Dev* 2009, 23:2639-2649.

35. Zareen N, 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.

36. 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.

37. Haussecker D, Huang Y, Lau A, Parameswaran P, Fire AZ, Kay MA: Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* 2010, 16:673-695.

38. Lee YS, Shiba Y, Malhotra A, Dutta A: A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev* 2009, 23:2639-2649.

39. Bogerd HP, Karnowski IW, Cai X, Shin J, Pohlers M, Gullen BR: A mammalian herpesvirus uses noncanonical expression and processing mechanisms to generate viral MicroRNAs. *Mol Cell* 2010, 37:135-142.

40. Diebel KW, Smith AL, van Dyk LF: Mature and functional viral miRNAs transcribed from novel RNA polymerase III promoters. *RNA* 2010, 16:170-185.

41. Koner V, Guevara C, Chen Y, Neuteboom S, Bockstein R, Tavitigian S, Lees E: The product of the candidate prostate cancer susceptibility gene ELAC2 interacts with the gamma-tubulin complex. *Int J Cancer* 2003, 104:283-288.

42. Noda D, Itoh S, Watanabe Y, Inamitsu M, Demniler S, Itoh F, Koike S, Danielpour D, Ten Dijke P, Kato M: ELAC2, a putative prostate cancer susceptibility gene product, potentiates TGF-beta1/Smad-induced growth arrest of prostate cells. *Oncogene* 2005, 25:5591-5600.

43. Mineri R, Pavelka N, Fernandez-Vizarra E, Riccardi-Castagnoli P, Zeviani M, Tinti V: How do human cells react to the absence of mitochondrial DNA? *PLoS* 2009, 4:65713.

44. Peng WT, Robinson MD, Mianmehn S, Krogan NJ, Cagney G, Morris Q, Davierwala AP, Grigull J, Yang X, Zhang W, Mitakakis N, Ryan OW, Datta N, Jojic V, Pal C, Canadini V, Richards D, Beattie B, Wu LF, Altshuler SJ, Roweis S, Frey BJ, Emil A, Greenblatt JF, Hughes TR: A panoramic view of yeast noncoding RNA processing. *Cell* 2003, 113:919-933.

45. Zhao W, Yu H, Li S, Huang Y: Identification and analysis of candidate fungal tRNA 3'-end processing endonucleases tRNase Zs, homologs of the putative prostate cancer susceptibility protein ELAC2. *EMBO J* 2010, 29:2727-2737.

46. Vogel A, Schilling O, Spah B, Marchfelder A: The tRNase Z family of proteins: physiological functions, substrate specificity and structural properties. *Biochem Biophys Res Commun* 2006, 386:1253-1264.

47. Janvi P, Robinson C: Mechanisms of protein import and routing in eukaryotic chloroplasts. *Curr Biol* 2004, 14:R1064-R1077.

48. Pfanner N, Neupert W: The mitochondrial protein import apparatus. *Annu Rev Biochem* 1990, 59:331-353.

49. Oommen A, Li XQ, Gegenheimer P: Cleavage specificity of chloroplast and nuclear tRNA 3'-processing nucleases. *Mol Biol Cell* 1992, 2:865-875.

50. Adams KL, Wendel JJ: Polyploidy and genome evolution in plants. *Curr Opin Plant Biol* 2005, 8:133-141.

51. Takahashi M, Takaku H, Nashimoto M: Regulation of the human tRNase Z gene expression. *FEBS Lett* 2008, 582:2532-2536.

52. Mohan A, Whyte S, Wang X, Nashimoto M, Levinger L: The 3' end CCA of mature tRNA is an antiodterminant for eukaryotic 3'-tRNase. *RNA* 1999, 5:245-256.

53. Nashimoto M: Distribution of both lengths and 5' terminal nucleotides of mammalian pre-tRNA 3' trailers reflects properties of 3' processing endonucleobonuclease. *Nucleic Acids Res* 1997, 25:1148-1154.

54. Pellegrini O, Nezzar J, Marchfelder A, Putzer H, Condron C: Endonucleolytic processing of CCA-less tRNA precursors by RNase Z in *Bacillus subtilis*. *EMBO J* 2003, 22:4534-4543.

55. Schiffer S, Rosch S, Marchfelder A: Recombinant RNase Z does not recognize CCA as part of the tRNA and its cleavage efficiency is influenced by acceptor stem length. *Biochem* 2003, 384:333-342.

56. Minagawa A, Iishi 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* 2006, 361:269-299.
Effects of conserved D/T loop substitutions in the tRNase Z substrate of Thermotoga maritima. RNA Biol 2008, 54:104-111.

The structure of the flexible arm of Thermotoga maritima tRNase Z differs from those of homologous enzymes. Acta Crystallogr Sect F Struct Biol Cyst Common 2007, 63:637-641.

Michaud M, Cognot V, Duchene AM, Marechal-Drouard L. A global picture of tRNAs in plant genomes. Plant J 2010, 66:80-93.

Pfitzinger H, Guillemaut F, Weil JH, Pillay DT. Adjustment of the tRNA population to the codon usage in chloroplasts. Nucleic Acids Res 1987, 15:1377-1381.

Ceballos-Chavez M, Mioque A. Sequence-dependent Cleavage Site Selection by tRNase Z from the Cyanobacterium Synechocystis sp. PCC 6803. J Biol Chem 2005, 280:33461-33469.

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.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGAS: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol 2011.

Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 2005, 21:2104-2105.

Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 2003, 19:1572-1574.

Ezraty B, Dahlgren B, Deutscher MP. The RNase Z homologue encoded by Escherichia coli elac gene is RNase BN. J Biol Chem 2005, 280:16542-16545.

Initiative TAG: Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 2000, 408:796-815.

Initiative TB: Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature 463:763-768.

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Li G, Lewis KL, Salzberg SL, Jones MR, Skelton RL, Murray JE, Chen C, Qian W, Shen J, Du P, Eustice M, Tong E, Tang H, Lyons E, Pauli RE, Michael TP, Wall K, Rice DW, Albert H, Wang ML, Zhu YJ, et al. The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). Nature 2008, 452:991-996.

Fan et al. BMC Evolutionary Biology 2011, 11:219 http://www.biomedcentral.com/1471-2148/11/219

Cite this article as: Fan et al. A survey of green plant tRNA 3’ end processing enzyme tRNase Zs, homologs of the candidate prostate cancer susceptibility protein ELAC2. BMC Evolutionary Biology 2011 11:219.