We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS

BOOK CITATION INDEX

INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Swinger RNAs in the Human Mitochondrial Transcriptome

Ganesh Warthi and Hervé Seligmann

Additional information is available at the end of the chapter

Abstract

Transcriptomes include coding and non-coding RNAs and RNA fragments with no apparent homology to parent genomes. Non-canonical transcriptions systematically transforming template DNA sequences along precise rules explain some transcripts. Among these systematic transformations, 23 systematic exchanges between nucleotides, i.e. 9 symmetric ($X \leftrightarrow Y$, e.g. $C \leftrightarrow T$) and 14 asymmetric ($X \rightarrow Y \rightarrow Z \rightarrow X$, e.g. $A \rightarrow T \rightarrow G \rightarrow A$) exchanges. Here, comparisons between mitochondrial swinger RNAs previously detected in a complete human transcriptome dataset (including cytosolic RNAs) and swinger RNAs detected in purified mitochondrial transcriptomic data (not including cytosolic RNAs) show high reproducibility and exclude cytosolic contaminations. These results based on next-generation sequencing Illumina technology confirm detections of mitochondrial swinger RNAs in GenBank's EST database sequenced by the classical Sanger method, assessing the existence of swinger polymerizations.

Keywords: swinger RNA, non-canonical transcription, mitogenome, systematic nucleotide exchange, blast analyses

1. Introduction

Transcription is an intracellular mechanism that produces RNA by DNA-dependant RNA polymerisation. RNAs coding for polypeptide chains are mRNAs translated by other transcription products, tRNAs and ribosomal RNAs. Some RNAs do not correspond to any DNA sequence in the genome, suggesting in some cases spontaneous emergence [1]. These RNAs remain usually unreported and are ignored. Similarly, proteomic data include numerous peptides that do not match canonical translation of predicted ORFs, but imply translation of stop

codons [2–8] by tRNAs with anticodons matching stops [9–11] or by tRNAs with expanded anticodons [12–14]. Assuming fusion of different transcripts explains the origins of some of these non-canonical RNAs [15]. Some human RNAs matching exons differ from their DNA by specific changes, called RDDs (RNA-DNA differences) [16]. RDDs can be single nucleotide substitutions or deletions [17–19], presumably resulting from post-transcriptional edition [20, 21]. Some short transcripts correspond to mitochondrial DNA at the condition that one assumes mono- or dinucleotide deletions after each transcribed nucleotide triplet [22, 23]. Formation of secondary structures by del-transformed sequences apparently downregulates del-transcription itself or its products, delRNAs [24].

Another type of systematic transformation consists of 23 systematic exchanges between nucleotides, 9 symmetric (X ↔ Y, e.g. A ↔ C,) [25, 26] and 14 asymmetric exchanges (X → Y → Z → X, e.g. A → C → G → A) [26, 27]. For example, in systematic transformation A ↔ C, nucleotide A is introduced in place of nucleotide C and vice versa. The two-headed arrow (↔) indicates that A and C replace each other during transcription. One-headed arrows (→) indicate asymmetric exchanges: in the example A → C → G → A, nucleotide A is systematically incorporated in place of every C; similarly, C replaces G and G replaces A during RNA polymerisation. Transcripts corresponding to systematic exchanges are called swinger RNAs. BLASTn analyses detect about 100 predicted swinger RNAs (longer than 100 nucleotides) in GenBank's EST database in addition to the (approximately) 10,000 canonical human mitochondrial RNAs in that database. Hence, about 1% of the human mitochondrial transcripts in GenBank's EST database correspond to 1 among 23 systematic nucleotide exchanges [25–28]. These systematic nucleotide exchanges (an expression that fits chemical contexts) are called bijective transformations in mathematical contexts [29–31]; swinger transcription fits biological contexts.

Mitogenomes are comparatively small, also because of the selection against multiple direct repeats [32–35] and invert repeats [15]: these form secondary structures that are frequently excised; such deletions are frequently deleterious. Vertebrate mitogenomes have densely packed coding and non-coding regions templating for RNAs. Non-canonical transformations greatly increase potential numbers of RNA products for single sequences: four and five RNA transcripts when assuming systematic deletions of mono- and dinucleotides for del-transcriptions, respectively, and 23 swinger RNAs when considering systematic nucleotide exchanges. Therefore, studies of swinger transformations focus on the human mitogenome, which is short (16,569 bp), hence reducing potential false-positive detections due to sheer genome size and because ample sequence data are available from several sources for this organism.

Note that swinger DNA has been detected (mainly corresponding to rRNA genes) for mitochondrial and nuclear sequences [36–38]. Hence, swinger RNAs result from canonical transcription of swinger-transformed DNA or swinger transcription of regular DNA [22]. Some mass spectra match predicted peptides translated from del- and swinger-transformed RNA [39–42]. Detection of chimeric RNAs, consisting of part regular, and part swinger-transformed contiguous sequences suggests that regular canonical and swinger-transformed RNA result

from single polymerisation events, probably by the same polymerase [43]. Peptides corresponding to such chimeric RNAs also occur [44].

Secondary structure formation by swinger-transformed sequences associates with swinger RNA detection [45], suggesting regulation of swinger RNA processing by secondary structures, as observed for canonical mitochondrial RNAs, i.e. tRNA punctuation [46].

Abundances of human mitochondrial swinger RNAs detected in GenBank's EST database [25, 26], originating from various sources using Sanger sequencing, are proportional to those detected in transcriptomic data produced by next-generation sequencing, Illumina technology [47]. Similarly, abundances and lengths of swinger RNAs detected in *Mimivirus*' transcriptome sequenced by 454 technology are proportional to those detected when using SOLID sequencing [01]. These analyses confirmed that swinger RNAs are not sequencing artefacts due to specific sequencing technologies, but data sources do not exclude contamination by cytosolic RNA. Here, we compare the previously described human mitochondrial swinger transcriptome [39] from a complete human transcriptome (including cytosolic RNAs) with the swinger transcriptome as detected in purified human mitochondrial lines [48]. Reproducibility of swinger RNA coverages of the human mitogenome would exclude sequencing artefacts and cytosolic contaminations as alternative explanations for hypothetical swinger RNAs. We predict (1) the detection of swinger RNAs from transcriptomic data extracted from purified mitochondrial lines and (2) high similarities between mitogenomic swinger RNA coverages described here and previously [39].

2. Materials and methods

2.1. Detection of swinger RNAs

We used GenBank's BLASTn ('somewhat similar sequences' with default alignment parameters) [49] for in-silico alignment searches between each of the 23 swinger-transformed versions of the human mitogenome (NC_012920) and transcriptomic data in GenBank's Sequence Read Archive (SRA) (SRX084350-SRX084355 and SRX087285), sequenced by RNA-Seq, Illumina HiSeq 2500 technology [48]. Alignments with more than 80% identity were recorded and used as a swinger RNA candidate for further analysis.

2.2. Mitogenomic gene coverage by swinger RNAs

Locations of detected swinger RNAs were recorded by mapping these RNAs on the human mitogenome. We analyse separately 17 mitogenomic regions: the D-loop, 2 ribosomal RNAs (12S and 16S), 13 protein-coding genes involved in the electron transport chain and the WANCY region (intragenic region between ND2 and CO1 that templates for tRNAs with cognate amino acids W, A, N, C and Y). Percentage coverages by detected swinger RNAs were calculated for each swinger transformation in each selected mitogenomic region and used for further statistical analyses.

3. Results and discussion

3.1. Swinger RNAs in the human mitochondrial transcriptome

Table 1 summarises results from BLASTn analyses of the purified mitochondrial transcriptome [48] for the 23 swinger-transformed versions of the human mitogenome. In total 4120 reads aligned with the 23 swinger-transformed versions of the human mitogenome,

	Read	Contigs	Id	Coverage	Read	Contigs	Id	Coverage
Regular	700	142	99.76	8479	400	69	100	7658
A ↔ C	181	42	93.94	1239	163	10	89.12	389
A ↔ G	448	5	92.20	184	6	2	87.26	97
A ↔ T	98	48	94.34	1376	186	17	90.63	593
C ↔ G	319	7	89.99	227	28	8	85.01	343
C ↔ T	338	51	90.59	1599	253	26	92.47	937
G ↔ T	435	35	92.43	1614	400	5	88.36	249
A ↔ C + G ↔ T	123	25	89.79	730	11	2	86.23	76
A ↔ G + C ↔ T	63	34	92.63	982	69	8	89.3	257
A ↔ T + C ↔ G	126	36	92.87	1112	97	11	94.8	395
A → C → G → A	80	25	93.81	**778**	21	12	90.59	439
A → C → T → A	160	52	92.99	1474	31	12	90.79	453
A → G → C → A	98	58	91.70	1729	43	17	93.13	573
A → G → T → A	98	39	91.29	1245	28	12	88.49	469
A → T → C → A	218	50	93.12	1578	363	20	92.13	810
A → T → G → A	84	29	94.08	873	12	5	90.21	190
C → G → T → C	122	43	94.15	1167	21	4	88.44	247
C → T → G → C	165	57	91.90	2058	126	17	86.42	791
A → C → G → T → A	140	45	92.33	1306	38	15	89.53	575
A → C → T → G → A	157	35	95.32	971	54	10	92.65	327
A → G → C → T → A	211	26	92.63	713	99	15	87.35	530
A → G → T → C → A	78	27	93.60	777	30	11	91.29	369
A → T → C → G → A	195	17	88.94	463	60	9	92.1	321
A → T → G → C → A	183	55	94.14	1874	115	16	94.78	601

Current data are from purified mitochondrial lines, previous data are from complete human transcriptome, including cytosolic and mitochondrial transcriptomes. Columns 2–5: current analyses. Columns are (1) swinger transformation (includes lack of transformation), (2) aligning read numbers, (3) contig numbers, (4) mean identity between reads and transformed mitogenome and (5) total mitogenomic coverage by all swinger contigs. Columns 6–9 indicate corresponding data in the same order for the previous study.

Table 1. Total human mitogenome coverage by detected swinger RNAs from current (2018) and previous (2016) [39] analyses of two different datasets sequenced by Illumina.

Figure 1. Total mitogenome coverages by swinger RNAs across the complete human mitogenome in previously analysed data [39] (y-axis) as a function of those obtained in current observations from purified mitochondrial lines.

producing 841 contigs. The highest detected identity between a theoretical mitogenome swinger transformation and a read was 95.32% for transformation A → C → T → G → A, and the lowest identity was 88.94% for transformation A → T → C → G → A. The overall identity averaged at 92.86%. A previous swinger analysis of other transcriptomic data [39] found swinger transformations A ↔ G and C ↔ T least and most frequent, respectively. Here, swinger transformation A ↔ G remains the least frequent; C ↔ T is the second most frequent, suggesting high reproducibility.

Total mitogenome coverages by swinger RNAs for each transformation were plotted as a function of corresponding coverages from a previous analysis published in 2016 [39]. Coverages are positively correlated (Pearson correlation coefficient r = 0.669, one-tailed p = 0.0002, **Figure 1**). Coverages for the purified mitochondrial line transcriptomes are systematically greater than for those for previous analyses (supplementary data and **Table 1**) [39].

3.2. Gene-level comparisons of swinger RNA coverages

Swinger RNA coverages of each of the 17 mitogenomic regions (D-loop, 2 rRNAs, 13 CDs and the WANCY region) are in **Tables 2** and **3**, for analyses of current and previous Illumina data [39], respectively. Pearson correlation coefficients between swinger coverages were calculated considering (1) genes, i.e. for each gene across the 23 different transformations, and (2) for each swinger transformation, across the 17 different mitogenomic regions.

Transformation	D-loop	12S	16S	ND1	ND2	W-Y	CO1	CO2	ATP8	ATP6	CO3	ND3	ND4L	ND4	ND5	ND6	Cytb
A ↔ C	6.4	5.6	9.1	0.0	15.1	25.0	5.8	0.0	9.2	3.7	7.3	11.8	0.0	8.3	7.9	26.7	3.7
A ↔ G	0.0	0.0	0.0	0.0	0.0	0.0	2.4	0.0	0.0	0.0	0.0	9.0	0.0	0.0	2.7	0.0	6.0
A ↔ T	9.4	2.8	8.3	3.2	15.9	11.5	1.7	3.1	22.2	0.0	7.7	7.5	8.4	20.4	10.6	20.2	2.5
C ↔ G	0.0	0.0	0.0	0.0	6.4	4.6	1.9	0.0	11.6	0.0	0.0	8.4	6.7	0.0	0.0	7.2	0.0
C ↔ T	13.9	2.5	9.9	11.3	8.8	0.0	0.1	1.5	28.5	13.5	6.8	9.5	0.0	5.4	17.9	41.5	11.4
G ↔ T	20.9	0.0	10.6	3.8	13.3	0.0	0.0	0.0	33.3	13.4	0.0	0.0	0.0	13.2	18.9	46.9	9.6
A ↔ C + G ↔ T	2.6	5.2	7.7	0.0	5.6	5.6	8.4	4.1	0.0	0.0	0.0	0.0	0.0	4.4	5.5	11.4	1.8
A ↔ G + C ↔ T	2.9	2.1	0.0	7.7	18.0	5.6	7.7	3.8	0.0	0.0	4.1	7.5	21.2	10.3	6.4	18.9	0.0
A ↔ T + C ↔ G	4.7	7.2	7.0	2.4	8.5	6.6	10.5	5.1	35.7	0.0	4.1	27.5	0.0	9.4	4.9	0.0	5.3
A → C → G → A	4.5	2.5	0.1	0.0	7.5	8.2	0.0	0.0	10.1	12.8	6.3	0.0	0.0	0.0	8.7	13.0	7.7
A → C → T → A	9.0	0.0	6.7	5.0	13.8	19.6	9.9	20.8	13.5	12.0	9.8	8.4	0.0	10.1	6.3	15.0	9.2
A → G → C → A	12.6	9.6	6.7	8.7	19.3	0.0	7.4	0.0	41.1	10.4	14.0	7.5	15.5	13.9	10.8	27.8	4.6
A → G → T → A	10.8	2.7	3.6	5.9	12.7	0.0	3.4	6.7	28.5	13.1	11.7	7.5	0.0	9.2	10.8	18.9	2.8
A → T → C → A	12.3	2.2	7.3	15.9	12.0	0.0	6.8	3.8	21.7	7.6	11.0	16.5	20.2	10.4	7.7	35.0	6.3
A → T → G → A	7.0	2.7	3.3	2.5	8.0	0.0	7.0	6.4	15.0	8.4	0.0	9.2	0.0	3.1	7.5	2.1	6.8
C → G → T → C	10.1	4.9	2.0	11.5	20.0	0.0	5.4	0.0	17.4	3.5	19.0	13.6	0.0	7.3	5.8	14.9	0.0
C → T → G → C	14.4	0.6	4.7	10.3	12.4	0.0	8.2	14.2	41.1	7.6	3.7	0.0	7.1	24.5	13.8	46.5	14.8
A → C → G → T → A	11.6	5.0	4.0	10.7	20.7	0.0	3.8	3.1	15.0	4.4	6.1	8.1	0.0	4.2	14.2	21.7	3.0
A → C → T → G → A	10.4	3.0	5.3	0.0	5.1	15.6	1.7	8.0	0.0	4.0	7.5	8.4	0.0	1.5	11.6	16.4	9.6
A → G → C → T → A	5.9	2.3	2.2	0.0	17.3	0.0	3.8	4.2	0.0	3.4	0.0	0.0	4.4	0.0	10.0	0.0	5.0
A → G → T → C → A	2.9	0.0	3.7	2.9	7.5	2.6	5.0	0.0	0.0	11.2	5.1	0.0	9.8	2.8	7.5	13.9	4.6
A → T → C → G → A	2.3	2.8	3.5	0.0	2.9	0.0	3.1	1.3	0.0	0.0	6.1	0.0	0.0	6.5	2.9	0.0	2.2
A → T → G → C → A	16.9	4.1	10.4	8.7	17.8	13.0	6.1	10.4	26.6	6.8	0.0	0.0	0.0	4.4	17.1	39.2	17.7

Table 2. Percentage coverage of mitogenomic regions by swinger RNAs in this study.

Transformation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	r	P
A ↔ C	4.1	0.0	0.0	0.0	3.1	0.0	0.0	0.0	14.5	0.0	0.0	9.5	9.4	7.0	3.8	8.4	0.0	0.21	0.418
A ↔ G	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.198	0.446
A ↔ T	8.9	4.1	4.2	0.3	7.7	0.0	1.6	0.0	17.9	13.0	0.0	0.0	0.0	7.5	4.4	0.0	0.0	0.366	0.148
C ↔ G	5.1	0.0	0.0	0.0	9.8	0.0	3.2	0.0	13.5	0.0	0.0	0.0	0.0	0.0	0.0	6.7	0.0	0.66	0.004
C ↔ T	9.4	5.9	7.0	9.4	7.6	0.0	0.0	6.1	15.5	0.0	0.0	9.8	0.0	4.4	6.9	33.0	2.6	0.869	0.000
G ↔ T	3.7	0.0	0.0	5.0	4.7	0.0	0.0	0.0	24.2	0.0	0.0	0.0	0.0	0.0	0.0	10.1	4.3	0.696	0.002
A ↔ C + G ↔ T	0.0	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0	0.0	0.0	9.8	0.0	0.0	0.0	0.0	0.0	−0.168	0.520
A ↔ G + C ↔ T	3.1	3.0	0.0	0.0	2.8	0.0	0.0	0.0	0.0	16.9	0.0	9.0	9.4	0.0	1.9	6.7	0.0	0.212	0.414
A ↔ T + C ↔ G	0.0	0.0	4.9	4.3	9.8	0.0	2.1	0.0	17.9	0.0	0.0	0.0	0.0	2.4	2.5	0.0	0.0	0.646	0.005
A → C → G → A	6.2	0.0	0.0	0.0	14.3	0.0	1.9	6.7	0.0	0.0	0.0	11.3	0.0	0.0	1.5	8.4	0.0	0.045	0.863
A → C → T → A	8.2	2.8	0.0	0.0	7.7	0.0	1.9	0.0	0.0	0.0	0.0	0.0	0.0	4.9	1.7	9.3	1.1	0.117	0.654
A → G → C → A	10.7	3.0	0.0	0.0	6.5	0.0	0.0	0.0	17.4	46.9	0.0	0.0	9.8	2.1	6.8	7.8	0.0	0.32	0.210
A → G → T → A	11.1	0.0	2.2	2.2	3.9	0.0	0.0	0.0	19.3	0.0	0.0	0.0	0.0	5.3	2.5	14.3	0.0	0.825	0.000
A → T → C → A	5.0	0.0	5.6	0.0	7.5	0.0	7.3	0.0	33.8	0.0	5.1	10.1	0.7	5.9	5.7	21.1	0.0	0.67	0.003
A → T → G → A	0.0	2.9	0.0	3.9	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.4	8.2	0.0	−0.263	0.308
C → G → T → C	4.5	4.5	1.8	3.1	0.0	0.0	0.0	0.0	0.0	0.0	3.6	0.0	0.0	0.0	0.0	6.9	0.0	0.335	0.189
C → T → G → C	7.9	0.0	2.2	4.8	3.8	12.0	2.3	0.0	30.0	0.0	0.0	0.0	0.0	0.0	7.0	40.2	6.6	0.833	0.000
A → C → G → T → A	8.8	0.0	2.2	1.4	6.6	7.4	0.0	0.0	24.2	0.0	0.0	0.0	0.0	4.4	6.0	10.1	0.0	0.571	0.017
A → C → T → G → A	5.4	2.5	4.1	0.0	0.0	8.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.2	12.8	0.0	0.765	0.000
A → G → C → T → A	4.5	0.0	4.7	0.0	3.1	0.0	2.1	4.4	19.3	0.0	9.3	9.2	0.0	0.0	7.5	0.0	0.0	−0.077	0.768
A → G → T → C → A	9.4	0.0	0.0	0.0	3.2	0.0	0.0	5.0	0.0	0.0	0.0	0.0	0.0	5.1	3.4	6.7	2.5	0.155	0.551

Transformation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	r	P
A → T → C → G → A	2.9	0.0	0.0	3.6	7.8	0.0	0.0	1.2	0.0	0.0	0.0	10.4	8.8	1.7	0.0	0.0	0.0	−0.255	0.323
A → T → G → C → A	14.6	0.0	0.0	0.0	9.1	0.0	2.7	0.0	16.9	0.0	0.0	0.0	0.0	2.9	8.6	13.3	0.0	0.767	0.000
r	0.535	0.142	0.269	0.289	0.063	−0.017	0.21	−0.276	0.678	0.03	0.087	−0.156	0.393	0.294	0.586	0.752	0.522		
P	0.008	0.519	0.215	0.182	0.776	0.939	0.337	0.203	0.000	0.892	0.693	0.477	0.064	0.174	0.003	0.000	0.011		

Columns 1–17 areas in **Table 2**. r and P (last two columns and last two rows, respectively) indicate linear Pearson correlation coefficients between coverages across mitogenomic region/swinger transformations, comparing data in **Tables 2** and **3** by rows and columns, respectively.

Table 3. Percentage coverage of mitogenomic regions by swinger RNAs from analyses of complete human transcriptomic data [39].

Most correlations are positive along both genewise (columns) and transformation-wise (rows) analyses when comparing **Tables 2** and **3** (last rows and last columns in **Table 3**). Focusing on transformations and comparing coverages across genes for each transformation, correlations are positive between **Tables 2** and **3** for 19 among 23 transformations (P = 0.00065, one-tailed sign test) with 10 correlations statistically significant at P < 0.05. Analyses at the gene level across transformations detect 14 among 17 positive correlations (P = 0.003, one-tailed sign test), and six correlations at the gene level have P < 0.05 (one-tailed).

Across genes, at the transformation level, the strongest correlation was observed for transformation C ↔ T (**Figure 2**) with Pearson r = 0.869 and one-tailed P = 0.0000029. Across transformations, at the gene level, the strongest correlation was observed for the gene ND6 with Pearson r = 0.752 and one-tailed P = 0.000017 with highest coverage at C ↔ T transformation in both datasets (**Tables 2** and **3**). In order to test whether swinger coverage has more transformation than gene-specific effect, we calculated the combined P value using Fisher's method to combine P values, for the 23 swinger transformations and, separately, for the 17 mitogenome regions. The method sums -2xln(Pi) where i ranges from 1 to k (k = 23 for transformations and k = 17 for genome regions/genes). This yield combined P = 5.7×10^{-21} for transformations and combined P = 1.93×10^{-7} for genes. This indicates a 3× stronger effect of transformation
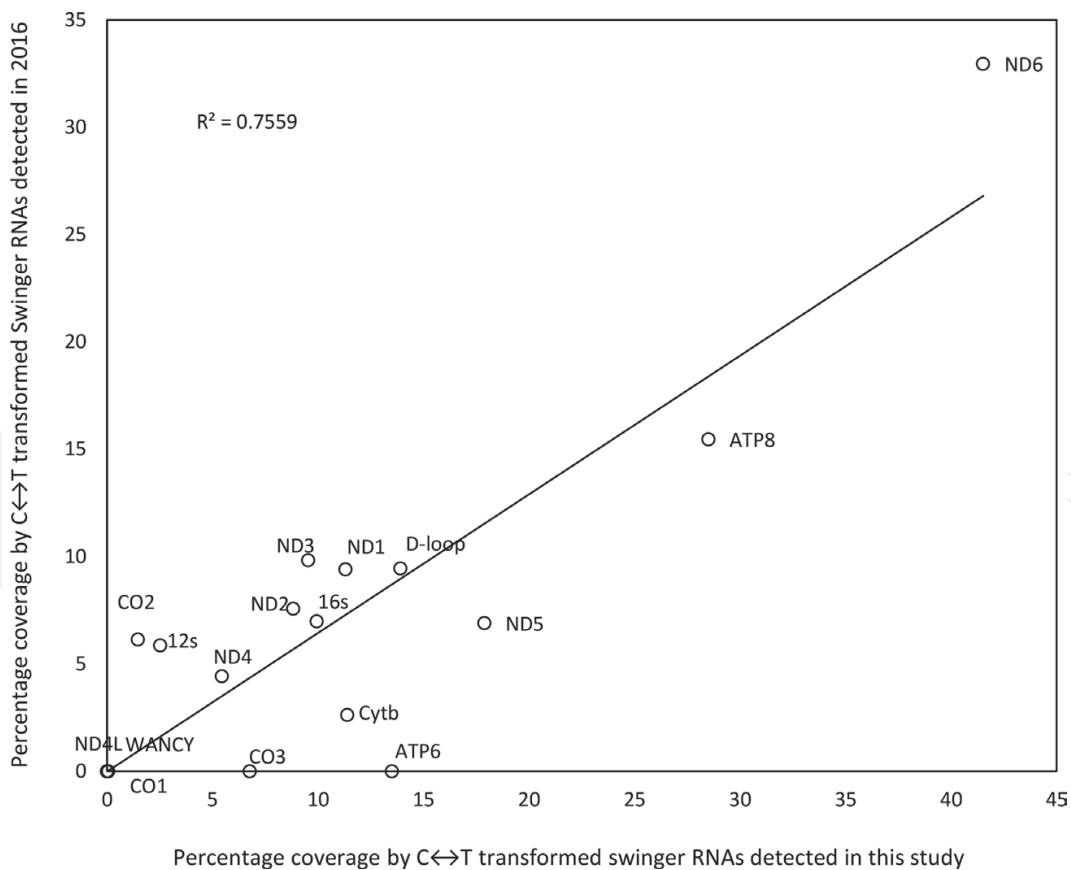
Figure 2. Percentage coverage of C ↔ T-transformed swinger RNAs across genes in this study as a function of their coverages in previously analysed data [39]. ND6 has the highest coverage among all transformations. Data from **Tables 2** and **3**.

across genes on coverage than a gene-specific effect across transformations. Hence, the most important unknown factor determines transformations. The genome region that is swinger-transcribed is important but secondary.

4. General conclusion

We find high reproducibility in swinger RNA coverage for the human mitogenome when comparing two independent transcriptomic datasets produced by Illumina sequencing. Positive correlations occur at each gene and transformation levels, reaffirming the reproducibility of the results, but are stronger at the transformation than gene level. The reproducibility of the swinger transcriptome in the giant virus *Mimivirus* and the ability to predict swinger RNA abundances from mathematical symmetry and error correcting principles [31, 50] together with present results from mitochondrial transcriptomes hint that swinger polymerizations are a universal phenomenon.

Acknowledgements

Author details

Ganesh Warthi[1]* and Hervé Seligmann[1,2,3]

*Address all correspondence to: g.warthi6791@gmail.com

1 Aix Marseille Univ, IRD, AP-HM, VITROME, IHU-Méditerranée Infection, Marseille, France

2 Aix-Marseille Univ, IRD, AP-HM, MEPHI, IHU-Méditerranée Infection, Marseille, France

3 The National Natural History Collections, The Hebrew University of Jerusalem, Jerusalem, Israel

References

[1] Seligmann H, Raoult D. Stem-loop RNA hairpins in giant viruses: Invading rRNA-like repeats and a template free RNA. Frontiers in Microbiology. 2018;**9**:101. Virology

[2] Seligmann H. Two genetic codes, one genome: Frameshifted primate mitochondrial genes code for additional proteins in presence of antisense antitermination tRNAs. Bio Systems. 2011;**105**(3):271-285. DOI: 10.1016/j.biosystems.2011.05.010

[3] Faure E, Delaye L, Tribolo S, Levasseur A, Seligmann H, Barthélémy RM. Probable presence of an ubiquitous cryptic mitochondrial gene on the antisense strand of the cytochrome oxidase I gene. Biology Direct. 2011;**6**:56. DOI: 10.1186/1745-6150-6-56

[4] Seligmann H. An overlapping genetic code for frame shifted overlapping genes in Drosophila mitochondria: Antisense antitermination tRNAs UAR insert serine. Journal of Theoretical Biology. 2012;**298**:51-76. DOI: 10.1016/j.jtbi.2011.12.026

[5] Seligmann H. Overlapping genetic codes for overlapping frameshifted genes in *Testudines*, and *Lepidochelys olivacea* as special case. Computational Biology and Chemistry. 2012;**41**:18-34. DOI: 10.1016/j.compbiolchem.2012.08.002

[6] Barthélémy RM, Seligmann H. Cryptic tRNAs in chaetognath mitochondrial genomes. Computational Biology and Chemistry. 2016;**62**:119-132. DOI: 10.1016/j.compbiolchem. 2016.04.007

[7] Seligmann H. Putative protein-encoding genes within mitochondrial rDNA and the D-Loop region. In: Lin Z, Liu W, editors. Ribosomes: Molecular Structure, Role in Biological Functions and Implications for Genetic Diseases. Nova Publishers; 2013. pp. 67-86. chapter 4

[8] Seligmann H. Phylogeny of genetic codes and punctuation codes within genetic codes. Bio Systems. 2015;**129**:36-43. DOI: 10.1016/j.biosystems.2015.01.003

[9] Seligmann H. Avoidance of antisense, antiterminator tRNA anticodons in vertebrate mitochondria. Bio Systems. 2010;**101**(1):42-50. DOI: 10.1016/j.biosystems.2010.04.004

[10] Seligmann H. Undetected antisense tRNAs in mitochondrial genomes? Biology Direct. 2010;**5**:39 DOI: 10.1186/1745-6150-5-39

[11] Seligmann H. Pathogenic mutations in antisense mitochondrial tRNAs. Journal of Theoretical Biology. 2011;**269**(1):287-296. DOI: 10.1016/j.jtbi.2010.11.007

[12] Seligmann H. Putative mitochondrial polypeptides coded by expanded quadruplet codons, decoded by antisense tRNAs with unusual anticodons. Bio Systems. 2012;**110**(2):84-106. DOI: 10.1016/j.biosystems.2012.09.002

[13] Seligmann H, Labra A. Tetracoding increases with body temperature in *Lepidosauria*. Bio Systems. 2013;**114**(3):155-163. DOI: 10.1016/j.biosystems.2013.09.002

[14] Seligmann H. Putative anticodons in mitochondrial tRNA sidearm loops: Pocketknife tRNAs? Journal of Theoretical Biology. 2014;**340**:155-163. DOI: 10.1016/j.jtbi.2013.08.030

[15] Yang W, J min W, ding BA, Ou-yang CY, Shen HH, Chirn WG, et al. Possible formation of mitochondrial-RNA containing chimeric or trimeric RNA implies a post-transcriptional and post-splicing mechanism for RNA fusion. PLoS One. 2013;**8**(10):e77016. DOI: 10.1371/journal.pone.0077016

[16] Li M, Wang IX, Li Y, et al. Widespread RNA and DNA sequence differences in the human transcriptome. Science. 2011;**333**(6038):53-58. DOI: 10.1126/science.1207018

[17] Bar-Yaacov D, Avital G, Levin L Richards L, Hachen B, Rebolledo J, Nekrutenk A, et al. RNA-DNA differences in human mitochondria restore ancestral form of 16S ribosomal RNA. Genome Research. 2013;**23**(11):1789-1796. DOI: 10.1101/gr.161265.113

[18] Hodgkinson A, Idaghdour Y, Gbeha JC, Grenier E, Hip-Ki V, Bruat JP, et al. High-resolution genomic analysis of human mitochondrial RNA sequence variation. Science. 2014;**344**(6182):413-415. DOI: 10.1126/science.1251110

[19] Chen C, Bundschuh R. Systematic investigation of insertional and deletional RNA-DNA differences in the human transcriptome. BMC Genomics. 2012;**13**(1):616. DOI: 10.1186/1471-2164-13-616

[20] Moreira S, Valach M, Aoulad-Aissa M, Otto C, Burger G. Novel modes of RNA editing in mitochondria. Nucleic Acids Research. 2016;**44**(10):4907-4919. DOI: 10.1093/nar/gkw188

[21] Wang IX, Core LJ, Kwak H, Brady L, Bruzel A, McDaniel L, et al. RNA-DNA differences are generated in human cells within seconds after RNA exits polymerase II. Cell Reports. 2014;**6**(5):906-915. DOI: 10.1016/j.celrep.2014.01.037

[22] Seligmann H. Codon expansion and systematic transcriptional deletions produce tetra-, pentacoded mitochondrial peptides. Journal of Theoretical Biology. 2015;**387**:154-165. DOI: 10.1016/j.jtbi.2015.09.030

[23] El Houmami N, Seligmann H. Evolution of nucleotide punctuation marks: From structural to linear signals. Frontiers in Genetics. 2017;**8**:36. DOI: 10.3389/fgene.2017.00036

[24] Seligmann H. Systematically frameshifting by deletion of every 4th or 4th and 5th nucleotides during mitochondrial transcription: RNA self-hybridization regulates delRNA expression. Bio Systems. 2016;**142-143**:43-51. DOI: 10.1016/j.biosystems.2016.03.009

[25] Seligmann H. Replicational mutation gradients, dipole moments, nearest neighbor effects and DNA polymerase gamma fidelity in human mitochondrial genomes. In: Stuart D, editor. The Mechanisms of DNA Replication. Rijeka, Croatia: InTech; 2013a. pp. 257-286. chapter 10

[26] Seligmann H. Polymerization of non-complementary RNA: Systematic symmetric nucleotide exchanges mainly involving uracil produce mitochondrial RNA transcripts coding for cryptic overlapping genes. Biosystems. 2013b;**111**:156-174

[27] Seligmann H. Systematic asymmetric nucleotide exchanges produce human mitochondrial RNAs cryptically encoding for overlapping protein coding genes. Journal of Theoretical Biology. 2013c;**324**:1-20

[28] Seligmann H. Coding constraints modulate chemically spontaneous mutational replication gradients in mitochondrial genomes. Current Genomics. 2012;**13**:37-54

[29] Almeida JS, Vinga S. Biological sequences as pictures – A generic two dimensional solution for iterated maps. BMC Bioinformatics. 2009;**10**:100. DOI: 10.1186/1471-2105-10-100

[30] Fimmel E, Giannerini S, Gonzalez DL, Strüngmann L. Circular codes, symmetries and transformations. Journal of Mathematical Biology. 2015;**70**(7):1623-1644. DOI: 10.1007/s00285-014-0806-7

[31] Michel CJ, Seligmann H. Bijective transformation circular codes and nucleotide exchanging RNA transcription. Bio Systems. 2014;**118**(1):39-50. DOI: 10.1016/j.biosystems. 2014.02.002

[32] Samuels DC. Mitochondrial DNA repeats constrain the life span of mammals. Trends in Genetics. 2004;**20**(5):226-229. DOI: 10.1016/j.tig.2004.03.003

[33] Samuels DC. Life span is related to the free energy of mitochondrial DNA. Mechanisms of Ageing and Development. 2005;**126**(10):1123-1129. DOI: 10.1016/j.mad.2005.05.003

[34] Samuels DC, Schon EA, Chinnery PF. Two direct repeats cause most human mtDNA deletions. Trends in Genetics. 2004;**20**(9):393-398. DOI: 10.1016/j.tig.2004.07.003

[35] Khaidakov M, Siegel ER, Shmookler Reis RJ. Direct repeats in mitochondrial DNA and mammalian lifespan. Mechanisms of Ageing and Development. 2006;**127**(10):808-812. DOI: 10.1016/j.mad.2006.07.008

[36] Seligmann H. Mitochondrial swinger replication: DNA replication systematically exchanging nucleotides and short 16S ribosomal DNA swinger inserts. Bio Systems. 2014;**125**:22-31. DOI: 10.1016/j.biosystems.2014.09.012

[37] Seligmann H. Species radiation by DNA replication that systematically exchanges nucleotides? Journal of Theoretical Biology. 2014;**363**:216-222. DOI: 10.1016/j.jtbi.2014.08.036

[38] Seligmann H. Sharp switches between regular and swinger mitochondrial replication: 16S rDNA systematically exchanging nucleotides A ↔ T+C ↔ G in the mitogenome of Kamimuria wangi. Mitochondrial DNA. 2016;**27**(4):2440-2446. DOI: 10.3109/19401736.2015.1033691

[39] Seligmann H. Translation of mitochondrial swinger RNAs according to tri-, tetra- and pentacodons. Bio Systems. 2016;**140**:38-48. DOI: 10.1016/j.biosystems.2015.11.009

[40] Seligmann H. Unbiased mitoproteome analyses confirm non-canonical RNA, expanded codon translations. Computational and Structural Biotechnology Journal. 2016;**14**:391-403. DOI: 10.1016/j.csbj.2016.09.004

[41] Seligmann H. Natural chymotrypsin-like-cleaved human mitochondrial peptides confirm tetra-, pentacodon, non-canonical RNA translations. Bio Systems. 2016;**147**:78-93. DOI: 10.1016/j.biosystems.2016.07.010

[42] Seligmann H. Natural mitochondrial proteolysis confirms transcription systematically exchanging/deleting nucleotides, peptides coded by expanded codons. Journal of Theoretical Biology. 2017;**414**:76-90. DOI: 10.1016/j.jtbi.2016.11.021

[43] Seligmann H. Swinger RNAs with sharp switches between regular transcription and transcription systematically exchanging ribonucleotides: Case studies. Bio Systems. 2015;**135**:1-8. DOI: 10.1016/j.biosystems.2015.07.003

[44] Seligmann H. Chimeric mitochondrial peptides from contiguous regular and swinger RNA. Computational and Structural Biotechnology Journal. 2016;**14**:283-297. DOI: 10.1016/j.csbj.2016.06.005

[45] Seligmann H. Swinger RNA self-hybridization and mitochondrial non-canonical swinger transcription, transcription systematically exchanging nucleotides. Journal of Theoretical Biology. 2016;**399**:84-91. DOI: 10.1016/j.jtbi.2016.04.007

[46] Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria. Nature. 1981;**290**(5806):470-474. DOI: 10.1038/290470a0

[47] Garzon R, Volinia S, Papaioannou D, et al. Expression and prognostic impact of lncRNAs in acute myeloid leukemia. Proceedings of the National Academy of Sciences of the United States of America. 2014;**111**(52):18679-18684. DOI: 10.1073/pnas.1422050112

[48] Mercer TR, Neph S, Dinger ME, et al. The human mitochondrial transcriptome. Cell. 2011;**146**(4):645-658. DOI: 10.1016/j.cell.2011.06.051

[49] Altschul SF, Madden SF, Schaeffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Research. 1991;**25**:3389-3402

[50] Seligmann H. Bijective codon transformations show genetic code symmetries centered on cytosine's coding properties. Theory in Biosciences. 2017:1-15