1-1-2018

Evolution of Eukaryal and Archaeal Pseudouridine Synthase Pus10.

Elisabeth Fitzek

Archi Joardar

Ramesh Gupta

Matt Geisler

This is a post-peer-review, pre-copyedit version of an article published in Journal of Molecular Evolution. The final authenticated version is available online at: http://dx.doi.org/10.1007/s00239-018-9827-y

Evolution of Eukaryal and Archaeal Pseudouridine Synthase Pus10

Elisabeth Fitzek[1,3], Archi Joardar[2,4], Ramesh Gupta[2], Matt Geisler[1*]

[1] Department of Plant Biology, Southern Illinois University, Carbondale, IL 62901-6509, USA

[2] Department of Biochemistry and Molecular Biology, Southern Illinois University, Carbondale, IL 62901-4413, USA

[3]Present address: Department of Biological Sciences, Northern Illinois University, IL 60115, USA

[4]Present address: Molecular and Cellular Biology, University of Arizona, Tucson, AZ 85721, USA

* Corresponding author:

Matt Geisler

Email: mgeisler@plant.siu.edu

Phone: (618) 453-3212

Fax: (618) 453-3441

Abstract

In archaea, pseudouridine (Ψ) synthase Pus10 modifies uridine (U) to Ψ at positions 54 and 55 of tRNA. In contrast, Pus10 is not found in bacteria, where modifications at those two positions are carried out by TrmA (U54 to m^5U54) and TruB (U55 to Ψ55). Many eukaryotes have an apparent redundancy; their genomes contain orthologs of archaeal Pus10 and bacterial TrmA and TruB. Although eukaryal Pus10 genes share a conserved catalytic domain with archaeal Pus10 genes, their biological roles are not clear for the two reasons. First, experimental evidence suggests that human Pus10 participates in apoptosis induced by the tumor necrosis factor–related apoptosis-inducing ligand. Whether the function of human Pus10 is in place or in addition to of Ψ synthesis in tRNA is unknown. Second, Pus10 is found in earlier evolutionary branches of fungi (such as chytrid *Batrachochytrium*) but is absent in all dikaryon fungi surveyed (Ascomycetes and Basidiomycetes). We did a comprehensive analysis of sequenced genomes and found that orthologs of Pus10, TrmA and TruB were present in all the animals, plants and protozoa surveyed. This indicates that the common eukaryotic ancestor possesses all the three genes. Next, we examined 116 archaeal and eukaryotic Pus10 protein sequences to find that Pus10 existed as a single copy gene in all the surveyed genomes despite ancestral whole genome duplications had occurred. This indicates a possible deleterious gene dosage effect. Our results suggest that functional redundancy result in gene loss or neofunctionalization in different evolutionary lineages.

Keywords: phylogeny, protein evolution, subfunctionalization, pseudogene, orthologs

Introduction

During protein synthesis, transfer RNA (tRNA) requires modifications at specific nucleotides to ensure the correct incorporation of amino acids to a growing peptide chain. A superfamily of pseudouridine synthases families can isomerize uridine (U) to pseudouridine (Ψ) at multiple positions in RNA post-transcriptionally (Kaya and Ofengand 2003). In this superfamily, members of TruA, TruB and TruD families modify tRNA; members of the RluA and RsuA families primarily involve in uridine modification of rRNA; members of Pus10 can modify tRNA at position U54 to Ψ (Gurha and Gupta 2008).

Both TruB (of bacterial origin) and Pus10 (of archaeal origin) can modify U to Ψ at tRNA position 55 *in vitro* (Blaby et al. 2011; Gurha and Gupta 2008; Roovers et al. 2006). In bacteria, TrmA methylates tRNA at position U54 to m^5U54 or ribo-T54 (ribothymidine) (Ny and Björk 1980). These three enzymes – TruB, TrmA, and Pus10 – all have eukaryal orthologs and thus, have potential overlapped roles in modification of uridines at position 54 (by TrmA and Pus10) and 55 (by TruB and Pus10) of tRNAs. Indeed, when crystal structures of Pus family members were compared, overlapping structural similarities of an aspartate residue as the catalytic amino acid in the active pocket were revealed and confirmed. The structure is preserved in all the eukaryal orthologs studied (Foster et al. 2000; Hoang 2004; Hoang et al. 2006; McCleverty et al. 2007; Pan et al. 2003; Sivaraman et al. 2002). Moreover, mutational analysis of Pus family members identified additional four amino acids and structural features – forefinger loop (FFL) and thumb loop – to be important for substrate recognition and binding (Chan and Huang 2009; Conrad et al. 1999; Gurha and Gupta 2008; Hamilton et al. 2005; Hur and Stroud 2007; Joardar et al. 2013; Kamalampeta et al. 2013; Spedaliere et al. 2000).

In this paper, the molecular evolution of Pus10 and resolution of functional redundancy between orthologs of Pus10, TruB, and TrmA in different eukaryotic lineages are described. Lineage-specific alterations in key Pus10 sequence features are shown: length of the FFL and substitution of amino acids in the catalytic region and thumb loop. Different evolutionary hypotheses are proposed to explain how neofunctionalization, subfunctionalization and gene loss have given different lineage-specific outcomes.

Materials and Methods

Collection of Data Sets

For the identification of Pus10-like proteins, only the species with whole sequenced and annotated genomes were retrieved from the National Center for Biotechnology Information (NCBI), the DOE Joint Genome Initiative (JGI), the Solanaceae Genomic Network (SGN), and the Ensembl genome browser (version 62) (Bombarely et al. 2011; Hubbard et al. 2009; McGinnis and Madden 2004). Sequence alignments were performed using the basic local alignment tool Protein BLAST (BLASTP; using default settings) in NCBI. The full length Pus10 sequences from *Homo sapiens* (NP_653310, PDB: 2V9K), *Arabidopsis thaliana* (NP_173466), *Methanocaldococcus jannaschii* (NP_247004) and *Pyrococcus furiosus* (NP_578868.1) were used as a query. Similarity Sequences with BLASTP score of 40 or better and with >25% coverage were initially retained. However, those that did not harbor the essential Asp (aspartic acid) in the conserved core catalytic sequence GREDXD (Joardar et al. 2013) or were found to be significantly better aligned to another known Pus protein (e.g.,, Cbf5, Pus4, etc.) were

excluded from later analysis. These were likely not functional orthologs of Pus10. Accession numbers for all the proteins obtained are in Supplementary Table 1.

Protein Alignment and Phylogenetic Analysis

Orthologous Pus10 protein sequences of 116 species were collected from Eukarya and Archaea. A weakly aligned Pus10 homolog was found for *Nanoarchaeum* but was excluded from the downstream phylogenetic analysis due to extreme divergence of the sequence and absence of key Pus10 sequence features. The N-terminal region of Pus10 showed high divergence specific to each taxonomic domain and kingdom, and resulted in very low or no bootstrap supported values. Thus, protein alignment and phylogenetic analyses were conducted using only the conserved C-terminal catalytic domain (e.g., Gly286-Asp528 in human Pus10). In this study, amino acid positioning of Pus10 refers to the corresponding sequence position in relation to *M. jannaschii*. Isolation of the C-terminal domain was done by multiple sequence alignment using ClustalX (version 2.0.12) of whole protein sequences followed by removing the N-terminus region (e.g., residues 1 to 286 in human Pus10) (Larkin et al. 2007; McCleverty et al. 2007). Maximum likelihood analyses of all taxa with respect to the C-terminal of orthologous Pus10 were calculated with the Whelan and Goldman (WAG + I + G) empirical substitution model (Whelan and Goldman 2001). The web server RAxML Black box (http://phylobench.vital-it.ch/raxml-bb/) was used for maximum likelihood analyses (Stamatakis 2006). Netwick files were uploaded into FigTree (version 1.3.1) (http://tree.bio.ed.ac.uk/software/figtree) and the following settings were altered. Line weight: 2x, cladogram was chosen as displayable transform version, node labels were chosen to display bootstrap values. Further features were selected: Node bars and

scale bar (altered to line weight of 2). Tree branches were colored to emphasize different kingdoms and the tree graphic was exported as a jpeg format.

Annotation of Phylogenetic Trees

The molecular phylogeny of the C-terminal Pus10 was annotated using the Interactive Tree Of Life (version 2.0) (iTOL; http://itol.embl.de/index.shtml) (Letunic and Bork 2011; Letunic and Bork 2016) featuring the key amino acids in the active pocket and structural features relevant for recognition and binding of tRNA substrates. Substitutions in the key Pus10 features were mapped onto a maximum likelihood tree and color coded. A separate species tree with 131 taxa (including those without detectable PUS10) was generated to illustrate the presence/absence pattern of Pus10 in relation to species ancestry and to determine potential gain/loss events (Supplementary Table 1). A reduced version with of the species tree was generated by acquiring taxonomy numbers from NCBI uploaded onto iTOL. The resulting species tree was color coded with respect to kingdom, except for phylum amoebozoa. Furthermore, presence/absence of Pus10 within clades was emphasized (Letunic and Bork 2011) (Refer to illustration in Fig. 5).

Homology Modeling and Structural Superimposition

A homology model of the protein structure of *M. jannaschii* Pus10 was calculated using 3D-JIGSAW (version 2.0; http://bmm.cancerresearchuk.org/~3djigsaw/) (Bates et al. 2001). The crystal structure of *H. sapiens* (PDB: 2V9K, *B* factor 38.6, resolution: 2.0 Å) was used as a template (McCleverty et al. 2007). Both the crystal structures of *H. sapiens* Pus10 and the homology-modeled protein structure of *M. jannaschii* were uploaded onto the Swiss-PDB

Viewer (version 4.1) and superimposed via the 'generate structural alignment tool' (Guex et al. 2009). Pus10 of *H. sapiens* and *M. jannaschii* are displayed as ribbon and each backbone of the 3D-structure was color coded. Conserved amino acids in all pseudouridine synthases were individually displayed, color coded, and labeled (refer to Fig. 3).

Results

Characterization of Key Conserved Structural Features of Pus10

Members of the Pseudouridine Synthase superfamily have a similar 3D structure and a conserved catalytic Asp (Mueller and Ferré-D'Amaré 2009). Other accessory domains with binding capacity and features (e.g., FFL and thumb loop) are conserved within sub-families and can distinguish between different members (Fig. 1). Crystal structure analysis of RluA and TruA identified FFL and thumb loop in the catalytic site as the 'pinch mechanism' holding the substrate in place (Hoang et al. 2006; Hoang and Ferré-D'Amaré 2001). Only three families (TruA, RluA and Pus10) have both the FFL and thumb loop, whereas other families have either loop individually (Fig. 1).

The THUMP-domain is unique to Pus10 within the pseudouridine synthase superfamily (Fig. 2). It is structurally similar to the THUMP domain of bacterial ThiI and present at the N-terminus, and likely involved in substrate recognition and binding (Aravind and Koonin 2001). Modification of the N-terminal THUMP-domain in Pus10 was observed throughout different eukaryotic lineages, and extensively so in plants and animals (Fig. 2). Eukaryotic Pus10 THUMP-domains contain large insertions to prevent detection by routine search of hidden

Markov models (HMM) for a THUMP-domain (when using PFAM PF02926 as template). But, all archaeal Pus10 THUMP-domains tested fit the PFAM model.

Structural pairwise alignment of bacterial ThiI (PDB: 2C5S) was made to the homology modeled THUMP-domain across eukaryal and archaeal lineages. The structural comparison analysis by the DALIlite (version 3.3) (Holm et al. 2008; Holm and Rosenstrom 2010) showed an overall high similarity Z-score (Table 1). The Z-scores of PUS10 from *P. furiosus* and *H. sapiens* are very high, 2.5 and 3.7, respectively. These scores indicate that although observed deletions in *P. furiosus* and insertions in *H. sapiens* could theoretically have affected the folding of the protein, this region still remains structurally similar to ThiI THUMP.

In Pus10 of *H. sapiens*, the zinc-binding site consists of four cysteine (Cys) residues at positions 21, 24, 109, and 112, and is thought to be involved in the maintenance of the native N-terminal structure (McCleverty et al. 2007). Multiple sequence alignment revealed the presence of the four Cys residues in nearly all Pus10 orthologs of Archaea and Eukarya. The exceptions were found in the fungi lineage. Here, all four Cys were present in early fungi Mucoromycotina and in several Chytridiomycota, but entirely absent in Microsporidia. Furthermore, the first Cys pair (Cys21 and 24) was absent in one chitrid fungus (*Allomyces macrogynus*). Interestingly, the latter was encountered in the following invertebrates; *Branchiostoma floridae*, *Daphnia pulex*, and *Caenorhabditis elegans*. Ala (alanine) substitutions of the second pair, C106/109A in *M. jannaschii* (equivalent to C109/C112 of human Pus10) showed a reduced Ψ54 modification, but Ψ55 modification was unaffected (Joardar et al. 2013). This suggests Pus10 proteins may retain partial (Ψ55) function despite the absence of the second Cys pair.

In the catalytic region, five amino acids (Asp275, Tyr339, Ile412, Lys413, Leu440 in *M. jannaschii*) were determined to be conserved throughout all pseudouridine synthase families

(Hamma and Ferré-D'Amaré 2006; McCleverty et al. 2007). Superimposition of a homology-modeled structure of *M. jannaschii* Pus10 onto that of *H. sapiens* Pus 10 revealed the same orientation and 3D positioning of the conserved amino acids present in the active pocket. This was observed throughout all 116 organisms surveyed (Fig. 3) with the exceptions of *Branchiostoma floridae* (Cephalochordata), *Daphnia pulex* (Brachipoda), *Manihot esculenta* and *Solanum lycopersicum*. A conserved consensus sequence of the catalytic site in the Pus10 orthologs centered on the catalytic Asp275 and 5 surrounding residues 'GRE**D**VD' (Fig. 1). Substitutions within this catalytic site have been observed throughout Archaea and Eukaryota (Fig. 4). In addition to the active pocket, two loops were experimentally confirmed to be of value for the pseudouridylation of tRNAs at positions U55 and U54. The FFL proved to contribute to modify U55 over U54. Therefore, the length of the loop was crucial, not the identities of the amino acids within the loop. The opposite was observed for the thumb-loop, where the amino acid identity (His and Arg) is of importance for pseudouridylation (Joardar et al. 2013).

Presence, Absence, and Copy Number of Pus10 Orthologs in Whole Genomes

Organisms with fully sequenced genomes were surveyed by BLASTP to identify Pus10 orthologs. In the identification of Pus10 orthologs, several eukaryotic and a few archaeal species were found to lack Pus10 and were placed on a species tree (Fig. 5). Four nodes of Pus10 loss were identified (Fig. 5). In fungi, all dikaryotes (Ascomycocota and Basidomycocota) lacked Pus10 but had TruB orthologs. The most parsimonious answer would be a single evolutionary event in which Pus10 was lost in the common ancestor of all Dikaryon fungi. Other fungal taxa, including Chitrids and Microsporidia, contain Pus10 orthologs (Fig. 5). Most of Archaea had Pus10 but only two genera lacked functional Pus10 genes (Fig. 5). In *Sulfolobus tokodaii*, *S.*

acidocaldarius, *S. solfataricus*, their most similar sequences (to Pus10) all lack the catalytic aspartate residue. The entire conserved 'GREDVD' sequence surrounding the catalytic aspartate residue is significantly altered (i.e., "PYSEPSDVR" in *S. acidocaldarius*). Although most of the residues upstream of this site are somewhat conserved. There is less conservation of the protein downstream of the catalytic site. Many of the other structural features thought to be needed to perform pseudouridylation are also missing or significantly divergent from those conserved between Archaea and Eukaryotes. The Pus10 sequence of *Nanoarchaeum* showed high overall divergence, with numerous substitutions throughout the sequence and resulted in low bootstrap values upon phylogenetic analysis. For these reasons, these two genera of Archaea were excluded from phylogenetic analysis. In four Archaeal orders and a few members within Desulfurococcales and Halobacteriales, shorter FFLs were observed. However, the majority of Archaea species in this study had Pus10 orthologs with completely intact Pus10 structural features (Fig. 4).

A surprising finding in all the genomes surveyed was the presence of only a single Pus10 gene per genome, despite the numerous genome duplications that have occurred in the evolution of animals, fungi and especially in plants (where genes frequently duplicate and subfunctionalize) (Koonin and Wolf 2010; Proost et al. 2011). This finding might point to a potential dosage effect or other harmful consequence for higher copy numbers of Pus10.

Pus10 in Archaea

A maximum likelihood tree was calculated for the C-terminal domain of the Pus10 in 44 archaeal species (Supplementary Fig. 1). The tree is comprised of 4 clades where the outer most clade is represented by members of the Thermococcales. *Crenarchaeum symbiosum* and

Nitrosopumilus maritimus (Thaumarchaeota). These are incorporated into the Methanococcales clade and show long branches, indicating high divergence and possible false placement due to long branch attraction or parallelism (Anderson and Swofford 2004). This sequence convergence indicates some possible alteration of the Pus10 catalytic function. When comparing key features of Pus10 protein sequences – FFL, catalytic 'GRED[V/X]D' and thumb loop with His and Arg – of Thaumarchaeota to other Archaea and Eukaryota, Pus10 of both *C. symbiosum* and *N. maritimus* show the same FFL size as that of *M. jannaschii*. This could explain the incorporation of *C. symbiosum* and *N. maritimus* into the clade of Methanococcales. However, the sequence of their catalytic site differs considerably. In addition, the thumb loop lacks a conserved histidine residue. Multiple sequence alignment of *Pyrobaculum arsenaticum*, *P. islandicum* and *Thermoproteus neutrophilus* also show modification in all structural features for Pus10.

The most important differences in Pus10 features were observed in Methanococcales and Thermococcales. Methanobacteriales, Halobacteriales and Methanococcales have insertions in the FFL, creating a longer loop than the other clades in Archaea (Fig. 4). Moreover, the Thermococcales clade, within this insert-containing group, had further changes in the length of the FFL and substitutions (His376Asn and Arg377Ser) in the thumb loop.

Pus10 in Fungi

Of the 50 fungi with fully sequenced genomes, a total of nine species were found with orthologous Pus10 sequences (Supplementary Fig. 2). All nine were found in the earliest diverging families of fungi, the Unikaryonidae and Nosematidae (Microsporidia), Mucoraceae (Zygomycota) and Batrachochytrium (Chytridiomycota). The related lineages Ascomycota (as

yeasts and sac fungi) and Basidiomycota (as mushroom fungi) did not have Pus10 orthologs in any of the 41 sequenced genomes surveyed, although other pseudouridine synthase genes (e.g., Pus4, Cbf5) were found (Fig. 5). This suggested that Pus10 was secondarily lost in the common ancestor of these fungal species.

Common sequence features among the Pus10 sequences of fungi included the presence of 10 amino acid deletions in the 11^{th} α-helix and 13^{th} α-helix C-terminal to the catalytic site (Supplementary Fig. 2) (McCleverty et al. 2007). Microsporidia are also missing the first three secondary structures (α4, α6, β1) which are critical to the THUMP-domain (Fig. 2). All early fungi species contain the same sequence length of the FFL, and three conserved amino acids were present in these Pus10 FFLs ('PxxxxGxxxxxS').

The catalytic site of Pus10 in *R. oryzae* and *M. circinelloides* showed conversion of the key second Asp 277 to Asn. Other minor substitutions in the catalytic site have also been observed (Fig. 4). The thumb-loop in all nine fungi species is highly conserved in its sequence, except for in *R. oryzae* and *M. circinelloides* (Gln, instead of His). This changes the composition of the thumb-loop from a basic amino acid to an acidic amino acid and could affect the binding of tRNA (Fig. 4). A TrmA ortholog was present in *B. dendrobatidis* but not in any other species of Microsporidian or Zygomycota (Table 2). Interestingly, TruB (Pus4) orthologs are not present in the earlier fungi lineages (Microsporidia, Chytridiomycote and Zygomycote) but are present in Ascomycotes (*S. cerevisiae*).

Pus10 in Plants

Arabidopsis thaliana has a total of 24 pseudouridine synthase-like proteins. Most are paralogous copies of TruA, indicating expansion of this family in plants. Only one copy of Pus10 and TruD were detected in *A. thaliana*, and in 18 other plant genomes.

Multiple sequence alignment showed the incorporation of Ile (isoleucine), instead of Val (valine), in the catalytic site of 'GREDID' in plants. Substitutions were observed in the following species: 'GREDLD' in *Cucumis sativus*; 'GREDMD' in the moss *Physcomitrella patens* and green algae *Ostreococcus tauri*; 'GREDAD' in *Chlamydomonas reinhardtii* (Fig. 4). A highly modified THUMP-domain was present in most plants. *Nicotiana tabacum* and *Solanum lycopersicum* (Solanaceae), *Carica papaya* and *Oryza sativa* all lack α4 and β1 of the THUMP-domain. In addition, *N. tabacum* and *S. lycopersicum* are also missing α6 and β2. *Sorghum bicolor* Pus10 uncharacteristically lacks the catalytic Asp and several amino acids of the catalytic site (Fig. 4), indicating it is a non-functioning protein, however other grasses all seem to have a functional Pus10 catalytic site.

The molecular phylogenetic analysis of Pus10 has the same topology as the phylogenetic tree of species, with the exception of *Ectocarpus siliculosus* (Heterokontophyta; brown algae) (Fig. 4). This resulted in a split of the two green algae representatives – *C. reinhardtii* and *O. tauri*, Chlorophyta – indicating possible genome fusion in the brown algae (Fig. 4). All plant genomes contain TruB and TrmA orthologs (N-terminus of TruB in *C. reinhardtii*). Most species within the plant kingdom have functional Pus10; but *C. reinhardtii* and *C. papaya* have modifications in the catalytic site that may have an effect on its function.

Pus10 in Animals and the Importance of THUMP Domain

THUMP is a RNA binding domain present in 17 protein architectures and found in over 3,587 proteins according to the PFAM database. Many of these proteins have different biological roles. For example, S-adenosyl methionine dependent methyltransferase, rodanese like proteins, FtsJ like methyltransferase, SpoU rRNA methylase family and cytidine zinc-binding region. In this study, three major insertions which potentially disrupt or significantly alter the functional THUMP-domain in Pus10 were identified (Fig. 2). All eukaryotic Pus10 THUMP-domains were larger than archaeal ones and included numerous lineage specific insertions between conserved portions of the THUMP-motif (Fig. 2); they make identification of the THUMP-domain by sequence alignment difficult. The first insertion was observed on the N-terminal side of the THUMP-domain and is found in all taxa except fungi and Archaea. A second insertion within the THUMP domain was observed to be present in animals and in some plants. The third insertion is found shortly before the FFL, it is conserved in higher eukaryotes and in several Protista (*Paramecium tetraurelia* and *Trichoplax adhaerens*). However, the insertions were absent in Archaea (Fig. 2). In this study, 42 animal species representing 32 different orders, a few animal-specific modifications in the THUMP-domain and key Pus10 structure features (conserved residues in the catalytic site, FFL and thumb-loop) were observed (Fig. 5). The C-terminal catalytic site is highly conserved in all animals with some minor differences in earlier lineages. A maximum likelihood tree showed that the Protista, Tricoplacia, and Dictyostelia, followed by representatives of Nematoda, generally show the most diverse sequences alterations of key Pus10 features (Supplementary Fig. 3). Nematodes, *Caenorhabditis elegans* and *Pristionchus pacificus*, have a shorter FFL and the substitution of His to Lys in the thumb-loop; same observations were made for the Diptera clade. However, *Tribolium castaneum* had ‘GREDFD’ in the catalytic motif/core instead of I412 in respect to *M. jannaschii* (refer to Fig. 5). Although

there is no experimental evidence, these modifications to the THUMP domain may potentially alter or weaken non-specific Pus10 RNA binding and potentially alter the range of substrate RNA.

In *H. sapiens*, Pus10 is also known as Downstream Of Bid (DOBI). It has been shown to be a downstream interactor of the Bid (BH3 interacting-domain death agonist) protein signal involved in TRAIL-induced apoptosis and release of cytochrome C from mitochondria (Aza-Blanc et al. 2003; Jana et al. 2017). TRAIL is tumor necrosis factor–related apoptosis-inducing ligand and TNF-related apoptosis-inducing ligand. Mammalian Pus10 is recognized and cleaved by caspase-3 or caspase-8, which may create the functional unit for TRAIL signaling (Park et al. 2009). To analyze the ancestry of Pus10 involvement in TRAIL-induced apoptosis and tRNA pseudouridylation, 17 animal species were tested for the presence of a) intact orthologs of redundant functioning pathways (TruB, TrmA) and b) essential interacting components of TRAIL-induced apoptosis, caspase-3 and caspase-8 (Table 3). TrmA as well as caspase-3 and caspase-8 genome were not detectable in *D. pulex*, suggesting TRAIL-induced apoptosis is absent and Pus10 is likely to perform pseudouridylation of U54 (Table 3). Caspase-3 and caspase-8 orthologs were found in *T. adhaerens*, a multicellular organism belonging to the phylum Placozoa which lacks the presence of organs (Table 3). This, together with the absence of caspase-3 and caspase-8 orthologs in all fungi and plants, indicates that TRAIL-induced apoptosis could have begun as early as the last common ancestor of animals with *T. adhearens*, at the base of the metazoan clade, approximately 635My ago.

Discussion

Evolutionary Resolution of Functional Redundancy in Pus10 and TruB

Pus10 is the most recently identified member of the pseudouridine synthase family and it has been experimentally shown to be crucial for tRNA pseudouridylation in Archaea (Gurha and Gupta 2008; Mueller and Ferré-D'Amaré 2009; Roovers et al. 2006). While there is significant similarity between eukaryal (*H. sapiens*) and archaeal (*M. jannaschii*) Pus10 orthologs, no remotely similar protein sequences were found in Bacteria. On the other hand, TrmA and TruB are absent in Archaea, but are present in Bacteria and Eukarya. This suggests that Pus10 must have originated in Archaea and is required to modify tRNA at both positions U55 and U54. The coexistence of Pus10, with parallel tRNA modifying genes TrmA and TruB, impacts the function of Pus10 in Eukarya. This situation may have also occurred in the observed species of *Sufolobus* which have genes similar to bacterial TruB and TrmA, possibly acquired through horizontal gene transfer (HGT). In this case, the presence of both of these genes might have lead to the loss of key amino acids involved in the pseudouridine synthase function, but not the complete loss of the Pus10 protein. In another Archaea *P. furiosus*, in which HGT has introduced bacterial RumA to modify tRNA position U54 (but not U55), may have resulted in a more specific loss of Pus10 function for position U54 only. In general, after a duplication event, an organism is less likely to maintain more than one copy of a gene involved in tRNA processing and this coincides with our study of Pus10 across all species examined. Redundancy of Pus10 with non-homologous proteins performing similar functions may have led to its functional diversification.

The most important differences in terms of potential for revealing the mechanism of function in Pus10 were observed in Methanococcales and Thermococcales (refer to Fig. 4). Structural comparison of *P. furiosus* to *M. jannaschii* showed these modifications are very likely relevant for hook formation and substrate binding, which may be related to the loss *of in vivo*

modification of U54 without affecting U55 modification (Joardar et al. 2013). Several *in vitro* generated mutants of *M. jannaschii* Pus10 at these residues lost Ψ54 activity, but not activity of Ψ55. *P. furiosus* Pus10, does have a weak *in vitro* ability to pseudouridylate at tRNA position 54, whereas *M. jannaschii* has a robust one (Gurha and Gupta 2008). An *in silico* mutagenesis of His376 to Asn and Arg377 to Ser in *M. jannaschii* shows that this change in the thumb-loop should increase the gap width between both loops (Joardar et al. 2013). A wider gap generated by changes to the thumb-loop, and/or shortening of the FFL could contribute to a reduction in the tight pinching of the tRNA substrate that might be required for modification of U54.

In the case of *P. furiosus*, a bacterial rRNA methyltransferase (in the subfamily RumA of Pus superfamily) that is known to be acquired through HGT, is responsible for U54 modification in tRNA (Urbonavičius et al. 2008). Thus, in *P. furiosus*, the presence of RumA likely resulted in the partial relaxation of selection pressure and subsequent loss of Pus10 Ψ54 synthase ability, whereas its Ψ55 activity is retained. These differences in sequence across Archaea may indicate the extent of possible subfunctionalization and uncoupling of Ψ55 and Ψ54 activity of Pus10 for the order Thermococcales.

Based on our findings, it is likely that only TruB and TrmA orthologs remain functionally unchanged. Pus4 (TruB ortholog) converts U55 to Ψ55, and Trm2 (TrmA ortholog) modifies U54 to ribo-T54 in tRNAs of eukaryotes (Becker et al. 1997; Nordlund et al. 2000; Nurse et al. 1995). A different pseudouridine synthase, Cbf5, performs pseudouridylation of rRNA in a guide-RNA dependent manner in both eukaryotes and Archaea. However, in eukaryotes, Cbf5 has also neofunctionalized to play a role during mitosis by binding to microtubules and centromeres (Blaby et al. 2011; Jiang et al. 1993; Rashid et al. 2006). Thus, like Cbf5, Pus10 may have neofunctionalized to signal TRAIL-induced apoptosis in humans, if not all animals

either in place of or in addition to the pseudouridylation function. Pus10 protein is nuclear localized, and translocates to the mitochondria upon cleavage by caspase-3, which in turn amplifies caspase-3 activity, creating a positive feedback loop that has a central role in apoptosis (Jana et al. 2017). However, some mammalian tRNAs have been shown to have Ψ at position 54 instead of ribo-T. This indicates the existence of a protein to provide Ψ synthase activity for U54 (Roe and Tsen 1977). The obvious candidate is Pus10 (Gurha and Gupta 2008). Thus, we hypothesize that in some eukaryotic lineages Pus10 and TrmA orthologs may both have subfunctionalized, partitioning U54 modification activity, each with its own specific subset of tRNA substrates. Furthermore, we observed in the multiple sequence alignment of all 131 taxa, three insertions in the THUMP-domain, which could alter RNA recognition. Interestingly, a few human mitochondrial $tRNA^{LEU}$ do contain Ψ at position 55 (Helm 2006). However, the presence of TruB in *H. sapiens* does not leave out the possibility that TruB performs the aforementioned pseudouridylation rather than Pus10 (Table 2). While it is clear that mammalian Pus10 is part of TRAIL-apoptosis it remains unclear if it continues tRNA-pseudouridylation at position 54 in the Mammalian clade, though it is the most obvious candidate. If so, it may still be capable to perform tRNA-pseudouridylation at position 55 as it may be impossible to uncouple this reaction from position 54 pseudouridylation.

A RsuA-like Ψ synthase protein is known to be a suppressor of *var2* mutant phenotype and plays a role in variegation, an apoptosis of the chloroplast, in *A. thaliana* (Yu et al. 2008). This suggested a potential TRAIL-like apoptosis pathway might exist in plants. However, no sequences like caspase-3 were found. Plants contain 3 metacaspases, which are involved in the so called 'deathosome', to regulate programmed cell death (Coll et al. 2010; Vercammen 2004; Vercammen et al. 2007). Metacaspases are also found in the kingdoms of Protozoa, fungi and

plants. Plants contain type I and II metacapsases, but not animal-like caspsases. Metacaspase-4 and metacaspase-9 (AtMC4, AtMC9) recognize positively charged amino acids such as Arg or Lys in the motifs of substrate peptides with the amino acids FR, GRR, GKR and VRPR and cleave after Arg or Lys (Vercammen 2004). Future experimental work will be needed to determine whether Pus10 plays a part of either a mitochondrial or plastid apoptotic-like machinery in plants.

Molecular Evolution of Pus10 in Eukaryota

In Eukarya, the common ancestor likely had three proteins from Bacteria and Archaea through genome fusion (Koonin 2010) that modified residues U54 and U55 of tRNA, which created a functional redundancy. Over time the modification and loss of Archaeal-origin Pus10 interacts with that of bacterial-origin TruB and TrmA orthologs. Surprisingly, unlike most redundancy created through gene duplication, this redundancy was not resolved through gene loss in most lineages (Sémon and Wolfe 2007). Individual eukaryotic lineages were examined and revealed lineage specific differences in the fate of these three proteins in the fungi, plants and animals.

A surprising finding in all the sequenced genomes of this study, was the presence of only a single Pus10 gene per genome, despite the numerous genome duplications that have occurred in the evolution of animals, fungi and especially in plants (where genes frequently duplicate and subfunctionalize) (Koonin and Wolf 2010; Proost et al. 2011). This finding might point to a potential dosage effect or other harmful consequence for higher copy numbers of Pus10.

In Fungi, TruB orthologs are not present in the earlier fungi lineages (Microsporidia, Chytridiomycote and Zygomycote) but are present in Ascomycotes. The most parsimonious

solution is that the three early-branching fungal lineages have lost both TruB and TrmA, but have retained Pus10 (Table 2), while in Dikaryon fungi (Ascomycota, Basidiomycota), the opposite order of gene loss is observed, TruB and TrmA are retained and Pus10 is lost. Thus, it seems that these two groups of fungi have undergone the opposite selection of the enzymes Pus10, TruB and TrmA that can modify uridine residues at positions 54 and 55 of tRNA. The pattern of gain and loss here clearly points to the random elimination of redundant enzymes for tRNA U54 and U55 modification, that could be due to the loss of selection pressure. Given the phylogenetic analysis, it is unlikely that this pattern was the result of HGTs (refer to Fig. 5). This simple explanation for resolving functional redundancy through random gene loss is however not the case in plants and animals which retained seemingly functional orthologs of all three genes.

A recent duplication event in angiosperms should have resulted in at least two copies of Pus10 and in the 11 eudicot genomes surveyed (Proost et al. 2011). However, it has been shown that the extra paralogous genes coding for signaling and metabolism (e.g., protein kinases) are often maintained in plants while paralogous genes involved in DNA repair and modification are typically thinned to a limited number through rapid gene loss (Thomas et al. 2006). Gene ontology analysis of a tetraploid genome revealed the preference of keeping or losing duplicated genes is frequently based on their biological role. Indeed, duplicated genes which are responsible for tRNA processing were less likely to be maintained in the genome.

Ala (alanine) substitutions in *M. jannaschii* differentially reduced Ψ54 activity while retaining most of wild type Ψ55 activity (Joardar et al. 2013). However, a triple mutant in *M. jannaschii* (Ile412A/Lys413A/Leu440A) showed no pseudouridine synthase activity at either site. These two species, along with *Sorghum bicolor* that is lacking a catalytic Asp, indicate that Pus10 activity is lost in some plant species. However, the overall presence of conserved,

seemingly functional Pus10 in all other plant species in this study would indicate a positive selection pressure for some function of this enzyme.

A RsuA-like pseudouridine synthase protein is known to be a suppressor of *var2* mutant phenotype and plays a role in variegation, an apoptosis of the chloroplast, in *A. thaliana* (Yu et al. 2008). This suggested a potential TRAIL-like apoptosis pathway might exist in plants. However, no sequences like caspase-3 and caspase-8 were found. Plants contain 3 metacaspases, which are involved in the so called 'deathosome', to regulate programmed cell death (Coll et al. 2010; Vercammen 2004; Vercammen et al. 2007). Metacaspases are also found in the kingdoms of Protozoa, fungi and plants. Plants contain type I and II metacapsases, but not animal-like caspsases. Metacaspase-4 and metacaspase-9 (AtMC4, AtMC9) recognize positively charged amino acid such as Arg or Lys in the motifs of substrate peptides with the amino acids FR, GRR, GKR and VRPR and cleave after Arg or Lys (Vercammen 2004). Future experimental work will be needed to determine whether Pus10 plays a part of either a mitochondrial or plastid apoptotic-like machinery in plants.

Conclusion

This study provides evidence of Pus10 subfunctionalization in Thermococcales (Archaea) and possibly in the Chytridiomycota. In both cases, TrmA/Trm2-like enzymes are present in the genome, indicating an alternate route for modification of U54 that co-occurs with mutations that reduce or eliminate Pus10 U54 modification. Apparently, methylation of U54 of tRNA to riboT-54 by proteins acquired through either HGT or descent can lead to functional diversity in Ψ54 synthase activity of Pus10. A gap in the catalytic site of Pus10 in *S. bicolor* (plant) was noticed which could indicate that Pus10 is on the verge to become a pseudogene in some plant lineages.

Acknowledgment

This work was supported by NIH grant GM55045 to R.G.

Figure and Table Legends

Fig. 1 Domain distribution of pseudouridine synthases. Pseudouridine synthase superfamily contains (six) families and their substrates. Domains are represented as boxes, loops are represented as ovals. Catalytic site and thumb-loop are represented by their consensus sequences. Conserved catalytic aspartate (D) across all pseudouridine synthases is marked in bold, X indicates any amino acid. Due to the focus of this paper, Pus10 and tRNA55 are in bold. TruD family = TruD, TruA family = TruA, TruB family = TruB, RsuA family = RluB and RsuA, RluA family = RluC and RluA. Members of RluA and RsuA families differ in their structural features. RluA is lacking two extensions compared to its family member RluC. RsuA is lacking the thumb-loop and the C-terminal extension compared to its member RluB (Mueller and Ferré-D'Amaré 2009).

Fig. 2 THUMP-domain modification of Pus10 in different lineages. Blue colored helices and arrows emphasize structural elements belonging to typical THUMP-domain and are observed in archaeal Pus10. Three insertions were observed in multiple sequence alignment across 116 taxa and are indicated with arrows. + means presence of an insertion; - means no insertion.

Fig. 3 Substrate specific amino acids conserved throughout all pseudouridine synthases. **A** superimposition of human Pus10 (grey) and of *M. jannaschii* (blue) generated via Swiss-PDB viewer version 4.0.1(Guex and Peitsch 1997). FFL and thumb-loop (indicated with black arrows) are located in the C-terminal part of the protein. **B** Back view of the C-terminus with close up

focusing on the catalytic core, which shows the catalytic Asp275 (white), Leu440 (orange), Tyr339 (blue), Ile412 (green) and Lys413 (pink).

Fig. 4 Summary of the Pus10 characteristics of the C-terminal across Eukarya and Archaea. Here we want to emphasize on the substitutions between clades and chose taxa (97 out of 111) that differ in their Pus10 characteristic. The maximum likelihood tree was calculated via RAxML black box, model WAG+G+I. The annotation was generated via iTOL. Forefinger-loop features are represented in a single-value bar chart to emphasize on the length differences: Forefinger-loop = 15 = insertion (longer), 10 = normal length based on *H. sapiens*, 5 = partial deletion (shorter). Presence/absence/substitution of amino acid are represented with a color gradient: presence of amino acid = 50, substitution = 25, absence = 0. The catalytic region: Y339 (yellow), I412 (orange), K413 (purple) and L440 (light blue). Catalytic site including the catalytic Asp (D) (green). Thumb-loop (focus on His and Arg presence; blue). All substitutions are displayed.

Pus10 key features are displayed as seen in multiple sequence alignment. Species labels were color coded based on their belonging to archaea (light green) and eukaryote (light blue).

Fig. 5 Presence and absence tree of Pus10 across representatives of tree of life. The species tree was generated via NCBI taxonomy accession numbers and modified in iTOL. Presence of Pus10 is color coded in blue (outer circle) whereas absence of Pus10 is indicated in white. Clades were color coded: Algae = light green, Amoebozoa = orange, Archaea = turquoise, planta = green, fungi = light pink, Protista = purple, Animalia = blue. The loss of four nodes across Eukarya and Archaea of Pus10 is indicated as red circles. The absence of Pus10 in bacteria is indicated with a brown circle. The presence of TrmA and TruB are indicated with yellow and green stars.

Table 1 Structural alignment with THUMP-domain via DALIlite

Table 2 Presence/absence of TruB, TrmA and Pus10

Table 3 Presence/ absence of TruB, TrmA, caspase-3 and caspase-8

Supplementary Materials

Supplementary Fig. 1 Maximum likelihood tree of archaeal Pus10.

Supplementary Fig. 2 Modifications specific to microsporidia (fungi).

Supplementary Fig. 3 Maximum likelihood tree of Pus10 represented across the animal kingdom.

Supplementary Table. 1 List of Pseudouridine synthase Pus10 accession numbers.

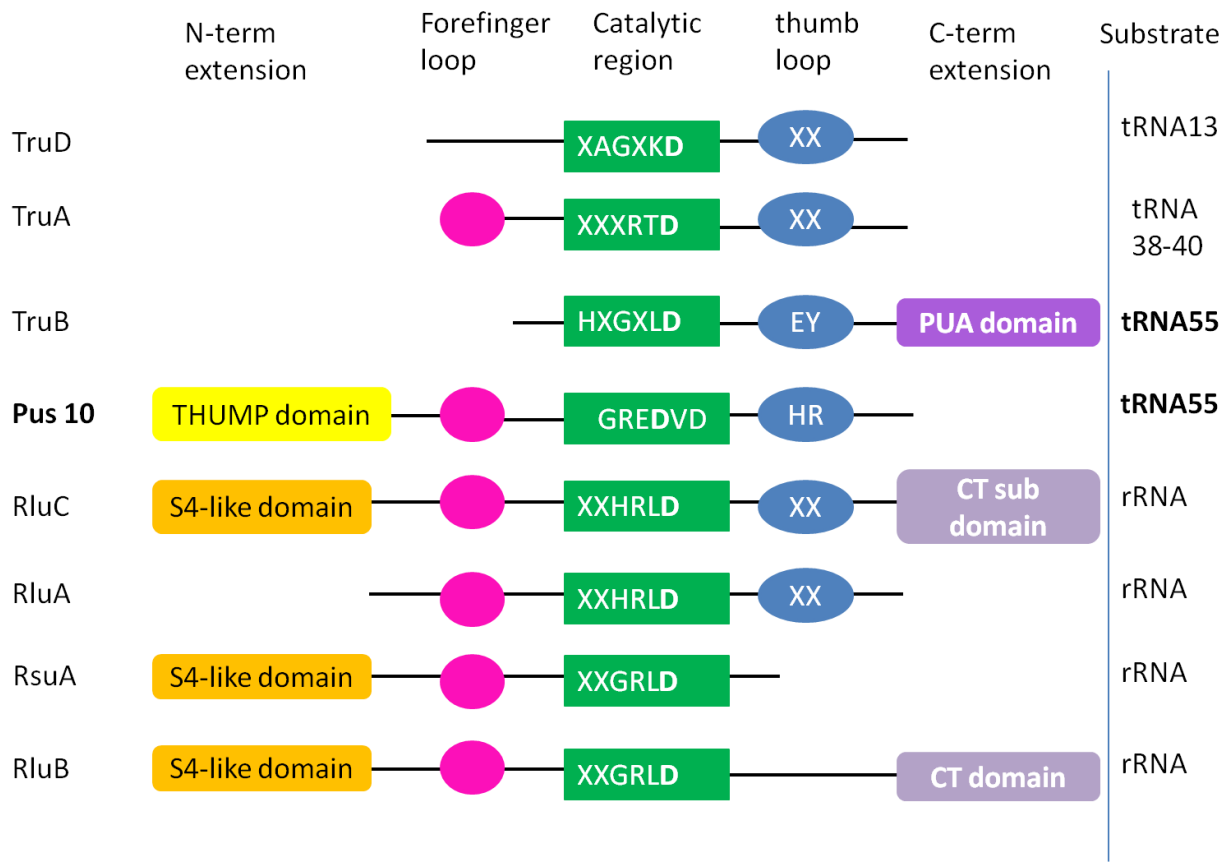
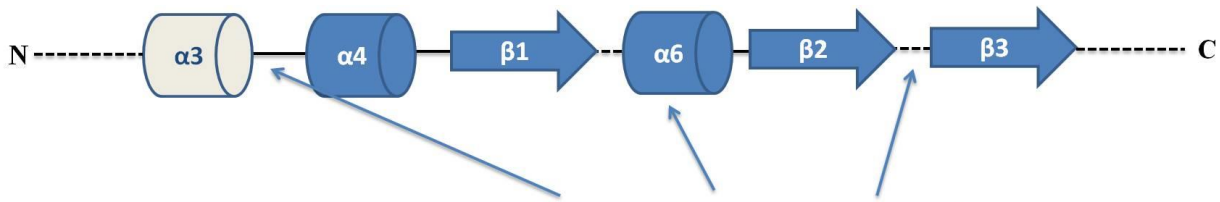


Fig. 1 Domain distribution of pseudouridine synthases. Pseudouridine synthase superfamily contains (six) families and their substrates. Domains are represented as boxes, loops are represented as ovals. Catalytic site and thumb-loop are represented by their consensus sequences. Conserved catalytic aspartate (D) across all pseudouridine synthases is marked in bold, X indicates any amino acid. Due to the focus of this paper, Pus10 and tRNA55 are in bold. TruD family = TruD, TruA family = TruA, TruB family = TruB, RsuA family = RluB and RsuA, RluA family = RluC and RluA. Members of RluA and RsuA families differ in their structural features. RluA is lacking two extensions compared to its family member RluC. RsuA is lacking the thumb-loop and the C-terminal extension compared to its member RluB (Mueller and Ferré-D'Amaré 2009).



Lineage	1st insertion	2nd insertion	3rd insertion
Animalia	+	+	+
Fungi	-	-	-
Planta	+	+	+
Protista	+	+	+
Archaea	-	-	-

Fig. 2 THUMP-domain modification of Pus10 in different lineages. Blue colored helices and arrows emphasize structural elements belonging to typical THUMP-domain and are observed in archaeal Pus10. Three insertions were observed in multiple sequence alignment across 116 taxa and are indicated with arrows. + means presence of an insertion; - means no insertion.

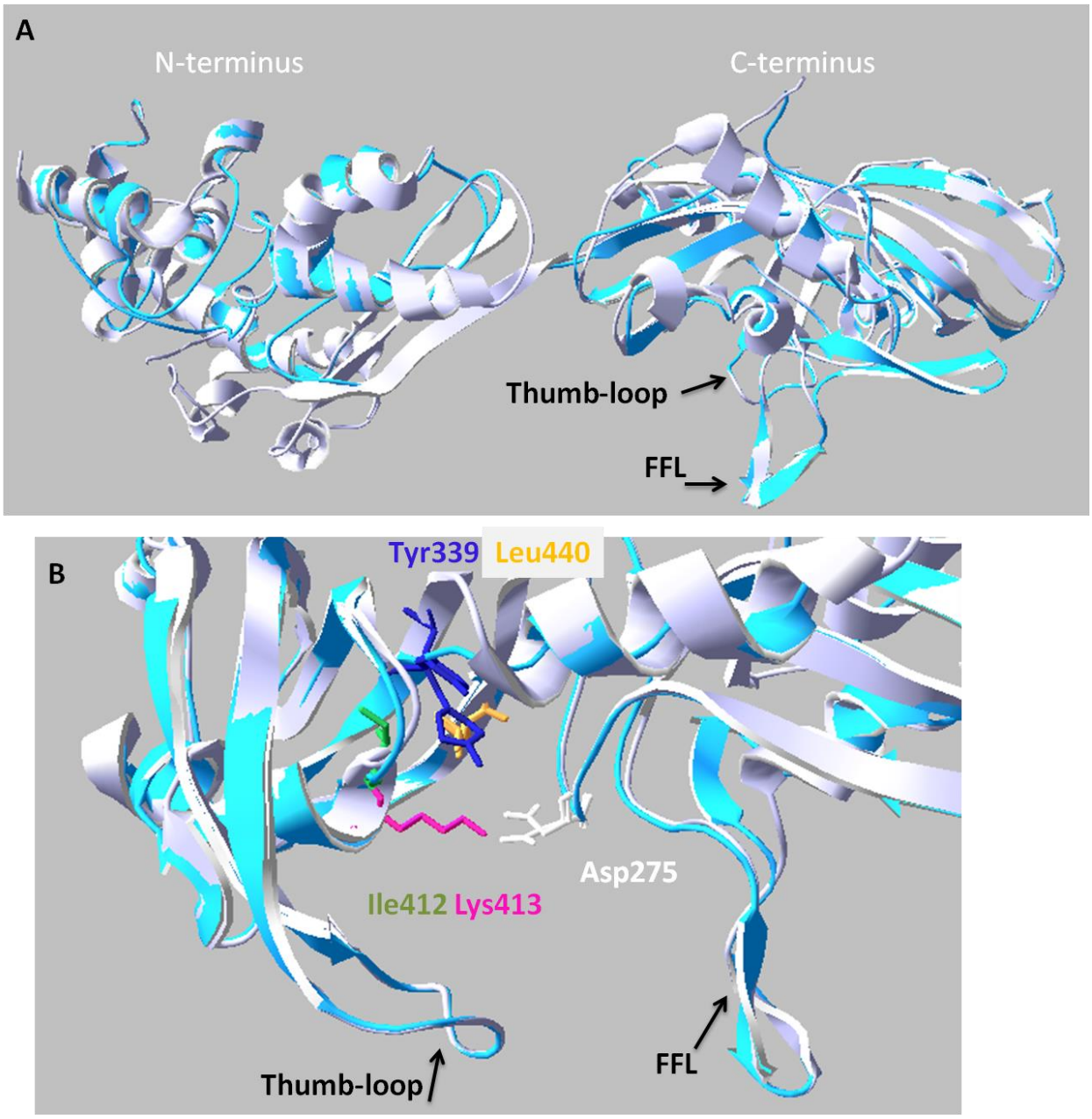


Fig. 3 Substrate specific amino acids conserved throughout all pseudouridine synthases. **A** superimposition of human Pus10 (grey) and of *M. jannaschii* (blue) generated via Swiss-PDB viewer version 4.0.1(Guex and Peitsch 1997). FFL and thumb-loop (indicated with black arrows) are located in the C-terminal part of the protein. **B** Back view of the C-terminus with close up

focusing on the catalytic core, which shows the catalytic Asp275 (white), Leu440 (orange), Tyr339 (blue), Ile412 (green) and Lys413 (pink).

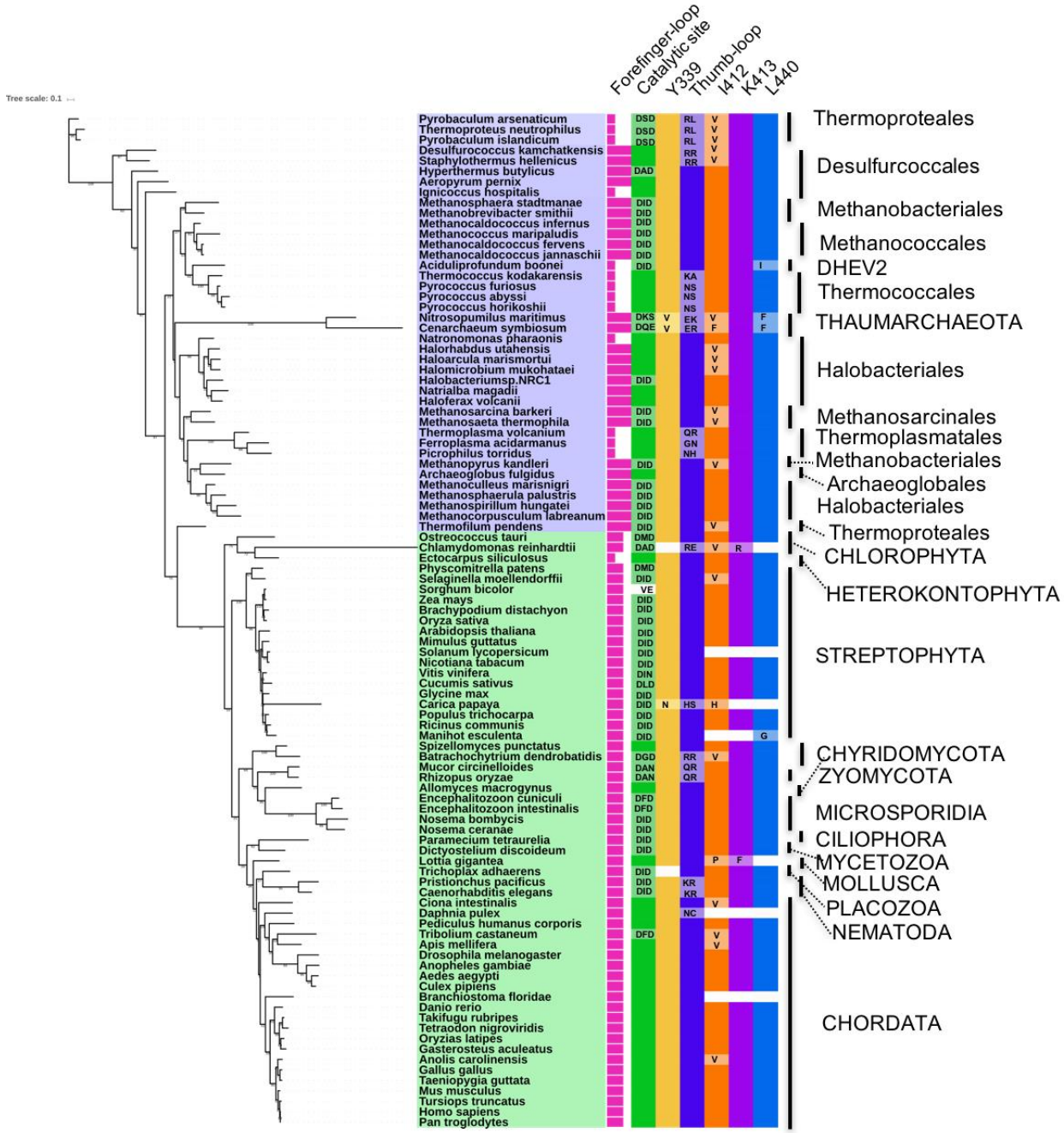


Fig. 4 Summary of the Pus10 characteristics of the C-terminal across Eukarya and Archaea. Here we want to emphasize on the substitutions between clades and chose taxa (97 out of 111) that differ in their Pus10 characteristic. The maximum likelihood tree was calculated via RAxML black box, model WAG+G+I. The annotation was generated via iTOL. Forefinger-loop features

are represented in a single-value bar chart to emphasize on the length differences: Forefinger-loop = 15 = insertion (longer), 10 = normal length based on *H. sapiens*, 5 = partial deletion (shorter). Presence/absence/substitution of amino acid are represented with a color gradient: presence of amino acid = 50, substitution = 25, absence = 0. The catalytic region: Y339 (yellow), I412 (orange), K413 (purple) and L440 (light blue). Catalytic site including the catalytic Asp (D) (green). Thumb-loop (focus on His and Arg presence; blue). All substitutions are displayed. Pus10 key features are displayed as seen in multiple sequence alignment. Species labels were color coded based on their belonging to archaea (light green) and eukaryote (light blue).

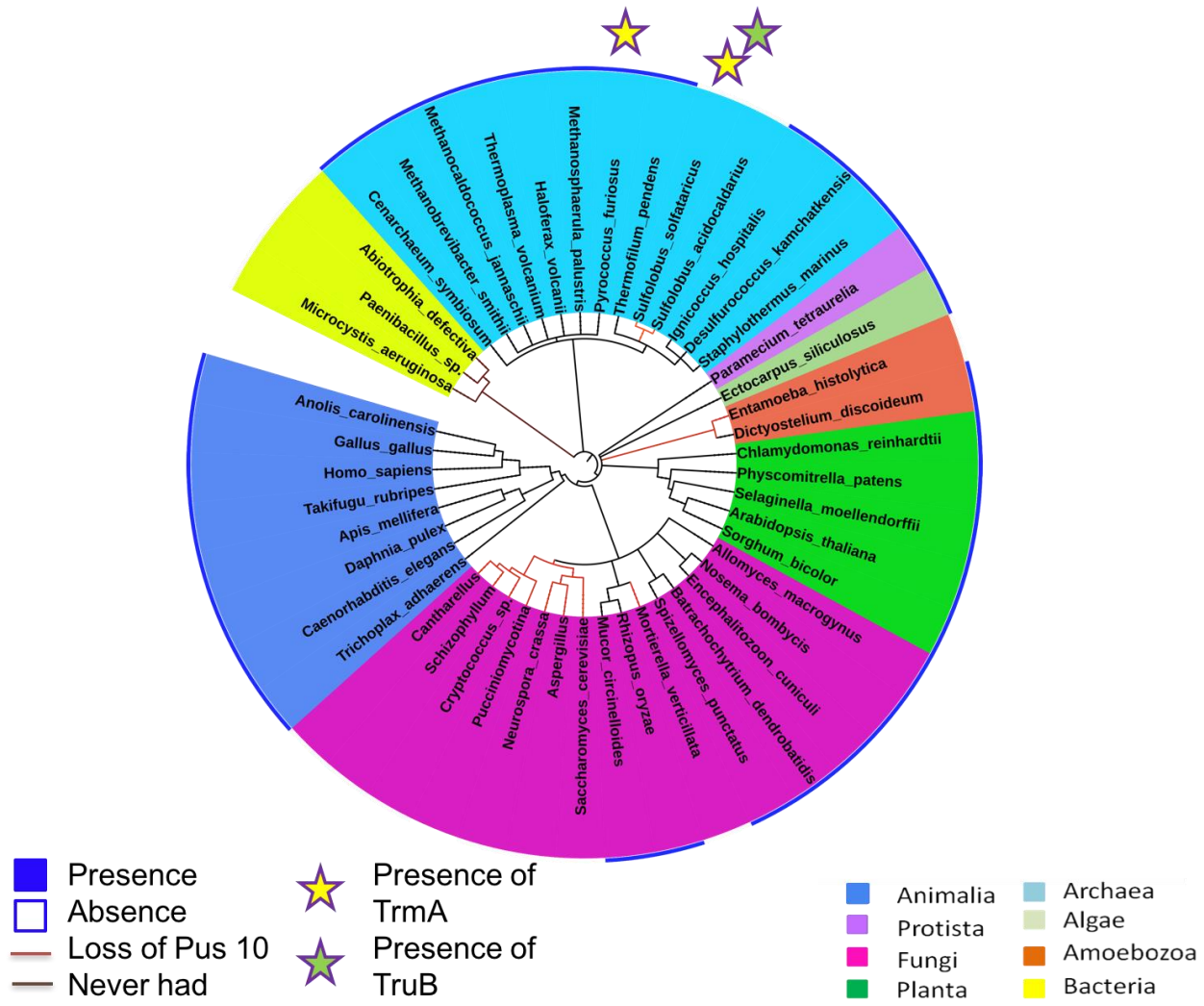


Fig. 5 Presence and absence tree of Pus10 across representatives of tree of life. The species tree was generated via NCBI taxonomy accession numbers and modified in iTOL. Presence of Pus10 is color coded in blue (outer circle) whereas absence of Pus10 is indicated in white. Clades were color coded: Algae = light green, Amoebozoa = orange, Archaea = turquoise, planta = green, fungi = light pink, Protista = purple, Animalia = blue. The loss of four nodes across Eukarya and Archaea of Pus10 is indicated as red circles. The absence of Pus10 in bacteria is indicated with a brown circle. The presence of TrmA and TruB are indicated with yellow and green stars.

Table 1 Structural alignment with THUMP-domain via DALIlite

Species	Lineage	Z-score
H. sapiens	Animal	3.7
D. melanogaster	Animal	3.7
C. elegans	Animal	2.8
H. volcanii	Archaea	3.3
M. jannaschii	Archaea	2.8
P. furiosus	Archaea	2.5
R. oryzae	fungi	4.3
N. bombycis	fungi	2.5
A. thaliana	plant	3.0
P. patens	plant	2.6

Table 2 Presence/absence of TruB, TrmA and Pus10

Species	TruB	Pus10	TrmA
Homo sapiens	NP_631908.1	NP_653310.2	NP_892029.2
Saccharomyces cerevisiae	EGA73247	ND	EDN59959
Zygomycota	ND	Yes	ND
Batrachochytrium dendrobatidis	ND	EGF78467	EGF83624.1
Microsporidians	ND	Yes	ND
Escherichia coli	TruB	ND	TrmA

ND = not detected by BLAST in the genome, Yes = present in all taxa sampled

Table 3 Presence/ absence of TruB, TrmA, caspase-3 and caspase-8

Species	TruB	TrmA (3BT7)	Caspase-3	Caspase-8
Homo sapiens	NP_631908.1	NP_892029.2	NP_116786.1	NP_001219.2
Gallus gallus	XP_421776.1	XP_415080.2	NP_990056.1	NP_989923.1
Taeniopygia guttata	XP_002186667.1	XP_002187418.1	XP_002191294.1	XP_002190143.1
Anolis carolinensis	XP_003223551.1	XP_003226338.1	ND	ND
Danio rerio	NP_001122159.1	NP_956223.1	NP_571952.1	NP_571585.2
Daphnia pulex	346482	ND	ND	ND
Branchiostoma floridae	XP_002607561.1	XP_002611903.1	XP_002610410.1	XP_002605135.1
Ciona intestinalis	XP_002130935.1	XP_002130635.1	XP_002131300.2	XP_002122848.1
Lottia gigantea	139266	121926	ND	ND
Paramecium tetraurelia	XP_001430104.1	XP_001442539.1	ND	ND
Trichoplax adhaerens	XP_002108027.1	XP_002115265.1	XP_002114146.1	XP_002114150.1
Dictyostelium discoideum	XP_640508.1	XP_629224.1	ND	ND
Drosophila melanogaster	NP_525120.1	NP_649083.2	NP_524551.2	NP_476974.1
Apis mellifera	XP_397244.4	XP_396538.2	XP_395697.2	XP_395697.2
Tribolium castaneum	XP_967291.1	XP_973242.1	XP_967501.2	XP_967501.2
Caenorhabditis elegans	NP_499370.1	NP_503353.2	ND	NP_001022453.1
Mucor circinelloides	113436	111833	ND	ND
Ostreococcus tauri	XP_003080670.1	XP_003078170.1	ND	ND
Physcomitrella patens	XP_001773683.1	XP_001766898.1	ND	ND
Oryza sativa	NP_001054428.2	EEC70100.1	ND	ND
Arabidopsis thaliana	NP_196950.2	NP_188767.2	ND	ND
Ricinus communis	XP_002521488.1	XP_002530289.1	ND	ND
Populus trichocarpa	EEF11588.1	XP_002314271.1	ND	ND

Number = accession number in GenBank, ND = not detected by BLAST of genome.

References

Anderson FE, Swofford DL (2004) Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA. Molecular Phylogenetics and Evolution 33:440

Aravind L, Koonin EV (2001) THUMP – a predicted RNA-binding domain shared by 4-thiouridine, pseudouridine synthases and RNA methylases. TRENDS in biochemical science 26:215

Aza-Blanc P, Cooper CL, Wagner K, Batalov S, Deveraux QL, Cooke MP (2003) Identification of Modulators of TRAIL-Induced Apoptosis via RNAi-Based Phenotypic Screening. Molecular Cell 12:627

Bates PA, Kelley LA, MacCallum RM, Sternberg MJE (2001) Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins: Structure, Function, and Bioinformatics 45:39

Becker HF, Motorin Y, Planta RJ, Grosjean H (1997) The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of Ψ55 in both mitochondrial and cytoplasmic tRNAs. Nucleic Acids Research 25:4493

Blaby IK, Majumder M, Chatterjee K, Jana S, Grosjean H, De Crécy-Lagard V, Gupta R (2011) Pseudouridine formation in archaeal RNAs: The case of Haloferax volcanii. RNA 17:1367

Bombarely A, Menda N, Tecle IY, Buels RM, Strickler S, Fischer-York T, Pujar A, Leto J, Gosselin J, Mueller LA (2011) The Sol Genomics Network (solgenomics.net): growing tomatoes using Perl. Nucleic Acids Research 39:1149

Chan CM, Huang RH (2009) Enzymatic characterization and mutational studies of TruD – the fifth family of pseudouridine synthases. Archives of Biochemistry and Biophysics 489:15

Coll NS, Vercammen D, Smidler A, Clover C, Van Breusegem F, Dangl JL, Epple P (2010) Arabidopsis Type I Metacaspases Control Cell Death. Science 330:1393

Conrad J, Niu L, Rudd K, Lane BG, Ofengand J (1999) 16S ribosomal RNA pseudouridine synthase RsuA of Escherichia coli: deletion, mutation of the conserved Asp102 residue, and sequence comparison among all other pseudouridine synthases. RNA 5:751

Foster PG, Huang L, Santi DV, Stroud RM (2000) The structural basis for tRNA recognition and pseudouridine formation by pseudouridine synthase I. Nat Struct Mol Biol 7:23

Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis:2714

Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis:162

Gurha P, Gupta R (2008) Archaeal Pus10 proteins can produce both pseudouridine 54 and 55 in tRNA. RNA 14:2521

Hamilton CS, Spedaliere CJ, Ginter JM, Johnston MV, Mueller EG (2005) The roles of the essential Asp-48 and highly conserved His-43 elucidated by the pH dependence of the pseudouridine synthase TruB. Archives of Biochemistry and Biophysics:322

Hamma T, Ferré-D'Amaré AR (2006) Pseudouridine Synthases. Chemistry & Biology 13:1125

Helm M (2006) Post-transcriptional nucleotide modification and alternative folding of RNA. Nucleic Acids Research 34:721

Hoang C (2004) Crystal structure of the highly divergent pseudouridine synthase TruD reveals a circular permutation of a conserved fold. RNA 10:1026

Hoang C, Chen J, Vizthum CA, Kandel JM, S. Hamilton C, Mueller EG, Ferré-D'Amaré AR (2006) Crystal Structure of Pseudouridine Synthase RluA: Indirect Sequence Readout through Protein-Induced RNA Structure. Molecular Cell:535

Hoang C, Ferré-D'Amaré AR (2001) Cocrystal Structure of a tRNA55 Pseudouridine Synthase: Nucleotide Flipping by an RNA-Modifying Enzyme. Cell 107:929

Holm L, Kääriäinen S, Rosenström P, Schenkel A (2008) Searching protein structure databases with DaliLite v.3. Bioinformatics 24:2780
Holm L, Rosenstrom R (2010) Dali server: conservation mapping in 3D. Nucleic Acids Research 38:W545
Hubbard TJP, Aken BL, Ayling S, Ballester B, Beal K, Bragin E, Brent S, Chen Y, Clapham P, Clarke L, Coates G, Fairley S, Fitzgerald S, Fernandez-Banet J, Gordon L, Graf S, Haider S, Hammond M, Holland R, Howe K, Jenkinson A, Johnson N, Kahari A, Keefe D, Keenan S, Kinsella R, Kokocinski F, Kulesha E, Lawson D, Longden I, Megy K, Meidl P, Overduin B, Parker A, Pritchard B, Rios D, Schuster M, Slater G, Smedley D, Spooner W, Spudich G, Trevanion S, Vilella A, Vogel J, White S, Wilder S, Zadissa A, Birney E, Cunningham F, Curwen V, Durbin R, Fernandez-Suarez XM, Herrero J, Kasprzyk A, Proctor G, Smith J, Searle S, Flicek P (2009) Ensembl 2009. Nucleic Acids Research 37:D690
Hur S, Stroud RM (2007) How U38, 39, and 40 of Many tRNAs Become the Targets for Pseudouridylation by TruA. Molecular Cell 26:189
Jana S, Hsieh AC, Gupta R (2017) Reciprocal amplification of caspase-3 activity by nuclear export of a putative human RNA-modifying protein, PUS10 during TRAIL-induced apoptosis. Cell Death & Disease 8:e3093
Jiang W, Middelton K, Yoon H-J, Fouquet C, Carbon J (1993) An Essential Yeast Protein, CBF5p, Binds In Vitro to Centromeres and Microtubules. Molecular and Cellular Biology 13:4884
Joardar A, Jana S, Fitzek E, Gurha P, Majumder M, Chatterjee K, Geisler M, Gupta R (2013) Role of forefinger and thumb loops in production of 54 and 55 in tRNAs by archaeal Pus10. RNA 19:1279
Kamalampeta R, Keffer-Wilkes LC, Kothe U (2013) tRNA Binding, Positioning, and Modification by the Pseudouridine Synthase Pus10. Journal of Molecular Biology 425:3863
Kaya Y, Ofengand J (2003) A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. RNA:711
Koonin EV (2010) The origin and early evolution of eukaryotes in the light of phylogenomics. Genome Biology 11:209
Koonin EV, Wolf YI (2010) Constraints and plasticity in genome and molecular-phenome evolution. Nature Reviews Genetics 11:487
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 (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947
Letunic I, Bork P (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Research 39:W475
Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Research 44:W242
McCleverty CJ, Hornsby M, Spraggon G, Kreusch A (2007) Crystal structure of Human Pus10, A novel pseudouridine synthase. J. Mol. Biol 373:1243
McGinnis S, Madden TL (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Research 32:W20
Mueller EG, Ferré-D'Amaré AR (2009) DNA and RNA Modification Enzymes. In: Grosjean H (ed) Pseuduridine formation the most common transglycosylation in RNA. Landes Bioscience, Austin, TX
Nordlund ME, Johansson JO, von Pawel-Rammingen U, Byström AS (2000) Identification of the TRM2 gene encoding the tRNA(m5U54)methyltransferase of Saccharomyces cerevisiae. RNA:844
Nurse K, Wrzesinski J, Bakin A, Lane BG, Ofengand J (1995) Purification, cloning, and properties of the tRNA Ψ55 synthase from Escherichia coli. RNA:102

Ny T, Björk GR (1980) Cloning and restriction mapping of the trmA gene coding for transfer ribonucleic acid (5-methyluridine)-methyltransferase in Escherichia coli K-12. Journal of Bacteriology 142:371

Pan H, Agarwalla S, Moustakas DT, Finer-Moore J, Stroud RM (2003) Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit. Proceedings of the National Academy of Sciences 100:12648

Park S-Y, Shin JN, Woo HN, Piya S, Moon AR, Seo Y-W, Seol D-W, Kim T-H (2009) DOBI is cleaved by caspases during TRAIL-induced apoptotic cell death. BMB reports 42:511

Proost S, Pattyn P, Gerats T, Van de Peer Y (2011) Journey through the past: 150 million years of plant genome evolution. The Plant Journal 66:58

Rashid R, Liang B, Baker DL, Youssef OA, He Y, Phipps K, Terns RM, Terns MP, Li H (2006) Crystal Structure of a Cbf5-Nop10-Gar1 Complex and Implications in RNA-Guided Pseudouridylation and Dyskeratosis Congenita. Molecular Cell 21:249

Roe B, Tsen H-Y (1977) Role of ribothymidine in mammalian tRNAPhe. Proc Natl Acad Sci 74:3696

Roovers M, Hale C, Tricot C, Terns MP, Terns RM, Grosjean H, Droogmans L (2006) Formation of the conserved pseudouridine at position 55 in archaeal tRNA. Nucleic Acids Research 34:4293

Sémon M, Wolfe KH (2007) Consequences of genome duplication. Current Opinion in Genetics & Development 17:505

Sivaraman J, Sauvé V, Larocque R, Stura EA, Schrag JD, Cygler M, Matte A (2002) Structure of the 16S rRNA pseudouridine synthase RsuA bound to uracil and UMP. Nature Structural Biology 9:353

Spedaliere CJ, Hamilton CS, Mueller EG (2000) Functional Importance of Motif I of Pseudouridine Synthases: Mutagenesis of Aligned Lysine and Proline Residues†. Biochemistry 39:9459

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688

Thomas BC, Pedersen B, Freeling M (2006) Following tetraploidy in an Arabidopsis ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. Genome Research 16:934

Urbonavičius J, Auxilien S, Walbott H, Trachana K, Golinelli-Pimpaneau B, Brochier-Armanet C, Grosjean H (2008) Acquisition of a bacterial RumA-type tRNA(uracil-54, C5)-methyltransferase by Archaea through an ancient horizontal gene transfer. Molecular Microbiology 67:323

Vercammen D (2004) Type II Metacaspases Atmc4 and Atmc9 of Arabidopsis thaliana Cleave Substrates after Arginine and Lysine. Journal of Biological Chemistry 279:45329

Vercammen D, Declercq W, Vandenabeele P, Van Breusegem F (2007) Are metacaspases caspases? The Journal of Cell Biology 179:375

Whelan S, Goldman N (2001) A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. Molecular Biology and Evolution 18:691

Yu F, Liu X, Alsheikh M, Park S, Rodermel S (2008) Mutations in SUPPRESSOR OF VARIEGATION1, a Factor Required for Normal Chloroplast Translation, Suppress var2-Mediated Leaf Variegation in Arabidopsis. The Plant Cell Online 20:1786