Gene 542 (2014) 248–257

Contents lists available at ScienceDirect

Gene

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

The relation between hairpin formation by mitochondrial WANCY tRNAs and the occurrence of the light strand replication origin in Lepidosauria

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

[a] Center for Ecological and Evolutionary Synthesis, Department of Bioscience, University of Oslo, Blindern, N-0316 Oslo, Norway
[b] National Natural History Museum Collections, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel
[c] Department of Life Sciences, Ben Gurion University, 84105 Beer Sheva, Israel
[d] Programa de Fisiología y Biofísica, Facultad de Medicina, Universidad de Chile, Casilla 70005, Santiago, Chile

ARTICLE INFO

Article history:
Received 30 May 2013
Received in revised form 27 December 2013
Accepted 17 February 2014
Available online 18 February 2014

Keywords:
Gene function duplication
Alternative gene function
Secondary structure prediction
Sphenodon
Covariance matrix
Gene interactions

ABSTRACT

Mitochondrial light strand DNA replication is initiated at light strand replication origins (OLs), short stem-loop hairpins formed by the heavy strand DNA. OL-like secondary structures are also formed by heavy strand DNA templating for the five tRNAs adjacent to OLs, the WANCY tRNA cluster. We tested whether natural OL absence associates with greater capacities for formation of OL-like structures by WANCY tRNA genes. Using lepidosaurian taxa (*Sphenodon*, lizards and amphisbaenids), we compared WANCY tRNA capacities to form OL-like structures between 248 taxa possessing an OL with 131 taxa without OL (from different families). On average, WANCY tRNA genes form more OL-like structures in the absence of a regular OL than in its presence. Formation of OL-like structures by WANCY tRNAs follows hierarchical patterns that may reduce competition between the tRNA's translational function and its secondary OL function: the rarer the tRNA's cognate amino acid, the greater the capacity to form OL-like structures. High OL-forming capacities for neighboring tRNAs are avoided. Because OL absence usually occurs in taxa with reduced genomes, increased formation of OL-like structures by WANCY tRNAs might result from selection for greater metabolic efficiency. Further analyses suggest that OL loss is one of the latest steps in genome reduction, and promotes the increase in formation of OL-like structures by WANCY tRNA genes in Lepidosauria.

1. Introduction

A major challenge of the postgenomic era is to make sense of the wealth of newly available molecular sequence information. Some DNA sequences have a presumably well-known, easily detected function. However, this is sometimes misleading, because it deflects from considering secondary functions, which may have biological importance. For example, many DNA regions that supposedly cannot code for proteins because they include stop codons seem to actually code for proteins, assuming induction of translational activity by suppressor or antitermination tRNAs that have anticodons that match stop codons (Seligmann, 2010a). The off frame regions of protein coding regions code for 'cryptic' overlapping protein coding genes according to a stopless parallel 'overlapping' genetic code, as suggested by comparative analyses of mitochondria from primates (Faure et al., 2011; Seligmann, 2011a, 2012a), *Drosophila* (Seligmann, 2012b) and turtles (Seligmann, 2012c). This system of overlapping genes by overlapping genetic codes increases the number of putatively mitochondrion-encoded proteins, without lengthening the genome. In fact, several additional mechanisms increase manifold the coding density of genes, such as tetracodons (Seligmann, 2012d, 2013a,b; Seligmann and Labra, 2013), 3′-to-5′ overlap coding (Seligmann, 2012e, 2013c,d) and permuting polymerization that produces transcripts with systematically exchanged nucleotides (Seligmann, 2013e,f). All this is in line with the description of a 'punctuation code' formed by 20 specific codons (the circular code regulates frame choice, Arquès and Michel, 1997), and that some codons regulate gene expression (Stergachis et al., 2013). This indicates that not all functions of DNA are known, and that some can still be detected. In fact, in the short mitochondrial vertebrate genome, the classical known role of tRNA genes is to template for tRNAs that function in translation, but DNA coding for tRNAs has other, less known functions (Giegé, 2008). Bioinformatic analyses of mitochondrial genomes suggest that DNA templating for tRNAs (tDNA) can function as additional mitochondrial light strand replication origin, OL (Seligmann, 2008, 2010b; Seligmann and Krishnan, 2006; Seligmann et al., 2006a,b). Similarly, Rhodakis et al. (2007) found that sequences other than the

Abbreviations: OL, light strand replication origin; WANCY region, the mitochondrial genome region that templates for tRNAs with cognates W (tryptophane), A (alanine), N (asparagine), C (cysteine) and Y (tyrosine).

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

http://dx.doi.org/10.1016/j.gene.2014.02.021
0378-1119/

recognized OL have an OL-like function, such as a short region in the coding sequence of ND3 in the mussel genus *Mytilus*.

1.1. Light strand replication initiation and tDNAs

The OL is a short heavy strand DNA stem-loop hairpin. The mitochondrial DNA polymerase gamma initiates mitochondrial light strand DNA replication in the OL vicinity (Wanrooij and Falkenberg, 2010). Two independent properties of mitochondrial tRNA genes correlate positively with estimations of the frequency at which light strand replication is initiated in the region of these specific tRNA genes: the tendency of these tRNAs to form OL-like secondary structures (Seligmann et al., 2006a), and the similarity between the sequences forming tRNA anticodon loops and those forming the loop of the 'real' OL (Seligmann, 2010a). The latter agree with the results by Fuste et al. (2010) on the requirement of polyT stretches in the OL loop, because both the OL loop and the tDNA anticodon loops predicted to function as OLs have polyT stretches.

Other properties of the classical OL region important to initiate DNA replication, include the presence of the presumed canonical motif 3′-GGCCG-5′ located near the OL (Hixson et al., 1986), as well as the presence in the loop of the OL, of a poly dT stretch (Fuste et al., 2010). Evidence indicates that similarities between the linear sequences of OL and anticodon loops are independent from the capacity of the complete tDNA to form OL-like secondary structures, and that similarities between tDNA and OL, according to different OL properties, predict OL function. Supporting this are results showing that the stability of hybridization between heavy strand tDNA and its expressed, complementary tRNA, decreases initiation of light strand replication at that tDNA (Seligmann, 2008). These observations agree with the hypothesis that the original role of tRNAs at the origins of life was for replication initiation (Maizels and Weiner, 1994), and, conversely, that RNA corresponding to the OL is expressed and loaded with amino acids (Yu et al., 2008).

The above described analyses used secondary structure predictions (Mfold, Zuker, 2003) that show formation of OL-like structures by heavy strand sequences of genes coding for mitochondrial tRNAs, notably the five tRNA genes adjacent to the OL (Seligmann, 2010a; Seligmann and Krishnan, 2006; Seligmann et al., 2006a,b). A total of six adjacent genes, the OL and the five tRNAs that form the WANCY region (tRNA Trp -W-, tRNA Ala -A-, tRNA Asn -N-, tRNA Cys -C- and tRNA Tyr -Y-) potentially function as OL in that region (Seligmann, 2010a; Seligmann et al., 2006a,b). The regular OL is located between tRNA Asn and tRNA Cys (besides some rare exceptions, i.e. in salamanders, Mueller et al., 2004). Fig. 1 presents a typical example of the regular OL and an OL-like structure formed by the heavy strand complement of the tDNA that templates for tRNA Trp in the western fence lizard, *Sceloporus occidentalis*. Note that the loop of the OL-like hairpin corresponds to sequences that form the anticodon loop in the tRNA cloverleaf structure. This is generally the case, as it has been shown in different taxa: for tRNA Tyr of the anguid arboreal alligator lizard *Abronia graminea* (Fig. 2 in Seligmann and Krishnan, 2006); for tRNA Asn of the dusky leaf monkey *Trachypithecus obscurus* (Fig. 1B in Seligmann et al., 2006a); for tRNA Thr of *Homo sapiens* (Figs. 1b and d in Seligmann et al., 2006b) and for tRNA Ser UCN of the Tibetan macaque, *Macaca thibetana* (Seligmann, 2010a).

1.2. Effects of multiple replication origins

Possessing several regions that function as OL has adaptive consequences for mitochondrial genomes. Each time the DNA duplex is opened for DNA or RNA syntheses (i.e., DNA replication and RNA transcription), the single stranded DNA is subject to enhanced mutation rates because single stranded DNA is more likely to mutate than in the protected duplex state. If replication starts at the regular OL, different regions remain single stranded for different periods of time. According to their position in relation to the OL, some regions remain a short time single stranded, and others a long time, and cumulate more mutations during replication (i.e., Hassanin et al., 2005; Tanaka and Ozawa, 1994; Reyes et al., 1998). A similar, yet not identical phenomenon, occurs during transcription, which makes that some regions mutate due to replication, and others, due to transcription (Seligmann, 2011b, 2012a, 2013f).

1.3. Working hypothesis and predictions

Analyses presented here use the WANCY system to address the following issues: 1. Does OL absence associate with greater capacities by tRNAs adjacent to the OL to function as OL? 2. If so, is it possible to

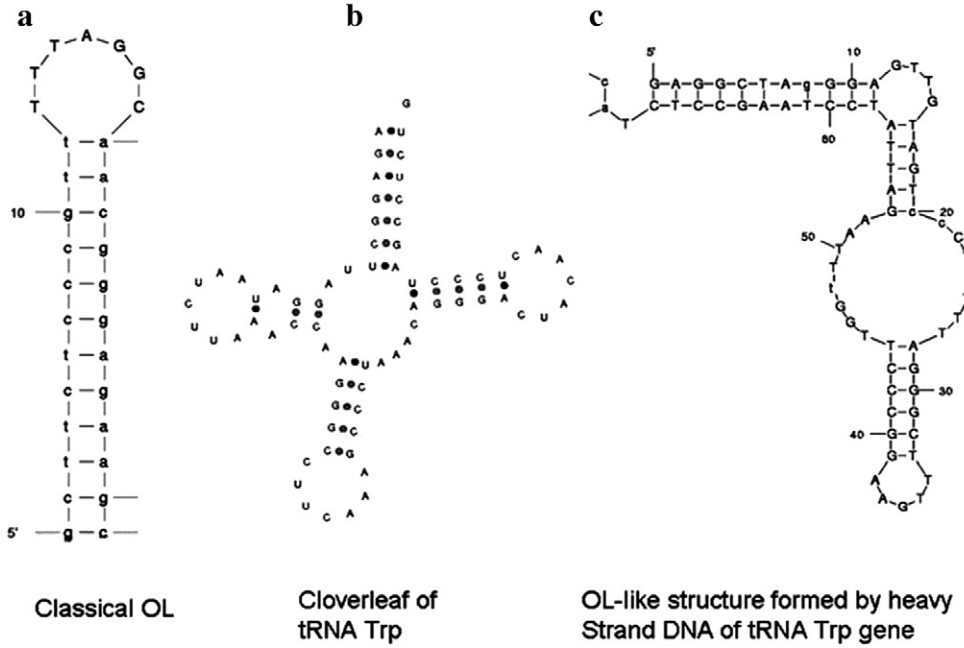


Fig. 1. Secondary structures in the western fence lizard, *Sceloporus occidentalis* (NC_005960 in NCBI (GenBank)). a—Classical OL formed by the heavy strand DNA between the genes coding for tRNAs Asn and Cys. b—Classical tRNA Trp cloverleaf structure functioning in translation and expressed by the light strand DNA. c—OL-like structure formed by the heavy strand DNA sequence that corresponds for that tRNA to the expressed sequence.

predict which tRNAs specialize as OLs? 3. Does OL loss drive increased formation of OL-like structures by adjacent tRNAs, or is OL loss enabled by that pre-existing capacity by WANCY tDNAs to form these OL-like structures?

Considering the vital OL function, we hypothesize that tDNAs surrounding the regular OL should form more OL-like structures in taxonomic groups lacking an OL than those possessing a regular OL. Second, in groups lacking a classical OL, the enhancement in OL-forming capacities should be proportional to the extent of OL formation by that tDNA in the presence of a classical OL. We also expect that the hierarchy between tDNAs for specializing for OL function should follow simple principles of trade-off between tRNA/tDNA functions. We expect negative associations between the primary function of the gene as tRNA in protein translation and its secondary function as OL during DNA replication.

We compared OL-forming capacities of tDNAs adjacent to the OL in lepidosaurian groups with and without a classical OL (see Table 1). Lepidosaurians are a monophylogenetic reptilian group which consists of Rhynchocephalians (as outer branch, with the tuatara, *Sphenodon*, Sphenodontidae, its only extant species), and the squamates, which include lizards, amphisbaenids and snakes. Analyses exclude the 'advanced' snakes, since their replication is peculiar, as their mitochondrial genomes possess two D-loops. For the sake of taxonomic uniformity in the analyses, we also excluded sequences from primitive snakes (Typhloidea: Leptotyphlopidae and Typhlopidae), which possess mitochondrial genomes with a structure like those found in lizards, but lack an OL in their WANCY region. Therefore, our analyses were restricted to lizards, amphisbaenids and *Sphenodon*. In lizards, the lack of the classical OL occurs occasionally in Amphisbaenia (two species, *Amphisbaena xera* and *Trogonophis wiegmanni*) and in some lizard families, Agamidae (Macey et al., 2000a,b), Gekkonidae (only in the genus *Cyrtodactylus*), Iguanidae (*Enyaloides laticeps*, *Phrynosoma mcallii* and *Uracentron flaviceps*), Scincidae (11 species) and Xantusiidae (one species), which are compared to other lizard groups.

Table 1
Capacity of the WANCY region to form OL-like structures and the general use of different amino acids coded by the WANCY region, in different Lepidosauria taxa.

Taxon	N	Mean percent of OL-like structures						Sum	Dist	Genome-wide usage of cognate				
		OL	Trp	Ala	Asn	Cys	Tyr			Trp	Ala	Asn	Cys	Tyr
OL absent														
Agamidae	88		23.22	2.53	15.03	21.34	19.49	80.01	0.38	26.1	76.0	31.8	6.7	30.0
Amphisbaenia	2		25.00	0.00	14.71	3.71	23.08	66.49						
Amphibaena xera	1		40.00	0.00	0.00	0.00	46.15	86.15						
Gekkonidae: *Cyrtodactylus*	20		22.81	16.70	7.54	23.53	33.25	103.83	0.16					
Iguanidae	3		22.94	6.67	6.67	19.44	4.76	60.48						
Enyalioides laticeps	1		22.22	20.00	0.00	25.00	0.00	67.22						
Phrynosoma mcallii	1		9.10	0.00	20.00	33.33	0.00	62.42						
Uracentron flaviceps	1		37.5	0.00	0.00	0.00	14.29	51.79						
Scincidae	12		16.72	7.59	11.51	11.40	17.45	64.67	−0.43					
Chalcides ocellatus	1		18.18	0.00	6.25	0.00	0.00	24.43						
Eumeces anthracinus	1		15.39	0.00	11.11	33.33	33.33	93.16						
Eumeces inexpectatus	1		27.27	0.00	27.27	0.00	23.08	77.62						
Eumeces skiltonianus	1		15.69	0.00	27.27	27.27	23.08	93.01						
Xantusiidae	5		57.29	0.00	5.00	31.71	18.89	112.89	−0.39	28.0	79.5	34.3	7.1	30.4
Sphenodontidae	1		12.50	7.69	0.00	75.00	0.00	95.19		31.3	65.2	36.4	6.6	31.4
OL present														
Agamidae	58	3.63	21.13	9.64	7.32	29.96	13.64	81.15	0.06	26.8	87.5	44.0	7.9	28.7
Amphisbaenia[a]	8	9.18	15.42	0.00	18.89	17.94	16.62	68.86	−0.05	26.5	84.6	32.3	7.7	33.1
Amphisbaena schmidti	1	11.35	25.00	0.00	37.50	0.00	11.11	73.61						
Anguidae	11	11.49	5.62	0.00	1.82	31.41	1.30	40.15	−0.25	27.4	67.8	40.1	8.2	30.3
Polychridae	20	11.55	26.54	1.74	8.80	2.88	21.05	61.00	0.34	26.2	70.3	41.0	7.9	29.6
Chamaeleonidae	9	7.77	37.22	7.41	2.38	24.22	2.62	73.85	−0.12	26.9	62.4	43.0	8.0	31.0
Cordylidae	3	11.07	16.24	0.00	27.78	0.00	4.76	48.78		28.0	71.4	36.4	8.9	27.2
Gekkonidae	20	10.11	17.60	3.33	20.75	15.15	12.24	69.07	**0.46**	27.6	74.9	36.5	8.8	29.4
Helodermatidae	2	10.19	4.55	0.00	0.00	10.00	0.00	14.55		26.9	64.1	38.0	8.1	33.2
Lacertidae	7	11.65	13.77	0.00	22.15	5.71	10.64	52.27	0.35	27.2	70.1	35.7	7.4	29.9
Liolaemidae: *Liolaemus*	36	12.25	47.54	0.00	2.04	3.08	28.43	81.09	**0.44**					
Phrynosomatidae: *Sceloporus*	16	11.07	27.15	0.00	7.99	3.70	26.23	65.07	0.38	27.3	74.9	37.8	7.9	29.3
Enyalioides laticeps	4	10.28	10.00	8.00	0.00	6.90	6.25	31.15						
Phrynosoma	8	11.72	23.36	0.00	3.54	22.21	7.50	56.61	−0.16					
Uracentron flaviceps	1	11.99	37.50	0.00	0.00	0.00	14.29	51.79						
Scincidae	29	6.8	19.19	3.23	14.44	13.41	16.16	66.44	−0.29	28.3	85.0	34.8	8.2	29.8
Chalcides ocellatus	1	6.43	16.17	0.00	6.25	0.00	0.00	22.92						
Eumeces anthracinus	1	6.68	15.39	0.00	11.11	33.33	33.33	93.16						
Eumeces inexpectatus	1	6.98	27.27	0.00	27.27	0.00	23.08	77.62						
Eumeces skiltonianus	1	6.68	11.11	0.00	27.27	0.00	50.00	88.38						
Shinisauridae	1	11.04	18.18	12.50	0.00	0.00	0.00	30.68		27.5	74.9	37.5	8.0	30.2
Teiidae	2	11.82	5.56	0.00	0.00	22.50	0.00	28.06						
Varanidae	19	12.49	12.43	25.51	4.62	40.15	9.53	92.24	−0.18	26.5	60.5	41.9	7.6	26.6
rOL-cognate usage										−0.07	−0.33	**−0.47**	**−0.60**	−0.43
Mean, all taxa	22	10.14	21.30	4.76	9.07	18.47	12.73	66.21	0.32	27.4	73.1	37.6	7.8	30.0
Mean taxa with OL	15	10.14	19.21	4.23	9.27	14.68	10.88	58.22	0.17	27.2	73.0	38.4	8.1	28.9
Mean taxa without OL	7		25.78	5.89	8.64	26.59	16.70	83.37	0.42	28.5	73.6	34.2	6.8	30.6
Δ taxa with and without OL			−6.57	1.66	−0.63	−11.92	−5.82	−25.15						

Mean percent of OL-like secondary structures formed by heavy strand sequences of tRNA genes for different lepidosaurian taxa (lizards, amphisbaenids and *Sphenodon*). N = number of species used to calculate averages for a taxon, OL = mean −ΔG of the stem-loop hairpin formed by the classical OL, Sum = sum of the percentages of the OL-like structures over all five tRNAs, Dist = Pearson correlation coefficient between pairwise correlations between OL formation and the distance between genes. Genome-wide usages of the cognate amino acids of the five tRNAs are also indicated when the complete sequence of at least one genome is available for that taxon. The last line indicates differences between mean OL-formation capacity by tRNAs for taxa with and without recognized OL. Bold indicates statistical significance at $P < 0.05$ according to one tailed t-tests for the last row, and for Dist.

[a] Includes members from all 4 amphisbaenian families.

It is important to consider that variation probably exists between species in the motifs initiating polymerase gamma binding (in the OL loop) and polymerization. These motifs have been determined for *Homo sapiens* (Hixson et al., 1986), but they are probably not universal, and probably vary even within Lepidosauria. Including them in the analyses would be more complete at the condition they are known for all the taxonomic groups included in this study. This would imply separate studies on each of these motifs for each taxon, which seems unpractical at this stage. Restricting the study to the secondary structure aspect enables to test whether the basic hypothesis of OL function by tRNAs is potentially viable according to that specific criterion. It should be seen as a first step towards more complete analyses that include the other criteria (binding region in the OL loop and polymerization initiation motif). Therefore, here we focus on a single tDNA property, the capacity to form OL-like secondary structures.

2. Materials and methods

We searched Genbank (NCBI; accessed January 2009 and sampling completed July 2013) for lizard, amphisbaenids and *Sphenodon* sequences of mitochondrial genes that normally flank the OL. For Amphisbaenia and each lizard family (or genus) and *Sphenodon*, tRNA sequences for the five tRNA genes flanking the OL (tRNA Trp, Ala, Asn, Cys and Tyr) were extracted using tRNAScan-Se (http://lowelab.ucsc.edu/tRNAscan-SE/, Lowe and Eddy, 1997; Schattner et al., 2005), to avoid inconsistencies between species in Genbank annotations.

Capacities to form OL-like structures were estimated by examining each of the alternative secondary structures predicted by Mfold (using the default mode in terms of temperature and ionic concentrations), for the heavy strand tDNA sequence of a given WANCY tRNA with stabilities down to half that of its most stable secondary structure (Krishnan et al., 2004, 2008; Seligmann and Krishnan, 2006; Seligmann et al., 2006a,b). Thereafter, we calculated percentages of secondary structures that are OL-like for all secondary structures (see Seligmann and Krishnan, 2006; Seligmann et al., 2006a,b). In order to decide that a structure formed by a tDNA is OL-like or not, we strictly used the following criteria. The structure has to consist of a single linear, non-branching stem. At one end of the stem are the sequence's 5′ and 3′ extremities, at the other end, a loop of undetermined size. Bulges or small loops (unpaired nucleotides) within the stem were tolerated, as long as these do not involve secondary (paired nucleotides) stems branching from the main stem. Hence, the stem-loop structure has to be linear. It is possible that these criteria were too restrictive and that other secondary structures, such as branching ones, or suggesting bidirectionality (see Fig. 2C in Seligmann and Krishnan, 2006, and 1C in Seligmann et al., 2006b where a tDNA sequence seems to form two OLs, each one directed in the opposite direction), also function as OLs. The criteria, however, were chosen because they enable objective and repeatable decisions upon visual examination of secondary structures.

Genbank entries frequently lack annotations for the OL. The lack of polyT stretch in the OL can be a cause for this lack of annotation. The cause probably varies between studies, but there is no doubt that in numerous cases the OL is not annotated but includes a polyT stretch (i.e., in Horsfield's tarsier, *Tarsius bancanus* (NC_002811)). However, in cases where the OL has been annotated, procedures determining its sequence differ among studies. Therefore, we determined the sequence forming the OL stem-loop hairpin using the following procedure: 1- Considering that for most vertebrates the OL is between tRNA Asn and tRNA Cys, we extracted the sequence between these two tRNAs, extended by about 15 nucleotides from the 3′ extremity of tRNA Asn and 15 nucleotides from the 5′ extremity of tRNA Cys. 2- The secondary structure formed by this sequence, as predicted by Mfold, was examined to determine the precise sequence forming the stem-loop structure. 3- The heavy strand of the latter sequence was used to predict the stability (ΔG) of the OL according to Mfold predictions. Stages 1 and 2 were necessary because the sequence forming the OL is not always precisely limited by the tRNA genes flanking the OL, it could be shorter than the inter-tRNA spacer, or expand into the tRNAs, as previously described (e.g., Macey et al., 1997). Usages of tRNAs are deduced from the complete genome usages of their corresponding amino acid cognates in mitochondrially-encoded protein coding genes for species from that family for which complete mitochondrial genome sequences are available.

Associations between variables were tested using Pearson correlations, in most cases one tailed. Sign tests were applied using the binomial distribution.

3. Results and discussion

3.1. Enhanced formation of OL-like structures in OL absence: tRNA Cys and tRNA Trp

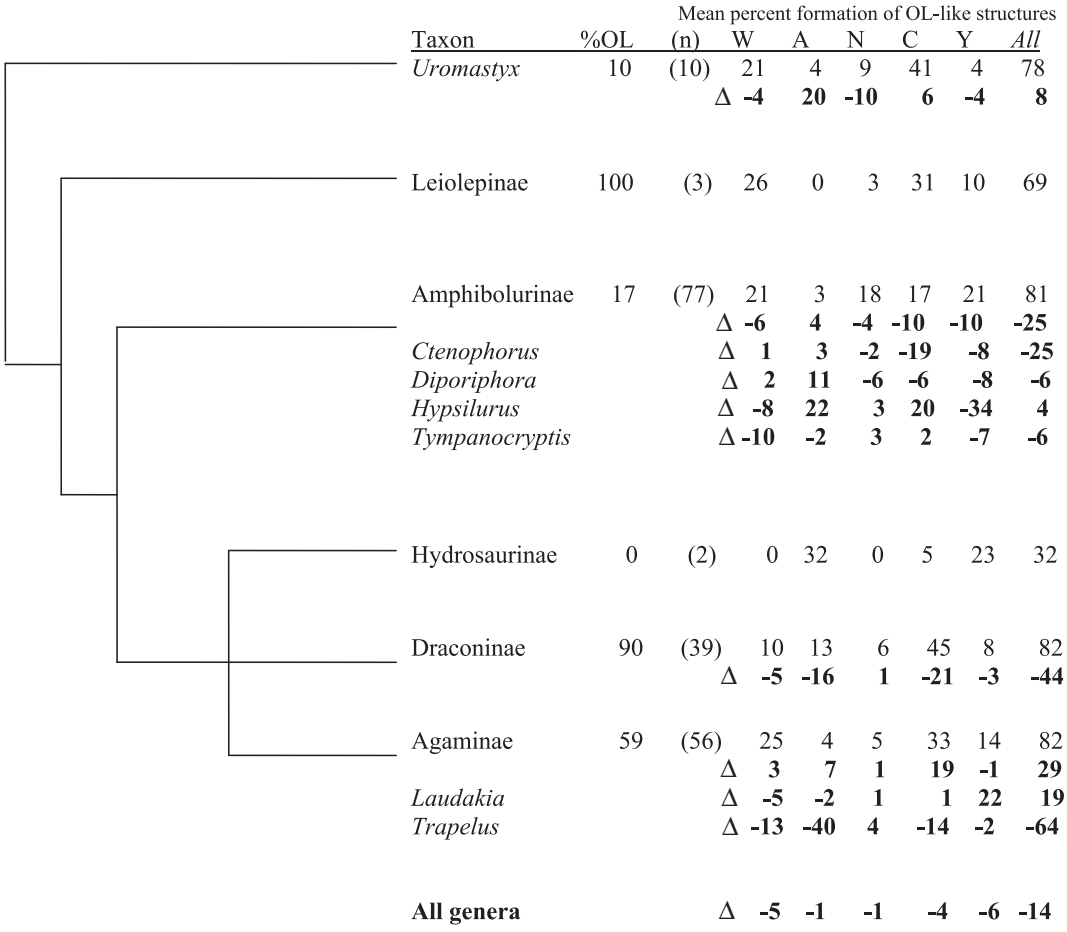
Among the 379 species examined, Mfold did not predict any OL-like structure for 44% of all tDNA sequences across all five tRNA species. In two unrelated taxa, *Heloderma suspectum* and *Anolis lionotus*, none of the five tDNAs forms OL-like structures. Overall, the distribution of OL-forming capacities seems random, with some indications of phyletic conservatism.

Table 1 presents the taxon-wide average percentage of OL-like structures for the heavy strand DNA of the five WANCY tRNAs, the mean −ΔG of OLs in these taxa, as well as data for single species or genera when sequences with and without OL exist for that specific species/genus (the comparisons within genus are discussed separately in a section below). The formation of OL-like structures by tDNAs is, as expected, on average larger in taxa lacking an OL than in the other taxa at four levels of taxonomic divergences, comparisons between families (Table 1, four last rows), within families, within genera, and within species (Table 1).

In general, tRNA Trp has among the five tRNAs the greatest capacity to form OL-like structures (Table 1). Data indicate that tRNA Cys has the highest difference between mean capacities to form OL-like structures in the presence versus the absence of OL. However, tRNA Trp has the greatest OL-forming capacity in four among seven groups lacking OL (57%) and six among 15 groups with OL (40%), and either the highest or second highest in six among seven groups lacking OL (86%) and 13 among 15 taxa with OL (87%). On the other hand, tRNA Ala forms rarely OL-like structures in the wide majority of taxa, with the exception of *Varanus*, where it evolved the capacity to form OL-like structures (discussed later). In addition, the difference in mean OL-formation capacity by tDNAs between taxa with and without OL is largest for OL formation summed over all five tRNAs. In *Sphenodon*, which lacks OL (Seutin et al., 1994), the situation was similar to other taxa (lizards and Amphisbaenia) lacking OL (Table 1).

3.1.1. OL-like structures with and without OL: comparisons within Agamidae

Agamidae is an extremely interesting lizard family, because it includes variation in OL presence/absence (Table 1), which enables more meaningful evolutionary tests to understand the implication of OL absence on tDNAs. The distribution of the presence/absence of OL in major agamid groups is shown in Fig. 2. As shown by Table 1, OL formation by tDNAs was, on average, weaker in taxa with an OL as compared to taxa lacking it, for all tDNAs. The effect was significant across all agamid groups using a sign test for tRNA Tyr (six among seven comparisons, $P = 0.03$, one tailed sign test). These results are an independent confirmation of the trends found across different lepidosaurian families; a negative association between OL presence and an enhanced capacity for formation of OL-like structures by WANCY tDNAs. Note that the differences between means of species with and without OL in Fig. 2 correspond to what is usually called phylogenetically independent contrasts (Felsenstein, 1985).



Taxon	%OL	(n)	Mean percent formation of OL-like structures W	A	N	C	Y	*All*
Uromastyx	10	(10)	21	4	9	41	4	78
			Δ -4	**20**	**-10**	**6**	**-4**	**8**
Leiolepinae	100	(3)	26	0	3	31	10	69
Amphibolurinae	17	(77)	21	3	18	17	21	81
			Δ -6	**4**	**-4**	**-10**	**-10**	**-25**
Ctenophorus			**Δ 1**	**3**	**-2**	**-19**	**-8**	**-25**
Diporiphora			**Δ 2**	**11**	**-6**	**-6**	**-8**	**-6**
Hypsilurus			**Δ -8**	**22**	**3**	**20**	**-34**	**4**
Tympanocryptis			**Δ -10**	**-2**	**3**	**2**	**-7**	**-6**
Hydrosaurinae	0	(2)	0	32	0	5	23	32
Draconinae	90	(39)	10	13	6	45	8	82
			Δ -5	**-16**	**1**	**-21**	**-3**	**-44**
Agaminae	59	(56)	25	4	5	33	14	82
			Δ 3	**7**	**1**	**19**	**-1**	**29**
Laudakia			**Δ -5**	**-2**	**1**	**1**	**22**	**19**
Trapelus			**Δ -13**	**-40**	**4**	**-14**	**-2**	**-64**
All genera			**Δ -5**	**-1**	**-1**	**-4**	**-6**	**-14**

Fig. 2. Phylogenetic reconstruction of the relationships among Agamid groups, according to Macey et al. (2000a,b). Mapped on it, is the distribution of the presence/absence of the OL, and the capacity of WANCY tDNAs to form OL-like structures among agamid subfamilies. Column numbers after taxon names are: % of species with an OL (total number of species examined between parentheses), % of OL-like structures formed by the heavy strand tDNA of each of the five WANCY tDNAs, and the sum of these five percentages. W: tRNA Trp; A: tRNA Ala; N: tRNA Asn; C: tRNA Cys; Y: tRNA Tyr. Bold indicates the subtraction (Δ) of the mean % for species without OL from the mean of those with OL in that taxon. When OL presence/absence occurs within a genus, the subtraction is also indicated for comparisons within that genus.

3.1.2. OL-like structures with and without OL: other comparisons

There are in total eight comparisons, one within genus *Amphisbaena*, three from Iguanidae and four from Scincidae, six among these are intraspecific variation in OL presence/absence for sequences believed from the same species (all cases from Scincidae and two from Iguanidae (*Enyalioides* and *Uracentron*)). There is no difference in formation of OL-like structures for tRNAs from sequences with and without OL for *Uracentron flaviceps*, *Eumeces anthracinus* and *Eumeces inexpectatus*, suggesting that the intraspecific variation in OL presence/absence is too recent for mutations to change formation of OL-like structures. Excluding these three cases, the percentage of OL-like structures increases in the majority of pairwise comparisons for tRNAs Trp, Ala and Cys, but decreases in the majority for Tyr, and increases versus decreases in equal numbers of comparisons for tRNA Asn (for the five remaining comparisons within species or genus levels). Considering pairwise comparisons for the sum of percentages of OL-like structures across all eight pairwise comparisons, values increase in all five comparisons that differ from zero (one tailed $P = 0.03$, sign test). This outcome overall independently confirms results from similar analyses for variation within genera in Agamidae (Section 3.1.1) and at family level (Table 1, Section 3.1).

A similar analysis using all pairwise comparisons, for Agamidae and cases from other families yields a general tendency for increase in OL-like structures associated with OL absence in the majority of tRNAs (exception tRNA Asn). This tendency is significant for the sum of percentages across all five tRNAs (one tailed $P = 0.037$, sign test, combining P values from Agamidae with those from this section using Fisher's method for combining P values (Fisher, 1950) yields a combined $P = 0.026$ for these two independent datasets).

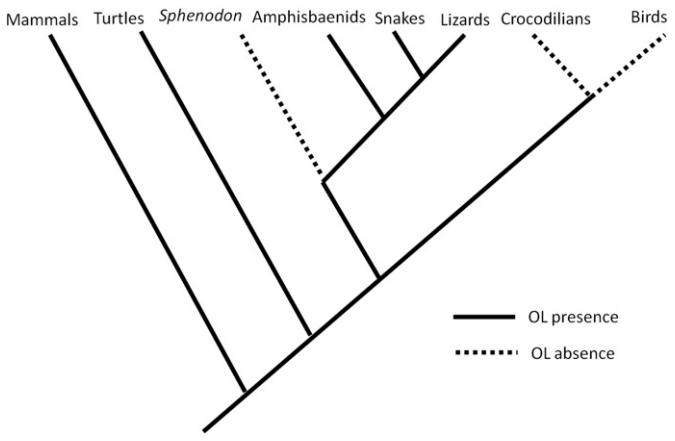


Fig. 3. Presence/absence of the mitochondrial light strand replication origin, OL, in the WANCY region of mitochondrial genomes mapped on the phylogeny of Amniota.

3.2. OL loss as a derived state in amniota

Is OL absence a derived state in Lepidosauria, or is it ancestral? This question is particularly justified because *Sphenodon* is basal to Lepidosauria and lacks an OL (Fig. 3). However, all bird and crocodile mitochondrial WANCY sequences available in GenBank by July 2013 lack an OL, so this is a generalized situation for the Archosauria. All snakes, excluding the primitive Typhloidea possess OL, Typhloidea lack OL, and almost all turtles, at the exception of two species, possess an OL. All mammal WANCY sequences include an OL, and more than 95% of amphibian and teleost (bony) fish species possess an OL. Hence, the most parsimonious evolutionary scenario is that OL presence is ancestral for higher vertebrates, and OL absence a derived state. However, for Sauria, two alternative scenarios seem a priori equally parsimonious: 1. OL loss for the common ancestor of Archosauria and Lepidosauria, and subsequent OL creation *de novo* at the same location in the WANCY region in Squamata (and several independent secondary losses as described in Table 1 and Fig. 2); 2. two independent OL losses for *Sphenodon* and Archosauria (and several independent secondary losses as described in Table 1 and Fig. 2).

Scenario 1 assumes two major events, one loss and one *de novo* production of OL, while scenario 2 assumes two major OL losses. Considering the possibility that evolutionary losses are reversible (Kohlsdorf and Wagner, 2006), parsimony does not distinguish between scenarios 1 and 2, unless one considers the probability to form *de novo* an OL at the exact same location within the WANCY region. Not only it is plausible that the probability of a loss is much greater than that of a gain (sequence deletion is simpler than *de novo* creation of a specific stem-loop structure), the plausibility of *de novo* creation of a new OL at the same location as the one that is lost is particularly unlikely. Hence, scenario 1 is much less probable than scenario 2. Therefore, OL presence as the ancestral state for Lepidosauria is the most likely scenario (Fig. 3).

3.3. What associates with OL presence/absence?

Results suggest that the function of initiating light strand replication is actually distributed among a number of genes, in a more or less consistent hierarchical way: recognized OL first, then tRNA Trp, Cys, Tyr, Asn and Ala (see Table 1). It is possible, as indicated by intraspecific comparisons in Table 1 and Fig. 2, that OL loss can occur without loss of replication origin function for the WANCY region, because the adjacent sequences can function as OL. One hypothesis is that loss of OL is a consequence of selection for reducing genome length. Indeed, as shown by Table 2, the mitochondrial genome of *Sphenodon* is the shortest among all complete lepidosaurian (excluding snakes) genomes available in Genbank. The only complete genome available for a xantusiid lizard, *Lepidophyma flavimaculatum*, is the second shortest in Table 2, and Agamid species without OL get the 3rd rank, followed closely by Agamid species with an OL. Hence, the absence of OL may be part of a general syndrome of genomic changes that includes a decrease in the mitochondrial genome length. According to Table 2, on average, agamid genomes possessing an OL are only 29 nucleotides longer than agamid genomes lacking an OL. This small difference, if meaningful, is about the length of the OL itself, hence agamid genomes with and without OL might differ in length only because of OL presence/absence. The only significantly shorter genome, the one of *Sphenodon*, reached this length by losing a long protein coding gene, which probably is exported to the nuclear chromosomes. This case is unique until now among vertebrate mitochondrial genomes, where, besides for *Sphenodon*, all 13 protein coding genes have always been observed. One could speculatively argue that these results suggest that besides exporting functionally important genes, OL loss is the last reduction occurring in the process of evolutionary reduction of mitochondrial genomes.

An alternative hypothesis involves developmental rates. These rates are inversely proportional to nuclear genome lengths (Chipman et al., 2001; Gregory, 2002; Sessions and Larson, 1987), probably because the rate of cell replication is partially determined by genome length. Therefore, selection for high developmental rates probably selects for fast replication, and hence shorter genomes. The same principle should hold for mitochondria, especially considering that mitochondrial genomes are more frequently replicated than nuclear chromosomes. Hence, the enhancement of OL-formation in adjacent tDNA could be an adaptation to maintain molecular processes of DNA replication after OL loss, as a result of pressures for genome size reduction, even considering that OL absence can barely account for the observed differences in genome lengths, as OLs are usually about 30 nucleotides long.

Table 2
Mean and standard deviation (sd) of the mitochondrial genome lengths in 71 lepidosaurian taxa (lizards, amphisbaenids and *Sphenodon*) according to families. N = number of species included in the analysis, by family.

Taxon	Length	Sd	N
Sphenodon[a]	15,181		1
Xantusiidae[a]	16,158		1
Shinisauridae	16,583		1
Agamidae, no OL	16,631	217	3
Agamidae, with OL	16,659	553	8
Amphisbaenia	16,703	418	8
Anguidae	16,748	1035	2
Helodermatidae	16,846		1
Cordylidae	17,184		1
Iguania[b]	17,283	698	10
Scincidae	17,407		1
Chamaeleonidae	17,837	725	12
Lacertidae	17,844	844	9
Gekkonidae	17,688	2757	11
Varanidae	18,000	723	2

[a] OL absent.
[b] Agamidae and Chamaelonidae excluded.

3.4. Is OL loss caused by formation of OL-like structures?

The association observed between OL presence/absence and capacities to form OL-like structures by WANCY tDNAs could result from tDNAs evolving after OL loss, or the opposite. If OL loss results from selection for shorter genomes, as suggested in the previous section, the increase in formation of OL-like structures by tDNAs is probably occurring after OL loss, as intraspecific comparisons indicate. The opposite, that greater capacity to form OL-like structures in WANCY tDNAs enables OL loss, is also possible. In the latter case, the greater capacity for OL formation evolved either by neutral drift, or under positive selection. Some evidence for positive selection is the fact that the formation of OL-like structures by WANCY tDNAs is associated with greater developmental stability in lizards (Seligmann and Krishnan, 2006) and in humans (Seligmann et al., 2006b). If one assumes that developmental stability is under positive selection, increased OL formation by tDNAs could have secondarily allowed the loss of the regular OL, under selection for shorter genomes. Hence, assuming that OL function is necessary, two potential evolutionary scenarios exist: 1. selection for OL loss, which drives afterwards an increase in formation of OL-like structures by WANCY tDNAs; and 2. increased formation of OL-like structures by WANCY tDNAs occurs due to drift or positive selection, which allows at a later stage, the loss of OL.

Distinguishing between these two potential evolutionary scenarios is not straightforward, and the scarcity of data does not yield clear cut answers to this. In addition, different scenarios may occur in different cases. The first scenario seems most parsimonious, as it implies two basic processes: the necessity for sequences with OL function, and selection in some taxa for shorter genomes. The second scenario assumes

that pressure associated with developmental stability would be primary (in historic order as well as importance), and the pressure associated with metabolic efficiency would be secondary. This is unlikely, but at least evidence exists that suggests effects of selection for developmental stability on mitochondrial sequences, other than the OL: developmental stability increases with rRNA chemical stability (Seligmann, 2006), and the density of off frame stop codons in protein coding sequences, which putatively decrease impacts of frameshifted protein synthesis after ribosomal slippage during mRNA-reading (Seligmann, 2007, 2010c). All these examples can be considered as special cases of selection for greater metabolic efficiency: ribosomal efficiency during protein synthesis (Seligmann, 2006, 2010c) and replicative efficiency (Seligmann and Krishnan, 2006; Seligmann et al., 2006b; and the results presented here). Hence, the second scenario is that the selection for greater metabolic efficiency that is associated with developmental stability is quantitatively more important than other components, such as those minimizing costs of DNA synthesis or the metabolic costs of amino acid synthesis (Akashi and Gojobori, 2002; Raiford et al., 2008; Seligmann, 2003, 2012f), and protein synthesis (Alves and Savageau, 2005; Brocchieri and Karlin, 2005). However, we consider this less likely, because the evidence cited above for cost minimization of metabolic processes is stronger and more numerous than that related to developmental stability.

The difference in formation of OL-like structures between pairs of taxa with and without OL correlates positively with the phylogenetic distance between these taxa for OL-formation by tDNAs Trp (see Fig. 4a), Ala, Asn and Tyr ($r = 0.50$, $P = 0.049$; $r = 0.27$, $P = 0.31$; $r = 0.18$, $P = 0.50$; and $r = 0.46$, $P = 0.073$, two tailed tests, respectively) and negatively for tRNA Cys ($r = -0.47$, $P = 0.068$, two tailed test, Fig. 4b). Summing the difference for all four tRNAs with a positive correlation (tRNAs Trp, Ala, Asn and Tyr) yields $r = 0.64$ (two tailed $P = 0.008$; Fig. 4c). A similar pattern is obtained if this sum includes tRNA Cys for which the correlation is negative ($r = 0.54$, two tailed $P = 0.03$). The ranked phylogenetic distances follow the principle that closest are sequences from the same species, then those from the same genus, followed by those of the same family, and so on. Overall, the difference in formation of OL-like structures apparently increases with time since OL loss (Fig. 4). The two points for the shortest evolutionary distances suggest that the difference was above zero just after OL loss, decreased, and then increased after that for most tRNAs (Figs. 4a and c). For tRNA Cys, the initial pattern of increased OL-formation right after OL-loss, and the decrease after this resembles what is observed in other tRNAs. The difference between tRNA Cys and other tRNAs resides in the fact that in the long term, the increase does not seem to occur in tRNA Cys, and that the decrease in OL-formation capacity continues.

These results indicate that the individuals that do not possess an OL, as compared to conspecifics that possess an OL, tend to have, on average, tRNAs with slightly greater capacities to form OL-like structures. This might result from a preexisting situation in these individuals which enabled OL loss, or that OL loss drives positive selection, among individuals lacking an OL, for greater capacities to form OL-like structures by adjacent tRNAs. The observation that in some cases (i.e. *U. flaviceps*, *E. anthracinus* and *E. inexpectatus*), the tRNA sequences of the individual lacking an OL have identical capacities for formation of OL-like structures as those from individuals with regular OL rather indicates that the second scenario, where mutations in tDNAs that increase formation of OL-like structures are favored by selection, is the simplest scenario that fits the results. Indeed, it also fits the observations that in some species (*C. ocellatus*, *E. skiltonianus*), formation of OL-like structures is identical in individuals with and without OL for four and two tRNAs, respectively. The fact that increased formation of OL-like structures in association with OL-absence occurs in all tRNAs only for one species where variation in OL presence/absence exists (*Enyalioides laticeps*), suggests that the scenario where preexisting high formation for OL-like structures in the individual that loses the OL is the less likely, or the rarest scenario.

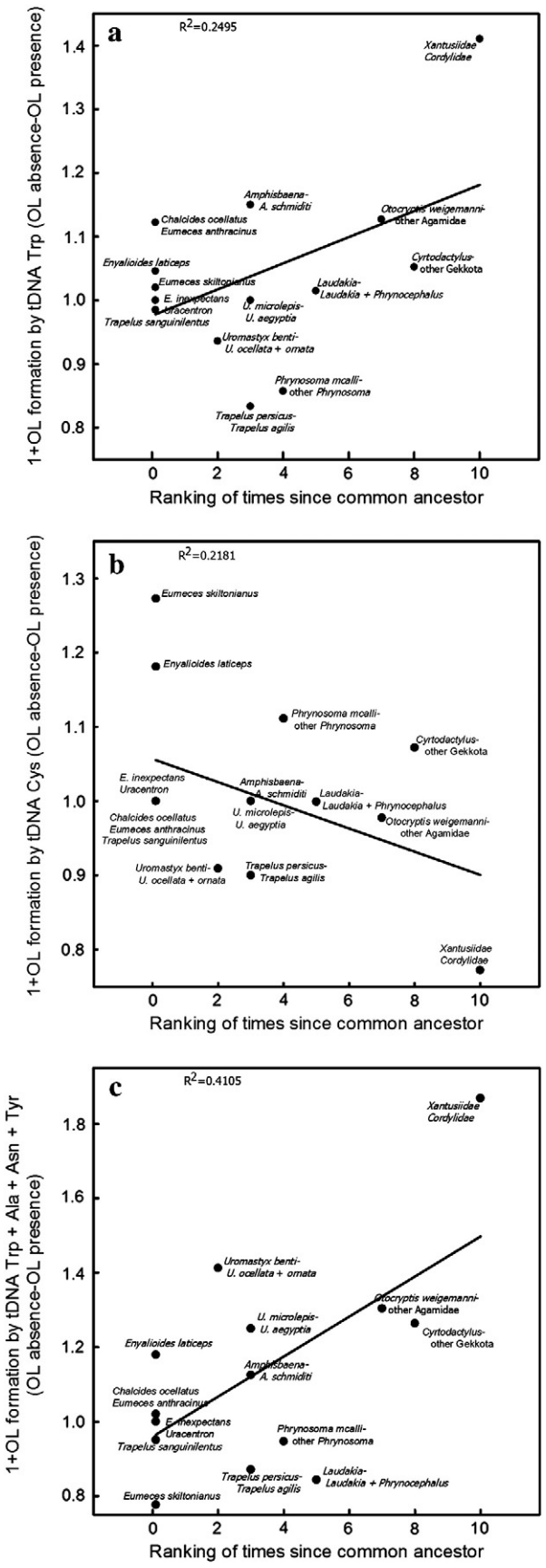


Fig. 4. Difference between formation of OL-like structures by WANCY tDNAs in taxa with and without OL, as a function of a ranking of times since their last common ancestor: a—tDNA Trp, b—tDNA Cys, and c—the sum of tRNAs Trp, Ala, Asn and Tyr.

3.5. Evolutionary variation in formation of OL-like structures by tDNAs

In Helodermatidae, OL function is more or less fulfilled only by the regular OL, while in most other groups in Table 1 several tRNAs have the capacity to form OL-like structures. In most tRNAs, variation exists in the extents of OL-forming capacities, independently of phylogenetic affiliations, suggesting multiple gains and losses of OL function for each tDNA. This principle is valid also within families, resembling results previously reported for primates (Seligmann et al., 2006a). In *Liolaemus*, for example, tDNA of tRNA Trp does not form OL-like structures in *L. lutzae*. However, 20% of the structures formed by that tDNA are OL-like in *L. dorbignyi* and *L. ruibali*, up to 83% in *L. chiliensis* and *L. tenuis* and 92% for *L. ornatus*. For 27 among 36 *Liolaemus* species included in this study, tRNA Ala forms no OL-like structures, but it forms 60% and 50% OL-like structures in *L. chaltin* and *L. fabiani*, respectively. In *Liolaemus*, the minimal OL-forming tendency is null also for tRNA Cys and tRNA Asn, and 14% for tRNA Tyr. The maximum is 63% for tRNA Cys, 83% for tRNA Tyr and 90% for tRNA Asn. Results suggest that the capacity to form OL-like structures is highly unstable from an evolutionary point of view because it varies widely, even among closely related species.

3.6. Tradeoffs for OL function

The tendency of a tRNA to function as OL may be predicted by three tRNA properties: 1. The relative importance of their primary function as tRNAs in protein synthesis, expecting tRNAs with rare cognates to form more OL-like structures. 2. If coding for the tRNA and the OL function is split between different DNA strands of the same tRNA gene. In this way, fewer constraints should be associated with each function than when the same DNA strand fulfills both functions. 3. Formation of several OL-like structures within a short range on the heavy strand DNA might create steric and other interactions that could perturb proper OL function by any of the potential alternative OL-like structures.

3.6.1. Property 1: Amino acid usage

This predicts negative correlations between the mitochondrial genome-wide usage of the cognate amino acid of a tRNA and formation of OL-like structures by its tDNA. For example, the only group where tRNA Ala has a relatively high capacity for formation of OL-like structures, Varanidae, has the lowest genome-wide usage for that amino acid among the Lepidosauria included here. This negative correlation was confirmed independently at two levels. First, comparisons of homologous tRNA genes in different families show that the family mean capacity to form OL-like structures is negatively correlated with family mean cognate amino acid usage in three among five tRNAs (tRNA Asn, $r = -0.47$, $P = 0.033$, tRNA Cys, $r = -0.60$, $P = 0.007$ and tRNA Tyr, $r = -0.43$, $P = 0.049$, one tailed tests). Although correlations were not significant in the two remaining tRNAs, they were also negative (tRNA Trp, $r = -0.07$, $P = 0.40$; tRNA Ala, $r = -0.33$, $P = 0.11$, one tailed tests). Second, comparing non-homologous tRNA species within the same taxon, tRNAs Trp and Cys, which have the highest capacities to form OL-like structures, have the rarest cognate amino acids, while tRNA Ala only rarely forms OL-like structures, and has the most common cognate amino acid among all WANCY tRNAs (Table 1).

3.6.2. Property 2: Coding strand and OL formation

Here, the prediction is that for tRNA genes whose light strand templates for the expressed tRNA, their heavy strand DNA is more free to evolve the secondary function as OL than in tRNA genes where the strand expressed as tRNA is the heavy strand tRNA, which hence has to fulfill both tRNA and OL functions. The only light strand-expressed tRNA in the WANCY region is tRNA Trp, which indeed, according to our results (see Table 1 and Fig. 2) has the highest capacity to form OL-like structures. The fact that in the absence of the regular OL, the increase in OL formation capacity is lower than for tRNA Cys might be because tRNA Trp has already a high capacity in the presence of an OL, and also because genome-wide Trp usage is greater than for Cys.

3.6.3. Property 3: Steric interactions

This principle, as the above mentioned property 2, is related to structural constraints on a single DNA strand. This predicts that the capacity to form OL-like structures for two genes should be inversely proportional to proximity between them. The Pearson correlation coefficient matrix between formation of OL-like structures by tRNAs, and with the -ΔG of the OL where it is present, is proportional to the distance between genes (column 'Dist' in Table 1). Though for half the families, the correlation in Dist is positive, the only significant cases are for positive correlations (*Liolaemus* and *Sceloporus*), which coincide with the largest sample sizes. For taxa with small samples, Dist does not differ statistically from zero, which is predictable because of enhanced effects of sampling error due to small sample sizes. These results indicate that the most positive correlations are usually between tRNA Trp and tRNA Tyr, which are at opposite extremities of the WANCY region, and the most negative correlations are between adjacent gene pairs.

3.7. Robustness of estimates of OL formation capacities

Data indicate that our estimates of the structure of correlation matrices between OL formation capacities are repeatable, as we get similar results for different families for which we analyzed a reasonable number of taxa (e.g., Liolaemidae, Phrynosomatidae, Polychridae). This repeatability suggests that the method used to estimate capacities of OL formation is robust, despite sampling biases due to relatively small samples (i.e., <200 species in a genus, see Steppan, 1997), and the lack of phylogenetic correction when calculating correlations. In addition, the percentage of OL-like structures among all secondary structures included, does not take into account the different stabilities of the various structures: if, for example, we obtained a single OL-like structure among 10 alternative secondary structures, our method does not distinguish between cases where this OL-like structure is the most versus the least stable one. Further calculations integrating over the ΔGs of the secondary structures are possible.

The issue of inaccuracies in secondary structure prediction as adding noise and hence not being responsible for the (repeatable) positive results described here, has been already discussed (Krishnan et al., 2008). The major point is that inaccuracies in predictions are very unlikely to produce similar results in independent cases. In addition, a previous study that compared the rough percentages such as those used here, and calculations integrating over ΔGs, showed similar results for both calculation methods, but that the rough estimates tend to fit slightly better predictions than the presumably more precise calculations (Seligmann and Krishnan, 2006). This might be because the precise calculations rely more heavily on parameters required for predicting ΔGs and the Boltzmann distribution of secondary structures. However, these parameters are usually not available: temperature, ionic concentrations, and interactions of the OL-forming sequences with unknown molecules that might enhance or prevent OL formation, or alter the stabilities of the OL-like structures. Even for a relatively well-known parameter, temperature, this can be problematic, as ranges and distributions of body temperatures have to be equally well-known for all the taxa considered, which is even more complicated when restrictions need to be included (e.g., season of activity, sex, reproductive condition). Even in 'homeotherms' body temperature varies between body parts (i.e., the body trunk vs. appendages). In this context, we suggest that in the absence of adequate precise information, the rough method used is most adequate for qualitative answers.

4. General conclusions

Formation of OL-like structures by tRNA genes is greater in lepidosaurian mitochondria lacking OL, interpreted as most frequently

resulting from OL loss. Though patterns in data are not conclusive, they suggest that this greater capacity usually develops after OL loss, rather than precludes it. Hence OL-loss would drive selection for greater capacities to form OL-like hairpins by tDNAs in the vicinity of the regular OL, and adjacent tDNAs compensate for OL absence. The increase in formation of OL-like structures by tDNAs includes complex patterns of interaction between tDNA genes, which involve an avoidance of competition between different tRNA genes and between different functions of a tRNA gene. This suggests that the evolution of secondary functions, and of the multi-gene complex that functions as OL is not random, but follows simple functional principles. In most multi-gene complexes, the various genes are not interchangeable, but code for different parts of the functional complex. In this particular case, the situation is simplified because any gene can function as an OL, which gives a special opportunity to study the evolution of gene interactions and their evolution in multi-gene complexes.

Conflict of interest

The authors have no conflict of interest.

Acknowledgments

We thank Thomas F. Hansen and all members of his lab at the University of Oslo for discussing results, and to one anonymous reviewer for comments. This project was partially funded by FONDECYT 1120181 to AL.

References

Akashi, H., Gojobori, T., 2002. Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. Proceedings of the National Academy of Sciences of the United States of America 99, 3695–3700.

Alves, R., Savageau, M.A., 2005. Evidence of selection for low cognate amino acid bias in amino acid biosynthetic enzymes. Molecular Microbiology 56, 1017–1034.

Arquès, D.G., Michel, C.J., 1997. A code in the protein coding genes. Biosystems 44, 107–134.

Brocchieri, L., Karlin, S., 2005. Protein length in eukaryotic and prokaryotic proteomes. Nucleic Acids Research 33, 3390–3400.

Chipman, A.D., Khaner, O., Haas, A., Tchernov, E., 2001. The evolution of genome size: what can be learned from anuran development? The Journal of Experimental Zoology 291, 365–374.

Faure, E., Dealaye, L., Tribolo, S., Levasseur, A., Seligmann, H., Barthélémy, R.M., 2011. Probable presence of an ubiquitous cryptic mitochondrial gene on the antisense strand of the cytochrome oxidase I gene. Biology Direct 6, 56.

Felsenstein, J., 1985. Phylogeneties and the comparative method. The American Naturalist 125, 1–15.

Fisher, R.A., 1950. Statistical Methods for Research Workers, 11th ed. Oliver and Boyd, Edinburgh and London 99.

Fuste, J.M., et al., 2010. Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. Molecular Cell 37, 67–78.

Giegé, R., 2008. Toward a more complete view of tRNA biology. Nature Structural & Molecular Biology 15, 1007–1014.

Gregory, T.R., 2002. Genome size and developmental parameters in the homeothermic vertebrates. Genome 45, 833–838.

Hassanin, A., Léger, N., Deutsch, J., 2005. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of metazoan, and consequences for phylogenetic inferences. Systematic Biology 54, 277–298.

Hixson, J.E., Wong, T.W., Clayton, D.A., 1986. Both the conserved stem-loop and divergent 5′-flanking sequences are required for initiation at the human mitochondrial origin of light-strand DNA-replication. The Journal of Biological Chemistry 261, 2384–2390.

Kohlsdorf, T., Wagner, G.P., 2006. Evidence for reversibility of digit loss: a phylogenetic study of limb evolution in *Bachia* (Gymnophthalmidae: Squamata). Evolution 60, 1896–1912.

Krishnan, N.M., Seligmann, H., Raina, S.Z., Pollock, D.D., 2004. Detecting gradients of asymmetry in site-specific substitutions in mitochondrial genomes. DNA and Cell Biology 23, 707–714.

Krishnan, N.M., Seligmann, H., Rao, B.J., 2008. Relationship between mRNA secondary structure and sequence variability in chloroplast genes: possible life history implications. BMC Genomics 9, 48.

Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Research 25, 955–964.

Macey, J.R., Larson, A., Ananjeva, N.B., Fang, Z., Papenfuss, T.J., 1997. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Molecular Biology and Evolution 14, 91–104.

Macey, J.R., Schulte, J.A., Larson, A., 2000a. Evaluating transtethys migration: an example using acrodont lizard phylogenetics. Systematic Biology 49, 233–256.

Macey, J.R., Schulte, J.A., Larson, A., 2000b. Evolution and phylogenetic information content of mitochondrial genomic structural features illustrated with acrodont lizards. Systematic Biology 49, 257–277.

Maizels, N., Weiner, A.M., 1994. Phylogeny from function: evidence from the molecular fossil record that tRNA originated in replication, not translation. Proceedings of the National Academy of Sciences of the United States of America 91, 6729–6734.

Mueller, R.L., Macey, J.R., Jaekel, M., Wake, D.B., Boore, J.L., 2004. Morphological homoplasy, life history evolution, an historical biogeography of plethodontid salamander inferred from complete mitochondrial genomes. Proceedings of the National Academy of Sciences of the United States of America 101, 13820–13825.

Raiford, D.W., Heizer, E.M., Miller, R.V., Akashi, H., Raymer, M.L., Krane, D.E., 2008. Do Amino acid biosynthetic costs constrain protein evolution in *Saccharomyces cerevisiae*? Journal of Molecular Evolution 67, 621–630.

Reyes, A., Gissi, C., Pesole, G., 1998. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. Molecular Biology and Evolution 15, 957–966.

Rhodakis, G.C., Cao, L.Q., Mizi, A., Kenchington, E.L.R., Zouros, E., 2007. Nucleotide content gradients in maternally and paternally inherited mitochondrial genomes of the mussel *Mytilus*. Journal of Molecular Evolution 65, 124–136.

Schattner, P., Brooks, A.N., Loewe, T.M., 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Research 33, W686–W689.

Seligmann, H., 2003. Cost minimization of amino acid usage. Journal of Molecular Evolution 56, 151–161.

Seligmann, H., 2007. Cost minimization of ribosomal frameshifts. Journal of Theoretical Biology 249, 162–167.

Seligmann, H., 2006. Error propagation across levels of organization: From chemical stability of ribosomal RNA to developmental stability. Journal of Theoretical Biology 242, 69–80.

Seligmann, H., 2008. Hybridization between mitochondrial heavy strand tDNA and expressed light strand tRNA modulates the function of heavy strand tDNA as light strand replication origin. Journal of Molecular Biology 379, 188–199.

Seligmann, H., 2010a. Mitochondrial tRNAs as light strand replication origins: similarity between anticodon loops and the loop of the light strand replication origin predicts initiation of DNA replication. Biosystems 99, 85–93.

Seligmann, H., 2010b. Avoidance of antisense, antiterminator tRNA anticodons in vertebrate mitochondria. Biosystems 101, 42–50.

Seligmann, H., 2010c. The ambush hypothesis at the whole-organism level: off-frame, 'hidden' stops in vertebrate mitochondrial genes increase developmental stability. Computational Biology and Chemistry 34, 80–85.

Seligmann, H., 2011a. Two genetic codes, one genome: frameshifted primate mitochondrial genes code for additional proteins in presence of antisense antitermination tRNAs. Biosystems 105, 271–285.

Seligmann, H., 2011b. Mutation patterns due to converging mitochondrial replication and transcription increase lifespan, and cause growth rate longevity-tradeoffs. In: Seligmann, H. (Ed.), DNA Replication: Current Advances. InTech. http://dx.doi.org/10.5772/24319.

Seligmann, H., 2012a. Coding constraints modulate chemically spontaneous mutational replication gradients in mitochondrial genomes. Current Genomics 13, 34–57.

Seligmann, H., 2012b. An overlapping genetic code for frameshifted overlapping genes in *Drosophila* mitochondria: antisense antitermination tRNAs UAR insert serine. Journal of Theoretical Biology 298, 51–76.

Seligmann, H., 2012c. Overlapping genetic codes for overlapping frameshifted genes in Testudines, and *Lepidochelys olivacea* as special case. Computational Biology and Chemistry 41, 18–34.

Seligmann, H., 2012d. Putative mitochondrial polypeptides coded by expanded quadruplet codons, decoded by antisense tRNAs with unusual anticodons. Biosystems 110, 84–106.

Seligmann, H., 2012e. Overlapping genes coded in the 3′-to-5′-direction in mitochondrial genes and 3′-to-5′ polymerization of non-complementary RNA by an 'invertase'. Journal of Theoretical Biology 315, 38–52.

Seligmann, H., 2012f. Positive and negative cognate amino acid bias affects compositions of aminoacyl-tRNA synthetases and reflects functional constraints on protein structure. BIO 2, 11–26.

Seligmann, H., 2013a. Putative anticodons in mitochondrial tRNA sidearm loops: pocketknife tRNAs? Journal of Theoretical Biology 340, 155–163.

Seligmann, H., 2013b. Pocketknife tRNA hypothesis: anticodons in mammal mitochondrial tRNA side-arm loops translate proteins? Biosystems 113, 165–176.

Seligmann, H., 2013c. Triplex DNA:RNA, 3′-to-5′ inverted RNA and protein coding in mitochondrial genomes. Journal of Computational Biology 20, 660–671.

Seligmann, H., 2013d. Systematic asymmetric nucleotide exchanges produce human mitochondrial RNAs cryptically encoding for overlapping protein coding genes. Journal of Theoretical Biology 324, 1–20.

Seligmann, H., 2013e. Polymerization of non-complementary RNA: systematic symmetric nucleotide exchanges mainly involving uracil produce mitochondrial RNA transcripts coding for cryptic overlapping genes. Biosystems 111, 156–174.

Seligmann, H., 2013f. Replicational mutation gradients, dipole moments, nearest neighbor effects and DNA polymerase gamma fidelity in human mitochondrial genomes. In: Stuart, D. (Ed.), The Mechanisms of DNA Replication. InTech. http://dx.doi.org/10.5772/51245.

Seligmann, H., Krishnan, N.M., 2006. Mitochondrial replication origin stability and propensity of adjacent tRNA genes to form putative replication origins increase

developmental stability in lizards. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution 306, 433–449.
Seligmann, H., Labra, A., 2013. Tetracoding increases with body temperature in Lepidosauria. Biosystems 114, 155–163.
Seligmann, H., Krishnan, N.M., Rao, B.J., 2006a. Possible multiple origins of replication in primate mitochondria: alternative role of tRNA sequences. Journal of Theoretical Biology 241, 321–332.
Seligmann, H., Krishnan, N.M., Rao, B.J., 2006b. Mitochondrial tRNA sequences as unusual replication origins: pathogenic implications for *Homo sapiens*. Journal of Theoretical Biology 243, 375–385.
Sessions, S.K., Larson, A., 1987. Developmental correlates of genome size in plethodontid salamanders and their implications for genome evolution. Evolution 41, 1239–1251.
Seutin, G., Lang, B.F., Mindell, D.P., Morais, R., 1994. Evolution of the WANCY region in amniote mitochondrial DNA. Molecular Biology and Evolution 11, 329–340.
Steppan, S.J., 1997. Phylogenetic analysis of phenotypic covariance structure. 1. Contrasting results from matrix correlation and common principal component analyses. Evolution 51, 571–586.
Stergachis, A.B., et al., 2013. Exonic transcription factor binding directs codon choice and affects protein evolution. Science 342, 1367–1372.
Tanaka, M., Ozawa, T., 1994. Strand asymmetry in human mitochondrial DNA mutations. Genomics 22, 327–335.
Wanrooij, S., Falkenberg, M., 2010. The human mitochondrial replication fork in health and disease. Biochimica et Biophysica Acta 1797, 1378–1388.
Yu, C.H., Liao, J.Y., Zhou, H., Qu, L.H., 2008. The rat mitochondrial Ori L encodes a novel small RNA resembling an ancestral tRNA. Biochemical and Biophysical Research Communications 372, 634–638.
Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31, 3406–3415.