Possible functionality of start and stop codons present at specific and conserved positions in animal mitochondrial genes specifying tRNA

Eric Faure, Roxane-Marie Barthelemy

International Journal of Zoology Studies
ISSN: 2455-7269; Impact Factor: RJIF 5.14
Received: 01-03-2019; Accepted: 06-04-2019
www.zoologyjournals.com
Volume 4; Issue 3; May 2019; Page No. 23-29

Possible functionality of start and stop codons present at specific and conserved positions in animal mitochondrial genes specifying tRNA

Eric Faure[1*], Roxane-Marie Barthélémy[2]
[1-2] Aix Marseille University, CNRS, Centrale Marseille, I2M, UMR 7373, 13453 Marseille, France

Abstract
Several metazoan mitochondrial-*trn* genes (specifying tRNA) exhibit nucleotide triplets at specific and conserved positions corresponding to stop codons (TAG/TAA) and/or start codons (ATG/ATA). The products of these genes, which bear one or two types of these codons are called ss-tRNAs (for stop/start). In this article, several examples of ss-*trn* genes bearing start or stop codons in metazoans (protostomes and deuterostomes) strongly suggest that these codons are functional. Moreover, the ss-*trn* genes found in apicoplast genomes likely have a mitochondrial origin, but the TAR and ATR triplets (R for purines) found in their sequences seem to play only a structural role. Future investigations should attempt to clarify the mechanisms of the transcription and maturation of ss-*trn* genes with their overlapping protein genes and the possible regulatory roles of stop and start codons found in *trn* sequences.

Keywords: mitochondrion, ss-tRNA, start codon, stop codon, overlapping genes

1. Introduction

Numerous trn genes (specifying tRNAs) bearing nucleotide (nt) triplets corresponding to stop or start codons at precise and conserved positions have been found in mitochondrial (mt) DNA of fungi and animals [1]. The products of these genes, which bear one or two types of these codons, are named ss-tRNAs (for stop/start) [1]. Most of the studies focus on DNA sequences; so, these codons are usually annotated TAR or ATR instead of UAR or AUR (R for purines). The last nucleotide of these triplets is the first involved in the 5'-D- or 5'-T-stem, so they are named TAR10 and ATR49, respectively (Fig. 1). A previous study identified TAR10 and ATR49 triplets in mt-trn genes of Platyhelminthes and of Hylidae (frogs), respectively; in these two taxa, these triplets play a crucial role during translation [1]. Moreover, another analysis showed that ss-trn genes are relatively frequently found in mtDNAs of the phylum chaetognatha and that TAR10 and ATR49 triplets could serve as the physiological initiation and stop codons in some cases [2]. In this present study, other examples in metazoan of ss-trn genes have been analysed, suggesting that trn genes overlapping with protein genes could play subtle roles in both transcription and translation.

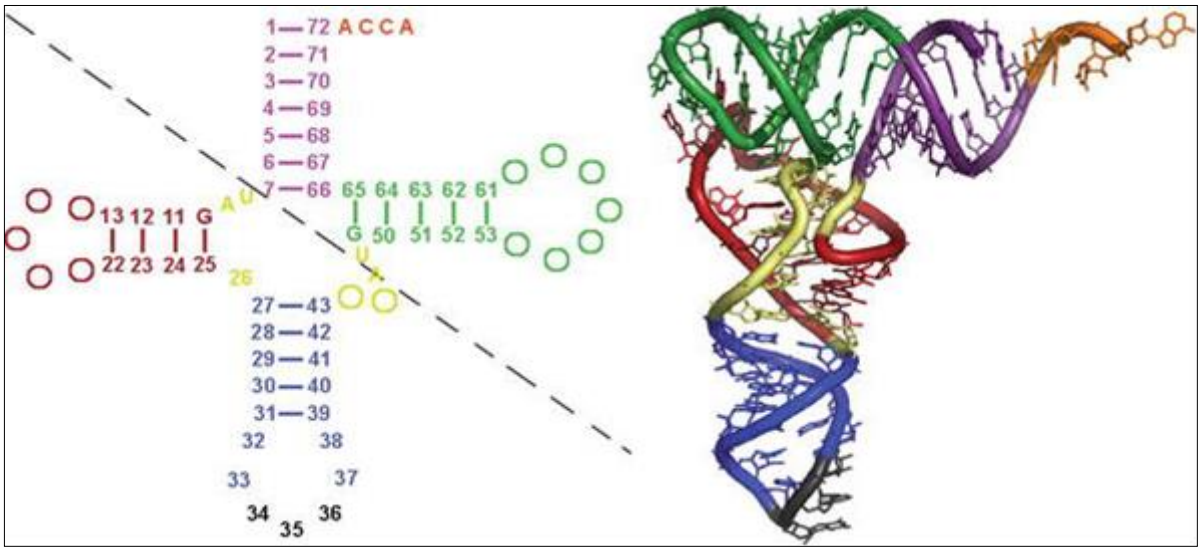

Fig 1

Fig 1: Typical cloverleaf secondary structure of a metazoan mt-ss-tRNA (left) with 3D image of an L-shaped tRNA (right). In 2D structure, the standard numbering was applied. The first two nucleotides of the variable region and those of the D-loops and T-loops were represented by circles. The diagonal dashed line indicates the approximate separation between the "top half" and the "cherry-bob"/"bottom half". Nucleotide types were given for UAG10 and AUG49

triplets, the discriminator base (preferentially an A), and the CCA tail at the 3′-end. Short lines connect nucleotides forming Watson-Crick pairing within stems. Colouring: acceptor-stem in purple, D-arm in red, anticodon-arm in blue with the anticodon in black, T-arm in green, and CCA tail in orange. The yellow segments represented respectively in descending order of size, the variable region (connector 2), the connector 1 and the nt 26. Figure adapted from Faure and Barthélémy [1] and Fig. 1B reproduced with the kind permission of Prof. N.R. Voss (Roosevelt University, Ill.).

2. Materials and Methods

Most of the research was performed in GenBank using the following keywords: "TAA stop codon is completed by the addition of 3' A residues to the mRNA", "alternative start codon" or "start codon not determined" and mitochondrion (or mitochondrial DNA) complete genome. Then, whether upstream (for start codon) or downstream (for stop codon) of the protein-encoding gene was a trn gene was assessed. When a trn gene was found, TAR10 or ATR49 triplets were searched, and the same investigation was then performed in conspecific mt-DNAs. Some tRNAs were detected using tRNAscan-Se 1.21 [3] or BLASTn searches. Multiple sequence alignments were made using crustal software [4], whereas secondary structures were predicted using tRNAscan-Se [3].

3. Results

3.1. ss-tRNAs in mt genomes of metazoan taxa

Following the strategy outlined in the Material and Methods, TAR10 or ATR49 triplets have been only found in mt-DNAs belonging to Metazoa. This finding could be explained by the fact that overlapping mt-trn genes have long been known in this taxon [1].

Regarding the putative use of the TAR8-10 triplet as a stop codon, all complete mt-genome sequences of two taxa were analysed (echinoderms and a superfamily of gastropoda). The first are deuterostomes, whereas the others are protostomia and they use different mt-genetic codes (n°9 and n°5, respectively) [5], but only TAG and TAA are used as stop codons, which avoids possible bias due to use of other types of termination codon.

Extant echinoderms (marine invertebrates) include two clades; the superclass Eleutherozoa [including the Asterozoa clade, which encompasses the classes Asteroidea (starfish or sea stars) and Ophiuroidea (brittle stars and basket stars), and the Echinozoa clade, which encompasses two classes Echinoidea (sea urchins) and Holothuroidea (sea cucumbers)] and class Crinoidea (feather stars and sea lilies). However, the phylogenetic position of the various clades within this phylum is still debated (reviewed in [6]). In this study, all the complete mt-genomes of the phylum Echinodermata were analysed (Table 1). The contiguous region, including the genes encoding or specifying ND4 and tRNA-His, respectively, is particularly interesting for this study. Indeed, in all the cases, the TAG (most often) or TAA (more rarely) triplets are present at the expected position. In the subphylum Eleutherozoa, the TAG triplet is the first potential stop codon of the upstream protein gene in all but one cases; indeed, in the Holothuria forskali trnH gene, a triplet TAG is present at the position 5 to 7, which directly precedes the TAG10 as the so called "tandem stop codon". In several cases, the authors who reported the sequences, presupposing that there could be no overlap between the two genes, considered that a TAA stop codon is completed by the addition of 3'-A residues to the mRNA; thus, the sequence of the nd4 gene would stop at the first T or TA immediately before the trn gene. In all but one cases the Asterozoa, the stop codon of the nd4 gene is upstream of the trn sequence. In addition, the distance between the stop codon and the 5'-end of the trn gene is very variable, and the stop codon and TAR10 triplets are in frame. In the sequence of an ophiuroid (Astrospartus mediterraneus), the TAG10 triplet is the first in frame stop codon, whereas the stop codon of the nd4 gene is always upstream of the trn sequence in the class Crinoidea. Moreover, in three out of four cases, the triplet at position 8-10 is TAA.

As it is hypothesized that Hemichordata have close relationship to the Echinodermata [7], the mt-region at the 3'-end of the nd4 gene has been studied in the first taxon. Four complete mt-genomes of Hemichordata are available in the GenBank database. However, the gene specifying tRNA-Ala follows those encoding ND4 and the stop codon is upstream of the trn sequence in Rhabdopleura compacta (NC_015649). In the three other mt genomes (Fig. 2), the nd4/trnH gene order is conserved. In the Saccoglossus kowalevskii mt-genome, the first stop codon is the triplet TAA (position 2-4), and this codon presents in the trn gene, which is in frame versus TAG10. In the two Balanoglossus sequences, TAR10 is the first in-frame stop codon. However, the low number of sequences does not allow us to conclude whether the stop codon role played by TAR10 of the trnH gene is an ancestral or a derived character state within echinoderms.

Table 1: Position of the first complete stop codon of the nad4 gene versus the following gene specifying tRNA-His in echinoderms. * sequences for which it had been considered that the TAA stop codon is completed by the addition of 3' A residues to the mRNA and the incomplete stop codon is upstream the tRNA sequence. § The first in-frame TAG triplet is in the tRNA sequence but upstream of TAG10. The tRNA of Luidia quinaria (NC_006664) was not been found by the authors who reported the sequence.

Table 1

Sub-phylum	Super-class	Class	Names of the species with accession numbers	TAR10 triplet	TAR10 as stop codon	Stop codon upstream TAR10
Eleuthe-rozoa	Echi-nozoa	Echinoidea	*Arbacia lixula* (NC_001770), *Diadema setosum* (NC_033522*), *Echinocardium cordatum* (NC_013881), *Echinometra mathaei* (NC_034767), *Echinothrix diadema* (NC_033523*), *Glyptocidaris crenularis* (NC_032365), *Heterocentrotus mammillatus* (NC_034768), *Loxechinus albus* (JX888466, NC_023770), *Mespilia globulus* (NC_034769), *Paracentrotus lividus* (NC_001572), *Salmacis bicolor rarispina* (KU302104*), *S. sphaeroides* (NC_033528*), *Sterechinus neumayeri* (NC_027063, KJ680295), *Strongylocentrotus droebachiensis*	TAG	X	

			(NC_009940, EU054306), *S. intermedius* (NC_023772), *S. pallidus* (NC_009941), *S. purpuratus* (NC_001453), *Temnopleurus hardwickii* (NC_026200*), *T. reevesii* (NC_033530*), *T. toreumaticus* (NC_033529*), *Tripneustes gratilla* (NC_034770, KY268294)			
		Holothu-roidea	*Apostichopus japonicus* (NC_012616, KP170616, KP170617, KP170618, GU557147, GU557148, FJ594963*, FJ594967*, FJ594968*, AB525437, AB525760, AB525761, FJ986223, EU294194), *Cucumaria miniata* (NC_005929*), *Holothuria scabra* (NC_027086*), *Parastichopus californicus* (NC_026727), *Parastichopus nigripunctatus* (NC_013432), *P. parvimensis* (NC_029699*), *Stichopus horrens* (NC_014454), *S.* sp. SF-2010 (NC_014452)	TAG	X	
			Holothuria forskali (NC_013884)§	TAG		TAG
	Aste-rozoa	Asteroidea	*Acanthaster planci* (NC_007788), *Asterias amurensis* (NC_006665), *Astropecten polyacanthus* (NC_006666), *Luidia quinaria* (NC_006664), *Patiria pectinifera* (NC_001627)	TAG		TAA
			Acanthaster brevispinus (NC_007789)	TAG		TAG
		Ophiuroidea	*Astrospartus mediterraneus* (NC_013878)	TAG	X	
			Ophiura lutkeni (NC_005930), *Ophiocomina nigra* (NC_013874), *Amphipholis squamata* (NC_013876), *Ophiopholis aculeata* (NC_005334)	TAG		TAA
			Ophiura albida (NC_010691)	TAG		TAG
Pelmatozoa		Crinoidea	*Antedon mediterranea* (NC_010692), *Phanogenia gracilis* (NC_007690)	TAA		TAA
			Neogymnocrinus richeri (NC_007689)	TAA		TAG
			Florometra serratissima (NC_001878)	TAG		TAA

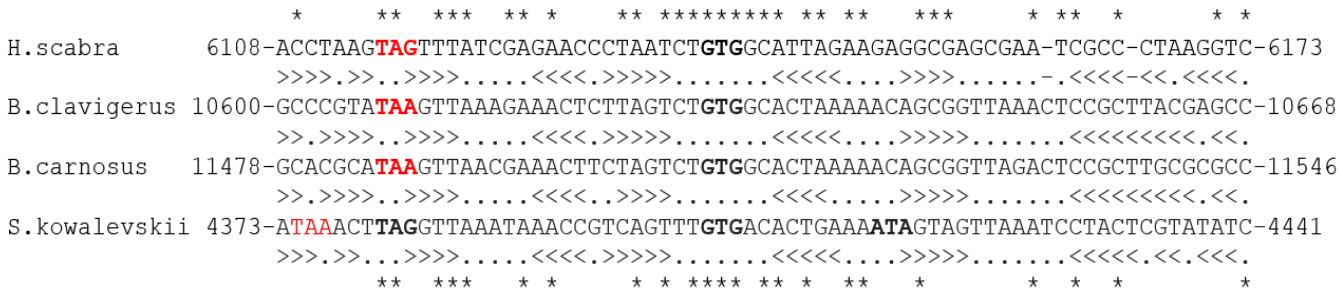

Fig 2

Fig 2: Alignment of primary and 2D sequences of the mt-trnH gene of Holothuria scabra (Echinodermata) (NC_027086) with the corresponding gene of three Hemichordata species: Balanoglossus clavigerus (NC_013877), Balanoglossus carnosus (NC_001887) and Saccoglossus kowalevskii (NC_007438). TAR10, anticodon and ATA triplets are indicated in boldface. The supposed stop codons of the nad4 gene are in noted red letters. Top and bottom asterisks (*) indicate only conserved nucleotides for H. scabra, B. clavigerus and B. carnosus sequences, as well as the four sequences.

All the complete mt-genomes of the superfamily Helicidae (Gastropoda) and a partial mt-genome (Euhadra herklotsi Z71696) were analysed (Table 2). According to the families, the trn gene that follows the cox2 gene is different. In Bradybaenidae, this gene specifies tRNA-Gly, whereas this gene specifies tRNA-Tyr in Geomitridae and Helicidae. Changes in trn gene order within Helicidae are well known [8]. The triplet TAG10 of the trn gene (trnG in one case and trnY in the other) that directly follows the cox2 gene is the first complete triplet that corresponds to a stop codon in only two out to 12 sequences. Moreover, in two cases, the first putative stop codon is located in the trn genes that are downstream of the trnG gene that directly follows cox2. For Aegista aubryana (NC_029419), this sequence is the trnH gene that directly follows the gene specifying tRNA-Gly (Table 2). In contrast, in the sequence of Aegista diversifamilia (NC_027584), the first putative stop codon is in the third trn gene, which specifies tRNA-Tyr. Some mt-trn genes of Pulmonata, an informal group of land snails and slugs, show mismatches in the acceptor stems predicted from their gene sequences, and the majority of these mismatches fall in regions where the trn genes overlap with adjacent downstream genes transcribe in the same direction [9]. The majority of mismatch cases in acceptor stems are replaced by Watson-Crick base pairs as a result of RNA editing [9]. This tRNA editing phenomenon appears to be a general feature of pulmonate gastropod mitochondria [10] but it concerns only the 3'-part of the acceptor stems after nucleotide 67 in the standard tRNA sequence and the discriminator bases. These findings do not implicate TAR10 and ATR49 in any way.

Table 2: Position of the first complete stop codon of the cox2 gene versus the following genes in the superfamily Helicidae (Gastropoda). * sequences for which it had been considered that the TAA stop codon is completed by the addition of 3' A residues to the mRNA and the incomplete stop codon is upstream the tRNA sequence. The symbols § or $ signify that the order of the genes is cox2-trnG-trnH-trnY and that the first triplet that can correspond to a stop codon is in the third or fourth genes, respectively.

Table 2

Family	Acc. n.	Type of amino acid carried by the tRNA	Type of triplet TAR10 triplet	TAR10 as putative stop codon	Putative stop codon upstream the *trn* gene	Putative stop codon in the *trn* gene but downstream TAR10	Putative stop codon downstream the *trn* gene
Bradybaenidae	*Euhadra herklotsi* (Z71696)	Gly	TAG	X			
	Dolicheulota formosensis (NC_027493), *Mastigeulota kiangsinensis* (NC_024935)	Gly	TAG		TAG		
	Aegista aubryana (NC_029419*)	Gly	TAG				TAA§
	Aegista diversifamilia (NC_027584*)	Gly	TAG				TAA$
Geomitridae	*Cernuella virgata* (NC_030723*)	Tyr	none			TAA	
	Helicella itala (KT696546*)	Tyr	TAG			TAG	
Helicidae	*Cepaea nemoralis* (NC_001816)	Tyr	TAG	X			
	Helix aspersa (JQ417195, JQ417196)	Tyr	TAG			TAA	
	Cylindrus obtusus (NC_017872), *Helix aspersa* (NC_021747*)	Tyr	TAG			TAG	

Concerning the putative start codon ATR49, taxa with "start codon not determined" for complete mt genomes in GenBank were analysed. Their number is however relatively low. Moreover, in some cases, the upstream gene encoded a protein or specified a rRNA and/or there was only one mention for a given taxa. Nevertheless, a significant example was identified in complete mt-DNAs within the protostomians (Ecdysozoa, insects).

In Auchenorrhyncha, a suborder of hemipteran insects, the cox2 gene is followed by those specifying tRNA-Leu2. In the latter gene at positions 47-49, the triplet ATR is always present and is mainly ATA (Table 3). Out of 50 sequences, the ATA49 triplet is the first putative in-frame start codon in 23 cases. If the ATR49 triplets are not considered, the first codon that could potentially allow the initiation of translation is positioned immediately after the 3' terminus of the trn gene in all but three cases. In addition, frequently, these putative start codons have been proposed by the authors who reported the sequences. In 22 cases, these codons correspond to an alternative start site even if some of them are known to be rare; however, analyses of protein sequences suggest that many of them might be used even more when the ATR49 is not in frame. The positioning of most of the putative start codons other than ATR49 would be in favour of a classical tRNA punctuation [11].

Table 3: Position of the first putative start codon of the cox2 gene versus the upstream gene specifying tRNA-Leu2 (TAA) in Auchenorrhyncha, a suborder of insects. In the last column *, § and $ indicate that according to the authors who reported the sequences of the putative start codon are respectively 1 nt and 9 nts downstream of the trn gene or overlaps by 1 nt with the latter.

Table 3

Superfamily	Family	Species and accession numbers.	Type of ATR49 triplet	ATR49 as first putative start codon	Putative start codon immediately after the *trn* gene unless otherwise noted
Cercopoidea	Aphrophoridae	*Philaenus spumarius* (NC_005944)	ATA		TTG
	Cercopidae	*Abidama producta* (NC_015799), *Aeneolamia contigua* (NC_025495), *Callitettix braconoides* (NC_025497), *C.* sp. (KY039124), *C. versicolor* (NC_020031), *Cosmoscarta bispecularis* (KP064511), *C.* sp. (MF621236)	ATA	X	ATG
		Callitettix biformis (NC_025496)	ATA		ATG
		Paphnutius ruficeps (NC_021100)	ATG		ATT
Fulgoroidea	Delphacidae	*Sogatella furcifera* (KC512915, NC_021417)	ATG	X	ATT
		Laodelphax striatella (JX880068), *L. striatellus* (NC_013706), *Nilaparvata bakeri* (NC_033388), *N. lugens* (JN563995, JN563996, JN563997, KC333653, KC333654, NC_021748), *N. muiri* (NC_024627)	ATG		ATT
	Flatidae	*Geisha distinctissima* (NC_012617)	ATA	X	ATA*
	Fulgoridae	*Lycorma delicatula* (EU909203, NC_012835)	ATA		ATA
		Laternaria candelaria (NC_019576)	ATA		ATG§
	Issinae	*Sivaloka damnosus* (NC_014286)	ATA		ATA
	Ricaniidae	*Ricania marginalis* (NC_019597), *R. speculum*	ATA	X	ATT

		(NC_031369)			
Membracoidea	Aetalionidae	*Darthula hardwickii* (NC_026699)	ATA	X	ATT
	Cicadellidae	*Durgades nigropicta* (NC_035684), *Macrosteles quadrilineatus* (NC_034781), *Nephotettix cincticeps* (NC_026977), *Phlogotettix* sp. (KY039135), *Tambocerus* sp. (KT827824), *Yanocephalus yanonis* (NC_036131)	ATA		ATA
		Idioscopus nitidulus (NC_029203), *Maiestas dorsalis* (NC_036296)	ATA	X	ATC
		Homalodisca vitripennis (NC_006899), *Japanagallia spinosa* (NC_035685), *Taharana fasciana* (NC_036015), *Typhlocyba* sp. (KY039138)	ATA	X	ATT
		Empoasca vitis (NC_024838)	ATA		GTG
		Bothrogonia ferruginea (KU167550)	ATA		ATA
		Japananus hyalinus (NC_036298)	ATA		ATG
		Drabescoides nuchalis (NC_028154)	ATA		ATA[$]
	Membracidae	*Leptobelus gazella* (NC_023219), *L.* sp. (JQ910984), *Tricentrus* sp. (KY039118)	ATA	X	ATA
		Entylia carinata (NC_033539)	ATA	X	ATT

In all the studied taxa, Blast analyses of the NCBI ESTs and SRA (Sequence Read Archive) databases were performed but none of the sequences are supportive of the proposed hypotheses. Thus, none of the sequences correspond to a transcript starting at ATR49 or terminating at TAR10 could be found. However, for each taxon, the percentage of transcripts of mitochondrial origin is relatively low even for Helicidae, and none of the transcripts have been annotated as corresponding to a protein gene of mt origin. Moreover, when the transcripts are annotated the number of fully matured transcripts is even lower.

3.2 The particular case of the plastid tRNAs of Apicomplexa

The cytoplasm of several Apicomplexa (a large phylum of parasitic protozoans) contains organelles, which are remnant (relict) chloroplasts called apicoplasts that are no longer able to perform photosynthesis. This vestigial plastid homologous with chloroplasts of algae suggests that Apicomplexa had a photosynthetic ancestry and were converted into parasitism early in the evolution of animals, perhaps for more than 500 million years [12]. The plastid genomes generally contain a great number of trn genes, i.e., the apicoplast genome of Plasmodium falciparum bears 33 trn genes, and 27 of them are grouped in four clusters [13]. All the 26 trn genes found in two sequences deposed in GenBank (Acc. n°X95275 and X95276 [14]) have been analysed (Table 4). The various types of trn genes can be differentiated according to the length of their variable regions (V-R). When the latter consists of 4 nts, which corresponds to the majority of the trn genes, 71% of the genes exhibit both TAR10 and ATR49 triplets. All other trn genes, i.e., trn genes with a V-R size greater than 4 nts, never harbour both types of triplets, and it has been previously shown that a number 4 nts in the V-R is a characteristic of most of the mt-ss-trn genes of fungi/metazoa [1]. It must also be noted that the apicomplexan mtDNAs with a size of ~6 kb are among the smallest known mtDNA genomes and do not apparently encode any tRNAs [15]. There is also a tendency for overall genome reduction associated with trn gene loss in parasitic species [13].

Alignments of apicomplexan trn-genes found in apicoplasts could suggest a mt origin (Fig. 3) especially given that DNA flux between the mitochondrion and the chloroplast has been reported; nevertheless, horizontal acquisition of mt-trn cannot be excluded. It was previously demonstrated that P. falciparum tRNAs share unique features; indeed, they are characterized by cytosolic-type secondary and tertiary structural features but also by mt-type nt content [16]. However, this combination might not be unique in the genomes of these plastids and deserves further investigation. In Plasmodium, polycistronic transcripts are cleaved by excision of tRNA sequences (tRNA punctuation processing); however, it is unclear whether this RNA processing is an ancestral or derived characteristic [17].

Table 4: Characteristics of trn genes of P. falciparum apicoplasts

V-R size	Numbers of *trn* genes	ss-*trn* genes with both TAR10 and ATR49	ss-*trn* genes with only TAR10	ss-*trn* genes with only ATR49	*trn* genes without both TAR10 and ATR49
4	14	10	3	1	0
5	5	0	5	0	0
6	2	0	2	0	0
>11	5	0	3	0	2

Fig 3

Fig 3: Alignments of 1D sequences and 2D structures of trn genes of the apicoplast of Plasmodium with closer trn sequences from other taxa. (A), alignment of the trnG gene of the apicoplast of Plasmodium falciparum (LN999985) with the mt-trnA of the fungi Magnusiomyces ingens (M.i. KJ459950) and trnG of an endosymbiont of the insect Pachyrhynchus infernalis (End. AP018160). Top and bottom asterisks (*) indicate only conserved nucleotides for P. falciparum and M. ingens sequences, respectively, and for the three sequences. (B), alignment of the trnQ gene of the apicoplast of Plasmodium vivax (P.v. LT635626) with the mt-trnQ of the Amoebozoa Balamuthia mandrillaris (B.m. KT030673) and the mt-trnQ of the Rhodophyta Hildenbrandia rubra (H.r. KF649304). Top and bottom asterisks (*) indicate only conserved nucleotides for P. vivax and B. mandrillaris sequences, respectively, for the three sequences. TAG10, anticodon and ATA49 are noted in bold.

4. Discussion

New examples of TAR10 or ATR49 potentially used as stop or start codons, respectively, have been reported in this article. To determine the position of the start and stop codons, authors who reported the sequences have made, among other things, alignments and Blast analyses, and we have checked the sequences using the same tools. Although tRNAs were not sequenced directly, the large number of sequences analysed for echinoderms and Auchenorrhyncha (a suborder of insects) strongly suggest that TAR10 or ATR49 can be used as codons. Concerning echinoderms, TAR10 would be functional in terms of translation almost exclusively in members of two classes (Echinoidea and Holothuroidea) of the superclass Echinozoa (28 out to 29 species). In the other classes or superclasses, only one possibly functional TAR10 has been found. The relationships between TAR10 as codons and phylogeny give credit to our hypothesis (in the opposite case, the distribution of stop codons would be more random). Analysis of sequences from the superfamily Helicidae (Gastropoda) shows that despite high levels of the TAR10 triplet, these three nts may rarely be used as a stop codon in a mt-trn gene (2 out to 12). The analysis of the use of the ATR49 triplet as a start codon at the level of a mt-trn gene of species belonging to a suborder Auchenorrhyncha shows that these three nts would be used as an initiation codon in approximately half the cases; however, the large number of possible alternative start codons creates many other opportunities for translation initiation.

Otherwise, the analysis of the trn genes of the apicoplast of P. falciparum strongly suggests that most of them have a mt-origin. However, as most of the trn genes are grouped in clusters, TAR10 or ATR49 cannot be used as stop or start codons, respectively. The role of these triplets in translation in ancestral mtDNAs of Apicomplexa is unknown; nevertheless, in apicoplast tRNAs, these triplets must have only a key structural role. Otherwise, the situation of plant mt-tRNAs is very complex; plants contain few "native" tRNAs expressed from true mt-trn genes. They frequently possess "chloroplast-like" trn genes inserted into their mt-DNAs; moreover, they compensate the loss of mt-trn genes by importing several nucleus-encoded tRNAs [118]. This notion could explain why ss-trn genes with both TAR10 and ATR49 triplets are extremely rare in plant mt-DNAs. Contrarily, in Apicomplexa, the trn genes are generally absent in mt-DNAs but several mt-ss-trn genes are present in apicoplast DNAs. Nonetheless, it must be emphasized that the apicoplasts also keep both bacterial and plant-like properties as they provide a target for both Tetracycline [19] and herbicidal drugs [20].

As previously suggested for chaetognatha, the use of TAR10 and ATR49 as stop or start codons, respectively, in ss-trn genes seem to be an exaptation [2]. These triplets would have first been selected for their role in the spatial structuring of tRNAs, and the use as a codon would have appeared much later. Moreover, as previously demonstrated [1] and evidenced here in apicoplast ss-trn genes, there is a correlation between the size of variable region of the 4 nts and the presence of the ATR49 triplets. Moreover, in mt-genomes, trn genes can play an analogous role to those of introns in the evolution of nuclear genomes via exon shuffling [21], and trn sequences can allow the promotion of non-homologous rearrangements between protein genes [22].

5. Conclusion

To date, the mechanisms of maturation of transcripts bearing ss-tRNA are unknown and must be investigated in the future. The maturation of polycistronic transcripts bearing TAR10 or ATR49 triplets will be likely implicated in complex regulatory pathways. The large number of ss-trn genes in animal mitochondria makes their systematic search in the algorithms predicting tRNAs indispensable; this would avoid, among other things, many annotation errors [2].

6. References

1. Faure E, Barthélémy R. True mitochondrial tRNA punctuation and initiation using overlapping stop and start codons at specific and conserved positions. In Mitochondrial DNA; Seligmann H, IntechOpen, Croatia, 2018, 3-29.
2. Faure E, Barthélémy R. Specific mitochondrial ss-tRNAs in phylum Chaetognatha. Journal of Entomology and Zoology Studies, 2019, in press.
3. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Research. 1997; 25(5):955-964.
4. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research. 1994; 22(22):4673-4680.
5. Elzanowski A, Ostell J. The genetic codes compiled at National Center for Biotechnology Information (NCBI), Bethesda, Maryland, USA, 2019. (https://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi). [Accessed: 2019-04-17]
6. Janies DA, Voight JR, Daly M. Echinoderm phylogeny including Xyloplax, a progenetic asteroid. Systematic Biology. Systematic Biology. 2011; 60(4):420-438.
7. Perseke M, Fritzsch G, Ramsch K, Bernt M, Merkle D, Middendorf M. *et al.* Evolution of mitochondrial gene orders in echinoderms. Molecular phylogenetics and evolution. Molecular Phylogenetics and Evolution. 2008; 47(2):855-864.
8. Gaitán-Espitia JD, Nespolo RF, Opazo JC. The complete mitochondrial genome of the land snail Cornu aspersum (Helicidae: Mollusca): intra-specific divergence of protein-coding genes and phylogenetic considerations within Euthyneura. PLoS One. 2013; 8(6):e67299.
9. Yokobori SI, Paabo S. tRNA editing in metazoans. Nature. 1995; 377(6549):490.
10. Yamazaki N, Ueshima R, Terrett JA, Yokobori S, Kaifu M, Segawa R. *et al*. Evolution of pulmonate gastropod mitochondrial genomes: comparisons of gene organizations of Euhadra, Cepaea and Albinaria and implications of unusual tRNA secondary structures. Genetics. 1997; 145(3):749-758.
11. Ojala D, Merkel C, Gelfand R, Attardi G. The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. Cell. 1980; 22(2 Pt 2):393-403.
12. Kalanon M, McFadden GI. Malaria, Plasmodium falciparum and its apicoplast. Biochemical Society transactions. 2010; 38(3):775-82.
13. Preiser P, Williamson DH, Wilson RJ. tRNA genes transcribed from the plastid-like DNA of Plasmodium falciparum. Nucleic Acids Research. 1995; 23(21):4329-4336.
14. Wilson RJ, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A. *et al*. Complete gene map of the plastid-like DNA of the malaria parasite Plasmodium falciparum. Journal of Molecular Biology. 1996; 261(2):155-172.
15. Gray MW, Lang BF, Burger G. Mitochondria of protists. Annual review of genetics. 2004; 38:477-524.
16. Pütz J, Giegé R, Florentz C. Diversity and similarity in the tRNA world: overall view and case study on malaria-related tRNAs. FEBS Letters. 2010; 584(2):350-358.
17. Nisbet RE, Kurniawan DP, Bowers HD, Howe CJ. Transcripts in the Plasmodium apicoplast undergo cleavage at tRNAs and editing, and include antisense sequences. Protist. 2016; 167(4):377-388.
18. Cognat V, Morelle G, Megel C, Lalande S, Molinier J, Vincent T. *et al*. The nuclear and organellar tRNA-derived RNA fragment population in Arabidopsis thaliana is highly dynamic. Nucleic Acids Research. 2017; 45(6):3460-3472.
19. Dahl EL, Shock JL, Shenai BR, Gut J, DeRisi JL, Rosenthal PJ. Tetracyclines specifically target the apicoplast of the malaria parasite Plasmodium falciparum. Antimicrobial Agents and Chemotherapy. 2006; 50(9):3124-3131.
20. Maréchal E, Cesbron-Delauw MF. The apicoplast: a new member of the plastid family. Trends in Plant Science. Trendsin Plant Science. 2001; 6(5):200-205.
21. Gilbert W. Why genes in pieces? Nature. 1978; 271(5645):501.
22. Brázda V, Laister RC, Jagelská EB, Arrowsmith C. Cruciform structures are a common DNA feature important for regulating biological processes. BMC molecular biology. 2011; 12:33.