Gene 475 (2011) 22–29

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Gene recruitment – A common mechanism in the evolution of transfer RNA gene families

Xiujuan Wang, Dennis V. Lavrov *

Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50011, USA

ARTICLE INFO

Article history:
Accepted 21 December 2010
Available online 30 December 2010

Received by I.B. Rogozin

Keywords:
Transfer RNA
Gene recruitment
Genetic code
Evolution

ABSTRACT

The evolution of alloacceptor transfer RNAs (tRNAs) has been traditionally thought to occur vertically and reflect the evolution of the genetic code. Yet there have been several indications that a tRNA gene could evolve horizontally, from a copy of an alloacceptor tRNA gene in the same genome. Earlier, we provided the first unambiguous evidence for the occurrence of such "tRNA gene recruitment" in nature — in the mitochondrial (mt) genome of the demosponge *Axinella corrugata*. Yet the extent and the pattern of this process in the evolution of tRNA gene families remained unclear. Here we analyzed tRNA genes from 21 mt genomes of demosponges as well as nuclear genomes of rhesus macaque, chimpanzee and human. We found four new cases of alloacceptor tRNA gene recruitment in mt genomes and eleven cases in the nuclear genomes. In most of these cases we observed a single nucleotide substitution at the middle position of the anticodon, which resulted in the change of not only the tRNA's amino-acid identity but also the class of the amino-acyl tRNA synthetases (aaRSs) involved in amino-acylation. We hypothesize that the switch to a different class of aaRSs may have prevented the conflict between anticodon and amino-acid identities of recruited tRNAs. Overall our results suggest that gene recruitment is a common phenomenon in tRNA multigene family evolution and should be taken into consideration when tRNA evolutionary history is reconstructed.

1. Introduction

Transfer RNA (tRNA) is a small RNA molecule that plays a central role in protein biosynthesis. Each tRNA carries a specific amino acid to the ribosome and recognizes one or several specific codons in mRNA, functioning as a liaison between DNA encoded genetic information and its expression in proteins (Crick, 1958). Based on their aminoacylation identity, tRNAs are subdivided into 20 amino acid accepting groups (alloacceptors), each of them comprising one to several tRNAs (isoacceptors) that translate synonymous codons (Saks et al., 1998). The aminoacylation reaction is catalyzed by two groups of proteins, class I and class II aminoacyl-tRNA synthetases (AARSs), unrelated in both sequence and structure (Ibba et al., 1997a; O'Donoghue and Luthey-Schulten, 2003). Ten alloacceptor tRNAs (for Met, Val, Ile, Leu, Cys, Glu, Gln, Arg, Trp, and Tyr) are always aminoacylated by class I AARSs while 9 (for Ala, His, Pro, Thr, Ser, Gly, Phe, Asp, and Asn) — by class II AARSs. Interestingly, tRNAs for Lys can be aminoacylated by either a class I or class II AARS (Ibba et al., 1997b). The number of tRNA genes varies from organism to organism. There are 86 tRNA genes in *Escherichia coli* DH10B, 286 in *Saccharomyces cerevisiae*, 298 in *Drosophila melanogaster*, 513 in human and 630 in *Arabidopsis thaliana* (Schattner et al., 2005). By contrast, mitochondrial genomes do not typically encode multiple isoacceptor tRNAs and are able to translate all codons with as few as 22 tRNAs (Marck and Grosjean, 2002). Although it is clear that tRNA gene families have diverged extensively among different organisms, the details of this evolution remain elusive.

The traditional view of tRNA evolution presumes that alloacceptor tRNAs coevolve with the genetic code while isoacceptor tRNA genes evolve by gene duplication from a common ancestor having the same amino-acid identity (Fitch and Upper, 1987; Wong, 1975; Xue et al., 2003). If this is indeed the case, then the modern phylogenetic relationships of alloacceptor tRNAs can be used to elucidate the evolution of the genetic code (Eigen et al., 1989; Fitch and Upper, 1987; Wong, 1975) and/or to infer phylogenetic relationships among early diverging groups of organisms (Kumazawa and Nishida, 1993; Xue et al., 2003). However, several previous studies reported an unexpectedly high similarity among alloacceptor tRNA genes in some modern organisms and suggested that at least some tRNAs could have evolved independently of the genetic code from duplicated genes for alloacceptor tRNAs (Burger et al., 1995; Cedergren et al., 1980). The latter hypothesis, named "gene recruitment", received support from experimental studies in *E. coli*, in which a $tRNA^{Arg}_{UCU}$ with a point mutation that changed its anticodon from UCU to UGU replaced the function of the $tRNA^{Thr}_{UGU}$ (Saks et al., 1998) and our analysis of alloacceptor tRNA genes in mitochondrial genomes of three

Abbreviation: mtDNA, mitochondrial genome.

* Corresponding author. 253 Bessey Hall, Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50011, USA. Tel.: +1 515 294 9091; fax: +1 515 294 1337.
E-mail address: dlavrov@iastate.edu (D.V. Lavrov).

0378-1119/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2010.12.009

demosponges (Lavrov and Lang, 2005). In the latter study, we provided several lines of evidence for the evolution of three tRNA genes in mtDNA of the demosponge *Axinella corrugata* from alloacceptor tRNA genes in the same genome and found a signature of gene recruitment in mitochondrial genomes of the green alga *Scenedesmus obliquus* and ichthyosporean *Amoebidium parasiticum*. Although these results indicate that some tRNA genes in modern organisms evolved by gene recruitment, the importance of this process in tRNA evolution remains uncertain and is the focus of the present study.

In order to evaluate the importance of gene recruitment in tRNA evolution, we analyzed two different datasets. First, we expanded our mitochondrial genome dataset to 21 species of demosponges representing all currently recognized orders in this group except two (Dendroceratida and Dictyoceratida) that experienced significant tRNA gene loss (Wang and Lavrov, 2008). Second, we analyzed all identified tRNA genes in the nuclear genomes of human, chimpanzee, and rhesus macaque (Chan and Lowe, 2009). Each of these datasets has its unique advantages along with some limitations. Demosponge mitochondrial genomes are a convenient system to study tRNA evolution due to their relatively small size (~19 kb), conserved set of tRNA genes, conventional secondary structures of encoded tRNAs, and low rate of sequence evolution (Wang and Lavrov, 2008). In addition, the set of tRNA genes in demosponges does not usually contain any redundant tRNAs; hence any observed change in tRNA anticodon identity likely represents a functional gene recruitment event rather than a deleterious mutation corrected by tRNA editing. By contrast, nuclear genomes contain multiple isoacceptor tRNA genes and, in some cases, multiple copies of each isoacceptor. Thus, at least in theory, they should be more prone to gene recruitment, although the functional significance of observed changes may remain uncertain.

2. Methods

2.1. Construction of mitochondrial and nuclear tRNA gene datasets

For the first part of this study, we analyzed 21 complete mitochondrial genome sequences of demosponges that did not undergo a significant loss of tRNA genes (Wang and Lavrov, 2008): *Agelas schmidti* (NC_010213), *Amphimedon compressa* (NC_010201), *Amphimedon queenslandica* (NC_008944), *Aplysina fulva* (NC_010203), *Axinella corrugata* (NC_006894), *Callyspongia plicifera* (NC_010206), *Ectyoplasia ferox* (NC_010210), *Ephydatia muelleri* (NC_010202), *Halisarca dujardini* (NC_010212), *Iotrochota birotulata* (NC_010207), *Negombata magnifica* (NC_010171), *Ptilocaulis walpersi* (NC_010209), *Topsentia ophiraphidites* (NC_010204), *Xestospongia muta* (NC_010211), *Oscarella carmela* (NC_009090), *Plakortis angulospiculatus* (NC_010217), *Chondrilla* aff. *nucula* CHOND (NC_010208), *Cinachyrella kuekenthali* (NC_010198), *Geodia neptuni* (NC_006990), *Suberites domuncula* (NC_010496), *Tethya actinia* (NC_006991). There were a total of 500 tRNA genes in these 21 demosponge mitochondrial genomes. For the second part, we analyzed nuclear tRNA genes of human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), and rhesus monkey (*Macaca mulatta*), three closely related species with well annotated genomes. The tRNA gene sequences were downloaded from the Genomic tRNA (GtRNA) database (Chan and Lowe, 2009) and filtered of pseudogenes and 100% identical copies. The final nuclear dataset consisted of 901 nuclear tRNA genes.

2.2. Sequence alignment and pairwise identity analysis

The mitochondrial and nuclear tRNA gene sequences were manually aligned based on the inferred tRNA secondary structures. For each dataset, pairwise sequence similarities between alloacceptor tRNAs within the same species and between isoacceptor tRNAs among different species were calculated as the percentages of matched nucleotides in the alignments. The distributions of such identities were plotted for each group of tRNAs using the R package (http://www.r-project.org/). The probabilities of observing high pairwise sequence identities for candidate tRNA gene recruitment cases were calculated in the R package based on normal distributions. Normal distribution assumptions were verified using Q–Q plots.

2.3. Phylogenetic analysis of mitochondrial and nuclear tRNA genes

Neighbor-joining analysis with 1000 bootstrap replicates was conducted for each dataset based on uncorrected ('p') pairwise distances using the PAUP* 4.0b10 program (Swofford, 2002). Additional phylogenetic analyses using other methods and/or models of sequence evolution (eg. corrected distances) were also conducted but produced less stable results likely due to the overparameterization problem in estimating model parameters from very short tRNA sequences (Sullivan and Joyce, 2005). Sequences of anticodons, variable-length portions of the extra loop, and introns (for nuclear tRNAs) were excluded from all phylogenetic analyses. For presentation purposes, both mitochondrial and nuclear tRNA gene phylogenies were arbitrarily rooted using *trnP* as an outgroup.

3. Results

3.1. Four new cases of alloacceptor tRNA gene recruitment in demosponge mitochondrial genomes

Mitochondrial genomes of most demosponges contain a conserved set of 24–25 tRNA genes comprised of two genes for arginine, isoleucine, leucine, and serine tRNAs, one or two genes for methionine tRNA, and a single gene for tRNA for every other amino acids (Wang and Lavrov, 2008). We analyzed tRNA gene sequences from 21 mitochondrial genomes of demosponges representing all but two recognized orders in this group. The neighbor-joining (NJ) tree generated using uncorrected p-distances showed that most of the equivalent tRNA genes (with the same amino-acid and anticodon identities) from different species form well-defined clades, an indication of their orthologous relationship. However, several genes appeared at unexpected positions on the tree — a potential indication of gene recruitment (Fig. 1). Similar results were obtained in phylogenetic analyses using other methods and/or models of sequence evolution, although overparameterization (Sullivan and Joyce, 2005) appeared to be a problem when more complex models of evolution were applied to such short sequences.

Although the results of a phylogenetic analysis provide an important indication of potential gene recruitments, an unexpected position of a gene on a phylogenetic tree may result from various factors (Brinkmann et al., 2005). Hence, we investigated these potential tRNA gene recruitment cases using two additional criteria identified in our previous study (Lavrov and Lang, 2005): 1) high sequence similarity between alloacceptor tRNA genes in the same genome and 2) a change in the position of the tRNA gene inferred to undergo gene recruitment relative to equivalent genes in other species.

A recent gene recruitment event should manifest itself through unexpectedly high sequence similarity between genes for alloacceptor tRNAs in the same genome and/or through unexpectedly low sequence similarity between genes for equivalent tRNAs among different genomes. In demosponges, the sequence similarities among equivalent tRNAs in different genomes are usually much higher than among alloacceptor tRNAs in the same genome and these measures form two largely non-overlapping normal distributions with means of 74.5% (standard deviation = 10.1%) and 46.5% (standard deviation = 7.5%), respectively (Fig. 2A). However, our analysis revealed the presence of outliers in both of these distributions, most of which correspond to tRNAs inferred to be involved in alloacceptor tRNA gene recruitment based on the results of the phylogenetic analysis. Nevertheless, some tRNA genes found at an unexpected position in