Journal of Theoretical Biology 387 (2015) 154–165

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Codon expansion and systematic transcriptional deletions produce tetra-, pentacoded mitochondrial peptides

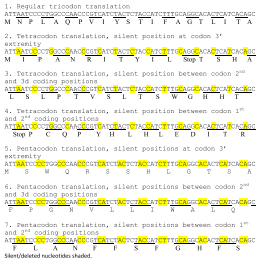
Hervé Seligmann

Unité de Recherche sur les Maladies Infectieuses et Tropicales Émergentes, Faculté de Médecine, URMITE CNRS-IRD 198 UMER 6236, Université de la Méditerranée, 13385 Marseille, France

HIGHLIGHTS

- RNAs match the mitogenome after systematic deletions every 3d nucleotide (dRNAs).
- Their translation produces peptides corresponding to tetra-and pentacodons.
- >1/3 peptides assume deletions at codon 3′ end: translation by expanded anticodons.
- Detected dRNAs and tetra-, pentapeptides map on the same mitogenome sequences.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 27 August 2015
Received in revised form
28 September 2015
Accepted 30 September 2015
Available online 9 October 2015

Keywords:
Systematic nucleotide exchange
Swinger DNA polymerization
Invertase
3′-to-5′ polymerization
Frameshift
Mass spectrometry
Expanded anticodon
Indels
Numt
Non-canonical RNA

ABSTRACT

Genes include occasionally isolated codons with a fourth (and fifth) silent nucleotide(s). Assuming tetracodons, translated hypothetical peptides align with regular GenBank proteins; predicted tetracodons coevolve with predicted tRNAs with expanded anticodons in each mammal, Drosophila and Lepidosauria mitogenomes, GC contents and with lepidosaurian body temperatures, suggesting that expanded codons are an adaptation of translation to high temperature. Hypothetically, continuous stretches of tetra- and pentacodons code for peptides. Both systematic nucleotide deletions during transcription, and translation by tRNAs with expanded anticodons could produce these peptides. Reanalyses of human nanoLc mass spectrometry peptidome data detect numerous tetra- and pentapeptides translated from the human mitogenome. These map preferentially on (BLAST-detected) human RNAs matching the human mitogenome, assuming systematic mono- and dinucleotide deletions after each third nucleotide (delRNAs). Translation by expanded anticodons is incompatible with silent nucleotides in the midst rather than at codon 3′ extremity. More than 1/3 of detected tetra- and pentapeptides assume silent positions at codon extremity, suggesting that both mechanisms, regular translation of delRNAs and translation of regular RNAs by expanded anticodons, produce this peptide subgroup. Results show that systematically deleting polymerization occurs, and confirm serial translation of expanded codons. Non-canonical transcriptions and translations considerably expand the coding potential of DNA and RNA sequences.

1. Introduction

Protein translation requires a complex set of molecules, from mRNAs, tRNAs and rRNAs, too complex for having existed at the dawn of molecular life. The hypothesis that primordial genomes were rRNA-like transcripts with subparts functioning also as mRNAs and tRNAs solves part of the conundrum (Root-Bernstein and Root-Bernstein, 2015). Modern translation cannot proceed in the absence of ribosomes, because classical triplet codon–anticodon interactions are too weak to enable peptide elongation and require ribosomal stabilization. Ribosomes are very complex molecules, suggesting that translational

E-mail address: varanuseremius@gmail.com

http://dx.doi.org/10.1016/j.jtbi.2015.09.030
0022-5193/

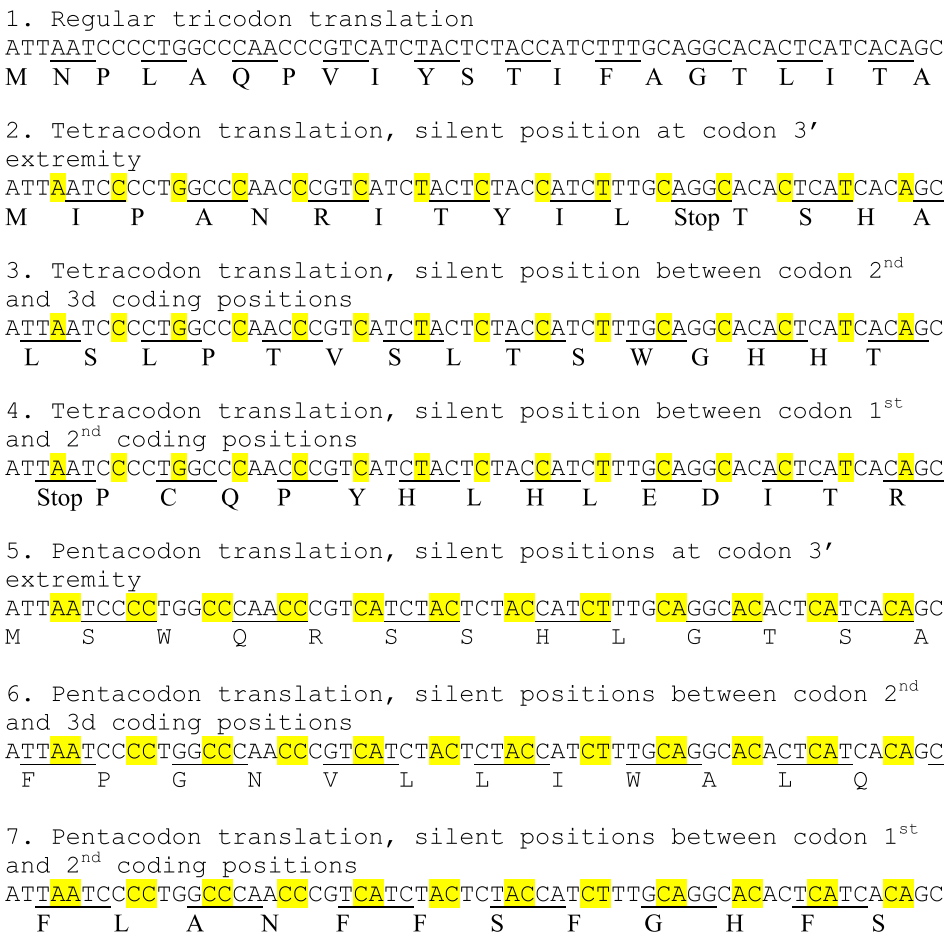


Fig. 1. Translation of the 5′ region of human mitochondrial gene ND2 according to regular tricodons (1), tetra- (2–4) and pentacodons (5–7). Peptides in 2 and 5 are compatible with translations by tRNAs with expanded anticodons and with translations by tRNAs with regular anticodons, of dRNAs where shaded nucleotides are systematically deleted by transcription that systematically deletes nucleotides. Peptides 3–4 and 6–7 are only compatible with translation by tRNAs with regular anticodons of dRNAs where shaded nucleotides were systematically deleted because expanded anticodons with silent positions positioned in the midst of the codon have not yet been described.

protein synthesis (unlike nonribosomal peptide synthesis, Kleinkauf and von Doehren, 1990; Stachelhaus and Marahiel 1995; Sieber and Marahiel, 2003; Strieker et al., 2010) could not have existed before the emergence of the complex ribosome. The assumption that primitive codons and anticodons were longer than three nucleotides, i.e. tetra- and/or pentacodons, potentially solves this thermodynamic paradox (Baranov et al., 2009), enables the selection of error minimizing tetracodons, the tessera (Gonzalez et al., 2012) and is compatible with evolutionary optimizations of the genetic code (Di Giulio et al., 2014).

Protein-encoding by expanded codons is supported by several observations. First, tRNAs with expanded anticodons suppress frameshift-inducing insertion mutations (Riddle et al., 1973; Sroga et al., 1992; Tuohy et al. 1992; Moore et al., 2000; Magliery et al., 2001; Anderson et al., 2002; Landweber, 2002; Dunham et al., 2007). These observations relate to isolated expanded codons, that can also be considered as part of the mechanisms for programmed frameshifts (Farabaugh 1996; Dinman, 2006, 2012; Russell and Beckenbach, 2008; Atkins and Björk, 2009; Melian et al., 2014). The hypothesis that modern genes encode for overlapping tetracoded genes is also supported by BLAST analyzes that detect alignments between regular Genbank proteins, and hypothetical peptides translated from mitogenomes assuming tetracodons. In addition, these alignment-predicted peptides coevolve with predicted mitochondrial tRNAs with expanded anticodons, in mammals, Drosophila (Seligmann, 2012a, 2013a, 2014a) and Lepidosauria (Seligmann and Labra 2013). Tetracoding appears as an adaptation to translation at high temperatures, as it associates with high GC contents (Seligmann, 2012a) and body temperature (Seligmann and Labra, 2013).

In addition to this translational tRNA-based mechanism, RNA editing by systematic nucleotide excision after each third nucleotide could produce 'delRNAs' (in short dRNAs) whose regular translation would correspond to tetra-decoding of the original, non-excised sequence. This assumes that polymerases occasionally enter modes where systematic deletions occur. The systematic nature of deletion-mutations would parallel previous observations of sequences where systematic nucleotide exchanges (rather than point mutations) produce 'swinger' sequences (mitochondrial swinger RNA (Seligmann 2012b, 2013b, 2013c, 2013d) and DNA (Seligmann, 2014b; Seligmann, 2015b); and nucleus-encoded eukaryotic swinger sequences (DNA, Seligmann 2014c; RNA, Seligmann, 2015b)). Swinger polymerization seems to be serial, systematic point mutations, as its frequency is proportional to rates of corresponding point mutations (Michel and Seligmann, 2014). The observation of chimaeric sequences, part corresponding to regular, and part to swinger polymerizations, strengthens the hypothesis that regular polymerases produce sequences transformed according to systematic changes (DNA, Seligmann, 2015b; RNA, Seligmann 2015a). The existence of swinger-transformed repeats of other mitogenomic regions suggests that systematic changes during polymerization produce new genetic information, occasionally integrated in genomes (Seligmann, 2015c).