Available online at www.sciencedirect.com
ScienceDirect
Journal of Theoretical Biology 243 (2006) 375–385
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Mitochondrial tRNA sequences as unusual replication origins: Pathogenic implications for *Homo sapiens*

Hervé Seligmann[a,*], Neeraja M. Krishnan[b], Basuthkar J. Rao[b]

[a]*Department of Evolution, Systematics and Ecology, The Hebrew University of Jerusalem, 91904, Israel*
[b]*Department of Biological Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400 005, India*

Received 6 February 2006; received in revised form 10 May 2006; accepted 27 June 2006
Available online 1 July 2006

Abstract

The heavy strand of vertebrate mitochondrial genomes accumulates deaminations proportionally to the time it spends single-stranded during replication. A previous study showed that the strength of genome-wide deamination gradients originating from tRNA gene's locations increases with their capacities to form secondary structures resembling mitochondrial origins of light strand replication (OL), suggesting an alternative function for tRNA sequences. We hypothesize that this function is frequently pathogenic for those tRNA genes that normally do not form OL-like structures, because this could cause excess mutations in genome regions unadapted to tolerate them. In human mitochondrial genomes, pathogenic tRNA variants usually form less OL-like structures than non-pathogenic ones in cases where the normal non-pathogenic tRNA variant can function as OL, as evolutionary analyses reveal. For tRNAs lacking the putative OL-like functioning capacity, pathogenic variants form more OL-like secondary structures, particularly structures that might invoke bi-directional replication (true for 14 among 21 tRNA species, $p<0.05$, sign test; significantly at $p<0.05$ (1 tailed test) for 7 tRNA species), but not more unidirectional replication invoking structures. Accounting for the functional cloverleaf-like structure-forming capacities of tRNAs yields similar results. Rare, non-pathogenic tRNA mutants tend to form more OL-like structures than the common, non-pathogenic ones, suggesting weak directional selection also among non-pathogenic variants. The duration spent single stranded by a region of the heavy strand (D_{ssH}) during replication, estimated by integrating over all regions that can function as OL in *Homo sapiens* mitochondrial genomes, increases with distance of that region from the Dloop. This suggests convergence of single-strandedness during replication and transcription, and explains conserved locations of tRNA species in mitochondrial genomes and bacterial operons. These locations minimize deamination costs only in anticodons and not in other tRNA regions, during replication and transcription. Therefore, putative functioning as OLs by tRNA sequences is normal at some locations and pathogenic at others.

Keywords: Functional duplication; Secondary structure; Substitution gradient; Mutational robustness; Cumulative error; Aging; Genome organization

1. Introduction

Causal relationships between genotype and phenotype remain usually unknown. Error propagation and error accumulation are of general interest in this context, because they can be quantified at all levels of organization, and predictions can be developed on how error at one level propagates to the next level, such as in protein synthesis (Blomberg et al., 1985; Blomberg, 1990; Johansson and Blomberg, 1995). Error propagation was probably a major factor influencing the first stages of origin of life (Blomberg, 1995). Estimates of error proneness at the level of the molecular machinery of cells, at least for the mitochondrion, are good predictors of developmental instability as measured at the whole organism level by the random components of bilateral asymmetry. This was true for repeatability estimates of protein synthesis (chemical stability of ribosomal RNA increases developmental stability, Seligmann, 2006) as well as of genome replication (secondary structural stability of the light strand replication origin increases developmental stability, Seligmann and Krishnan, 2006). They follow the principle that morphological asymmetry is the direct result of differences

*Corresponding author. Tel./fax: +972 2 6585876.
E-mail address: hselig1@yahoo.com (H. Seligmann).

0022-5193/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2006.06.028

in developmental rates of bilateral traits of the two sides. Variation in energetic outputs by mitochondria on the two sides, directly affects growth and differentiation, and hence, causes random differences in the resulting morphologies. These studies describe correlations between molecular and whole organism levels based on variation between different species. The present analyses show similar associations by comparing individuals from the same species (*Homo sapiens*). They show that repeatability of mitochondrial DNA replication associates with cumulative, developmental disorders at the whole organism level, in the form of various human pathologies associated with mutations in mitochondrial tRNAs. The corroboration at the level of population genetics of the same principles as those observed at micro-evolutionary levels in primates strengthens the hypothesized mechanisms that link molecular and macroscopic repeatabilities. It also emphasizes how the combination of methods from bioinformatics and evolutionary biology can reveal molecular mechanisms with potential medical applications.

We consider that this parallel between pathogenic disorders and developmental instability is justified by numerous associations discovered between morphological inaccuracies and various disorders, including indicators of decreased fitness (Thornhill and Moller, 1997; Moller, 1997, 1999, also see Appendix 1 in supplementary material). Our hypothesized mechanism suggests that mutations in the DNA sequences coding for mitochondrial tRNAs alter their propensities to form secondary structures that resemble regular mitochondrial origins of replication of the light strand (OL), thereby altering replication patterns, and causing disorders (see sections below).

A notable feature of the approach we adopt is the use of evolutionary variation among primates in order to indicate which mitochondrial tRNAs function in these species as origins of replication, and which do not. Seligmann et al. (2006) show that in the primate mitochondrial genomes, where a tRNA forms OL-like structures, a deamination gradient starting at the location of that tRNA exists, while in those primate species where the same homologous tRNA does not form such structures, no such gradient exists. (These gradients are a direct result of the directional mode of replication, see explanations below.) This correlation between structural propensity at a tRNA, and the strength of a deamination gradient across the genome starting at that tRNA, suggests coevolution. We consider that the tRNA for which such a coevolution exists, is used as OL, at least in some species and, therefore, we hypothesize that its use as OL within the frame of primate evolution is evidence that this function does not cause major disruptions. Hence, we predict that: (1) mutations that decrease the structural propensity of tRNA sequences to form origins of replication are pathogenic in those tRNA species that can normally function as OL (as indicated by coevolutionary analyses between secondary structure of tRNA and deamination gradients), and (2) mutations that increase the structural propensity of tRNA sequences to form origins of replication are pathogenic in those tRNA species that normally cannot function as OL. This approach is an original application of the principles indicated by the Kluge–Kerfoot phenomenon (Kluge and Kerfoot, 1973; Pierce and Mitton, 1979), which predicts that a correspondence exists between variation between different species and variation within the same species. Our results highlight the advantages of integrating information from population genetics and evolutionary processes, in the same analyses.

1.1. General background on mitochondrial DNA replication

In vertebrate mitochondria, replication of the 'lagging' light strand usually starts when the 'leading' heavy strand replication fork reaches a 30 base pair long stretch that forms a linear stem-loop hairpin structure, called the OL (Clayton, 1982). Mutations, mainly transitions resulting from hydrolytic deaminations (cytosine→thymine and adenine→hypoxanthine, which later gets converted to guanine) accumulate proportionally to the time spent single stranded by the heavy strand until the light strand replication fork synthesizes the complementary light strand. Therefore, gradients of increasing substitution frequencies exist across mitochondrial genomes, proportionally to the time spent single stranded by a site, which is determined by its location versus the OL (Krishnan et al., 2004a, b). This OL is flanked on either side by 2 and 3 tRNA genes, respectively, which also sometimes form OL-like structures (Seligmann et al., 2006; Seligmann and Krishnan, 2006). Evidences indicate that OL-flanking tRNA genes can also sometimes function as OLs: (1) developmental instability, as estimated by fluctuating asymmetry in scalation traits in several lizard families (Amphisbaenidae, Anguidae and Polychritidae), decreases with stability of their mitochondrial OL's secondary structure, and also decreases with the capacity of the OL-flanking tRNA genes to fold as OL-like structures (Seligmann and Krishnan, 2006); (2) strengths of deamination gradients existing across different primate genomes are frequently proportional to the OL-like structure-forming capacities of the tRNA sequences that mark the origin of these gradients in these species (Seligmann et al., 2006); (3) suppressive mutant *Neurospora* spp. mitochondrial plasmids that carry insertions at the major 5′ end, corresponding to a mitochondrial tRNA, suggest that these tRNAs play a role in replication of the plasmids by reverse transcription, leading to over-production of plasmid transcripts, thereby outcompeting mitochondrial DNA and impairing growth (Akins et al., 1989) and (4) the active sites of class II amino acyl tRNA synthetase and those of the gamma polymerase, the sole polymerase replicating vertebrate mitochondrial DNA (Bolden et al., 1977; Kaguni, 2004), are homologous (Carrodeguas et al., 1999, 2001; Fan et al., 1999; Carrodeguas and Bogenhagen, 2000). Strengthening these lines of evidence are also independent findings that replication of mitochondrial light strands occurs at multiple origins, although apparently independently of