BioSystems 114 (2013) 155–163

Contents lists available at ScienceDirect

BioSystems

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

Tetracoding increases with body temperature in Lepidosauria

Hervé Seligmann[a,b,c,*], Antonieta Labra[b,d]

[a] National Natural History Museum Collections, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel
[b] Center for Ecological and Evolutionary Synthesis, Department of Bioscience, University of Oslo, Blindern, N-0316 Oslo, Norway
[c] Department of Life Sciences, Ben Gurion University, 84105 Beer Sheva, Israel
[d] Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Chile

ARTICLE INFO

Article history:
Received 11 July 2013
Received in revised form 4 September 2013
Accepted 5 September 2013

Keywords:
Suppressor tRNAs
Genetic code expansion
Synonymous codon usage
tRNA cloverleaf structure
Developmental stability

ABSTRACT

Codons expanded by a silent position (quadruplet or tetracodons) may solve the conundrum that at life's origins, the weak tricodon–anticodon interactions could not promote translation in the absence of complex ribosomes. Modern genomes have isolated tetracodons resulting from insertion mutations. Some bioinformatic analyses suggest that tetracoding stretches overlap with regular mitochondrial protein coding genes. These tetragenes are probably decoded by (antisense) tRNAs with expanded anticodons. They are GC-rich, which produce stronger basepairs than A:T interactions, suggesting expression at high temperatures. The hypothesis that tetracoding is an adaptation to high temperatures is tested here by comparing predicted mitochondrial tetracoding in Lepidosauria (lizards, amphisbaenia, and *Sphenodon*), in relation to body temperature, expecting more tetracoding in species with high body temperature. The association between tRNAs with expanded anticodons and tetracoding previously described for mammals and *Drosophila* is confirmed for Lepidosauria. Independent evidence indicates that tetracoding increases with body temperature, supporting the hypothesis that tetracoding is an adaptation for efficient translation when conditions (temperature) make triplet codon-anticodons too unstable to allow efficient protein elongation.

1. Introduction

The origin of life is inherently connected with the early formation of the complex molecular machinery that codes and produces proteins. Ribosome-free translation of mRNAs seems impossible because interactions between the triplet codon and the tRNA's matching anticodon are too weak for efficient peptide elongation (Baranov et al., 2009). This point is even more extreme when a thermophilic origin of the genetic code is assumed (Di Giulio, 2000, 2003). However, it seems unlikely that at the origins of life, structures as complex as ribosomes were available. This conundrum seems circumvented if codons were longer, by at least one nucleotide, and if the ancestral genetic code was based on quadruplet codons or tetracodons (Baranov et al., 2009). Hypothetically, this ancestral genetic code used a subset of 64 tetracodons, called the tesserae, among the 256 potential tetracodons, and chosen based on principles of codon symmetry and error prevention (Gonzalez et al., 2012). This seems justified because error prevention is more important than usually assumed (Warnecke and Hurst, 2011). In fact, error prevention explains the modern genetic code's structure at the level of impacts of substitution mutations (e.g. Di Giulio, 1989; Haig and Hurst, 1991; Szathmary and Zintzaras, 1992; Freeland and Hurst, 1998; Freeland et al., 2000; Woese et al., 2000; Sella and Ardell, 2006; Novozhilov et al., 2007), deletion mutations (Jestin and Kempf, 1997), amino acid misinsertions due to tRNA misloading and codon-anticodon mismatch (Seligmann, 2010a, 2011a, 2012a), protein folding (Guilloux and Jestin, 2012) and missense translation (Seligmann and Pollock, 2003, 2004; Itzkovitz and Alon, 2007; Seligmann, 2007, 2010b, 2012a; Pienaar and Viljoen, 2008).

Some tRNAs have the ability to decode tetracodons, and were first discovered as frameshift mutation suppressors (Riddle and Carbon, 1973; Sroga et al., 1992; Tuohy et al., 1992). It has been suggested that these have the ability to read occasional 'extra' nucleotides in mitochondrial genes of birds and turtles (Mindell et al., 1998), and other Metazoas (review in Seligmann, 2012b). In addition, in ciliate genomes, tetracodons are very frequent (euplotids, Klobutcher and Farabaugh, 2002; Klobutcher, 2005). The mechanisms for tetracodon decoding seem multiple, and are not straightforward (Atkins and Bjork, 2009), as tRNAs with expanded anticodons are in some cases involved (Walker and

* Corresponding author at: Unité de Recherche sur les Maladies Infectieuses et Tropicales Émergentes, Faculté de Médecine, URMITE CNRS-IRD 198 UMR 6236, Université de la Méditerranée, Marseille, France. Tel.: +33 491 324375.
E-mail addresses: varanuseremius@gmail.com (H. Seligmann), a.l.lillo@bio.uio.no (A. Labra).

0303-2647/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.biosystems.2013.09.002

Frederick, 2006), but also some tRNAs with regular anticodons have that ability of decoding tetracodons (O'Connor et al., 1989; Dunham et al., 2007). Some comparative analyses of mitochondrial genomes suggest that large loops of tRNA sidearms are involved in tetradecoding, putatively through crossovers between the anticodon and a sidearm, or because tRNA halves also function in translation (Seligmann, 2013a). This hypothesis of anticodons in tRNA sidearms is in line with comparative analyses of split tRNA genes (Di Giulio, 1992, 1995, 1999, 2004, 2006, 2008a,b, 2009, 2012, 2013), the existence of armless tRNAs such as in roundworm Enoplea mitochondria (Jühling et al., 2012), and the fact RNA transcribed from the mitochondrial light strand replication origin forms a stem-loop hairpin that is aminoacylated and whose 3′ extremity is extended by the standard CCA extension of mitochondrial tRNAs (Yu et al., 2008). Expanded loops in some tRNA sidearms might be the answer why some regular tRNAs have the ability to read tetracodons.

The possibility of excision of the extra nucleotide, in some cases, cannot be excluded. This would not imply tRNA decoding, but some unknown mechanism of RNA editing. This mechanism is irrelevant to the working hypothesis developed here, as it would not involve interactions between tetracodons and tetra-anticodons.

The natural history of tetracoding has not yet been studied, and the occurrence of tetragenes is not yet confirmed by direct experiments. However, the computational evidence is coherent and in artificial setups, tetracoding actually occurs: tRNAs with expanded anticodons have been used for biotechnological applications to insert unnatural amino acids in proteins (Moore et al., 2000a,b; Magliery et al., 2001; Wang et al., 2001, 2010; Anderson et al., 2002; Rodriguez et al., 2007; Chen and Schindlinger, 2010; Neumann et al., 2010). Note that ribosomes were also artificially selected to improve tetradecoding (Wang et al., 2007, 2010; Neumann et al., 2010).

The hypothesis that natural encoding of polypeptides by presumably continuous stretches of quadruplet codons was only recently tested by comparative analyses of mitochondrial genomes of mammals and *Drosophila* (Seligmann, 2012b). Tessera tetracodons are overrepresented in the predicted mammal mitochondrial tetragenes, a verification of the tesserae hypothesis (Gonzalez et al., 2012). The tesserae hypothesis is tailored to the vertebrate genetic code due to its inherent symmetry properties (Gonzalez et al., 2012). It is unclear to what extents it is adequate for other genetic codes. These predicted tetragenes are relatively GC-rich as compared to the rest of the genome (Seligmann, 2012b). This suggests that tetragenes are adapted for efficient translation at high temperatures, and is in line with the rationale that four pairs of interacting nucleosides yield more stable codon–anticodon interactions than three pairs of interacting nucleosides.

The number of predicted tetragenes per mitochondrial genome coevolves with numbers of predicted antisense tRNAs with expanded anticodons in the same genome, for each a sample of mammal and of *Drosophila* mitochondria (Seligmann, 2012b). Similarly, numbers of predicted antisense tRNAs with expanded anticodons predict the frequency of isolated quadruplet codons within regular mitochondrial protein coding genes of turtles (Seligmann, 2012c), and in some other groups, notably birds (Seligmann, 2012b).

Computational and comparative analyses suggest translational activity by antisense tRNAs (Seligmann, 2010c). These include the observation that antisense tRNAs with anticodons matching stop codons are avoided (Seligmann, 2010d), that predicted antisense tRNA anticodon numbers increase with genomic codon usages (Seligmann, 2010c, 2013a, 2011), that predicted numbers of antisense tRNA properties matching translational activity increase with observed antisense tRNA abundances (in *Drosophila* mitochondria, Seligmann, 2012d), and that associations exist between mutation pathogenicity and their effects on the antisense, rather than the sense, tRNA's cloverleaf folding capacity (Seligmann, 2011c). However, antisense tRNA translational activity has not yet been demonstrated by direct observations (Brzezniak et al., 2011). Nevertheless, besides the coevolution between tetragenes and antisense tRNAs with expanded anticodons (Seligmann, 2012b), antisense tRNAs with anticodons matching stops coevolve with overlapping genes that include stop codons and could not be expressed otherwise (Faure et al., 2011; Seligmann, 2011d, 2012c,d,e, 2013b). This suggests that several genetic coding systems coexist in parallel to the regular, recognized coding system, putatively adapted for specific conditions that remain unknown, but would associate with the expression of antisense tRNAs. Further results also indicate that systematic nucleotide permutations, by exchanging during polymerization (or RNA editing) between nucleosides, reveals unsuspected coding potential in mitochondrial genomes. Empirical results indicate that symmetric nucleoside exchanges (i.e. $A \leftrightarrow C$, Seligmann, 2013c) are more frequent than those involving the potentially more complex asymmetric exchanges (i.e. $A \rightarrow C \rightarrow G \rightarrow A$, Seligmann, 2013d).

Tetragenes would be an additional coding system increasing the coding density of polynucleotide sequences. The fact that tetracodon hybridization with tetra-matching tetra-anticodons, yields more stable duplexes, suggests that this system is adapted for expression at high temperatures, where tricodons are less efficient due to their less stable interactions with anticodons. The association between tetracoding and increased GC contents (Seligmann, 2012b) is a further indication that tetracoding is an adaptation of translation to high temperatures.

The working hypothesis examined here is that tetracoding increases with body temperatures. The adequate combination of complete mitochondrial genome and body temperature data exists for Lepidosaurians (lizards, amphisbaenias and *Sphenodon*), an ectothermic group where body temperatures vary among taxa. The simple prediction that tetracoding increases with body temperature is tested using species comparisons within this group, and the coevolution of tetracoding with antisense tRNAs possessing expanded anticodons is independently tested for Lepidosaurians, considering the previous information for mammals and *Drosophila* (Seligmann, 2012b).

2. Materials and methods

The complete Lepidosaurian mitochondrial genomes (lizards, amphisbaenians, and *Sphenodon*) available in Genbank in November 2012 were listed. The literature was searched for mean field body temperatures for each of these species. Table 1 presents the species for which both complete genomes and mean field body temperature data are available.

2.1. Expanded anticodons in antisense tRNAs

Each mitochondrial genome listed in Table 1 was searched for antisense tRNAs with expanded anticodon loops, as was done previously (Seligmann, 2012b). Each genome was analyzed by the online available software tRNAscan-SE (http://lowelab.ucsc.edu/tRNAscan-SE/, Lowe and Eddy, 1997; Schattner et al., 2005), setting the cut off score for COVE at −20. COVE scores are log ratios of capacities to form the classical cloverleaf secondary structure in the focal sequence, versus random sequences, hence negative values for the structural component of COVE indicate lower capacities to form cloverleaf structures than random sequences. All tRNAs matching the 22 regular mitochondrial tRNA sequences were extracted, and each of these sequences was again analyzed by tRNAscan-SE, setting the COVE cut off