

Article

Characterization of Aminoacyl-tRNA Synthetases in Chromerids

Abdoallah Sharaf ^{1,2}, Ansgar Gruber ¹ , Kateřina Jiroutová ¹ and Miroslav Oborník ^{1,3,*}

¹ Institute of Parasitology, Biology Centre, Czech Academy of Sciences, 370 05 České Budějovice, Czech Republic

² Genetics Department, Faculty of Agriculture, Ain Shams University, Cairo 11241, Egypt

³ Faculty of Science, University of South Bohemia, 370 05 České Budějovice, Czech Republic

* Correspondence: obornik@paru.cas.cz

Received: 1 July 2019; Accepted: 28 July 2019; Published: 31 July 2019



Abstract: Aminoacyl-tRNA synthetases (AaRSs) are enzymes that catalyze the ligation of tRNAs to amino acids. There are AaRSs specific for each amino acid in the cell. Each cellular compartment in which translation takes place (the cytosol, mitochondria, and plastids in most cases), needs the full set of AaRSs; however, individual AaRSs can function in multiple compartments due to dual (or even multiple) targeting of nuclear-encoded proteins to various destinations in the cell. We searched the genomes of the chromerids, *Chromera velia* and *Vitrella brassicaformis*, for AaRS genes: 48 genes encoding AaRSs were identified in *C. velia*, while only 39 AaRS genes were found in *V. brassicaformis*. In the latter alga, ArgRS and GluRS were each encoded by a single gene occurring in a single copy; only PheRS was found in three genes, while the remaining AaRSs were encoded by two genes. In contrast, there were nine cases for which *C. velia* contained three genes of a given AaRS (45% of the AaRSs), all of them representing duplicated genes, except AsnRS and PheRS, which are more likely pseudoparalogs (acquired via horizontal or endosymbiotic gene transfer). Targeting predictions indicated that AaRSs are not (or not exclusively), in most cases, used in the cellular compartment from which their gene originates. The molecular phylogenies of the AaRSs are variable between the specific types, and similar between the two investigated chromerids. While genes with eukaryotic origin are more frequently retained, there is no clear pattern of orthologous pairs between *C. velia* and *V. brassicaformis*.

Keywords: *Chromera velia*; *Vitrella brassicaformis*; Aminoacyl tRNA synthetase (AaRS); evolution; protein localization; chloroplast; mitochondrion; nucleus

1. Introduction

Chromerids are single-celled photosynthetic apicomonads associated with corals [1–5]. Phylogenetic analyses have shown that these algae are closely related to Apicomplexa, confirming the hypothesis that apicomplexan parasites originate from a photosynthetic alga [1,6–8]. Only two chromerid species, *Chromera velia* and *Vitrella brassicaformis*, representing two distinct lineages, have been formally described so far [1,2]. They do not form sister species and they are considerably different in morphology, ultrastructure, pigmentation, organellar genomes, respiratory chains, and life cycles [1,2,5,8–12].

Chromerid genomes can be used to trace the evolutionary transition from a photosynthetic ancestor to one of the most successful groups of eukaryotic parasites [7,13,14]. Chromerid plastid genomes differ substantially in their topology, gene content, the variability of the genes they encode, and even in the genetic code, they use [4,5,10,11,15]. Chromerids possess highly reduced photosystems, which also contain several undescribed protein subunits [16]. Complete metabolic pathways, such

as sugar metabolism, assimilation of nitrate and sulfite, and photosynthesis-related processes, take place in chromerid plastids [7,17]. Phylogenetic analyses [1,9,15,18] and the use of non-canonical coding for tryptophan in *C. velia* [1,15] suggest that chromerid plastids share a common ancestry with the apicomplexan relic plastid [10]. This relationship is also evident in the similar arrangement of heme biosynthesis in chromerids and apicomplexans: δ -aminolevulinic acid (ALA) is synthesized in the mitochondrion (C4 pathway) and then it is exported to the plastid [19,20]. This metabolic curiosity qualifies chromerids as the only known phototrophs that synthesize chlorophyll from glycine rather than from glutamate [19]. Chromerid mitochondrial genomes have been described as a set of heterogeneous linear molecules [8,10], like the ones found in dinoflagellates [10,21–24]. Another feature shared by chromerids and dinoflagellates is the oligoadenylation of short mito-rRNA fragments that do not seem to be spliced together [10,24]. In contrast with other myzozoans that have a conserved mitochondrial gene set consisting of three genes coding for *cox1*, *cox3*, and *cyb*, the mitochondrial genome of *C. velia* encodes only two of these proteins, with *cyb* missing. Consequently, *C. velia* has a reduced respiratory chain, lacking not only complex I like other myzozoans, but also complex III [8,10,25]. This means that the mitochondrial genome of *C. velia* has the most reduced coding capacity of those studied so far. Recently, a mitochondrion with virtually the same respiratory chain as that of *C. velia*, but lacking any genome, was discovered in the dinoflagellate *Amoebophyra* [26].

Aminoacyl-tRNA synthetases (AaRSs) ligate amino acids to specific tRNAs according to the genetic code [27]. Around 18–20 archaeal and bacterial types of AaRSs have been identified; additional AaRSs were identified in eukaryotes because they contain intracellular compartments, such as mitochondria and plastids, with their own translational activity [28–31]. The compartment-specific AaRS isoforms are often encoded by different nuclear genes [28]. Recently, several gene duplications, horizontal gene transfers, and gene losses of AaRSs have been described, as a result of the availability of the complete genome sequences [32]. Generally, each compartment in which translation takes place needs a complete set of AaRSs for the synthesis of proteins.

Aminoacyl-tRNA synthetases (AaRSs) can be divided into two classes each with 10 enzymes, based on the conserved sequence motifs, quaternary structure, and aminoacylation function [33,34]. Class I enzymes have two consensus sequence motifs, while class II enzymes have three such motifs [33]. All enzymes show the characteristics typical for one of the above-mentioned classes, except Alanyl-tRNA synthetase (AlaRS), Phenylalanyl-tRNA synthetase (PheRS), and Glycyl-tRNA synthetase (GlyRS) (class II): in *Escherichia coli*, AlaRS contains only two of the motifs typical for class II AaRSs, while GlyRS contains only one. PheRSs possess the three typical class II motifs, but in contrast to all other class II AaRSs, they bind amino acids on the 2'-hydroxyl group (OH) of the tRNA terminal adenosine (A76) [35]. In general, AaRSs show divergent oligomeric structures: the prokaryotic GlyRSs are heterotetramers ($\alpha_2\beta_2$) [35–37], while the eukaryotic enzymes display a homodimeric (α_2) structure [35,38,39]. The evolutionary origins of these differences remain unclear.

There are alternative mechanisms known to be involved in the loading of tRNAs in addition to the direct loading of a specific amino acid onto its tRNA by AaRSs. Asparagine and glutamine could be misloaded onto tRNAs with the anticodons for aspartic acid and glutamic acid, respectively. When this occurs, amidotransferases can specifically convert the loaded tRNA to restore the amino acid-anticodon fit [40]. Interestingly, the absence of AaRS-encoding genes has also been reported [41–45].

Previously, AaRSs genes were automatically annotated in the genomes of *C. velia* and *V. brassicaformis* by Woo et al. [7]. Gile et al. [46] provided detailed annotations of AaRSs genes and the localization of the gene products in the diatoms *Phaeodactylum tricorutum* and *Thalassiosira pseudonana*, as well as in the cryptophyte *Guillardia theta*. Furthermore, annotations of AaRSs genes in protists are contained in the EuPathDB database [47]. The focus of this study was an exploration of the evolutionary origins of the genes encoding the chromerid Aminoacyl-tRNA synthetases. We employed a method based on Hidden Markov models (HMM) of conserved protein motifs [32] for the identification of the AaRS, and with this approach, corrected the previously published annotations for a number of genes (see results for details). The evolutionary histories of AaRSs could be used to

trace the cellular evolution because of their metabolic universality and their essential role in protein synthesis [48]. AaRSs commonly have irregular evolutionary patterns due to gene duplications, high levels of sequence divergence, and horizontal gene transfers [49], but some AaRS trees show clear ancestral relationships [48]. Since in chromerids, no Aminoacyl-tRNA synthetases are plastid or mitochondrially encoded, they are all encoded by nuclear genes, and translated in the cytosol, with posttranslational targeting to the translationally active organelles.

2. Material and Methods

2.1. Gene Identification and Model Assessment

The genomes of the chromerids (*C. velia* and *V. brassicaformis*) and a closely related apicomplexan (*Toxoplasma gondii*) were searched for AaRSs with the BLASTp tool in the Eukaryotic Pathogen database (<https://eupathdb.org>) [7,47], using previously characterized AaRS amino acid sequences as queries [32]. We validated the selected gene models by checking the transcription of the 5'-end of the gene model by BLAST against the identified genes from *C. velia* and *V. brassicaformis* transcriptomes [7]. Conserved protein motifs in each of the AaRS loci were identified by the Hidden Markov model (HMM)-based tool for the AaRS coding sequence detection with a cut-off of 5 out of 10 motifs [32,50]. Applying the former analysis pipeline on the Swiss-Prot annotated AaRSs showed that this pipeline could capture and annotate 99.8% of the AaRSs [32].

2.2. Localization Prediction

We predicted the protein N-termini (starting with a methionine residue) preceding the conserved AaRS domain; also tested the gene model for possible gene fusion using the genome browser (JBrowse) in the CryptoDB database (<https://cryptodb.org/cryptodb/>) [47]. Targeting predictions of endoplasmic reticulum (ER) signal peptides were performed using SignalP 3.0 [51] and SignalP 4.1 [52], the bipartite targeting sequences (BTS) were searched by ASAFind [53], transit peptides by TargetP 1.1 [54], and mitochondrial transit peptides were predicted using iPSORT [55], WoLF PSORT [56], and Mitoprot [57] (Table 1 and Supplementary Table S1A). Dual targeting by an ambiguous targeting presequence was assumed for those AaRSs predicted to possess localization signals for several different subcellular compartments.

Table 1. Summary of the subcellular localization and origin of the identified Aminoacyl-tRNA synthetases (AaRS) in chromerids (*Chromera velia* and *Vitrella brassicaformis*). Four letters are used to describe the predicted localization of AaRSs (Cyto. = Cytoplasm; Nucl. = Nucleus; Mito. = Mitochondria; Plas. = Plastid; Endo. = Endoplasmic reticulum) and their origins (Euka. = Eukaryotic; Mito. = Mitochondrial; Plas. = Plastidal; Bact./Orga. = Bacterial/Organellar; Uncl. = Unclear), while the number of transcripts is between parenthesis if it was found.

AaRSs	<i>Chromera velia</i>			<i>Vitrella brassicaformis</i>		
	Gene	Localization	Origin	Gene	Localization	Origin
ArgRS	cvel_7363	Cyto. Nucl.	Bact./Orga.	vbra_19912	Nucl.	Plas.
	cvel_22194	Cyto.	Euka.			
	cvel_3008	Mito.	Plas.			
CysRS	cvel_18990	Cyto.	Euka.	vbra_21787	Cyto.	Euka.
	cvel_1412	Plas.	Bact./Orga.	vbra_7006	Mito.	Bact./Orga.
GluRS	cvel_31597	Plas.	Bact./Orga.	vbra_21912	Plas.	Bact./Orga.

Table 1. Cont.

AaRSs	<i>Chromera velia</i>			<i>Vitrella brassicaformis</i>		
	48			39		
	Gene	Localization	Origin	Gene	Localization	Origin
GlnRS	cvel_8110	Cyto.	Mito.	vbra_11768	Cyto.	Mito.
	cvel_1112	Cyto.	Mito.	vbra_19782	Mito.	Bact./Orga.
	cvel_874	Mito.	Bact./Orga.			
LeuRS	cvel_7809	Plas. Mito.	Bact./Orga.	vbra_9619	Plas. Mito.	Bact./Orga.
	cvel_16973	Cyto.	Euka.	vbra_9824	Nucl.	Euka.
IleRS	cvel_28423	Plas.	Euka.	vbra_7770	Cyto.	Euka.
	cvel_28875	Cyto.	Euka.	vbra_11931	Endo.	Euka.
MetRS	cvel_21221	Cyto.	Euka.	vbra_1113	Plas.	Mito.
	cvel_7165	Plas. Mito.	Mito.	vbra_13101	Cyto.	Euka.
TrpRS	cvel_7641	Cyto.	Euka.	vbra_16713	Mito.	Euka.
	cvel_20758	Plas. Mito.	Plas.	vbra_20646	Endo.	Plas.
TyrRS	cvel_7146	Plas.	Mito.	vbra_8727	Endo.	Mito.
	cvel_4070	Cyto.	Euka.	vbra_13490	Cyto.	Euka.
ValRS	cvel_6889	Nucl.	Euka.	vbra_14431	Plas.	Euka.
	cvel_16754	Cyto.	Euka.	vbra_19209	Cyto.	Euka.
	cvel_9091	Cyto.	Euka.			
AlaRS	cvel_11640	Nucl.	Bact./Orga.	vbra_6631	Nucl.	Euka.
	cvel_19371	Mito.	Euka.	vbra_13386	Mito.	Bact./Orga.
	cvel_21018	Nucl.	Euka.			
AspRS	cvel_29834	Cyto.	Euka.	vbra_13004	Mito.	Euka.
	cvel_4486	Cyto.	Euka.	vbra_6695	Plas.	Euka.
	cvel_13271	Plas.	Euka.			
AsnRS	cvel_14360	Cyto.	Uncl.	vbra_5336	Plas. Mito.	Euka.
	cvel_7405	Mito.	Euka.	vbra_5330	Mito.	Euka.
	cvel_23534	Cyto.	Euka.			
GlyRS	cvel_30153	Cyto.	Euka.	vbra_22902	Cyto.	Euka.
	cvel_8554	Cyto.	Euka.	vbra_14962	Plas. Mito.	Bact./Orga.
HisRS	cvel_21705	Cyto.	Euka.	vbra_20528	Plas.	Euka.
	cvel_3311(2)	(2) Plas.	Euka.	vbra_6931	Cyto.	Euka.
LysRS	cvel_9397	Mito.	Mito.	vbra_6201	Cyto.	Euka.
	cvel_20258	Cyto.	Euka.	vbra_4678	Nucl.	Mito.
	cvel_7175	Cyto. Nul.	Euka.			
α -PheRS	cvel_9563	Plas. Mito.	Uncl.	vbra_15842	Cyto. Nucl.	Uncl.
	cvel_5984 (3)	(3) Cyto.	Euka.	vbra_3529	Nucl.	Euka.
β -PheRS	cvel_22325	Cyto.	Euka.	vbra_20531	Cyto.	Euka.
ProRS	cvel_26651	Cyto.	Uncl.	vbra_17540	Cyto.	Euka.
	cvel_17161	Cyto.	Euka.	vbra_8238	Endo.	Uncl.
	cvel_28365	Cyto.	Euka.	vbra_2397	Cyto.	Euka.
SerRS	cvel_4888	Nucl.	Euka.	vbra_7841	Nucl.	Mito.
	cvel_18368	Endo.	Mito.			
ThrRS	cvel_8549	Plas. Mito.	Mito.	vbra_13128	Nucl.	Euka.
	cvel_11445	Cyto.	Euka.	vbra_14881	Plas.	Mito.

2.3. Molecular Phylogeny

Amino acid sequences of all 20 Aminoacyl-tRNA synthetases from plants, chlorophytes, rhodophytes, alveolates (apicomplexans and ciliates), stramenopiles (oomycetes and diatoms), metazoans, fungi, kinetoplastids, amoebozoans, parabasalids, and diplomonads were downloaded from the NCBI GeneBank (www.ncbi.nlm.nih.gov) and specialized databases (www.genome.jgi-psf.org; www.paramecium.cgm.cnrs-gif.fr; www.merolae.biol.s.u-tokyo.ac.jp; www.ensembl.org; www.arabidopsis.org). The available bacterial homologs were obtained from the prokaryotic AaRS database (www.bioinf.bio.uth.gr/AaRS/#/), with special emphasis on those prokaryotes that are supposed to be involved in endosymbiotic events, such as cyanobacteria (the supposed ancestor of primary plastids) and α -proteobacteria (the supposed ancestor of mitochondria). The sequences were aligned using the MAFFT software [58]; ambiguously aligned sequences and gaps were excluded from further analysis. The alignments were tested using Protest [59,60] to specify the appropriate amino acid substitution model. The maximum likelihood method, as implemented in PhyML (www.atgc-montpellier.fr/phyml) [61,62], and Bayesian inference (Mr. Bayes; www.mrbayes.sourceforge.net) [63], were used to construct the phylogenetic trees.

3. Results and Discussion

Gene Identification and in Silico Localization of Gene Products

BLAST search and motif screening of the total predicted proteomes of the two chromerids (*C. velia* and *V. brassicaformis*) and their close relative, the apicomplexan *T. gondii*, resulted in the identification of 48, 39, and 31 AaRS genes, respectively, with some of the previous annotations corrected (Table 1 and Supplementary Table S1A). Certainly, this is substantially less than the expected number (roughly 60) necessary for providing AaRSs to all three translationally active compartments—the cytosol, plastid, and mitochondrion—via exclusive transport mechanisms that provide only one gene product to only one intracellular location. AaRSs are generally conserved; therefore, it is unlikely that any additional enzymes were missed during our search due to sequence diversity. In the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, and in the cryptophyte *Guillardia theta*, our genome search and motif screening of previously identified AaRSs revealed 17.5% misidentified genes in the diatom *Thalassiosira pseudonana*, 5% in the diatom *Phaeodactylum tricornutum*, and 7% in the cryptophyte *Guillardia theta* compared to Gile et al. [46] (Table 1 and Supplementary Table S1A). For example, the genes Jgi_51430 and Jgi_42226 were identified as cytoplasmic Glu-tRNA synthetases (GluRSs) in *P. tricornutum* and *T. pseudonana* by Gile et al. [46]; however, our motif search revealed that these genes contain three GluRS specific motifs in *T. pseudonana* and only two such motifs in *P. tricornutum*, while they possess six (*T. pseudonana*) and five (*P. tricornutum*) GlnRS specific motifs instead (Supplementary Table S1A). Other examples include the plastid-localized Jgi_24163 and periplastidal Jgi_68323 proteins, identified as the α -subunit of PheRS (α -PheRS) in *T. pseudonana* and *G. theta* [46]. Again, our motif search revealed that they contain four and three motifs for *T. pseudonana* and *G. theta*, respectively, which is below the standard cut-off value [32], meaning the genes probably encode other proteins than AaRSs (Supplementary Table S1A).

The chromerids each have a different number of genes coding for AaRSs. Nine (45%) AaRSs in *C. velia* are encoded by three distinct genes; whereas, ten (45%) are encoded by two distinct genes. Interestingly, only GluRS was encoded by a single gene (Figure 1a and Table 1). Our premise assumes that each compartment with translational activity needs an AaRS for each amino acid. Since all AaRSs in chromerids are nuclear-encoded [7,8,15], and only nine AaRSs are present in three copies in the genome, as well as five that are targeted to all three compartments, the remaining AaRSs must be dually or even multiply targeted to the cytosol, the mitochondrion, and the plastid to satisfy the need for a complete set of AaRS for each of the three subcellular compartments in which translation takes place.

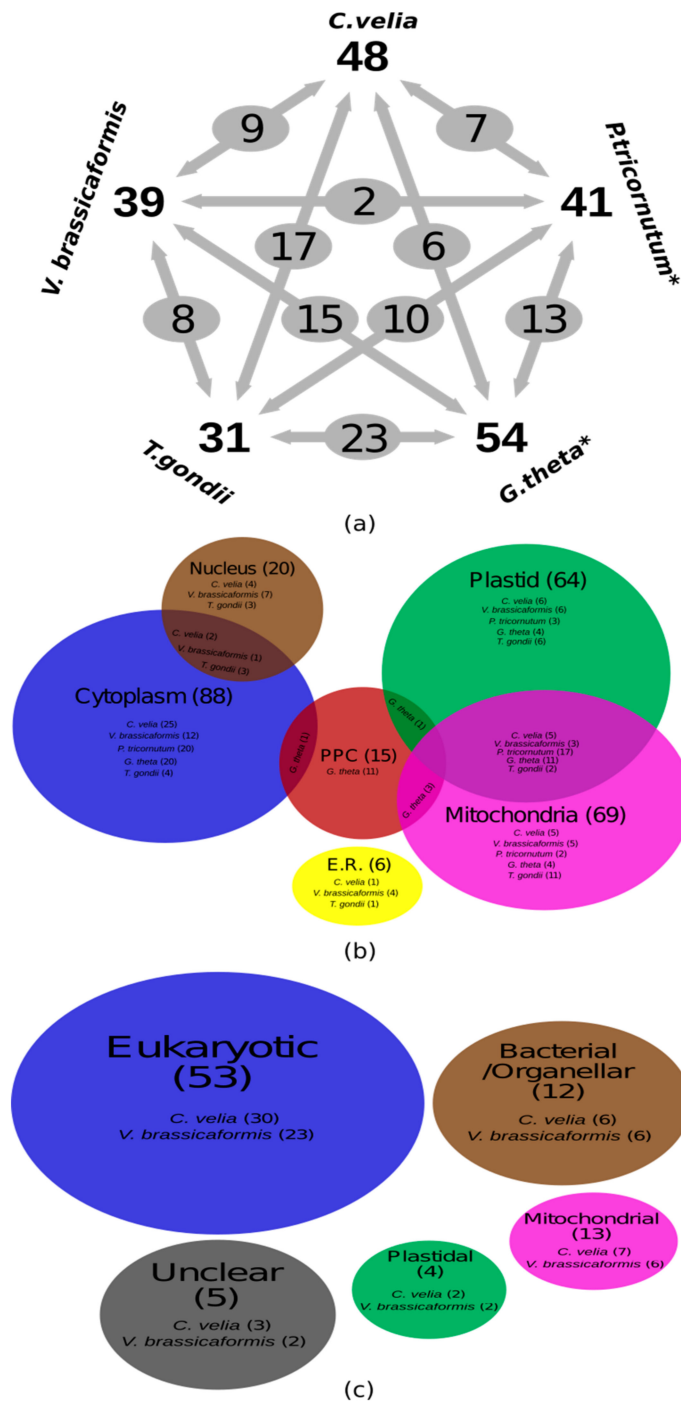


Figure 1. (a) Numbers of Aminoacyl-tRNA synthetases in the chromerids (*Chromera velia* and *Vitrella brassicaformis*), in comparison to a diatom (*Phaeodactylum tricornutum*), a cryptophyte (*Guillardia theta*), and the apicomplexan *Toxoplasma gondii*; (b) Euler diagram representing the protein subcellular localization results of Aminoacyl-tRNA synthetases (AARS) in the chromerids, *P. tricornutum*, *G. theta* and *T. gondii*; (c) Euler diagram representing the evolutionary origins of the chromerid AARSs.

Alternatively, the activity of some AARSs could be replaced by a tRNA-dependent amino acid transformation mechanism [40]. In this study, we found that both *C. velia* and *V. brassicaformis* possess two genes coding for Glutamyl-tRNA^{Gln} amidotransferase (Glu-AdT). Both *Chromera*'s (*cvel_23134*) and *Vitrella*'s (*vbra_8651*) Glu-AdT were predicted to be targeted to the plastid (Supplementary Table S1B), in which Glu-tRNA^{Gln} is most likely transamidated to produce Gln-tRNA^{Gln}. Moreover, two plastid-targeted Aspartyl/Glutamyl-tRNA^{Asn/Gln} amidotransferases (Asp/Glu-AdT) were identified

in *C. velia* (cvel_28674 and cvel_12310), and one plastid-targeted Asp/Glu-AdT was identified in *V. brassicaformis* (vbra_10654). This could compensate for the lack of Asparaginyl-tRNA (AsnRS) and Glutaminyl-tRNA (GlnRS) synthetases in the plastids of both investigated chromerids (Supplementary Table S1B), producing Asn-tRNA^{Asn} or Gln-tRNA^{Gln} through the transamidation of misacylated Asp-tRNA^{Asn} or Glu-tRNA^{Gln}, respectively.

We identified many AaRSs putatively localized to the nucleus. Since tRNAs are formed in the nucleus, these enzymes could be involved in the transport of tRNAs to the cytoplasm. The localization of AaRSs to the nucleus has been shown in numerous previous studies on nuclear aminoacylation of tRNAs [64–66]. Recently, at least 13 active AaRSs were found in purified mammalian cell nuclei [66,67]. In general, nucleus-targeted AaRSs are known to play a role in tRNA maturation and tRNA export control. The human nuclear enzyme, MetRS seems to be related to the biogenesis of rRNA in nucleoli, in addition to its catalytic activity in protein synthesis in the cytoplasm [66,68].

A quarter of the AaRSs (LeuRS, MetRS, TrpRS, PheRS and ThrRS) in *C. velia* were predicted to be targeted to all three compartments (Figure 1a and Table 1). In *C. velia*, GlyRS and ProRS lack any targeting presequence, which suggests they are only localized to the cytoplasm; the same has been experimentally proven for human GlnRS [69] (Table 1 and Supplementary Table S1A). Ten AaRSs (GluRS, AspRS, CysRS, GlyRS, HisRS, IleRS, ProRS, SerRS, TyrRS, and ValRS) do not possess a mitochondrial transit peptide in *C. velia* (Figure 1a). This is a common scenario known from ArgRS and LeuRS in *Arabidopsis* [70,71], GluRS and GlnRS in *T. pseudonana*, and GlnRS in both *P. tricornutum* and trypanosomatids [46,72,73] (Supplementary Table S1). We did not detect the BTS necessary for plastid targeting in 45% of the AaRSs in *C. velia*; it was missing from AlaRS, ArgRS, AsnRS, LysRS, GluRS, ValRS, GlyRS, ProRS, and SerRS (Figure 1a and Table 1). The same has been observed in *Arabidopsis*, with GluRS, GlyRS, TrpRS, and TyrRS lacking chloroplast transit peptides [70] (Supplementary Table S1A). BTS presequences are also absent in AlaRS, AspRS, IleRS, and ValRS in *Guillardia theta* (Figure 1a). On the other hand, almost complete sets of plastid-targeted AaRSs were found in the diatoms, *T. pseudonana* and *P. tricornutum* [46], except PheRS in *T. pseudonana* (Figure 1a and Supplementary Table S1A). Cytoplasmic AlaRS and GluRS were absent from both chromerids, only plastid GluRS was identified. Finally, SerRS was predicted to be targeted to the cytoplasm, nucleus, and endoplasmic reticulum (Endo.) (Figure 1a and Supplementary Table S1).

In *V. brassicaformis*, 17 (85%) of the AaRSs were encoded by two genes; whereas, ArgRS and GluRS were only encoded by a single gene (Figure 1a and Table 1). PheRS was encoded by three genes, two codings for the α -subunit and one for the β -subunit. Dual targeting by an ambiguous targeting sequence was identified in AsnRS, GlyRS, LeuRS, and α -PheRS. Like in *C. velia*, nucleus-localized AlaRS, ArgRS, LeuRS, LysRS, α -PheRS, SerRS, and ThrRS were also found in *V. brassicaformis*. Only GlyRS was predicted to be targeted to all three compartments, employing a dual-targeting mechanism in one of the two predicted genes (vbra_14962) (Figure 1a and Table 1). In contrast to *C. velia*'s GlnRS, GlyRS, ProRS, and TrpRS, none of the *V. brassicaformis* AaRSs were targeted only to the cytoplasm (Figure 1a and Table 1). Interestingly, in *V. brassicaformis*, AlaRS, AsnRS, AspRS, CysRS, GlnRS, GlyRS, LeuRS, and TrpRS were identified as mitochondrial-targeted proteins, while AlaRS, ArgRS, CysRS, GlnRS, IleRS, LysRS, PheRS, ProRS, SerRS, TrpRS, and TyrRS do not seem to have plastid-targeting signals (Figure 1a and Table 1). Furthermore, AlaRS, ArgRS, AsnRS, AspRS, GluRS, LeuRS, ThrRS, and TrpRS were identified to be localized in the organelles. Finally, IleRS, ProRS, TrpRS, and TyrRS were predicted to be localized to the endoplasmic reticulum (Figure 1a and Table 1).

GlyRSs are a special class of Aminoacyl-tRNA synthetases because they have variable functional properties and divergent oligomeric structures [35,74]. Prokaryotic GlyRSs recognize tRNA molecules with a U73 discriminator base; whereas, the eukaryotic ones recognize them with A73 [35,39,74,75]. This suggests that the interaction between the structure and function differs between prokaryotic and eukaryotic glycation systems. The oligomeric structure of most AaRSs is conserved, and sequence comparisons reveal significant similarities. In contrast, GlyRSs show high structural divergence, which leads to the non-conservation of the oligomeric structure [35]. The subunit structure of GlyRSs is not

conserved in prokaryotes: two oligomeric types of GlyRS are found in nature (see above). The α_2 type has been identified in all three kingdoms of life, while the $\alpha_2\beta_2$ type is only found in bacteria and chloroplasts [35].

Only two α_2 -type enzymes were identified in *C. velia*, both providing cytoplasmic functions. Conversely, *V. brassicaformis* contained enzymes of the α_2 -type (vbra_22902) and the fused $(\alpha\beta)_2$ -type (vbra_14962), functioning in all three compartments (Figure 1a and Table 1). The fused $(\alpha\beta)_2$ -type enzyme was previously found in *Arabidopsis thaliana* and functions in both mitochondria and chloroplasts [76]. GlyRSs α_2 -type (vbra_22902) and the fused $(\alpha\beta)_2$ -type (vbra_14962) in *V. brassicaformis* show low amino acid sequences (17.3 %), but both genes contain 10 GlyRS functional motifs, which support the importance of motif scanning during the gene identification pipeline [32].

4. Phylogenetic Analyses

We performed phylogenetic analyses of all 20 AaRSs from chromerid algae. Similarity networks for the whole set of Aminoacyl-tRNA synthetases reflect the tree topology of each Aminoacyl-tRNA synthetase [77] (Figure 2). Out of the 21 computed trees (PheRS has two subunits), only the β -PheRS tree fully corresponds to the classical three-domain pattern that is postulated for eukaryotic enzymes. The archaeal and eukaryotic AaRSs were well separated from the bacterial group, in agreement with previous studies [48]. The remaining trees show two or more divergent bacterial clades (Figure 2). Furthermore, in both chromerids, endosymbiotic gene transfer events were only observed in ArgRS and TrpRS. Interestingly, for LeuRS and TyrRS there are two unrelated groups, likely the result of independent evolution of the eukaryotic and prokaryotic versions of these two enzymes (Figure 2). The remarkable dissimilarities of various regions in bacterial and eukaryotic LeuRS and TyrRS have also been observed in previous studies [48,78].

AaRSs represents an example of the true ancestral paralogous evolved from gene duplication, but it is also a Horizontal Gene Transfer (HGT) leading to pseudoparalogy [79]. Out of nine cases for which *C. velia* contained three genes of a given AaRS (Table 1), we identify gene duplication in seven *C. velia*'s AaRSs (GlnRS, ValRS, AlaRS, AspRS, GlyRS, LysRS, and SerRS) while AsnRS and PheRS are more likely pseudoparalogs (Supplementary Figure S1 and File sf1). In contrast, no gene duplication was identified in *V. brassicaformis*'s AaRSs.

Bacterial PheRS consists of an α and β subunit encoded by the pheS and pheT genes, canonically located on the same operon [80]. Mitochondrial PheRS is a fusion of the N-terminal part of the α -subunit and the C-terminal part of the β -subunit [81]. We concatenated the two subunits into one sequence. Nevertheless, we constructed phylogenetic trees for the α - and β -subunits separately, to ensure that the PheRS tree topology is not affected by mixed signals from the two PheRS subunits. Analyzing the subunits individually results in the same topology as for the concatenated sequences (Supplementary Figure S1). We also constructed a phylogenetic tree from the combined sequences, representing both GlyRSs, types α_2 and $\alpha_2\beta_2$. The similarity network and the phylogenetic tree confirmed that the enzymes of type $\alpha_2\beta_2$ are only present in bacteria, resulting in a disconnected network and the formation of a clade that is separate from enzyme type α_2 (Figure 2, Supplementary Figure S1 and File sf1). All three domains of life have the α_2 -type enzyme, including both chromerids (Figure 2, Supplementary Figure S1 and File sf1).

In general, *C. velia* and *V. brassicaformis* AaRSs show similar evolutionary patterns (Figure 1b and Table 1), with the exception of AgrRS, AsnRS, and GlyRS. Seven (14.6%) of the identified genes in *C. velia* and six (15.4%) genes in *V. brassicaformis* are mitochondrial in origin (Figure 1b and Table 1). In *C. velia* and *V. brassicaformis* respectively, 30 out of the 48 (62.5%) and 23 out of the 39 (59%) identified AaRSs originate from the eukaryotic nucleus. Six AaRSs show bacterial/organellar origin in both chromerids (Table 1). We were unable to specify the origin of three (6.2%) and two (5.1%) of the AaRSs from *C. velia* and *V. brassicaformis*, respectively (Table 1 and Supplementary File sf1). Since chromerids appear to be closely related to apicomplexan parasites, we also looked at the origins of AaRSs in the apicomplexan *T. gondii*. In this parasite, 21 (67.7%) of the identified AaRS genes originate

from the eukaryotic nucleus, while 3 and 4 AaRSs display bacterial/organelar and mitochondrial origin, respectively. We were unable to specify the origin of two (6.4%) AaRS genes from *T. gondii* (Supplementary Figure S1 and File sf1). This apicomplexan and the chromerids have homologous evolutionary patterns in all of their AaRSs except GluRS, supporting the close evolutionary relationship between apicomplexans and chromerids (Supplementary Figure S1 and File sf1).

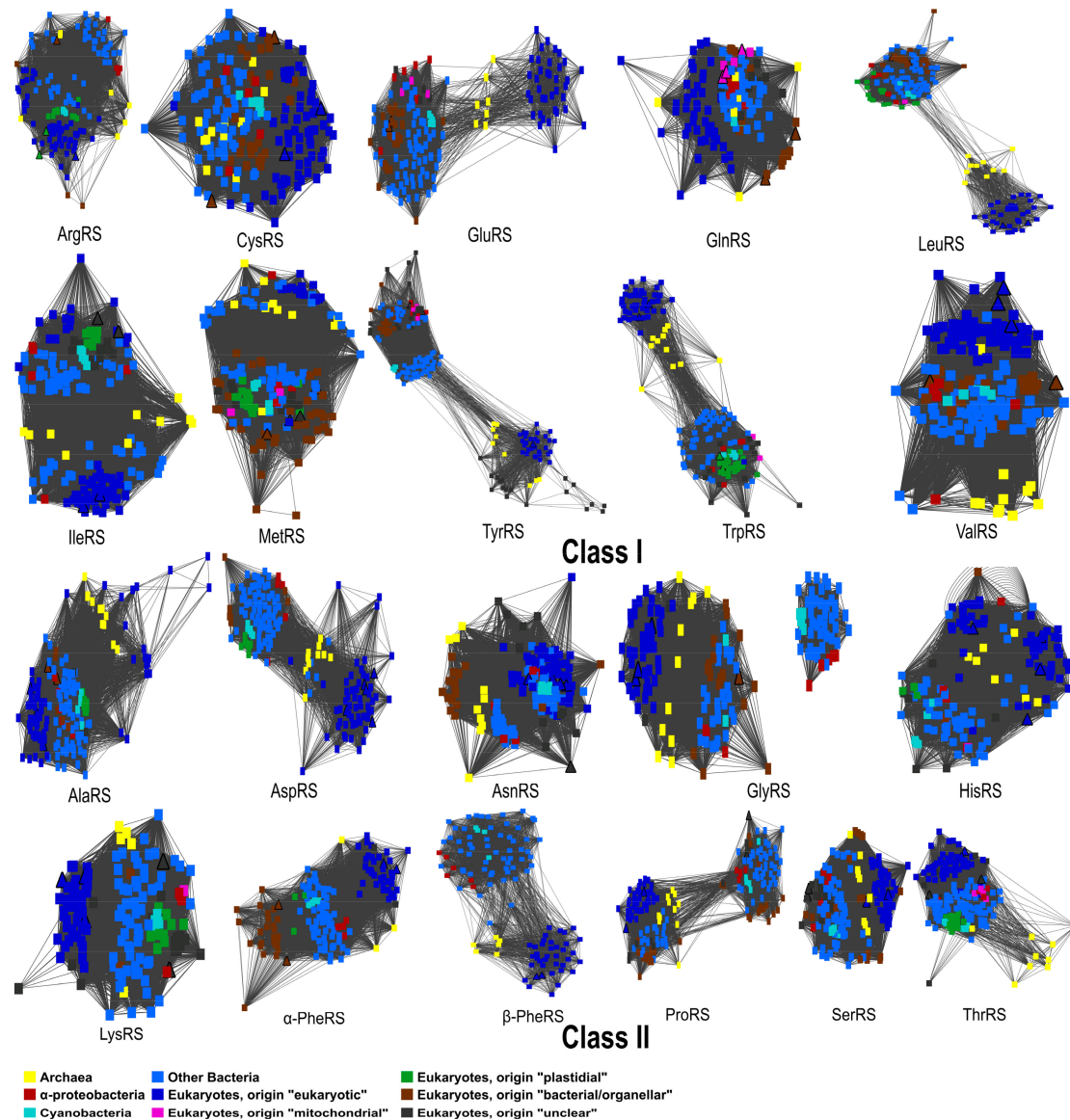


Figure 2. A similarity network for each sequence dataset of the 20 different Aminoacyl-tRNA synthetases shows the sequence similarity and reflects the phylogenetic tree topology of each Aminoacyl-tRNA enzyme.

5. Conclusions

This study identified sequences of Aminoacyl-tRNA synthetases (AaRSs) in chromerids, along with their putative protein localization and evolutionary origins. We identified 48 and 39 AaRS genes in *C. velia* and *V. brassicaformis*, respectively, representing the full set of 20 AaRS types. Five AaRSs were predicted to be targeted to all three compartments (mitochondrion, plastid, and cytosol) by an ambiguous targeting sequence that likely leads to the dual-targeting of these enzymes in *C. velia*; whereas, only two such AaRSs were predicted in *V. brassicaformis*. We identified five (*C. velia*) and eight (*V. brassicaformis*) nucleus-localized AaRSs in the chromerids, which we assume to be first targeted to the

nucleus and then to the cytoplasm. GlnRS and AsnRS were absent from the plastids of both chromerids; their activity is likely restored by tRNA-dependent amino acid transformation mechanisms.

The α_2 -type of GlyRS was identified in *C. velia*; whereas, both the α_2 -type and the fused ($\alpha\beta$)₂-type of GlyRS were found in *V. brassicaformis*. We identified gene duplications of seven AaRSs (GlnRS, ValRS, AlaRS, AspRS, GlyRS, LysRS, and SerRS) in *C. velia*, while no AaRS gene duplication was found in *V. brassicaformis*. Both LeuRS and TyrRS have two disconnected eukaryotic and prokaryotic groups, which could be a result of the independent evolution of the two versions of these enzymes. In both chromerids, ArgRS and TrpRS were shown to be acquired from endosymbionts by endosymbiotic gene transfer; all other genes are of eukaryotic origin with proteins targeted to the various compartments. The tree topologies suggest that numerous gradual losses of pseudoparalogs occurred in eight enzymes (ArgRS, CysRS, IleRS, MetRS, AlaRS, HisRS, LysRS, and ThrRS).

Given that *C. velia* and *V. brassicaformis* are such closely related organisms, the number of differences in their AaRS genes is much higher than we would expect. The intracellular targeting of AaRSs is independent of their evolutionary origin. There is no clear pattern of orthologous pairs that are retained in both organisms. Instead, gene duplications, gene losses, and changes in targeting signals account for the required activities of tRNA synthesis in the different cellular compartments.

Supplementary Materials: The following supplementary materials are available online at <http://www.mdpi.com/2073-4425/10/8/582/s1>. Table S1: (A) Summary of the protein subcellular localization of the identified Aminoacyl-tRNA synthetases (AaRSs) in the chromerids (*Chromera velia* and *Vitrella brassicaformis*), diatoms (*Phaeodactylum tricornerutum* and *Thalassiosira pseudonana*), cryptophyte (*Guillardia theta*), apicomplexan (*Toxoplasma gondii*), archeplastida (*Arabidopsis thaliana*), and opisthokont (*Homo sapiens*), previous gene identification for (*) chromerids by Woo et al., 2015 [7], (**) diatoms and the cryptophyte *Guillardia theta* by Gile et al., 2015 [46] and (***) *Toxoplasma gondii* in EuPathDB [47] were inserted. (B) Summary of the subcellular protein localization of the identified amidotransferase genes in the chromerids, Figure S1: Phylogenetic supertrees of all identified Aminoacyl-tRNA synthetases (AaRSs) from chromerids and other domains of life, based on protein sequences, branch support values are given in % (maximum likelihood/Bayesian inference), Supplementary File sf1: Interpretation of all identified Aminoacyl-tRNA synthetases (AaRS) phylogenetic supertrees.

Author Contributions: M.O. conceived the project and acquired funding. A.S., A.G. and K.J. performed the analyses. A.S. drafted the manuscript. All authors revised the first draft and read and approved the final version of the manuscript.

Funding: This work was supported by the Czech Science Foundation (project 15-17653S) and ERDF/ESF Centre for research of pathogenicity and virulence of parasites (No.CZ.02.1.01/0.0/0.0/16_019/0000759).

Acknowledgments: The authors would like to thank MetaCentrum/CERIT-SC for providing the computational cloud for performing the bioinformatic analyses and Rachel Kolísko, for language editing of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Moore, R.B.; Oborník, M.; Janouškovec, J.; Chrudimský, T.; Vancová, M.; Green, D.H.; Wright, S.W.; Davies, N.W.; Bolch, C.J.S.; Heimann, K.; et al. A photosynthetic alveolate closely related to apicomplexan parasites. *Nature* **2008**, *451*, 959–963. [[CrossRef](#)] [[PubMed](#)]
2. Oborník, M.; Modrý, D.; Lukeš, M.; Černotíková-Stříbrná, E.; Cihlář, J.; Tesařová, M.; Kotabová, E.; Vancová, M.; Prášil, O.; Lukeš, J. Morphology, Ultrastructure and Life Cycle of *Vitrella brassicaformis* n. sp., n. gen., a Novel Chromerid from the Great Barrier Reef. *Protist* **2012**, *163*, 306–323. [[CrossRef](#)] [[PubMed](#)]
3. Cumbo, V.R.; Baird, A.H.; Moore, R.B.; Negri, A.P.; Neilan, B.A.; Salih, A.; van Oppen, M.J.H.; Wang, Y.; Marquis, C.P. *Chromera velia* is endosymbiotic in larvae of the reef corals *Acropora digitifera* and *A. tenuis*. *Protist* **2013**, *164*, 237–244. [[CrossRef](#)]
4. Janouškovec, J.; Horák, A.; Barott, K.L.; Rohwer, F.L.; Keeling, P.J. Environmental distribution of coral-associated relatives of apicomplexan parasites. *ISME J.* **2013**, *7*, 444–447. [[CrossRef](#)] [[PubMed](#)]
5. Füßy, Z.; Oborník, M. Chromerids and Their Plastids. In *Advances in Botanical Research*; Academic Press: Cambridge, MA, USA, 2017; Volume 84, pp. 187–218. ISBN 9780128026519.
6. McFadden, G.I.; Waller, R.F. Plastids in parasites of humans. *Bioessays* **1997**, *19*, 1033–1040. [[CrossRef](#)]

7. Woo, Y.H.; Ansari, H.; Otto, T.D.; Linger, C.M.K.; Olisko, M.K.; Michálek, J.; Saxena, A.; Shanmugam, D.; Tayyrov, A.; Veluchamy, A.; et al. Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. *eLife* **2015**, *4*, e06974. [[CrossRef](#)] [[PubMed](#)]
8. Flegontov, P.; Michálek, J.; Janouškovec, J.; Lai, D.H.; Jirků, M.; Hajdušková, E.; Tomčala, A.; Otto, T.D.; Keeling, P.J.; Pain, A.; et al. Divergent mitochondrial respiratory chains in phototrophic relatives of apicomplexan parasites. *Mol. Biol. Evol.* **2015**, *32*, 1115–1131. [[CrossRef](#)] [[PubMed](#)]
9. Oborník, M.; Janouškovec, J.; Chrudimský, T.; Lukeš, J. Evolution of the apicoplast and its hosts: From heterotrophy to autotrophy and back again. *Int. J. Parasitol.* **2009**, *39*, 1–12. [[CrossRef](#)]
10. Oborník, M.; Lukeš, J. The organellar genomes of *Chromera* and *Vitrella*, the phototrophic relatives of apicomplexan parasites. *Annu. Rev. Microbiol.* **2015**, *69*, 129–144. [[CrossRef](#)]
11. Oborník, M.; Lukeš, J. Cell Biology of Chromerids: Autotrophic Relatives to Apicomplexan Parasites. In *International Review of Cell and Molecular Biology*; Academic Press: Cambridge, MA, USA, 2013; Volume 306, pp. 333–369.
12. Oborník, M.; Kručínská, J.; Esson, H. Life cycles of chromerids resemble those of colpodellids and apicomplexan parasites. *Perspect. Phycol.* **2016**, *3*, 21–27. [[CrossRef](#)]
13. Janouškovec, J.; Tikhonenkov, D.V.; Burki, F.; Howe, A.T.; Kolísko, M.; Mylnikov, A.P.; Keeling, P.J. Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10200–10207. [[CrossRef](#)] [[PubMed](#)]
14. Füßy, Z.; Oborník, M. Reductive Evolution of Apicomplexan Parasites from Phototrophic Ancestors. In *Evolutionary Biology: Self/Nonself Evolution, Species and Complex Traits Evolution, Methods and Concepts*; Springer: Cham, Switzerland, 2017; pp. 217–236.
15. Janouškovec, J.; Horák, A.; Oborník, M.; Lukeš, J.; Keeling, P.J. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10949–10954. [[CrossRef](#)] [[PubMed](#)]
16. Sobotka, R.; Esson, H.J.; Koník, P.; Trsková, E.; Moravcová, L.; Horák, A.; Dufková, P.; Oborník, M. Extensive gain and loss of photosystem I subunits in chromerid algae, photosynthetic relatives of apicomplexans. *Sci. Rep.* **2017**, *7*, 13214. [[CrossRef](#)] [[PubMed](#)]
17. Füßy, Z.; Faitová, T.; Oborník, M. Subcellular compartments interplay for carbon and nitrogen allocation in *Chromera velia* and *Vitrella brassicaformis*. *Genome Biol. Evol.* **2019**. [[CrossRef](#)] [[PubMed](#)]
18. Ševčíková, T.; Horák, A.; Klimeš, V.; Zbránková, V.; Demir-Hilton, E.; Sudek, S.; Jenkins, J.; Schmutz, J.; Pribyl, P.; Fousek, J.; et al. Updating algal evolutionary relationships through plastid genome sequencing: Did alveolate plastids emerge through endosymbiosis of an ochrophyte? *Sci. Rep.* **2015**, *5*, 10134. [[CrossRef](#)]
19. Kořený, L.; Sobotka, R.; Janouškovec, J.; Keeling, P.J.; Oborník, M. Tetrapyrrole Synthesis of Photosynthetic Chromerids Is Likely Homologous to the Unusual Pathway of Apicomplexan Parasites. *Plant Cell* **2011**, *23*, 3454–3462. [[CrossRef](#)] [[PubMed](#)]
20. van Dooren, G.G.; Kennedy, A.T.; McFadden, G.I. The Use and Abuse of Heme in Apicomplexan Parasites. *Antioxid. Redox Signal.* **2012**, *17*, 634–656. [[CrossRef](#)] [[PubMed](#)]
21. Patron, N.J.; Waller, R.F.; Archibald, J.M.; Keeling, P.J. Complex protein targeting to dinoflagellate plastids. *J. Mol. Biol.* **2005**, *348*, 1015–1024. [[CrossRef](#)]
22. Slamovits, C.H.; Saldarriaga, J.F.; Larocque, A.; Keeling, P.J. The Highly Reduced and Fragmented Mitochondrial Genome of the Early-branching Dinoflagellate *Oxyrrhis marina* Shares Characteristics with both Apicomplexan and Dinoflagellate Mitochondrial Genomes. *J. Mol. Biol.* **2007**, *372*, 356–368. [[CrossRef](#)]
23. Nash, E.A.; Nisbet, R.E.R.; Barbrook, A.C.; Howe, C.J. Dinoflagellates: A mitochondrial genome all at sea. *Trends Genet.* **2008**, *24*, 328–335. [[CrossRef](#)]
24. Waller, R.F.; Jackson, C.J. Dinoflagellate mitochondrial genomes: Stretching the rules of molecular biology. *BioEssays* **2009**, *31*, 237–245. [[CrossRef](#)] [[PubMed](#)]
25. Waller, R.F.; Kořený, L. Plastid Complexity in Dinoflagellates: A Picture of Gains, Losses, Replacements and Revisions. In *Advances in Botanical Research*; Academic Press: Cambridge, MA, USA, 2017; Volume 84, pp. 105–143. ISBN 9780128026519.
26. John, U.; Lu, Y.; Wohlrab, S.; Groth, M.; Janouškovec, J.; Kohli, G.S.; Mark, F.C.; Bickmeyer, U.; Farhat, S.; Felder, M.; et al. An aerobic eukaryotic parasite with functional mitochondria that likely lacks a mitochondrial genome. *Sci. Adv.* **2019**, *6*, eaav1110. [[CrossRef](#)] [[PubMed](#)]

27. Zhou, X.L.; Du, D.H.; Tan, M.; Lei, H.Y.; Ruan, L.L.; Eriani, G.; Wang, E.D. Role of tRNA amino acid-accepting end in aminoacylation and its quality control. *Nucleic Acids Res.* **2011**, *20*, 8857–8868. [[CrossRef](#)] [[PubMed](#)]
28. Chien, C.-I.; Chen, Y.-W.; Wu, Y.-H.; Chang, C.-Y.; Wang, T.-L.; Wang, C.-C. Functional Substitution of a Eukaryotic Glycyl-tRNA Synthetase with an Evolutionarily Unrelated Bacterial Cognate Enzyme. *PLoS ONE* **2014**, *9*, e94659. [[CrossRef](#)] [[PubMed](#)]
29. Carter, C.W., Jr. Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. *Annu. Rev. Biochem.* **1993**, *62*, 715–748. [[CrossRef](#)] [[PubMed](#)]
30. Giege, R.; Sissler, M.; Florentz, C. Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* **1998**, *26*, 5017–5035. [[CrossRef](#)] [[PubMed](#)]
31. Schimmel, P.R.; Söll, D. Aminoacyl-tRNA Synthetases: General Features and Recognition of Transfer RNAs. *Annu. Rev. Biochem.* **1979**, *48*, 601–648. [[CrossRef](#)]
32. Chaliotis, A.; Vlastaridis, P.; Mossialos, D.; Ibba, M.; Becker, H.D.; Stathopoulos, C.; Amoutzias, G.D. The complex evolutionary history of aminoacyl-tRNA synthetases. *Nucleic Acids Res.* **2017**, *45*, 1059–1068. [[CrossRef](#)]
33. Eriani, G.; Delarue, M.; Poch, O.; Gangloff, J.; Moras, D. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **1990**, *347*, 203–206. [[CrossRef](#)]
34. Ludmerer, S.W.; Schimmel, P. Gene for yeast glutamine tRNA synthetase encodes a large amino-terminal extension and provides a strong confirmation of the signature sequence for a group of the aminoacyl-tRNA synthetases. *J. Biol. Chem.* **1987**, *262*, 10801–10806.
35. Mazauric, M.H.; Keith, G.; Logan, D.; Kreutzer, R.; Giegé, R.; Kern, D. Glycyl-tRNA synthetase from *Thermus thermophilus*—wide structural divergence with other prokaryotic glycyl-tRNA synthetases and functional inter-relation with prokaryotic and eukaryotic glycylation systems. *Eur. J. Biochem.* **1998**, *251*, 744–757. [[CrossRef](#)] [[PubMed](#)]
36. Ostrem, D.L.; Berg, P. Glycyl transfer ribonucleic acid synthetase from *Escherichia coli*: Purification, properties, and substrate binding. *Biochemistry* **1974**, *13*, 1338–1348. [[CrossRef](#)] [[PubMed](#)]
37. Surguchov, A.P.; Surguchov, I.G. Two Enzymically Active Forms of Glycyl-tRNA Synthetase from *Bacillus brevis* Purification and Properties. *Eur. J. Biochem.* **1975**, *54*, 175–184. [[CrossRef](#)] [[PubMed](#)]
38. Kern, D.; Giegé, R.; Ebel, J.P. Purification and some properties of alanyl- and leucyl-tRNA synthetases from baker's yeast. *BBA Sect. Nucleic Acids Protein Synth.* **1981**, *653*, 83–90. [[CrossRef](#)]
39. Shiba, K.; Schimmel, P.; Motegi, H.; Noda, T. Human glycyl-tRNA synthetase. Wide divergence of primary structure from bacterial counterpart and species-specific aminoacylation. *J. Biol. Chem.* **1994**, *269*, 30049–30055. [[PubMed](#)]
40. Curnow, A.W.; Hong, K.W.; Yuan, R.; Kim, S.I.; Martins, O.; Winkler, W.; Henkin, T.M.; Söll, D. Glu-tRNA^{Gln} amidotransferase: A novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proc. Natl. Acad. Sci. USA* **1997**, *22*, 11819–11826. [[CrossRef](#)] [[PubMed](#)]
41. Becker, H.D.; Kern, D. *Thermus thermophilus*: A link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc. Natl. Acad. Sci. USA* **1998**, *22*, 12832–12837. [[CrossRef](#)]
42. Ibba, M.; Curnow, A.W.; Söll, D. Aminoacyl-tRNA synthesis: Divergent routes to a common goal. *Trends Biochem. Sci.* **1997**, *22*, 39–42. [[CrossRef](#)]
43. Leinfelder, W.; Zehelein, E.; Mandrandberthelot, M.; Bock, A. Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature* **1988**, *331*, 723–725. [[CrossRef](#)]
44. Sauerwald, A.; Zhu, W.; Major, T.A.; Roy, H.; Palioura, S.; Jahn, D.; Whitman, W.B.; Yates, J.R.; Ibba, M.; Söll, D. RNA-dependent cysteine biosynthesis in archaea. *Science* **2005**, *307*, 1969–1972. [[CrossRef](#)]
45. Sheppard, K.; Yuan, J.; Hohn, M.J.; Jester, B.; Devine, K.M.; Söll, D. From one amino acid to another: tRNA-dependent amino acid biosynthesis. *Nucleic Acids Res.* **2008**, *36*, 1813–1825. [[CrossRef](#)] [[PubMed](#)]
46. Gile, G.H.; Moog, D.; Slamovits, C.H.; Maier, U.G.; Archibald, J.M. Dual organellar targeting of aminoacyl-tRNA synthetases in diatoms and cryptophytes. *Genome Biol. Evol.* **2015**, *7*, 1728–1742. [[CrossRef](#)] [[PubMed](#)]
47. Warrenfeltz, S.; Basenko, E.Y.; Crouch, K.; Harb, O.S.; Kissinger, J.C.; Roos, D.S.; Shanmugasundram, A.; Silva-Franco, F. EuPathDB: The eukaryotic pathogen genomics database resource. In *Eukaryotic Genomic Databases*; Humana Press: New York, NY, USA, 2018; Volume 1757, pp. 69–113.
48. Woese, C.R.; Olsen, G.J.; Ibba, M.; Soll, D. Aminoacyl-tRNA Synthetases, the Genetic Code, and the Evolutionary Process. *Microbiol. Mol. Biol. Rev.* **2000**, *64*, 202–236. [[CrossRef](#)] [[PubMed](#)]

49. Brown, J.R. Ancient horizontal gene transfer. *Nat. Rev. Genet.* **2003**, *4*, 121–132. [[CrossRef](#)] [[PubMed](#)]
50. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME Suite. *Nucleic Acids Res.* **2015**, *43*, 39–49. [[CrossRef](#)] [[PubMed](#)]
51. Bendtsen, J.D.; Nielsen, H.; Von Heijne, G.; Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **2004**, *340*, 783–795. [[CrossRef](#)] [[PubMed](#)]
52. Petersen, T.N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **2011**, *8*, 785–786. [[CrossRef](#)]
53. Gruber, A.; Rocop, G.; Kroth, P.G.; Armbrust, E.V.; Mock, T. Plastid proteome prediction for diatoms and other algae with secondary plastids of the red lineage. *Plant J.* **2015**, *81*, 519–528. [[CrossRef](#)]
54. Emanuelsson, O.; Brunak, S.; von Heijne, G.; Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* **2007**, *2*, 953–971. [[CrossRef](#)]
55. Bannai, H.; Tamada, Y.; Maruyama, O.; Nakai, K.; Miyano, S. Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* **2002**, *18*, 298–305. [[CrossRef](#)]
56. Horton, P.; Park, K.J.; Obayashi, T.; Fujita, N.; Harada, H.; Adams-Collier, C.J.; Nakai, K. WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* **2007**, *35*, 585–587. [[CrossRef](#)] [[PubMed](#)]
57. Claros, M.G.; Vincens, P. Computational Method to Predict Mitochondrially Imported Proteins and their Targeting Sequences. *Eur. J. Biochem.* **1996**, *241*, 779–786. [[CrossRef](#)] [[PubMed](#)]
58. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
59. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. ProtTest-HPC: Fast selection of best-fit models of protein evolution. *Bioinformatics* **2011**, *27*, 1164–1165. [[CrossRef](#)] [[PubMed](#)]
60. de Juan, D.; Pazos, F.; Valencia, A.; Evaluation, P.M.; Rehbein, P.; Schwalbe, H.; Jones, D.T.; Buchan, D.W.A.; Cozzetto, D.; Pontil, M.; et al. ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics* **2015**, *21*, 2104–2105.
61. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)] [[PubMed](#)]
62. Stamatakis, A.; Ludwig, T.; Meier, H. RAxML-III: A fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **2005**, *21*, 456–463. [[CrossRef](#)] [[PubMed](#)]
63. Huelsenbeck, J.P.; Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **2001**, *17*, 754–755. [[CrossRef](#)]
64. Hopper, A.K.; Martinis, S.A.; Schimmel, P.; Kisselev, L.; Wolfson, A.; Melton, D.A.; De Robertis, E.M.; Cortese, R.; Popenko, V.I.; Wolfe, C.L.; et al. Nuclear functions charge ahead. *Science* **1998**, *282*, 2003–2004. [[CrossRef](#)]
65. Lund, E. Proofreading and Aminoacylation of tRNAs Before Export from the Nucleus. *Science* **1998**, *282*, 2082–2085. [[CrossRef](#)]
66. Szymanski, M.; Deniziak, M.A.; Barciszewski, J. Aminoacyl-tRNA synthetases database. *Nucleic Acids Res.* **2001**, *29*, 288–290. [[CrossRef](#)] [[PubMed](#)]
67. Nathanson, L.; Deutscher, M.P. Active aminoacyl-tRNA synthetases are present in nuclei as a high molecular weight multienzyme complex. *J. Biol. Chem.* **2000**, *275*, 31559–31562. [[CrossRef](#)] [[PubMed](#)]
68. Ko, Y.G.; Kang, Y.S.; Kim, E.K.; Park, S.G.; Kim, S. Nucleolar localization of human methionyl-tRNA synthetase and its role in ribosomal RNA synthesis. *J. Cell Biol.* **2000**, *149*, 567–574. [[CrossRef](#)] [[PubMed](#)]
69. Kaminska, M.; Havrylenko, S.; Decottignies, P.; Le Maréchal, P.; Negrutskii, B.; Mirande, M. Dynamic organization of aminoacyl-tRNA synthetase complexes in the cytoplasm of human cells. *J. Biol. Chem.* **2009**, *284*, 13746–13754. [[CrossRef](#)] [[PubMed](#)]
70. Duchene, A.-M.; Giritch, A.; Hoffmann, B.; Cognat, V.; Lancelin, D.; Peeters, N.M.; Zaepfel, M.; Marechal-Drouard, L.; Small, I.D. Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16484–16489. [[CrossRef](#)] [[PubMed](#)]
71. Ito, J.; Batth, T.S.; Petzold, C.J.; Redding-Johanson, A.M.; Mukhopadhyay, A.; Verboom, R.; Meyer, E.H.; Millar, A.H.; Heazlewood, J.L. Analysis of the *Arabidopsis* cytosolic proteome highlights subcellular partitioning of central plant metabolism. *J. Proteome Res.* **2011**, *10*, 1571–1582. [[CrossRef](#)]

72. Araiso, Y.; Huot, J.L.; Sekiguchi, T.; Frechin, M.; Fischer, F.; Enkler, L.; Senger, B.; Ishitani, R.; Becker, H.D.; Nureki, O. Crystal structure of *Saccharomyces cerevisiae* mitochondrial GatFAB reveals a novel subunit assembly in tRNA-dependent amidotransferases. *Nucleic Acids Res.* **2014**, *42*, 6052–6063. [[CrossRef](#)]
73. Frechin, M.; Duchêne, A.-M.; Becker, H.D. Translating organellar glutamine codons: A case by case scenario? *RNA Biol.* **2009**, *6*, 31–34. [[CrossRef](#)]
74. Freist, W.; Logan, D.T.; Gauss, D.H. Glycyl-tRNA synthetase. *Biol. Chem. Hoppe Seyler* **1996**, *377*, 343–356.
75. Hipps, D.; Shiba, K.; Henderson, B.; Schimmel, P. Operational RNA code for amino acids: Species-specific aminoacylation of minihelices switched by a single nucleotide. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5550–5552. [[CrossRef](#)]
76. Duchêne, A.M.; Peeters, N.; Dietrich, A.; Cosset, A.; Small, I.D.; Wintz, H. Overlapping Destinations for Two Dual Targeted Glycyl-tRNA Synthetases in *Arabidopsis thaliana* and *Phaseolus vulgaris*. *J. Biol. Chem.* **2001**, *276*, 15275–15283. [[CrossRef](#)] [[PubMed](#)]
77. Halary, S.; McInerney, J.O.; Lopez, P.; Baptiste, E. EGN: A wizard for construction of gene and genome similarity networks. *BMC Evol. Biol.* **2013**, *13*, 146. [[CrossRef](#)] [[PubMed](#)]
78. Brindefalk, B.; Viklund, J.; Larsson, D.; Thollesson, M.; Andersson, S.G.E. Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases. *Mol. Biol. Evol.* **2007**, *24*, 743–756. [[CrossRef](#)] [[PubMed](#)]
79. Makarova, K.S. Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. *Nucleic Acids Res.* **2005**, *33*, 4626–4638. [[CrossRef](#)] [[PubMed](#)]
80. Ribas de Pouplana, L.; Brown, J.R.; Schimmel, P. Structure-based phylogeny of class IIa tRNA synthetases in relation to an unusual biochemistry. *J. Mol. Evol.* **2001**, *53*, 261–268. [[CrossRef](#)] [[PubMed](#)]
81. Roy, H.; Ling, J.; Alfonzo, J.; Ibba, M. Loss of editing activity during the evolution of mitochondrial phenylalanyl-tRNA synthetase. *J. Biol. Chem.* **2005**, *280*, 38186–38192. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).