Journal of Molecular Evolution (2021) 89:484–493
https://doi.org/10.1007/s00239-021-10019-z

ORIGINAL ARTICLE

The Evolutionary Fate of Mitochondrial Aminoacyl-tRNA Synthetases in Amitochondrial Organisms

Gabor L. Igloi[1]

Received: 28 May 2021 / Accepted: 26 June 2021 / Published online: 12 July 2021
© The Author(s) 2021

Abstract
During the endosymbiotic evolution of mitochondria, the genes for aminoacyl-tRNA synthetases were transferred to the ancestral nucleus. A further reduction of mitochondrial function resulted in mitochondrion-related organisms (MRO) with a loss of the organelle genome. The fate of the now redundant ancestral mitochondrial aminoacyl-tRNA synthetase genes is uncertain. The derived protein sequence for arginyl-tRNA synthetase from thirty mitosomal organisms have been classified as originating from the ancestral nuclear or mitochondrial gene and compared to the identity element at position 20 of the cognate tRNA that distinguishes the two enzyme forms. The evolutionary choice between loss and retention of the ancestral mitochondrial gene for arginyl-tRNA synthetase reflects the coevolution of arginyl-tRNA synthetase and tRNA identity elements.

Keywords Aminoacyl-tRNA synthetase · Identity elements · Amitochondrial · Mitosome · Protozoa · Arginyl-tRNA

Introduction

The evolution of eukaryotes was dictated by the endosymbiotic acquisition of an α-proteobacterial cell and its subsequent development to the aerobic energy-generating mitochondrial organelle. Over time, the genome of the engulfed prokaryote became reduced with many of the gene products required for mitochondrial metabolism becoming-encoded by genes transferred to the nucleus whereas some were retained to permit organelle-specific protein synthesis. Among the genes transferred from the endosymbiont to the nucleus were those of the family of aminoacyl-tRNA synthetases whose products are then imported to participate in mitochondrial protein biosynthesis. However, their subsequent presence in modern eukaryote cells is very variable (Igloi 2020b). The mitochondrial gene of arginyl-tRNA synthetase, for example, has either been retained in the nucleus (e.g., in Metazoans), eliminated such that the cytoplasmic form is imported into the organelle (e.g., in higher plants), or has itself replaced the ancestral nuclear gene (e.g., in Fungi; Karlberg et al. 2000; Brindefalk et al. 2007; Furukawa et al. 2017)). This evolutionary variation within cells performing oxidative phosphorylation also raises the question as to the fate of mitochondrial aminoacyl-tRNA synthetases, in general, and of arginyl-tRNA synthetase, in particular, in the genome of modern amitochondrial, anaerobic eukaryotes. As is common usage, the product from the ancestral nuclear gene is referred to as the "cytoplasmic" enzyme, whereas the translation product of the nuclear-encoded mitochondrial gene is termed "mitochondrial" enzyme.

Arginyl-tRNA synthetase is responsible for the accurate attachment of L-arginine to its cognate tRNA in the first step leading to ribosomal protein chain extension. The specific recognition of tRNA by its cognate aminoacyl-tRNA synthetase is determined by identity elements positioned within the tRNA structure. Identity elements have been well-studied after being initially proposed (Loftfield et al. 1968) and experimentally verified (Schulman and Pelka 1989; McClain 1993). In the case of arginyl-tRNA synthetase detailed kinetic investigation of yeast transcript variants (Sissler et al. 1996) and a study of yeast tRNA mutants conferring lethality in vivo (Geslain et al. 2003b) showed that C35 (the 2nd base of the anticodon) was the major identity element with minor contributions from U/G36. In the hamster cytoplasmic system (Guigou and Mirande 2005) and in plants (Aldinger

Handling Editor: **Alan Christensen**.

✉ Gabor L. Igloi
igloi@biologie.uni-freiburg.de

[1] Institute of Biology III, University of Freiburg, Schaenzlestr. 1, 79104 Freiburg, Germany

et al. 2012) additionally position A20 in the D-loop emerged as an essential determinant. A20 is also a major site of recognition in bacteria (Tamura et al. 1992; McClain 1993; Shimada et al. 2001a) and its tight binding to the enzyme has been confirmed by crystallography (Stephen et al. 2018) but has been ruled out as such in yeast, both by transcript studies (Liu et al. 1999) and by the viability of mutants in vivo (Geslain et al. 2003b).

Evolutionary aspects of identity element divergence between bacteria and yeast have been discussed (Nameki et al. 1995). The phylogenetic reconstruction of yeast nuclear genes showed that a small group of aminoacyl-tRNA synthetases, including arginyl-tRNA synthetase, occurred by gene duplication from the mitochondrial lineage followed by loss of the homologue from the cytoplasmic lineage (Karlberg et al. 2000).A similar phylogenetic clustering of the nuclear-encoded yeast cytoplasmic and mitochondrial arginyl-tRNA synthetases has been noted (Brindefalk et al. 2007; Furukawa et al. 2017). It is apparent that the yeast enzyme and hence its mode of tRNA recognition originates from an ancient gene replacement of the nuclear gene by that of the mitochondrial gene. As a result of a similar gene replacement, the single valyl-tRNA synthetase gene in *Homo sapiens* provides a product with both cytosolic and mitochondrial functions and is of mitochondrial origin (Brown and Doolittle 1995; Wolf et al. 1999). A duplication of mitochondrial threonyl-tRNA synthetase and of alanyl-tRNA synthetase in eukaryotes with displacement of the ancestral eukaryotic form has also been proposed (Doolittle and Handy 1998; Wolf et al. 1999; Chihade et al. 2000).

Mammalian mitochondrial tRNA recognition has been reviewed in detail (Salinas-Giegé et al. 2015; Krahn et al. 2020) and that of arginyl-tRNA synthetase examined in insects (Igloi and Leisinger 2014) but in view of the bizarre tRNA structures of metazoan tRNAs (with reduced or missing D stem-loops, making the location of position 20 uncertain) this can not necessarily be extrapolated to conventional tRNA structures in protists. The recognition of mitochondrial tRNAs in non-metazoans having canonical secondary structures appears to follow the "universal identity rules" (Salinas-Giegé et al. 2015). tRNA recognition in protists has, with few exceptions (Cela et al. 2018) not been subjected to extensive study. However, an indication that the identity rules as far as $tRNA^{Arg}$ is concerned may extend to non-eumetazoans can be gained from the fact that in a heterologous system, the (non-metazoan, mitochondrial-like) yeast arginyl-tRNA synthetase, where position 20 is not a recognition element, is able to aminoacylate both poriferan mitochondrial $tRNA^{Arg}$ isoacceptors possessing either A20 and U20 (Igloi and Leisinger 2014). It is tempting, then to use the presence or absence of the tRNA-A20 nucleotide in eukaryotes possessing canonical tRNA structures, to distinguish between the recognition by cytoplasmic and mitochondrial enzymes. Thus, organisms featuring solely $tRNA^{Arg}$A20 could survive with either the cytoplasmic or mitochondrial enzyme whereas others with non-A20 or a mixture including A20 would of necessity require the less stringent recognition by the mitochondrial enzyme which might lead to the evolutionary elimination the primordial nuclear gene.

Sequence analysis and comparisons of arginyl-tRNA synthetase has revealed a clear and characteristic distinction between the cytoplasmic and the mitochondrial forms of the enzyme. In the N-terminal region the cytoplasmic enzyme has a conserved domain recognizable by a GDYQ-like motif (whose function remains to be established). This motif is also found in some bacteria, including *E. coli* which relies on the tRNA-A20 recognition element and may be a relic of the ancestral prokaryotic host of the proto-endosymbiont. It is absent from the prokaryotic suborder Cystobacterineae from which the mitochondrial arginyl-tRNA synthetase has been derived (Igloi 2020a). The mitochondrial enzyme, on the other hand, has, when aligned with the cytoplasmic version a barely discernible GDYQ-like motif but a five-amino acid deletion followed by an MSTR-like sequence nearer the C terminus (referred to here as Δ5MSTR). This region is part of the conserved signature sequence motif KMSK which is a characteristic of the class I aminoacyl-tRNA synthetases, in general (but, exceptionally, is degenerate in arginyl-tRNA synthetases) (Sekine et al. 2001). It constitutes a part of the catalytic site in three dimensions. Although these distinctive features have been confirmed in hundreds of species from numerous phyla (Igloi 2020b), one is aware of the pitfalls in making global generalizations since it is likely that evolutionary niches with diverging characteristics will be found.

Exceptionally, in some protists gene loss in mitochondria has proceeded to completion giving rise to mitochondrion-related organelles (MRO) lacking all DNA and, hence, mitochondrial protein expression (Makiuchi and Nozaki 2014). Having eliminated mitochondrial protein synthesis, the question concerning the fate of the now redundant ancestral mitochondrial aminoacyl-tRNA synthetase genes arises. In order to trace the ancestral source of the arginyl-tRNA synthetase gene in amitochondrial organisms, the arginyl-tRNA synthetase encoded by their genomes has been examined and classified as originating from the lost mitochondrial genome or as being the retained nuclear species. Since the coevolving cognate tRNA must respond in terms of identity elements to changes in the nuclear-encoded arginyl-tRNA synthetase, a comparison of the nature of the identity element at position 20 of the corresponding nuclear $tRNA^{Arg}$, shows a trend relating the tRNA identity element to the subcellular source of the arginyl-tRNA synthetase gene.

Table 1 Summary of mitosomal organisms analysed for their arginyl-tRNA synthetase (ArgRS)

Clade or Higher Classification	Phylum	Class/Order	Species	Predicted ArgRS Type	tRNA N20
Alveolata	Apicomplexa	Eucoccidiorida	*Cryptosporidium parvum*	Cyto	A
			Gregarina niphandrodes	Cyto	A
	Dinophyceae	Syndiniales	*Amoebophrya ceratii*	Cyto	A
Amoebozoa		Archamoeba	*Entamoeba histolytica*	Mito	A
			*Mastigamoeba balamuthi**	?	A
Eumetazoa	Cnidaria	Bivalvulida	*Henneguya salminicola*	Cyto	A, C*
Fungi	Chytridiomycota	Neocallimastigales	*Neocallimastix sp*	Mito	A
	Microsporidia		*Anncaliia algerae*	Mito	C
			Edhazardia aedis	Mito	U
			Encephalitozoon intestinalis	Mito	C, U
			Enterocytozoon bieneusi	Mito	C, U
			Hepatospora eriocheir	Mito	A
			Nematocida parisii	Mito	C, U
			Nosema ceranae	Mito	C, U
			Ordospora colligata	Mito	C, U
			Pseudoloma_neurophilia	Mito	C
			Spraguea lophii	Mito	C
			Trachipleistophora hominis	Mito	C, U
			Vavraia culicis	Mito	C, U
			Vittaforma corneae	Mito	C, U
Metamonada	Fornicata	Diplomonadida	*Giardia lamblia*	Cyto	A
			Kipferlia bialata	Cyto	A
			Retortamonas cf. caviae	Cyto	A
			Spironucleus salmonicida	Cyto	A
			Trepomonas sp	Cyto	?
	Preaxostyla	Oxymonad	*Monocercomonoides sp*	Mito	A
			Streblomastix strix	Mito	A, U
	Parabasalia	Tritrichomonadida	*Tritrichomonas foetus*	Mito?	A
			Histomonas meleagridis	Mito?	?
Rhizaria	Endomyxa	Ascetosporea	*Mikrocytos mackini*	Mito?	?

Columns to the right indicate the predicted ancestral type and the identity element at position 20 of the cognate $tRNA^{Arg}$

*For *Henneguya salminicola*, and Mastigamoeba, see text. Uncertain predictions or missing tRNA data are shown by "?" Cyto = cytoplasmic, Mito = mitochondrial

Methods

Protein sequences were manually extracted from whole genome shotgun (WGS) and transcriptome (TSA) databases following TBLASTN analysis using sequences from closely related organisms or phyla. Protein sequences were derived from genomic sequence hits using the FGENESH + protein-based gene prediction algorithm (Solovyev 2008). This relies on the accessibility of genome-specific parameters which are lacking for most protists so that the data from closely related organisms was used. Considering the extent of N-terminal variability between arginyl-tRNA synthetases from different species (Igloi 2019), this section of the predicted protein sequence must be treated with caution until confirmation from the corresponding transcriptome sequence becomes available. The alternative genetic code for ciliates was used for translating the Trepomonas (Keeling and Doolittle 1997) and Streblomastix (Keeling and Leander 2003) genes. Alignments were performed with the MAFFT server (Katoh and Standley 2013) using the E-INS-I parameters and depicted in GENEDOCv2.7 (Nicholas and Nicholas 1997) with similarity groups enabled.

tRNA sequences were mined manually from the corresponding genome or transcriptome collections by BLASTN and/or with tRNAscan-SE (Chan and Lowe 2019).

A

```
                                                                                 tRNA20 Type
Amoebophrya_sp(Dinophyceae)                       :FGDYQCNNAMDLFSQYRGDVE : TVAPAGFISVRINRKW   A    Cyto
Cryptosporidium_parvum(Apicomplexa)               :FGDYQCNSCMTIFSQFKGKLE : EVAPQGFITIKLSDEF   A    Cyto
Gregarina_niphandrodes(Apicomplexa)               :FGDFQCNNAMGLFKKFKSSLP : SVAPAGFVTVKLDPRW   A    Cyto
Henneguya_salminicola(Cnidaria)                   :EFDYQNSDAMKISKLLSKKDV : DISNTGFINIKLKTIY   A    Cyto
Giardia_lamblia(Metamonada)                       :YGHYQFNSTMGIAKLIKDLHY : NIA-GSFINIVLTPEY   A    Cyto
Retortamonas_cf_caviae(Metamonada)                :FGDFQVNNTMPIMKLIKELKV : APA-GIFVNIFLSASY   A    Cyto
Spironucleus_salmonicida(Metamonada)              :FGHYQLNNTMKIQKQLQMLDI : TVA-AVFVNITLSDKF   A    Cyto
Trepomonas_sp(Metamonada)                         :FGHYQINNVMQISKLLKANGY : AGA-AIFVNITLSDQF   ?    Cyto
Kipferlia_bialata(Metamonada)                     :YGDYQVNNAMGIFKTLKQHTQ : TIA-GGFINLRVSETY   A    Cyto
Entamoeba_histolytica(Amoebozoa)                  :LGDF----SVPVAKIIRQMKT : TVT-GPFLNFAINKSA   A    Mito
Neocallimastix_sp(Chytridiomycota)                :FGDY----CIPIARLRLQ--- : NSV-GMFLNFSISKEK   A    Mito
Anncaliia_algerae(Microsporidia)                  :QSDY----TLEVVKFSKN--- : FTH-GPVVFIVIKKYE   C    Mito
Edhazardia_aedis(Microsporidia)                   :NGDF----SIHETRLKPY--- : LYT-NKSLYFQVDKKS   U    Mito
Pseudoloma_neurophilia(Microsporidia)             :KSDF----TLQLGKIAKQ--- : IKQ-GLAHHITLNKQK   C    Mito
Trachipleistophora_hominis(Microsporidia)         :KSDF----TIQLAKVLHS--- : SKS-KLCYNVKINKAK   C,U  Mito
Vavraia_culicis(Microsporidia)                    :KSDF----TIQLAKVLHS--- : SKS-KLCYNVKIDKTR   C,U  Mito
Spraguea_lophii(Microsporidia)                    :SADI----TLQLAKMYKN--- : EVA-GTIFKIRLNKKK   C    Mito
Encephalitozoon_intestinalis(Microsporidia)       :KPNV----TLFLIKISSS--- : AIR-GPSVCFDINKKA   C,U  Mito
Ordospora_colligata(Microsporidia)                :KQDV----TLFVIKISNT--- : SLK-GPSISFDVNKKV   C,U  Mito
Enterocytozoon_bieneusi(Microsporidia)            :KKDF----IIFLNKITND--- : ELQ-ETSILFNLNKKK   C,U  Mito
Vittaforma_corneae(Microsporidia)                 :NSDF----LLFLNKVSSS--- : ACH-DTHLSFNLNKKM   C,U  Mito
Hepatospora_eriocheir(Microsporidia)              :KSEF----VIWFTKLSNE--- : QAS-NNRLSIYIDKRK   A    Mito
Nosema_ceranae(Microsporidia)                     :KPHC----TLFLPKLELQ--- : TTF-KSSISFNINLIP   C,U  Mito
Nematocida_parisii(Microsporidia)                 :NGDF----IVPFSQLTRD--- : EVI-KQMLCIYIDVKG   C,U  Mito
Mastigamoeba_balamuthi(Amoebozoa)                 :AADI----ALPCFKLAQK--- : VAS-GPFLNLTVAPAD   A    Mito
Mikrocytos_mackini(Rhizaria)                      :NFDF----SILSFPLEPF--- : NVS-KKYVNIALNPS-   ?    ?
Histomonas_meleagridis(Metamonada)                :VGDYACNVSKYFAFLKKEKKA : SAA-NAYLNLKVERGE   ?    ?
Tritrichomonas_foetus(Metamonada)                 :AGDFSANVPRYLATLPKEEKL : EAT-GPYLNFRVNRGY   A    ?
Monocercomonoides_sp(Metamonada)                  :KFDL----AFSIHNRLPQGSK : EGAGTGFLNVILNRQE   A    Mito
Streblomastix_strix(Metamonada)                   :LRDV----SFPLHSIGNEKAE : ILQ-KGFLNIKLNRAE   A,U  Mito

Saccharomyces_cerevisiae                          :RGDL----LIPIPR----LRI : EAN-GPFIQFFNPQFL   C,U  Mito
                                                  66    GDYQ REGION     78 104 A20 REGION 119
```

B

```
                                                                                                Type
Amoebophrya_sp(Dinophyceae)                   : TKCEHMGFGVVQGEDGK-----KFKTRSGDTVKLEMLLDE : Cyto
Cryptosporidium_parvum(Apicomplexa)           : TRLFFIGFGVIQGEDGK-----KFKTRSGDVVKLTELIDE : Cyto
Gregarina_niphandrodes(Apicomplexa)           : TRLDHCGLGLVCGDDGK-----KFKTRSGEVVKLVDLLDE : Cyto
Henneguya_salminicola(Cnidaria)               : SRVEHVGFGLVLGEGRK-----KLKTRSGETVKLRELLDE : Cyto
Giardia_lamblia(Metamonada)                   : NQAVHVPFGVVLGNDGK-----RLKTRAGETVKLEDLLDE : Cyto
Retortamonas_cf_caviae(Metamonada)            : IVAEHVPFGVVLGDDGK-----RLKTRSGETVKLMDLLKE : Cyto
Spironucleus_salmonicida(Metamonada)          : NQAQHIGFGVVLGEDGK-----RLKTRSGETVKLQELLNE : Cyto
Trepomonas_sp(Metamonada)                     : AQAQHCGFGVVVGEDGK-----KLKTRSGETV-LQELMDE : Cyto
Kipferlia_bialata(Metamonada)                 : AKATHVGFGLVCGPDGK-----RYRTRDGRTEKLQDLLKN : Cyto
Entamoeba_histolytica(Amoebozoa)              : KECYHISFGMVMG----------MSTRRGTAVFLQDILDD : Mito
Neocallimastix_sp(Chytridiomycota)            : PQCHHINFGLVKSKDGQ------MSTRKGTVVFLEEILDA : Mito
Anncaliia_algerae(Microsporidia)              : DIITHVEYGLILG----------MSTREGKVKFLEDVIEE : Mito
Edhazardia_aedis(Microsporidia)               : TSFEHVNYGMVNG----------MSTRKGDVVFLEDIIEI : Mito
Pseudoloma_neurophilia(Microsporidia)         : CKLVHIGFGMVKG----------FSTRKGNLVFLEDIINE : Mito
Trachipleistophora_hominis(Microsporidia)     : VELEHIGFGMVNG----------LSTRKGNLVFLEDIIDA : Mito
Vavraia_culicis(Microsporidia)                : LELEHISFGMVNG----------LSTRKGNLVFLEDIIDA : Mito
Spraguea_lophii(Microsporidia)                : NKLVHVNFGMVKG----------MSSRKGTVVFLEDIIKE : Mito
Encephalitozoon_intestinalis(Microsporidia)   : SKFLHVSYGLVAG----------MSTRAGKAQFLEDIIQE : Mito
Ordospora_colligata(Microsporidia)            : SKFQHVNYGLVAG----------MSTREGKVEFLEDILDE : Mito
Enterocytozoon_bieneusi(Microsporidia)        : EKFLHVNYGLVKG----------MSTREGTVHFIDDVIDL : Mito
Vittaforma_corneae(Microsporidia)             : SMFEHVNYGSVKG----------MSTRGGTVHFLEDIIQT : Mito
Hepatospora_eriocheir(Microsporidia)          : SKLRHVNYGLVTG----------MSTRAGNVHFIEDVISR : Mito
Nosema_ceranae(Microsporidia)                 : DNLEHVQYGMVSG----------MSTRKGTVKFLEDIIDI : Mito
Nematocida_parisii(Microsporidia)             : ALFLHINFGMVKG----------MSTRKGTAVSLSDIIDV : Mito
Mastigamoeba_balamuthi(Amoebozoa)             : CSLRHVAFGVMLLWDDESKSWGKGRTRSGGVMLLREVLEE : ?
Mikrocytos_mackini(Rhizaria)                  : KNCLHLPYGQVVLESG------KMSSRLNNSIAFTQLKKE : ?
Histomonas_meleagridis(Metamonada)            : RPQFHMAYELVKLSENGKIK--RMSGRKGYYILADTLYND : ?
Tritrichomonas_foetus(Metamonada)             : RQQFHMAYELVKLSEKGKIK--RMSGRKGYYILADTLYDD : ?
Monocercomonoides_sp(Metamonada)              : EKQWHLSYDLVRLVNC------RMSGRRGIYVTADEVYQE : Mito
Streblomastix_strix(Metamonada)               : EKQWHLSYELVRLVGR------RMSGRRGIFVTADEVYKE : Mito

Saccharomyces_cerevisiae                      : KDLQHVNFGMVQG----------MSTRKGTVVFLDNILEE : Mito
                                              395                 △5MSTR-region    424
```

Fig. 1 Multiple sequence alignments of arginyl-tRNA synthetase protein sequences derived from genomic or transcriptome data. **A**. Alignment of the GDYQ-like signature region and of the proposed tRNA nucleotide A20 binding-region (bold) within the arginyl-tRNA synthetase of mitosomal organisms. Columns to the right give the nucleotide at position 20 of the cognate tRNA and the ancestral source of the enzyme. **B**. Alignment of the MSTR-like signature region within the arginyl-tRNA synthetase of mitosomal organisms. Column to the right gives ancestral source of the enzyme. For orientation, the sequence of the cytoplasmic *S. cerevisiae* protein and its numbering is appended. Shading colours denote 100% identity (red), 80–100% identity (green) and 60–80% identity (blue). Cyto and Mito indicates cytoplasmic and mitochondrial characteristics of the enzyme, respectively (Color figure online)

Results

From 6 phylogenetic clades, 30 organisms that are known to possess the MRO (mitosomes or hydrogenosomes) and for which BLAST searches have been able to identify either annotated or similarity-defined arginyl-tRNA synthetase genes have been compiled. Their putative translation products were classified (Table 1; Online Resource 1) as being of mitochondrial or of nuclear origin on the basis of the sequence motifs described above and aligned (Fig. 1). Following a highly variable N-terminus, the most obvious N-terminal region of similarity tends to commence close to the typical GDYQ-like domain in products from the nuclear-encoded gene. The genomes of the organisms were searched for as many different $tRNA^{Arg}$ isoacceptors as possible and the nature of the base at N20 was examined and summarised (Table 1).

Amoebozoa

Amoebozoa are a major taxonomic group most of whose members possess a mitochondrial genome and consequently perform mitochondrial protein biosynthesis. Their nuclear genome encodes both cytoplasmic and mitochondrial forms of the enzyme (Igloi 2020b). Nevertheless, certain Archamoeba are amitochondrial. The parasitic Entamoeba species are remarkable with their mitochondria being reduced to mitosomes and lacking organelle protein biosynthesis. In its multiple protein alignment with other amoebozoa the Archamoeba form a striking outgroup (Igloi 2019). The sequence of the protein quite clearly indicates a mitochondrial origin. It not only has a merely rudimentary cytoplasmic N-terminal GDYQ-motif but an explicit mitochondrial Δ5MSTR signature (Fig. 1), which is characteristic for fungal and all 492 metazoan mitochondrial arginyl-tRNA synthetases. The Δ5MSTR region is confirmed by EST data (Acc. No. CX094868). The gene for this protein was evidently transferred to the nucleus prior to loss of the endosymbiotic genome. A transfer of mitochondrial genes to the nucleus of *E. histolytica* and other amitochondrial organisms has been previously documented (Clark and Roger 1995; Hashimoto et al. 1998; Tovar et al. 1999; Bakatselou et al. 2003; Luque et al. 2008). The mitochondrial enzyme will recognize all four nuclear $tRNA^{Arg}$ isoacceptors that possess A20 (Table 1). The tRNA genes are more likely to have originated from the ancestral endosymbiotic host since all available sequences of aerobic amoebozoa mitochondrial-encoded $tRNA^{Arg}$ possess U20.

In contrast, the free-living Archamoeba *Mastigamoeba balamuthi* is also amitochondrial (Gill et al. 2007). Its arginyl-tRNA synthetase is exceptional as it cannot be unambiguously classified according to the criteria used for hundreds of other eukaryotic species (Igloi 2020b). It possesses a barely recognizable GDYQ-like domain but its Δ5MSTR signature which aligns poorly with other cytoplasmic Amoebozoa arginyl-RNA synthetases (not shown) is highly atypical but with some similarity to the cytoplasmic form. The structure of this region is, nevertheless, confirmed by EST data (Acc.No. EC698286). Furthermore, all three nuclear-encoded $tRNA^{Arg}$ isoacceptors have nucleotide A20 and would be recognized by both enzyme types.

Apicomplexa

Within the Apicomplexans, Cryptosporidium (Hikosaka et al. 2013) and Gregarina (Toso and Omoto 2007) possess only mitosomes. Their genomes carry the cytoplasmic form of the arginyl-tRNA synthetase and correspondingly, all $tRNA^{Arg}$ carry the A20 identity element.

Of the other classes within the Alveolata, Dinophyceae, are represented by the endoparasite *Amoebophrya ceratii,* which has lost its mitochondrial genome while retaining the organelle itself (John et al. 2019). It uses the cytoplasmic arginyl-tRNA synthetase in conjunction with the A20 identity element.

Eumetazoa

The Cnidarian *Henneguya salminicola* is the sole representative of a eumetazoan anaerobe to date that survives in its host without requiring a mitochondrion (Yahalomi et al. 2020). Its cytoplasmic arginyl-tRNA synthetase has been detected in both WGS and TSA databases and the genomic data provides sequences corresponding to single copies of four $tRNA^{Arg}$ isoacceptors. Interestingly, $tRNA^{Arg}_{UCG}$ is reported with C20 (Acc.No. SGJC01004408). Without intending to cast unfounded doubt on this sequence, one should point out that it has close similarity to the genomic sequence from its host *Oncorhynchus kisutch* (Acc. No. MPKV02000020) and the related *Salmo salar* (Acc. No. AGKD04837854), both with C20, but, remarkably, the equivalent *S. salar* TSA (Acc.

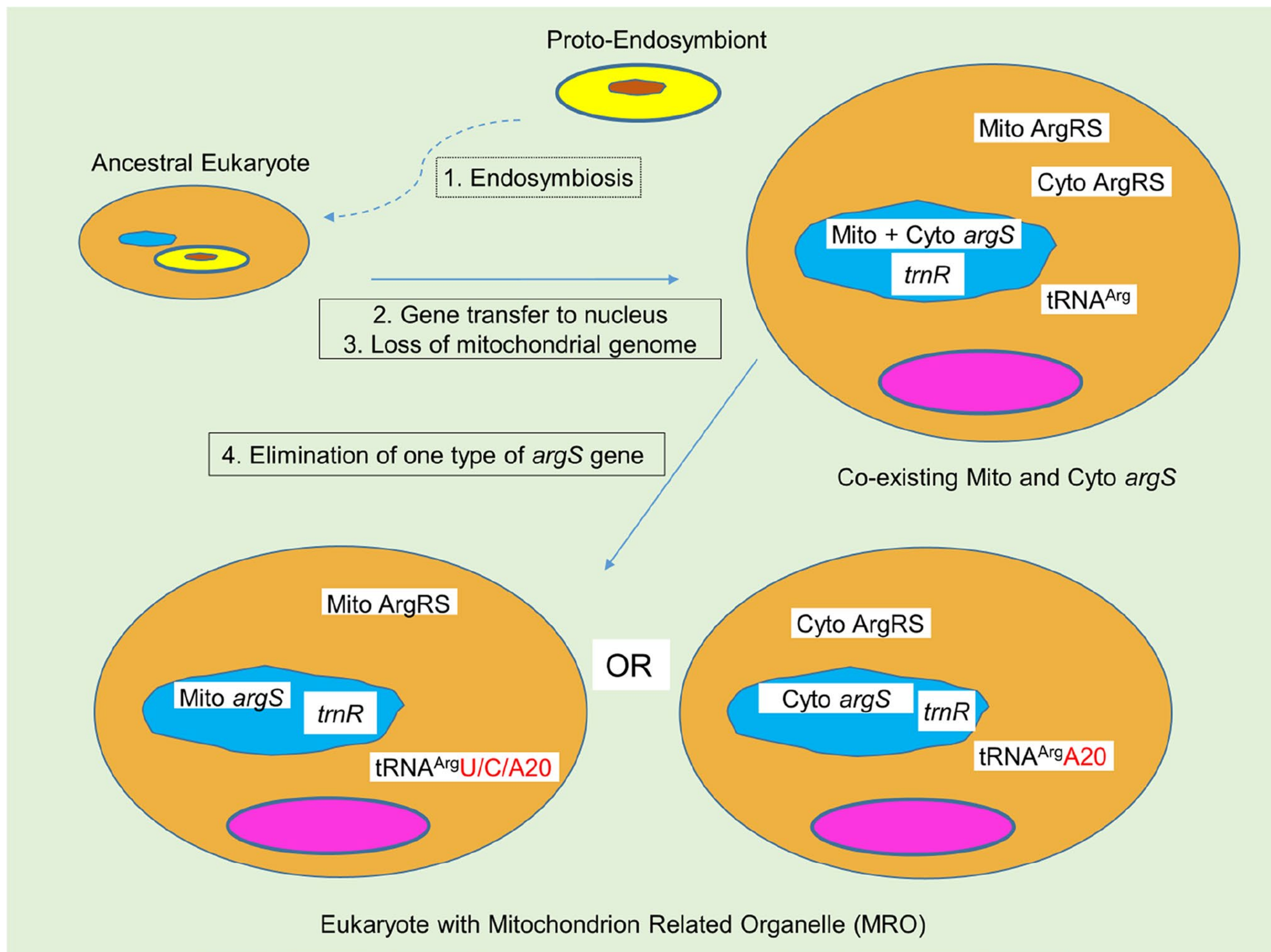

Fig. 2 Model for the evolutionary option between retention and loss of the ancestral mitochondrial arginyl-tRNA synthetase provided by the proto-endosymbiont (yellow oval, with brown genome) in mitosomal organisms. The residual mitosome is shown in purple; the nuclear genome in blue. The genes for arginyl-tRNA synthetase and $tRNA^{Arg}$ are symbolized by *argS* and *trnR*, respectively. Their gene products are Mito ArgRS, Cyto ArgRS and $tRNA^{Arg}$ with the identity element given in red (Color figure online)

No. GBRB01035513) is identical to the genomic sequence except for having A20. This may be indicative of a rare but not unknown C-to-A editing event (Smith et al. 1997; Paul et al. 2020) that would make it accessible for recognition by the cytoplasmic enzyme.

Fungi

Fungi, in general, are characterised by having replaced the ancestral nucleus-encoded version of arginyl-tRNA synthetase gene by the mitochondrial gene (Igloi 2020b). This trend is retained in the mitosomal *Neocallimastix sp* (Chytridiomycota) (Van Der Giezen et al. 1997) but, as pointed out previously for Chytridiomycota in general (Igloi 2020b), with a reduced deletion at the Δ5MSTR region. The group of related Microsporidia have lost the mitochondrial genome but have retained its organelle-derived arginyl-tRNA synthetase (Fig. 1A and B). The genomic tRNA isoacceptors of all 13 species listed, have either A, C or U at position 20 requiring recognition by the less discriminating mitochondrial enzyme.

Metamonada

All three phyla within the Metamonada harbour mitosomal organisms. The arginyl-tRNA synthetases, of Fornicata, represented here by five genera are, according to the sequence criteria, clearly of nuclear origin.

Within the Oxymonad order, the product of the *argS* in Streblomastix (Keeling and Leander 2003) has lost the GDYQ-feature, whereas the mitochondrial Δ5MSTR signature is evident, albeit with a reduced deletion. It is dependent on the recognition of the $tRNA^{Arg}$U20-encoded by the organism (Table 1) so that one can classify it as mitochondrial

with some confidence. Alignment with the sequence from the other available Oxymonad in the compilation, the mitosomal *Monocercomonoides sp* (Karnkowska et al. 2016) shows high similarity (with 56% identity overall, not shown) and although the tRNA isoacceptors all possess A20, the enzyme has been classed as originating from the ancestral mitochondrion. Similarly, the two closely related amitochondrial candidates from the Tritrichomonadida order, *Tritrichomonas foetus* (Lindmark and Muller 1973) and *Histomonas meleagridis* (Mielewczik et al. 2008) are highly similar in their protein alignments; 46% identity (not shown). The GDYQ-domain is barely discernible and the MSTR-feature resembles more closely that of the mitochondrial enzyme, although again the characteristic five-amino acid deletion is missing. In Tritrichomonas the cognate genomic $tRNA^{Arg}$ has the A20 identity element which is recognizable by both enzyme types. No $tRNA^{Arg}$ information from Histomonas is currently available.

Rhizaria

For the single example whose mitochondrion has undergone loss of it genome, *Mikrocytos mackini* (Burki et al. 2013), the multiple sequence alignment resembles in its signature regions the mitochondrial enzyme. Unfortunately, its transcriptome has, as yet, failed to reveal any $tRNA^{Arg}$ candidate sequences.

Discussion

Endosymbiosis led to the transfer of most of the genes from the chromosome of the proto-symbiont prokaryote to the chromosome of the host proto-eukaryote. Genes required for the metabolism of modern-day mitochondria were compulsorily retained within the nuclear genome with the translation products being imported to the organelle. Redundant house-keeping genes that had nuclear equivalents performing identical functions could be eliminated. Such genes included those for the family of aminoacyl-tRNA synthetases. This however meant an evolutionary choice between retaining the host or the symbiont gene. In the case of eumetazoans, whose non-canonical, truncated mitochondrial tRNAs (Watanabe 2010) are not substrates for the cytoplasmic enzymes (Kumazawa et al. 1991), the genes for both mitochondrial and cytoplasmic forms of the enzymes had of necessity to be retained. However, in protists, having canonical mitochondrial tRNA structures (Gray et al. 1998) retention of the mitochondrial form is optional, depending on availability of matching tRNA identity elements (and the requirement for a mitochondrial signal peptide). For amitochondrial organisms, on the other hand, for which all components of the protein synthesising machinery are located in the cytoplasm, it is not immediately obvious which evolutionary option regarding elimination of aminoacyl-tRNA synthetase genes has been taken. One possibility, investigated here, is that differences in identity element recognition might be a key aspect in matching the retained enzyme with the genomic tRNA isoacceptors.

In this model (Fig. 2), following endosymbiosis, gene transfer from the proto-endosymbiont to the host nucleus and loss of the mitochondrial genome, an evolutionary pathway leads to the elimination of either one of the redundant arginyl-tRNA synthetases. The choice may involve the nature of the identity element present in the expressed tRNA isoacceptors. For this enzyme position 20 in the tRNA is the critical difference between recognition by the two types of enzyme. Whereas the mitochondrial form is indifferent to the nature of the nucleotide at position 20, the nuclear-encoded cytoplasmic synthetase is strictly dependent on the presence of A20. Consequently, the cognate tRNA with U20 or C20 (G20 has not been documented in any $tRNA^{Arg}$, to date) demands the aminoacylation activity of the mitochondrial enzyme, permitting the elimination of the nuclear form. This is exemplified by the genome of Microsporidia. If all $tRNA^{Arg}$ isoacceptors of an organism have A20 then they will be recognized by both forms of the enzyme and may retain either the mitochondrial form, as in Entameoba, or the cytoplasmic version, as in Apicomplexa. Remarkable in this respect is that Metamonada seem to be divided into two groups; the Fornicata, retaining the cytoplasmic enzyme and Preaxostyla/ Parabasalia that have opted for the mitochondrial form.

The interaction of the A20 identity element with arginyl-tRNA synthetase of bacteria has been examined in crystallographic detail (Shimada et al. 2001a). It was proposed that the amino acid equivalent to Q111 in yeast is replaced by N in bacteria. This is involved in A20 recognition and appeared to be phylogenetically invariant in the arginyl-tRNA synthetases from all organisms which possessed A20 in their $tRNA^{Arg}$. The accumulation of eukaryotic data in the intervening years has shown that N at this position, although common, is by no means universal in non-metazoans with T being the preferred replacement (Igloi 2019). In the case of the mitosomal organisms an examination of the aligned sequences with the presence or absence of A20 in the cognate tRNA reveals a degree of agreement with the requirement for asparagine for tRNA-A20 interaction in both cytoplasmic and mitochondrial enzymes. The Microsporidia enzymes binding to U20 or C20 possess a wide variety of amino acids at the critical position (Fig. 1). In yeast, mutagenesis of F109 and/or Q111 to alanine had no effect on the viability of the organism or on the enzyme activity in vitro (Geslain et al. 2003a). Hence the mitochondrial enzymes are insensitive at this position as far as U20 or C20 are concerned but, with the exception of Hepatospora,

possess N when tRNA-A20 needs to be bound (Fig. 1). On the other hand, the major deviation from the combination of N in the protein and A20 in tRNA is seen in Apicomplexa and Dinophyceae with T or S replacing N. (Fig. 1). However, the structural study of bacterial arginyl-tRNA synthetase showed that mutagenic replacement of N by D, for instance, had little influence on the kinetics of the enzyme in vitro (Shimada et al. 2001a) so that other amino acids, such as S or T might also be able to provide the H-bond that enables the critical interaction.

The nature of aminoacyl-tRNA synthetases in mitosomal organisms has been addressed in studies applied to the evolution of these enzymes with some data emerging, notably, in the case of Giardia. For alanyl-tRNA synthetase the sequences from all mitochondrion-containing eukaryotes, are of the mitochondrial type (Chang et al. 2012). However, the *alaRS* gene in the amitochondriate protists *Giardia lamblia* and *Trichomonas vaginalis* appears to have evolved from archaeal (nuclear) origin (Chihade et al. 2000; Bunjun et al. 2000; Chang et al. 2012). Similarly, Giardia tryptophanyl-tRNA synthetase is more similar to its cytoplasmic form (Arakaki et al. 2010). On the other hand, the analysis of Giardia valyl-tRNA synthetase clearly indicates that lateral gene transfer occurred between the ancestral genomes of extant mitochondria and the ancestral nuclear genome of Giardia (Brown and Doolittle 1995; Hashimoto et al. 1998) leading to the replacement of the ancestral nuclear gene.

The decision of which will be lost must then depend on additional criteria which provide an evolutionary advantage. Viability can, evidently, be achieved by either pathway. *E. histolytica* with $tRNA^{Arg}$A20 has clearly retained the mitochondrial form whereas numerous other organisms with the same identity element from a variety of phyla have opted for the cytoplasmic version. One may speculate that with a more flexible, less stringent set of identity elements, the mitochondrial enzyme can react more easily to changes/mutations in its substrate. On the other hand, relaxed identity elements can lead to mutational tRNA misrecognition (Igloi and Leisinger 2014) and translational error propagation. Indeed there are indications that mitochondrial protein synthesis is less stringent in terms of accuracy (Shimada et al. 2001b; Lue and Kelley 2005; Roy et al. 2005) and a case in point is the high frequency of mistranslation that has been reported in microsporidia (Melnikov et al. 2018). The multiple identity elements required by the cytoplasmic enzyme (at least A20 and C35 Shimada et al. 2001a; Guigou and Mirande 2005; Aldinger et al. 2012), while making it more dependent on the tRNA structure may thereby enhance the accuracy of translation.

The concept discussed here relies on the major assumption that tRNA identity elements that are required for accurate recognition by the corresponding aminoacyl-tRNA synthetases have been co-evolutionarily conserved since the endosymbiotic acquisition of mitochondria. When it comes to evolution, it is tempting but hazardous to make global generalizations; the samples numbers in question cannot be representative of all evolutionary niches. However, the assumption regarding co-evolutionary conservation is generally admitted. Indeed, the statement that a "general trend is the global conservation throughout evolution of the identity set for a given amino acid" (Bonnefond et al. 2005) has not been questioned. Furthermore, the current evidence suggests that mitochondrial tRNAs with canonical structure (e.g., those found in plants and fungi) seem to follow the so-called universal identity rules (Salinas-Giegé et al. 2015). Additionally, the model presented here relies on the tRNA identity elements differing between cytoplasmic and mitochondrial pairs. The need for this distinction in eukaryotes performing mitochondrial protein synthesis has been elucidated in detail (Lovato et al. 2001) for alanyl-tRNA synthetase but is generally applicable.

Within the limited sample size of amitochondrial organisms, the evolutionary choice between loss and retention of the ancestral mitochondrial gene for arginyl-tRNA synthetase reflects the coevolution of arginyl-tRNA synthetase and tRNA identity elements. How this model responds to the additional complexity provided by the mitochondrial and plastid genomes and their corresponding aminoacyl-tRNA synthetase/tRNA pairs in aerobic non-metazoans remains to be established.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00239-021-10019-z.

Acknowledgements This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Aldinger CA, Leisinger A-K, Igloi GL (2012) The influence of identity elements on the aminoacylation of tRNA Arg by plant and

Escherichia coli arginyl-tRNA synthetases. FEBS J 279:3622–3638. https://doi.org/10.1111/j.1742-4658.2012.08722.x
Arakaki TL, Carter M, Napuli AJ et al (2010) The structure of tryptophanyl-tRNA synthetase from Giardia lamblia reveals divergence from eukaryotic homologs. J Struct Biol 171:238–243. https://doi.org/10.1016/j.jsb.2010.04.010
Bakatselou C, Beste D, Kadri AO et al (2003) Analysis of genes of mitochondrial origin in the genus Entamoeba. J Eukaryot Microbiol 50:210–214. https://doi.org/10.1111/j.1550-7408.2003.tb00119.x
Bonnefond L, Giegé R, Rudinger-Thirion J (2005) Evolution of the tRNATyr/TyrRS aminoacylation systems. Biochimie 87:873–883. https://doi.org/10.1016/j.biochi.2005.03.008
Brindefalk B, Viklund J, Larsson D et al (2007) Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases. Mol Biol Evol 24:743–756. https://doi.org/10.1093/molbev/msl202
Brown JR, Doolittle WF (1995) Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. Proc Natl Acad Sci USA 92:2441–2445. https://doi.org/10.1073/pnas.92.7.2441
Bunjun S, Stathopoulos C, Graham D et al (2000) A dual-specificity aminoacyl-tRNA synthetase in the deep-rooted eukaryote Giardia lamblia. Proc Natl Acad Sci U S A 97:12997–13002. https://doi.org/10.1073/pnas.230444397
Burki F, Corradi N, Sierra R et al (2013) Phylogenomics of the intracellular parasite mikrocytos mackini reveals evidence for a mitosome in rhizaria. Curr Biol 23:1541–1547. https://doi.org/10.1016/j.cub.2013.06.033
Cela M, Paulus C, Santos MAS et al (2018) Plasmodium apicoplast tyrosyl-tRNA synthetase recognizes an unusual, simplified identity set in cognate tRNATyr. PLoS ONE 13:e0209805. https://doi.org/10.1371/journal.pone.0209805
Chan PP, Lowe TM (2019) tRNAscan-SE: searching for tRNA genes in genomic sequences. In: Methods in molecular biology. Humana Press Inc., pp 1–14
Chang C-PP, Tseng Y-KK, Ko C-YY, Wang C-CC (2012) Alanyl-tRNA synthetase genes of Vanderwaltozyma polyspora arose from duplication of a dual-functional predecessor of mitochondrial origin. Nucleic Acids Res 40:314–322. https://doi.org/10.1093/nar/gkr724
Chihade JW, Brown JR, Schimmel PR, de Pouplana LR (2000) Origin of mitochondria in relation to evolutionary history of eukaryotic alanyl-tRNA synthetase. Proc Natl Acad Sci U S A 97:12153–12157. https://doi.org/10.1073/pnas.220388797
Clark CG, Roger AJ (1995) Direct evidence for secondary loss of mitochondria in Entamoeba histolytica. Proc Natl Acad Sci 92:6518–6521. https://doi.org/10.1073/pnas.92.14.6518
Doolittle RF, Handy J (1998) Evolutionary anomalies among the aminoacyl-tRNA synthetases. Curr Opin Genet Dev 8:630–636. https://doi.org/10.1016/S0959-437X(98)80030-0
Furukawa R, Nakagawa M, Kuroyanagi T et al (2017) Quest for ancestors of eukaryal cells based on phylogenetic analyses of aminoacyl-tRNA synthetases. J Mol Evol 84:51–66. https://doi.org/10.1007/s00239-016-9768-2
Geslain R, Bey G, Cavarelli J, Eriani G (2003a) Limited set of amino acid residues in a class Ia aminoacyl-tRNA synthetase is crucial for tRNA binding. Biochemistry 42:15092–15101. https://doi.org/10.1021/bi035581u
Geslain R, Martin F, Camasses A, Eriani G (2003b) A yeast knockout strain to discriminate between active and inactive tRNA molecules. Nucl Acids Res 31:4729–4737. https://doi.org/10.1093/nar/gkg685
Gill EE, Diaz-Triviño S, Barberà MJ et al (2007) Novel mitochondrion-related organelles in the anaerobic amoeba Mastigamoeba balamuthi. Mol Microbiol 66:1306–1320. https://doi.org/10.1111/j.1365-2958.2007.05979.x
Gray MW, Lang BF, Cedergren R et al (1998) Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res 26:865–878. https://doi.org/10.1093/nar/26.4.865
Guigou L, Mirande M (2005) Determinants in tRNA for activation of arginyl-tRNA synthetase: evidence that tRNA flexibility is required for the induced-fit mechanism. Biochemistry 44:16540–16548. https://doi.org/10.1021/bi051575h
Hashimoto T, Sánchez LB, Shirakura T et al (1998) Secondary absence of mitochondria in Giardia lamblia and Trichomonas vaginalis revealed by valyl-tRNA synthetase phylogeny. Proc Natl Acad Sci U S A 95:6860–6865. https://doi.org/10.1073/pnas.95.12.6860
Hikosaka K, Kita K, Tanabe K (2013) Diversity of mitochondrial genome structure in the phylum Apicomplexa. Mol Biochem Parasitol 188:26–33. https://doi.org/10.1016/j.molbiopara.2013.02.006
Igloi GL (2019) Compilation and alignment of eukaryotic arginyl-tRNA synthetases 2018. In: Mendeley Data. https://data.mendeley.com/datasets/ts4jbw9nft/1. Accessed 3 Sep 2019
Igloi GL (2020a) Molecular evidence for the evolution of the eukaryotic mitochondrial arginyl-tRNA synthetase from the prokaryotic sub-order cystobacterineae. FEBS Lett 594:951–957. https://doi.org/10.1002/1873-3468.13665
Igloi GL (2020b) Gene organization and phylum-specific attributes of eukaryotic arginyl-tRNA synthetases. Gene Reports 20:100778. https://doi.org/10.1016/j.genrep.2020.100778
Igloi GL, Leisinger A-K (2014) Identity elements for the aminoacylation of metazoan mitochondrial tRNA Arg have been widely conserved throughout evolution and ensure the fidelity of the AGR codon reassignment. RNA Biol 11:1313–1323. https://doi.org/10.1080/15476286.2014.996094
John U, Lu Y, Wohlrab S et al (2019) An aerobic eukaryotic parasite with functional mitochondria that likely lacks a mitochondrial genome. Sci Adv. https://doi.org/10.1126/sciadv.aav1110
Karlberg O, Canbäck B, Kurland CG, Andersson SGE (2000) The dual origin of the yeast mitochondrial proteome. Yeast 1:170–187. https://doi.org/10.1002/1097-0061(20000930)17:3%3c170::AID-YEA25%3e3.0.CO;2-V
Karnkowska A, Vacek V, Zubáčová Z et al (2016) A eukaryote without a mitochondrial organelle. Curr Biol 26:1274–1284. https://doi.org/10.1016/j.cub.2016.03.053
Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780. https://doi.org/10.1093/molbev/mst010
Keeling PJ, Doolittle WF (1997) Widespread and ancient distribution of a noncanonical genetic code in diplomonads. Mol Biol Evol 14:895–901. https://doi.org/10.1093/oxfordjournals.molbev.a025832
Keeling PJ, Leander BS (2003) Characterisation of a Non-canonical genetic code in the oxymonad Streblomastix strix. J Mol Biol 326:1337–1349. https://doi.org/10.1016/S0022-2836(03)00057-3
Krahn N, Fischer JT, Söll D (2020) Naturally occurring tRNAs with non-canonical structures. Front Microbiol. https://doi.org/10.3389/fmicb.2020.596914s
Kumazawa Y, Himeno H, Miura K, Watanabe K (1991) Unilateral aminoacylation specificity between bovine mitochondria and eubacteria. J Biochem 109:421–427
Lindmark DG, Muller M (1973) Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate Tritrichomonas foetus, and its role in pyruvate metabolism. J Biol Chem 248:7724–7728. https://doi.org/10.1016/s0021-9258(19)43249-3
Liu W, Huang Y, Eriani G et al (1999) A single base substitution in the variable pocket of yeast tRNA(Arg) eliminates species-specific aminoacylation. BBA 1473:356–362

Loftfield RB, Eigner EA, Nobel J (1968) Inter-phylogenetic specificity in the bonding of amino acids to tRNA. Biol Bull 135:181–192. https://doi.org/10.2307/1539625

Lovato MA, Chihade JW, Schimmel P (2001) Translocation within the acceptor helix of a major tRNA identity determinant. EMBO J 20:4846–4853. https://doi.org/10.1093/emboj/20.17.4846

Lue SW, Kelley SO (2005) An aminoacyl-tRNA synthetase with a defunct editing site. Biochemistry 44:3010–3016. https://doi.org/10.1021/bi047901v

Luque I, Riera-Alberola ML, Andújar A, Ochoa De Alda JAG (2008) Intraphylum diversity and complex evolution of cyanobacterial aminoacyl-tRNA synthetases. Mol Biol Evol 25:2369–2389. https://doi.org/10.1093/molbev/msn197

Makiuchi T, Nozaki T (2014) Highly divergent mitochondrion-related organelles in anaerobic parasitic protozoa. Biochimie 100:3–17. https://doi.org/10.1016/j.biochi.2013.11.018

McClain WH (1993) Rules that govern tRNA identity in protein synthesis. J Mol Biol 234:257–280. https://doi.org/10.1006/jmbi.1993.1582

Melnikov SV, Rivera KD, Ostapenko D et al (2018) Error-prone protein synthesis in parasites with the smallest eukaryotic genome. Proc Natl Acad Sci 115:E6245–E6253. https://doi.org/10.1073/pnas.1803208115

Mielewczik M, Mehlhorn H, Al-Quraishy S et al (2008) Transmission electron microscopic studies of stages of Histomonas meleagridis from clonal cultures. Parasitol Res 103:745–750. https://doi.org/10.1007/s00436-008-1009-1

Nameki N, Asahara H, Tamura K et al (1995) Similarities and differences in tRNA identity between Escherichia coli and Saccharomyces cerevisiae: evolutionary conservation and divergence. Nucleic Acids Symp Ser 205–206

Nicholas KB, Nicholas HBJ (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by author

Paul D, Ansari AH, Lal M, Mukhopadhyay A (2020) Human brain shows recurrent non-canonical microRNA editing events enriched for seed sequence with possible functional consequence. Non-Coding RNA 6:21. https://doi.org/10.3390/ncrna6020021

Roy H, Ling J, Alfonzo J, Ibba M (2005) Loss of editing activity during the evolution of mitochondrial phenylalanyl-tRNA synthetase. J Biol Chem 280:38186–38192. https://doi.org/10.1074/jbc.M508281200

Salinas-Giegé T, Giegé R, Giegé P (2015) TRNA biology in mitochondria. Int J Mol Sci 16:4518–4559. https://doi.org/10.3390/ijms16034518

Schulman L, Pelka H (1989) The anticodon contains a major element of the identity of arginine transfer RNAs. Science 246(4937):1595–1597. https://doi.org/10.1126/science.2688091

Sekine SI, Shimada A, Nureki O et al (2001) Crucial role of the high-loop lysine for the catalytic activity of arginyl-tRNA synthetase. J Biol Chem 276:3723–3726. https://doi.org/10.1074/jbc.C000756200

Shimada A, Nureki O, Goto M et al (2001a) Structural and mutational studies of the recognition of the arginine tRNA-specific major identity element, A20, by arginyl-tRNA synthetase. Proc Natl Acad Sci USA 98:13537–13542. https://doi.org/10.1073/pnas.231267998

Shimada N, Suzuki T, Watanabe K (2001b) Dual Mode Recognition of Two Isoacceptor tRNAs by Mammalian Mitochondrial Seryl-tRNA Synthetase. J Biol Chem 276:46770–46778. https://doi.org/10.1074/jbc.M105150200

Sissler M, Giegé R, Florentz C (1996) Arginine aminoacylation identity is context-dependent and ensured by alternate recognition sets in the anticodon loop of accepting tRNA transcripts. EMBO J 15:5069–5076

Smith HC, Gott JM, Hanson MR (1997) A guide to RNA editing. RNA 3:1105–1123

Solovyev V (2008) Statistical Approaches in Eukaryotic Gene Prediction. Handbook of Statistical Genetics, 3rd edn. John Wiley & Sons Ltd, UK, pp 97–159

Stephen P, Ye S, Zhou M et al (2018) Structure of Escherichia coli Arginyl-tRNA Synthetase in Complex with tRNAArg: pivotal role of the D-loop. J Mol Biol 430:1590–1606. https://doi.org/10.1016/j.jmb.2018.04.011

Tamura K, Himeno H, Asahara H et al (1992) In vitro study of E.coli tRNA(Arg) and tRNA(Lys) identity elements. Nucl Acids Res 20:2335–2339

Toso MA, Omoto CK (2007) Gregarina niphandrodes may lack both a plastid genome and organelle. J Eukaryot Microbiol 54:66–72. https://doi.org/10.1111/j.1550-7408.2006.00229.x

Tovar J, Fischer A, Clark CG (1999) The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite Entamoeba histolytica. Mol Microbiol 32:1013–1021. https://doi.org/10.1046/j.1365-2958.1999.01414.x

Van Der Giezen M, Sjollema KA, Artz RRE et al (1997) Hydrogenosomes in the anaerobic fungus Neocallimastix frontalis have a double membrane but lack an associated organelle genome. FEBS Lett 408:147–150. https://doi.org/10.1016/S0014-5793(97)00409-2

Watanabe K (2010) Unique features of animal mitochondrial translation systems. The non-universal genetic code, unusual features of the translational apparatus and their relevance to human mitochondrial diseases. Proc Jpn Acad Ser B Phys Biol Sci 86:11–39

Wolf YI, Aravind L, Grishin NV, Koonin EV (1999) Evolution of aminoacyl-tRNA synthetases–analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. Genome Res 9:689–710. https://doi.org/10.1101/gr.9.8.689

Yahalomi D, Atkinson SD, Neuhof M et al (2020) A cnidarian parasite of salmon (Myxozoa: Henneguya) lacks a mitochondrial genome. Proc Natl Acad Sci U S A 117:5358–5363. https://doi.org/10.1073/pnas.1909907117