The evolution of aminoacyl-tRNA synthetases in chromerids

Abdoallah Sharaf[1, 2], Kateřina Jiroutová[1] and Miroslav Oborník[1, 3, 4]

[1] Biology Centre (ASCR)· Institute of Parasitology, 37005, České Budějovice, Czech Republic, [2] Genetic Dept. Faculty of Agriculture, Ain Shams University, Cairo, 11241, Egypt, Faculty of Science, [3] University of South Bohemia, 37005 České Budějovice, Czech Republic, [4] Institute of Microbiology, Opatovický mlýn, Třeboň, Czech Republic.

Introduction

Chromerids are algae possessing a complex plastid surrounded by four membranes. These algae are the closest known phototrophic relatives to apicomplexan parasites [1]. The focus of this study is the evolutionary origin of the genes for chromerid aminoacyl-tRNA synthetases (aaRS), key components of the protein translation machinery that catalyze two basic reactions: (1) activation of amino acids via the formation of aminoacyl adenylates and (2) linking the activated amino acid to the cognate tRNA. They are required in every compartment where translation takes place. Chromerids have three such compartments: the plastid, the mitochondrion and the cytosol. Because of their ubiquity, conservation, specificity, and defined interactions in protein synthesis, the aaRSs represent important keys for resolving early cellular evolution [2]. Although canonical patterns have been partially eroded by duplication, divergence, and horizontal gene transfers [3], traces of ancestral relationships are still evident in many aaRS trees [2].

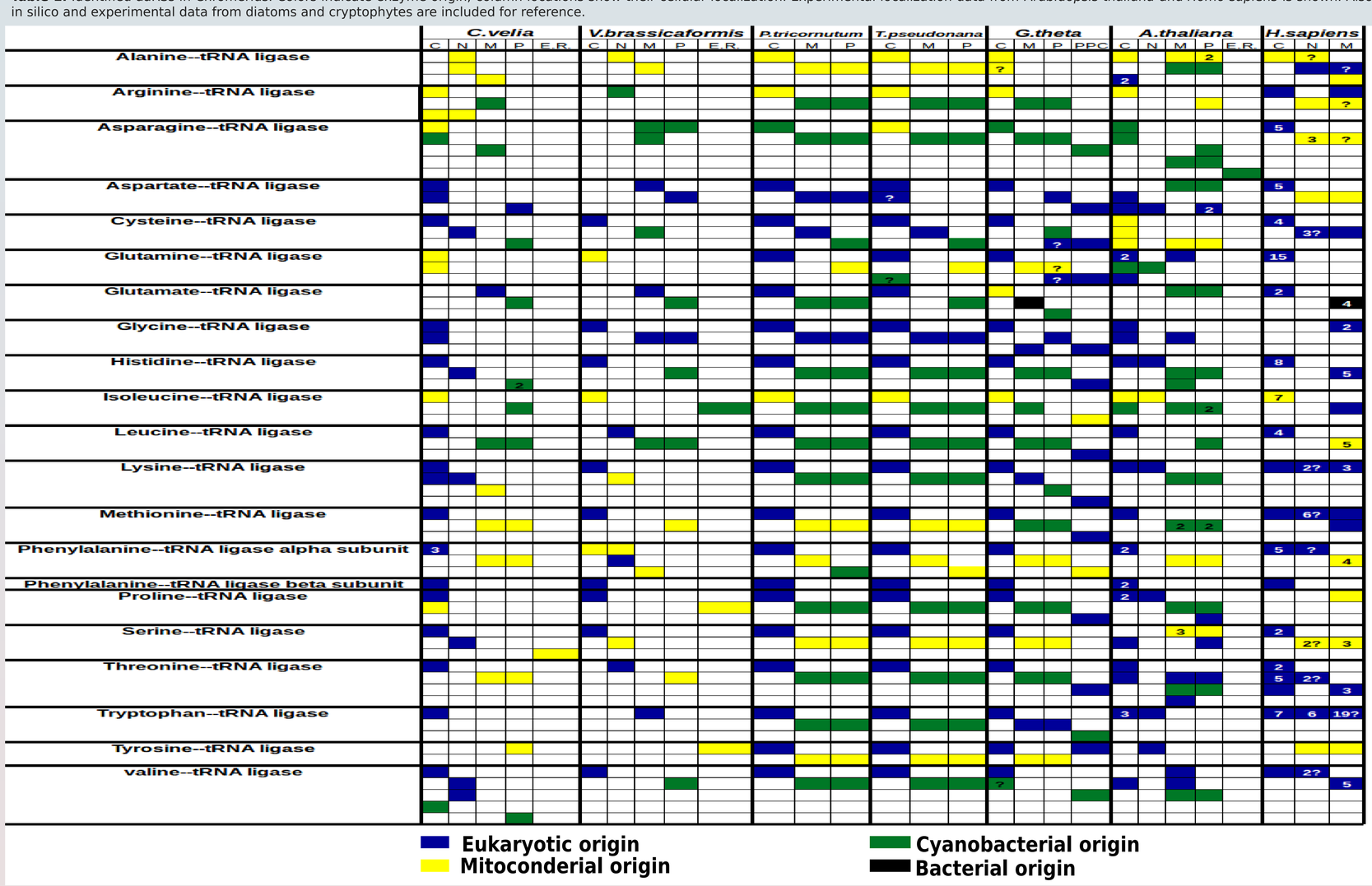
Methods

The total predicted proteins of *Chromera velia* and *Vitrella brassicaformis* [4] were searched for aaRSs with BLASTp using previously characterized aaRS amino acid sequences from eukaryotes, bacteria, and archaea as queries. We predicted the intercellular localization of the enzymes using SignalP 3.0, SignalP 4.1, ASAFind, TargetP 1.1, iPSORT, WoLF PSORT and Mitoprot.

Phylogenetic trees were constructed using Maximum likelihood [6] and Bayesian [7] methods based on amino acid sequences of the enzymes, Sequences were retrieved from public databases.

Results

Table 1. Identified aaRSs in Chromerids. Colors indicate enzyme origin; column locations show their cellular localization. Experimental localization data from *Arabidopsis thaliana* and *Homo sapiens* is shown. Also, in silico and experimental data from diatoms and cryptophytes are included for reference.



Conclusions

Our BLAST search of the total predicted proteins of *C. velia* and *V. brassicaformis* identified 50 and 38 aaRSs loci respectively (Table 1). Although this is substantially fewer than the number (roughly 60) required to provide the three compartments (cytosol, plastid, and mitochondrion) with their own unique proteins, aaRSs are highly conserved and it is unlikely that any were missed by our search.

Forty-five percent of *C. velia*'s aaRSs are encoded by three distinct loci, whereas 35% of aaRSs are encoded by two distinct loci. Interestingly, both tryRS and trpRS are encoded by only one locus and valRS is encoded by five loci. In contrast, 70% of the *V. brassicaformis* aaRSs are encoded by just two distinct loci. Only asnRS was encoded by three loci. ArgRS, GlnRS, TrpRS and tysRS are encoded by only one locus. PheRS was encoded by four loci, three for alpha subunit and one for subunit beta. Dual targeting were identified in three aaRSs (AsnRS, GlyRS, LeuRS and pheRS subunit alpha).

Phylogenetic trees of alaRS, AspRS, CysRS, GluRS, LysRS, SerRS, TyrRS and valRS conform to the classical 3-domain pattern, proving a good separation of the eukaryotic aaRSs from the bacterial group [2]. Only gluRS in *C.velia* and *V. brassicaformis* show the same gene numbers and evolutionary pattern. All identified genes for alaRS, GluRS and tyrRS in *C. velia* and *V. brassicaformis* had a mitochondrial origin, While aspRS, GlyRS, TrpRS and pheRS subunit beta had a eukaryotic origin. Finally, endosymbiotic gene transfer events were observed in argRS, AsnRS, CysRS, GluRS, HisRS, IleRS, LeuRS and valRS for both chromerids.

References

1. Oborník and Lukeš. 2013.
2. Woese *et al*. 2000.
3. Brown *et al.* 2003.
4. Woo *et al.* 2015.
5. Stamatakis, 2006.
6. Lartillot *et al.* 2009.

Acknowledgements

This work was support by the Czech Science Foundation (GACR 15-17643S)