Does the evolutionary history of aminoacyl-tRNA synthetases explain the loss of mitochondrial tRNA genes?

André Schneider

The importation of cytosolic tRNAs is required for protein synthesis in the mitochondria of the wide variety of eukaryotes that lack a complete set of mitochondrial tRNA genes. The evolutionary history of the process, however, is still enigmatic. The analysis presented here suggests that the loss of distinct mitochondrial tRNA genes was not random and that it might be explained by the differential capabilities of mitochondrial aminoacyl-tRNA synthetases to charge imported eukaryotic-type tRNAs with amino acid.

The absence of a variable number of apparently essential mitochondrially encoded tRNA genes is widespread among eukaryotes. The prokaryotic progenitor of mitochondria had a complete set of tRNA genes. The lack of mitochondrial tRNA genes is, therefore, a derived trait that is generally compensated for by importing the corresponding cytosolic tRNAs. Interestingly, imported tRNAs originate from nuclear genes encoding cytosolic, eukaryotic-type tRNAs involved in cytosolic translation. The extensive distribution among eukaryotes of mitochondrial tRNA importation suggests that tRNA import must have an adaptive advantage and a polyphyletic origin[1].

The large number of complete mitochondrial DNA sequences[2] allows prediction of the codons that cannot be read by endogenous mitochondrial tRNAs and therefore must be decoded by imported ones. The prediction might be limited by putative unconventional tRNA genes, which are difficult to detect, by as yet unknown decoding rules or by tRNA editing. However, experimental evidence suggests that for most organisms, tRNA import is correctly predicted[1,3,4]. It is clear though that the predicted imported tRNAs represent only the cases where, owing to the loss of the endogenous gene, import has been evolutionarily fixed. Examples of the import of apparently non-essential tRNAs have also been described[5,6].

Frequency of tRNA import

Figure 1 shows the amino acids most frequently predicted to be decoded by imported tRNAs in a group of 27 different organisms. All species known to lack at least one mitochondrial tRNA gene were included. For closely related organisms showing identical mitochondrial tRNA gene content (e.g. *Trypanosoma brucei* and *Leishmania*), only one was included in the analysis. I analyzed a total of 206 predicted import events. If in the same organism the import of more than one isoacceptor for a single amino acid is predicted, it was only scored once. Unexpectedly, I observed a distribution that is not random (Fig. 1). Whereas the import frequency of the $tRNA^{Thr}$ was high (22 out of 27 species), $tRNA^{Met}$ and $tRNA^{Trp}$ were predicted to be imported in only 3 and 4 species, respectively.

Imported tRNAs have the problem that they are of eukaryotic origin but must function in the prokaryotic-type translation system of mitochondria. Eukaryotes have a cytosolic and a mitochondrial set of aminoacyl-tRNA synthetases (AARSs), which, for the most part, do not overlap. Whereas their genes are nuclear, the evolutionary origin of organellar AARSs is expected to be from the mitochondrial progenitor, whose closest living relative is *Rickettsia prowazekii*[7].

Involvement of AARSs in the evolution of tRNA import

I propose that one factor that determined the observed import frequency of a given tRNA, was whether it could be charged with amino acid by the corresponding mitochondrial AARS. The proposed hypothesis makes two testable predictions. First, the relative similarities between eukaryotic and mitochondrial AARSs should correlate positively with the import frequencies of their corresponding tRNAs. Second, each AARS recognizes specific identity elements on its respective tRNA substrate[8]. It is therefore expected that the similarity of the identity elements between eukaryotic tRNAs and the bacterial-type tRNAs of the mitochondria should also correlate positively with their observed import frequencies. To test these predictions, it is best to look at a situation resembling the one in ancestral mitochondria as closely as possible. To do this, I used AARS and tRNA sequences from *R. prowazeckii*. For the same reason, the eukaryotic sequences were chosen from organisms with a complete set of mitochondrial tRNA genes (*Homo sapiens*, *Saccharomyces cerevisiae*).

Figure 2a shows that the first prediction holds true, there is indeed a positive correlation between the import frequency of specific tRNA isoacceptors and the relative sequence identity between the corresponding eukaryotic and rickettsial AARSs ($R = 0.679$; $P = 0.003$). I excluded $tRNA^{Gln}$ and $tRNA^{Asn}$ from the analysis because no rickettsial GlnRS and AsnRS exist; these tRNAs are indirectly charged in a two-step process involving the transamidation of misacylated tRNAs[9].

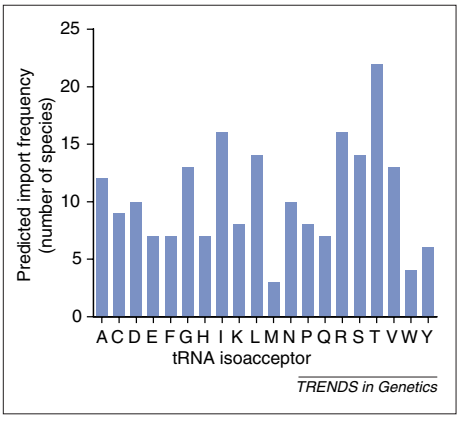

Fig. 1. Predicted import frequency of different tRNA isoacceptors (amino acids are indicated by the one letter code) in 27 different species (14 protozoans and algae: *Acanthamoeba castelanii, Cafeteria roenbergensis, Chlamydomonas eugmatos, Chondrus crispus, Nephroselmis olivacea, Ochromonas danica, Paramecium aurelia, Pedinomonas minor, Phytophthora infestans, Plasmodium falciparum, Porphyra purpurea, Reclinomonas americana, Tetrahymena pyriformis, Trypanosoma brucei*; 4 plants: *Arapidopsis thaliana, Beta vulgaris, Marchantia polymorpha, Secendesmus obliquus*; 1 fungus and 2 slime molds: *Allomyces macrocygnus, Dictyostelium discoideums, Physarum polycephalum*; 6 invertebrates: *Cepaea memoralis, Crassostrea gigas, Echinococcus multiocularis, Metridium senile, Paragonimus westerman, Sarcophyton glaucum).* Import of essential tRNA genes was predicted using the annotated sequences found on The Organelle Genome Megasequencing Program database[2,14], taking into account the different variants of the genetic code[11].

 0168-9525/01/$ – see front matter PII: S0168-9525(01)02439-8

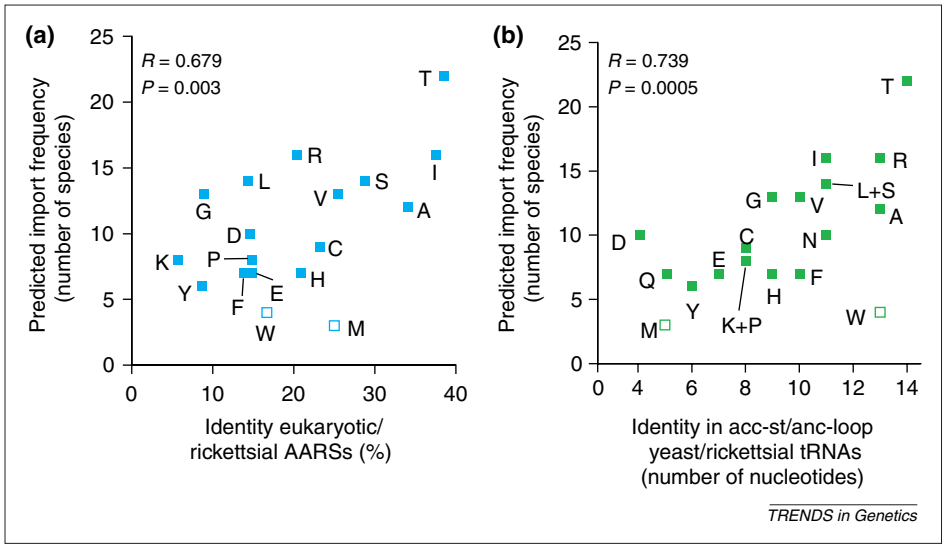

Fig. 2. tRNA import frequency correlates with the capability of mitochondrial AARSs to charge imported tRNA. (a) Percentage identity between corresponding AARSs of eukaryotes and *R. prowazekii* were plotted against the predicted import frequency (shown in Fig. 1) of their substrate tRNAs. The identities were determined using sequence information in the aminoacyl-tRNA synthetases database[15] and the BLOSUM30 alignment program (http://www.expasy.ch/tools/sim-prot.html). In cases where the overlap between the compared AARSs was smaller than the rickettsial protein, they were corrected for the length of the bacterial protein. All percentages are means between independent pairwise comparisons of the *S. cerevisiae/R. prowazekii* and human/*R. prowazekii* proteins. If a AARS activity consists of more than one subunit (e.g. PheRS), or if more than one AARS exists for a single isoacceptor (e.g. SerRS in *S. cerevisiae*), both were included in the analysis and the mean value was used. LysRS and GlyRS of eukaryotes are not obviously phylogenetically related with their rickettsial counterparts[16], the low identity values observed could therefore reflect the presence of some convergent feature present in both proteins. The $tRNA^{Met}$ and $tRNA^{Trp}$ (open squares) were not included in the calculation of the *R* and *P* values for reasons discussed in the text. (b) The number of nucleotides shared between identity elements of the corresponding yeast and rickettsial tRNAs were plotted against their predicted import frequency. The discriminator nucleotide, the first five base pairs of the acceptor stem and the anticodon loop excluding the anticodon itself (acc-st/anc-loop) were compared (maximal possible number of identical nucleotides is 15). In cases where more than one isoacceptor exists in the two organisms, the value for the most similar pair of tRNAs was used. $tRNA^{Met}$ and $tRNA^{Trp}$ (open squares) were excluded for the same reasons as in the analysis above.

Indeed, import frequencies of $tRNA^{Gln}$ and $tRNA^{Asn}$ are quite low as expected from this scenario (Fig. 1). To calculate the *R* and *P* values, I excluded $tRNA^{Met}$ and $tRNA^{Trp}$, for reasons explained below.

What about the second prediction, can it be verified as well? It is difficult to compare identity elements between different tRNA isoacceptors, because they can be quite complex. A survey of all known identity elements, however, shows that, for the large majority of tRNA isoacceptors, they are confined to three regions: the discriminator nucleotide, the first five base pairs of the acceptor stem and the anticodon loop[8]. I therefore compared these regions between rickettsial and yeast tRNAs. The anticodon itself was excluded from the comparison because it must be identical by definition. The graph in Fig. 2b shows that, as predicted, there is a good positive correlation ($R = 0.739$; $P = 0.0005$) between import frequency and similarity of the identity elements. As before, $tRNA^{Met}$ and $tRNA^{Trp}$ are outliers, which can be explained as follows.

Only $tRNA^{Met}$ carrying a formylated methionine functions in prokaryotic-type translation initiation[10]. The limiting factor for the observed low frequency import might therefore not be aminoacylation, but formylation of the already charged $tRNA^{Met}$. This reaction is catalyzed by the methionyl-tRNA transformylase, an enzyme that does not exist in the eukaryotic cytosol and therefore cannot formylate $tRNAs^{Met}$ of eukaryotic origin. Mitochondria of most organisms included in the analysis decode the standard stop codon AGU as tryptophan requiring a $tRNA^{Trp}$ carrying a UCA anticodon[11]. Such a $tRNA^{Trp}$, however, would be harmful when present in the cytosol providing a possible explanation for the low frequency import of $tRNA^{Trp}$ without the need to imply a role for a divergent TrpRS.

Conclusions

In summary, this study shows that the evolutionary history of mitochondrial tRNA import can be retraced by comparing AARSs and tRNAs from *R. prowazeckii*, the closest extant relative of mitochondria, with the corresponding homologs of the eukaryotic cytosol. In species having an incomplete set of mitochondrial tRNA genes, the chance that a certain tRNA gene was lost correlates with the capability of the AARSs of the ancestral mitochondria to charge imported eukaryotic-type tRNAs. There is no evidence for a direct role of AARSs in import of essential mitochondrial tRNAs[12,13]. This analysis therefore supports the idea that the loss of organellar tRNAs genes is determined by the capability of the imported eukaryotic-type tRNA to function in bacterial-type translation and not by the appearance of a selective import mechanism.

Acknowledgements

I thank Thomas Steinger, Dieter Meyer and Louis-Felix Bersier for help with the statistical analysis and Adrian Streit and Tibor Vellai for critical review of the manuscript. This study was supported by grants 31-056825.99 and 4037-55154 from the Swiss National Foundation and by a fellowship of the 'Prof. Dr Max Cloëtta' Foundation.

References

1 Schneider, A. and Marechal-Drouard, L. (2000) Mitochondrial tRNA import: are there distinct mechanisms? *Trends Cell Biol.* 10, 509–513
2 Shimko, N. *et al.* (2001) GOBASE: the organelle genome database. *Nucleic Acids Res.* 29, 128–132
3 Akashi, K. *et al.* (1997) Accumulation of nuclear-encoded threonyl-tRNA(AGU) in mitochondria of the liverwort *Marchantia polymorpha*. *Biochim. Biophys. Acta* 1350, 262–266
4 Rusconi, C.P. and Cech, T.R. (1996) Mitochondrial import of only one of three nuclear-encoded glutamine tRNAs in *Tetrahymena thermophila*. *EMBO J.* 15, 3286–3295
5 Akashi, K. *et al.* (1998) Coexistence of nuclear DNA-encoded valine tRNA(AAC) and mitochondrial DNA-encoded valine tRNA(UAC) in mitochondria of a liverwort *Marchantia polymorpha*. *Nucleic Acids Res.* 26, 2168–2172
6 Tarassov, I. and Martin, R. (1996) Mechanisms of tRNA import into yeast mitochondria: an overview. *Biochimie* 78, 502–510
7 Andersson, S.G.E. *et al.* (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* 396, 133–140
8 Giege, R. *et al.* (1998) Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* 26, 5017–5035
9 Ibba, M. and Soll, D. (2000) Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* 69, 617–650
10 RajBhandary, U.L. (1994) Initiator transfer RNAs. *J. Bacteriol.* 176, 547–552
11 Knight, R.D. *et al.* (2001) Rewriting the keyboard: evolvability of the genetic code. *Nature Rev. Genet.* 2, 49–58
12 Mireau, H. *et al.* (2000) Expression of *Arabidopsis thaliana* mitochondrial alanyl-tRNA synthetase is not sufficient to trigger mitochondrial import of alanyl-tRNA in yeast. *J. Biol. Chem.* 275, 13291–13296

13 Nabholz, C.E. *et al.* (1997) *Leishmania tarentolae* contains distinct cytosolic and mitochondrial glutaminyl-tRNA synthetase activities. *Proc. Natl. Acad. Sci. U. S. A.* 94, 7903–7908
14 Gray, M.W. *et al.* (1998) Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Res.* 26, 865–878
15 Szymanski, M. *et al.* (2001) Aminoacyl-tRNA synthetases database. *Nucleic Acids Res.* 29, 288–290
16 Woese, C.R. *et al.* (2000) Aminoacyl-tRNA synthetase, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.* 64, 202–236

André Schneider
Dept of Biology/Zoology,
University of Fribourg,
Chemin du Musee 10, CH-1700 Fribourg,
Switzerland.
e-mail: andre.schneider@unifr.ch

Meeting Report

Well developed: the SDB reaches 60

Cecilia B. Moens

The 60th Annual Meeting of the Society for Developmental Biology was held at the University of Washington, Seattle, WA, USA, from 18 to 22 July 2001.

The central question that occupies developmental biologists is how the information encoded in the genome of the fertilized egg brings about the unfolding of the adult body plan. At the recent meeting of the Society for Developmental Biology, the parallel approaches of genetics, cell biology, biochemistry and experimental embryology were brought to bear on the many forms this question takes. A few particularly interesting presentations are highlighted here. Although the vast majority of the work used the small group of traditional model organisms famous for their genetic and/or experimental tractability, innovative new approaches using systems such as the bat and the stickleback fish are allowing researchers to examine how genetic changes lead to the evolution and development of new morphologies.

Co-receptors in developmental signaling events

Recently, the essential roles of putative co-receptors in developmental signaling have been demonstrated, both in vertebrates and invertebrates. In vertebrates, Cripto/Oep/Frl-1, a membrane-anchored protein with homology to the epidermal growth factor (EGF), is required for signalling by Nodal, a member of the transforming growth factor β (TGFβ) family (reviewed in Ref. 1). In flies, the glypicans Dally and Dally-like, both heparin sulfate proteoglycans, modulate Wnt signaling by limiting the diffusion or degradation of ligand (reviewed in Ref. 2). Two talks described how these putative co-receptors function during vertebrate development,

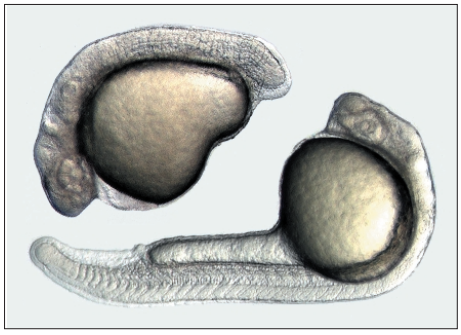

Fig. 1. Dramatic shortening of the embryonic mutant body (top) compared with wild type (bottom) at day one of development is a consequence of the zebrafish *knypek* gene mutation. *knypek* encodes a heparan sulfate proteoglycan of the glypican family that promotes noncanonical Wnt signaling to mediate cell polarization underlying gastrulation convergent-extension movements. Photograph courtesy of Lila Solnica-Krezel.

and how this function could be modulated by glycosylation.

During gastrulation, a complex series of cell movements bring about the narrowing and lengthening of the embryonic axis. Lilianna Solnica-Krezel (Vanderbilt University, Nashville, TN, USA) described a zebrafish mutant, *knypek*, in which cells undergoing these convergent-extension movements fail to orient themselves appropriately, resulting in an embryo with a short, broad axis. Positional cloning of *knypek* showed that it encodes a glypican related to Dally. The *knypek* phenotype is similar to that of another zebrafish mutant, *silberblick*, in which a Wnt (Wnt11) that functions in the noncanonical Wnt pathway is inactivated. As originally described in *Drosophila*, Wnt signaling that occurs through β-catenin accumulation is known as the 'canonical' pathway, whereas 'noncanonical' Wnt signaling is independent of β-catenin (reviewed in Ref. 3). Solnica-Krezel and colleagues demonstrated both that loss of *knypek* function enhances the *silberblick* phenotype and that injection of *knypek* mRNA facilitates the rescue of *silberblick* embryos by *wnt11* mRNA injection. These results implicate heparan sulfate proteoglycans in noncanonical Wnt signaling in vertebrates. Processes controlled by canonical Wnt signaling, meanwhile, appear unaffected in *knypek* mutants. In *Drosophila*, *dally* and *dally-like* function in the canonical *wingless* pathway. Heparan sulfate proteoglycans might thus affect different signaling pathways in a context-dependent manner. The molecular mechanisms of Knypek action – whether Knypek binds Wnt ligand or receptor directly, and how such an interaction affects signaling – remain to be elucidated.

Previous work in a number of vertebrate systems showed that Nodal signaling is involved in mesoderm induction in a pathway that requires both a type-1 activin receptor, Alk4, and the co-receptor Cripto. Malcom Whitman (Harvard Medical School, Boston, MA, USA) and colleagues have shown that Cripto interacts directly with the Alk4 receptor, and that this interaction is required for Nodal responsiveness in *Xenopus*. In his talk, Whitman noted that during normal *Xenopus* development, competence to respond to Nodal signaling does not necessarily correlate with the onset of either *cripto* or *alk4* expression, suggesting that post-transcriptional modification of one of these molecules could control when, and possibly where, responses to Nodal signaling can occur. Whitman, with colleagues at BioGen, found that the extracellular EGF motif of Cripto is glycosylated at a consensus site for *O*-linked fucosylation. Fucosylation of the EGF motifs of the Notch receptor, catalyzed by the Fringe glycosyltransferase, was recently shown to modulate this important developmental signaling pathway (reviewed in Ref. 4). When the fucosylation site in Cripto is mutated, Nodal signaling is

 0168-9525/01/$ – see front matter PII: S0168-9525(01)02488-X