Received: 13 January 2023 | Revised: 7 May 2023 | Accepted: 11 May 2023

DOI: 10.1002/arch.22026

RESEARCH ARTICLE

Archives of INSECT BIOCHEMISTRY AND PHYSIOLOGY WILEY

Differential sense and antisense expression profiles of *Syrista parreyssi* (Hymenoptera: Cephidae) mitochondrial transcripts

Habeş Bilal Aydemir[1] | **Merve Nur Aydemir**[1] | **Ertan Mahir Korkmaz**[2]

[1]Department of Molecular Biology and Genetics, Faculty of Science and Letters, Tokat Gaziosmanpaşa University, Tokat, Turkey

[2]Department of Molecular Biology and Genetics, Faculty of Science, Sivas Cumhuriyet University, Sivas, Turkey

Correspondence
Habeş Bilal Aydemir, Department of Molecular Biology and Genetics, Faculty of Science and Letters, Tokat Gaziosmanpaşa University, Tokat, Turkey.
Email: bilal.aydemir@gop.edu.tr and hbilalaydemir@hotmail.com

Abstract

The transcription of the mitogenome shows a unique pattern that is both similar to and different from the nuclear and bacterial patterns. Mitochondrial transcription generates five polycistronic units from three promoters in *Drosophila melanogaster*, and different expression levels of genes were observed in both different and, interestingly, the same polycistronic units in *D. melanogaster*. This study was conducted to test this phenomenon in the mitogenome of *Syrista parreyssi* (Hymenoptera: Cephidae). RNA isolation and DNase digestion were performed using only one whole individual, and real-time polymerase chain reaction analyses were performed with complementary DNAs of 11 gene regions using gene-specific primers. It was found that the expression level of each gene exhibited differences from each other, and some genes (e.g., *cox* genes, and rrnS) were interestingly expressed at significant levels in the corresponding antisense chain. Additionally, the mitogenome of *S. parreyssi* was found to have the capacity to encode 169 additional peptides from 13 known protein-coding genes, most of which were encoded in antisense transcript units. One of the unique findings was a potential open reading

frame sequence that was potentially encoded in the antisense rrnL gene and included a conserved *cox3* domain.

KEYWORDS
antisense transcript, differential gene expression, Hymenoptera, lncRNAs, mitochondrial gene expression, *Syrista parreyssi*

Research Highlights

- Mitochondrial expression profiles of *Syrista parreyssi* are different between genes and gene clusters which composed related protein complexes.
- Mitochondrial antisense transcripts seem to mediate to produce regulatory or translated RNA molecules.
- The mitogenome of *S. parreyssi* has the capacity of encoding 169 additional peptides from 13 known protein-coding genes, most of which encoded in antisense transcript units.

1 | INTRODUCTION

The mitochondrial genome (mitogenome) is mostly highly conserved in animals. This genome has been derived from the alpha-proteobacteria symbiont, the source of mitochondria in all eukaryotes, and it exhibits bacterial properties. Symbiont ancestor of mitochondria has transferred too much genes to the nucleus and/or there has been too much gene loss from the endosymbiont genome during the evolutionary process. Therefore, the mitogenome is highly differentiated from its bacterial origin (Gray, 1989). Bilateral animal mitogenomes usually consist of 37 "known" genes: 13 protein-coding genes (PCGs) that are parts of the electron transport chain, two ribosomal RNA genes (rRNAs) and 22 transfer RNA genes (tRNAs) which are collectively responsible for the transcription of the PCGs (Osigus et al., 2013). In addition to the relatively conserved gene content, the mitogenome also includes one or more noncoding regions. These regions are primarily functional the binding sites for proteins involved in genomic replication or transcription (Aydemir et al., 2023). Despite the limited content of gene (mRNA, tRNA, and rRNA genes) and regulatory sequences (intergenic and noncoding regions) of the mitogenome and its economization pattern, noncoding sequences are transcribed in mitochondrial transcription.

In most animal species, the mitochondrial DNA (mtDNA) is transcribed from both chains: if a chain more often encodes genes, it is called a "heavy chain," while if it encodes less, it is then called "light chain." The mitogenome is also transcribed using a specialized system. In mammals, there are two alternative promoters (HSP1 and HSP2) in the heavy chain, and one promoter (LSP) in the D-loop of the light chain (Asin-Cayuela & Gustafsson, 2007; Scarpulla, 2008). In *Drosophila*, RNA synthesis begins from the second promoter of the heavy chain and third promoter of the two light chain (Stewart & Beckenbach, 2009). The respective gene blocks located in the heavy and light chains are mostly paired (Berthier et al., 1986; Torres et al., 2009). The production of the major transcripts of the initial polycistronic units causes the formation of nontranslated transcripts to a large extent. Based on detailed DNA sequencing studies (Brzezniak et al., 2011; Lopez Sanchez et al., 2011), excessive mRNAs are largely controlled by post-transcriptional processes (Mercer et al., 2011; Torres et al., 2009) and arises as products of nonstandard processes. This approach is particularly linked to the secondary structure of tRNAs because the

regions handled by nonstandard processes are associated with RNA secondary structure clusters in mRNAs (Mercer et al., 2011).

"In addition to the known genes, various types of coding and noncoding RNAs are also commonly found in mitochondria (Lung et al., 2006). Two types of short RNAs are generated from the regions downstream of the 5′−3′ processing sites of tRNA genes (Mercer et al., 2011). A small control region called 7S RNA, which is a polyadenylated RNA with an unknown function, is transcribed from the initial L-chain transcript. Three large noncoding RNAs (lncRNA) have been found in mouse and human mitochondria, and they are known to play a role in regulating mitochondrial gene expression (Mercer et al., 2011; Rackham et al., 2011). In addition to regulated antisense transcripts, it has been reported that normal human cells can be transformed into tumor cells through the reverse complementarity of the rrnL-binding region (Burzio et al., 2009). Interestingly, mitochondria also encode additional peptides, called mitochondria-derived peptides (MDPs), which are not involved in oxidative phosphorylation (OXPHOS) (Lee et al., 2013, 2016). A short open reading frame (ORF) sequence containing 24 amino acids, called Humanin, which is thought to exhibit neuroprotective properties, has been detected in the rrnL gene, while MOTS-c is encoded in the rrnS gene. However, the expression and function of these peptides are still unknown (Alexeyev, 2017; Clayton et al., 2000).

Although the gene sequences and transcription patterns have been rearranged in most insect species, they generally maintain a circular arrangement, as shown in Figure 1 (Cameron, 2014). Therefore, it can be concluded that mammalian and invertebrate mitogenomes differ significantly in their transcription patterns. Although mitochondrial transcription and post-transcriptional RNA maturation have been extensively studied in *Drosophila* and sea urchins, further research is necessary to understand the variation in transcription profiles among different invertebrates (for more detailed information, see Cantatore et al., 1990; Polosa et al., 2007; Torres et al., 2009)."

Excessive differences in both functional and structural levels between expected and observed mitochondrial mature transcripts may result from variations in the stability of transcriptional and post-transcriptional processes. According to the tRNA punctuation model, the transcript maturation process may also influence the steady-state levels of mRNAs. As a result, gene expression levels may vary in the same mitochondrial complex (Torres et al., 2009).

This study employed real-time polymerase chain reaction (PCR) with gene-specific primers (GSPs) to quantitatively analyze complementary DNA (cDNA) samples from isolated total RNAs of *Syrista parreyssi* (Hymenoptera: Symphyta). The aim was to examine the gene regions expressed in the light and heavy chain of the mtDNA of this insect species and analyze the polycistronic transcript units. The expression levels of the gene regions were interpreted using the expression of the *gapdh* gene as the reporter gene. In addition, potential ORFs that could be encoded in the mitogenome were determined through bioinformatics analysis with ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/).

2 | MATERIALS AND METHODS

2.1 | Sample collection

Adult individuals of *S. parreyssi* were collected from Sivas (N: 3872090 E: 3727634) on May 26, 2015 and put directly into the RNAlater solution. They were stored at −20°C in this solution before use.

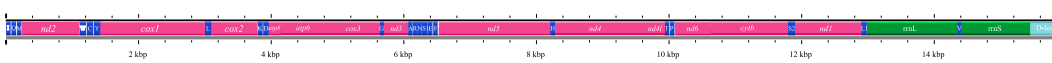

FIGURE 1 A representation of the insect mitogenomes with minor changes (genes encoded in minority chains are underlined, modified from Cameron 2014).

2.2 | RNA isolation

RNA isolation was carried out using the Qiagen RNeasy Mini Kit following the manufacturer's recommended protocol. First, the washed and dried tissue was treated with lysis buffer containing 2-mercaptoethanol. The sample was homogenized using a rotor-stator for 2 min at maximum speed, and then centrifuged at 2600*g* for 5 min at room temperature. The resulting supernatant was transferred to a clean RNase-free 1.5 mL tube and mixed with an equal volume of 70% ethanol, and the mixture was transferred to spin columns and centrifuged. Subsequently, the column was washed with 700 µL washing buffer and centrifuged at 12,000*g* for 15 s at room temperature. The column was further washed with absolute ethanol after DNase treatment. The RNA-attached column membrane was centrifuged at 12,000*g* for 1 min at room temperature, and the dry RNA was finally dissolved in 100 µL RNase-free water.

2.3 | First-strand cDNA synthesis

1 µL of 2 pmol GSP, designed for 11 different gene regions (Supporting Information: Table S1), was combined with 1 µL of 1 ng to 5 µg total RNA and 1 µL of 10 mM dNTP mix solution. The final volume of the mixture was adjusted to 12 µL. The mixture was then incubated at 65°C for 5 min, and 4 µL of 5 × first-strand buffer and 2 µL of 0.1 M DTT were added on ice. After incubation for 2 min at 42°C, 1 µL (200 units) of Superscript II Reverse Transcriptase was added to a final volume of 20 µL. The mixture was incubated at 42°C for 50 min and then inactivated at 70°C for 15 min.

2.4 | Real-Time PCR

Real-time PCR was conducted using a reaction mixture containing 5 µL of 2 × Supermix (Sso Advanced Universal SYBR Green Supermix), 300 nM of each GSP, and 1 µg of cDNA in a total volume of 10 µL. The PCR thermal cycle involved a denaturation step for cDNA at 98°C for 30 s, followed by an amplification denaturation at 98°C for 15 s and annealing at 50°C for 30 s for 40 cycles. Each sample was analyzed in triplicate for both sense and antisense transcripts, and the Biorad CFX Connect Real-Time PCR detection system was used for amplification, detection, and data analysis. The quantification cycle (Cq) values were interpreted as follows: a Cq value between 15 and 25 indicated a successful reaction and a high expression level of the target nucleotide sequence, while a Cq value between 25 and 35 indicated intermediate expression levels and a successful reaction. A Cq value below 15 or above 45 indicated an unsuccessful reaction, which could be due to environmental contamination or the presence of only a small number of target nucleotide sequences (Schmittgen & Livak, 2008).

2.5 | Detection of putative mitochondrial ORFs

To make sure that expressed antisense transcripts are translated to any protein or they are ncRNAs, we analysed three datasets of *S. pareyyssi* (NCBI accession number: KX907847) by NCBI ORF Finder. The datasets contain (i) total mitogenome, (ii) monocistronic PCGs and rRNA genes, (iii) polycistronic transcript units. The novel ncRNA candidates were screened by NCBI ORF Finder to detect possible ORFs using invertebrate mitochondrial code and "ATG" and/or alternative initiation codons priors. The ORFs less than 75 nt in length were ignored and all degenerate nucleotides (Y, W, R, K, D, M) were converted to N base. Then, the ORFs were searched for homologs using blastp against the nonredundant protein sequence database (nr) at NCBI.

Five possible transcription units are recognized to identify the correct transcriptional restriction for ORF Finder analysis. The transcription unit 1 comprises *nd2-cox1-cox2-atp6-atp8-cox3* and *nd3* genes, and is located between the bases of 1–6442. The transcription unit 2 contains *nd5*, *nd4* and *nd4L* genes, and is located between the bases of 6443–9929, while transcription unit 3 comprising *nd6* and *cytB* genes is located between bases 9930–11,904. The transcription unit 4 comprising *nd1-16S rRNA* and *12S rRNA* genes is located between 11,959 and 15,446 bases; and finally, transcription unit 5 which contains only rRNA genes is found between bases 12,916–15,446.

3 | RESULTS

3.1 | The results of real-time PCR

Real-time PCR experiments were conducted to detect sense and antisense transcripts of several mitochondrial genes, including rrnS, *cox1*, *cox2*, *cox3*, and *nd5*. Results showed that only sense transcripts were significantly detected for rrnL, *cytB*, *nd1*, *nd3*, and bicistronic units *nd4/nd4L* and *atp6/atp8*, while the levels of antisense transcripts for these genes were low or undetectable (Cq > 35) (Supporting Information: Table S2). This could be due to the lack of antisense transcripts or technical issues with the real-time PCR reactions. Real-time PCR reactions were not performed for *nd2* and *nd6* genes due to the partial sequence information in the NCBI database (KX907847) and the unavailability of primers for these regions (Supporting Information: Table S1).

The average Cq value for the *gapdh* reporter gene was 19.11. Cq values varied across mitochondrial transcripts, both within and between sense and antisense transcripts (Figure 2). The lowest Cq value was observed for the sense transcript of rrnS gene (Cq = 14.11), while the highest was observed for the antisense transcript of *nd5* gene (Cq = 30.33). The Cq values of sense transcripts ranged from 21.21 (rrnL gene) to 29.46 (*nd5* gene). The Cq values of sense transcripts of cox genes were mostly similar: 23.46 for *cox1*, 24.75 for *cox2* and 24.20 for *cox3*. For the sense transcripts of *nd* genes, except for *nd5* gene, the Cq values ranged from 25.61 (*nd4/nd4L* gene) to 28.71 (*nd3* gene). The Cq values of the sense transcripts of *cytB* and *atp6/atp8* were 27.71 and 29.13 on average, respectively. The Cq value of the antisense transcript of rrnS gene was higher than that of the sense transcript (23.46). For the antisense transcripts of *cox* genes, the mean Cq values were 26.60, 19.73, and 27.48 for *cox1*, *cox2*, and *cox3* gene, respectively.

In terms of sense expressions of the mitochondrial genes in *S. parreyssi*, the most highly expressed gene was rrnS, while the lowest expression profile was observed for *nd5* gene. The order of gene expression for the mitochondrial genes was rrnS > rrnL > *cox1* > *cox3* > *cox2* > *nd4/nd4L* > *cytB* > *nd1* > *nd3* > *atp6/atp8* > *nd5*. In

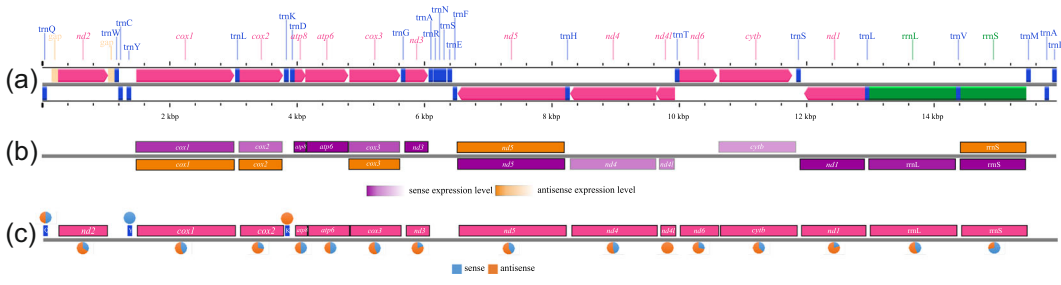

FIGURE 2 (a) The mitogenome scheme of *Syrista parreyssi*. (b) mRNA and rRNA expression profiles of the mitochondrial genes and sense and antisense expression profiles of monocistronic transcription units of *S. parreyssi*. (c) Sense and antisense ORFs were predicted in ORF Finder in genes of *S. parreyssi*. mRNA, messenger RNA; rRNA, ribosomal RNA.

contrast, for antisense expressions of the mitochondrial genes, the order was *cox2* > rrnS > *cox1* > *cox3* > *nd5*. Interestingly, the antisense expression level of *cox2* gene (19.73) was higher than the sense expression level (24.75).

3.2 | The results of ORF Finder

As a result of the ORF Finder analyses, a total of 182 putative ORFs longer than 75 nucleotides were identified in the mitogenome, of which 13 were functional PCGs encoded by the mitochondrial genome (Supporting Information: Table S3). ORF 120 was the largest putative protein sequence, measuring 173 amino acids in length, while ORF 10, 36, 56, 59, 85, 125, and 142 were the smallest, measuring only 24 amino acids in length. Of these putative ORFs, 18 were located in the *nd5* gene, and 17 were potentially encoded within the *cox1* and *cytB* genes. Additionally, the rrnL gene was capable of coding 16, while *cox3* and *nd4* genes were capable of coding 15 possible ORFs. Eight putative ORFs were identified in *nd6* and *cox2* genes, seven in the *atp6*, *nd2*, and rrnS genes, six in the *nd1* and *nd3* genes, three in *atp8* and *nd4L* genes, two in trnQ and one in trnK and trnY genes. Among these ORFs, 12 were found to contain intergenic regions (IG) with a part of gene sequences, while 6 of them contained consecutive gene regions without IG (Supporting Information: Table S3). Notably, 78 ORFs seemed to have sense expression potential, depending on the orientation of the gene from which it is encoded, while 100 of them were antisense. However, ORF107, ORF80, ORF29, and ORF 152 sequences located in the chain-replacement region could not be evaluated for the presence of sense or antisense expression (Supporting Information: Table S3). Except for the PCG ORFs, most of the ORFs were potentially coded from the antisense chain of related gene sequences (Figure 2).

In blastp results, the amino acid sequences of 29 potential ORFs, except the formal 13 PCGs, were aligned with any sequences. Fourteen of which were aligned with the alternative reading frame ORFs of the gene where they are located, and as expected, these ORFs did not have significant similarities with the related gene of *S. pareyyssi*. Thirteen of the 29 ORFs were aligned with "hypothetical" or "uncharacterized" proteins of organisms with high diversities from bacteria to insects (Table 1).

One of the unique findings was the observation of ORF27 in the antisense chain of rrnL. ORF27 was aligned with a domain of cox3 that consisted of the conserved "MSNHFSRNHHGFGFASLXHXHFVDIVXLFL" amino acid sequence, measuring 30 aa in length. This residue or "spare" fragment exhibited 86% identity to 93% of the mitochondrial *cox3* gene of Janus megamaculatus (Hymenoptera: Cephidae), which is closely related to *S. parreyssi*. ORF27 also contained a portion of rrnL domain 5, which is a highly conserved and peptidyl transferase region of rRNA (Mears et al., 2006), and domain 6 in the large mature major gene chain. Another remarkable finding was obtained from ORF95, which potentially encoded the sense chain of rrnL. The amino acid sequence of ORF95 was aligned with uncharacterized AAEL017413-PA protein of *Aedes aegypti* (Diptera), which was the most highly expressed gene in larval and adult mitochondria.

4 | DISCUSSION

To test the hypothesis that mitochondria transcribe a full chain despite their tendency towards economization, we quantitatively confirmed the transcription of antisense and noncoding RNAs in *S. parreyssi* using Real-time PCR analysis. In addition to the 13 formally recognized protein coding genes, we discovered that the mitochondrial genome can encode 169 more peptides by performing ORF Finder analyses. Recently, novel protein coding genes, named MDPs, have been identified in other species (Lee et al., 2013; Lee et al., 2015; Lee et al., 2016), providing further support for this finding. Alternatively, these potential sequences could act as a template for lncRNAs.

The highest level of sense expression was observed in the rRNA genes (Figure 2), which may be due to the constant need for high levels of expression since these genes are involved in ribosome structure

TABLE 1 The blastp results of potential ORFs obtained from ORF Finder.

ORF	Start	Stop	aa lenght	Location	Strand	Sense or antisense	Blastp search
30	257	1030	258	nd2	Heavy	Sense	ARO34957.1 NADH dehydrogenase subunit 2 [Syrista parreyssii]
62	1449	3011	520	cox1	Heavy	Sense	ARO34958.1 cytochrome c oxidase subunit I [Syrista parreyssii]
63	3096	3776	226	cox2	Heavy	Sense	ARO34959.1 cytochrome c oxidase subunit II [Syrista parreyssii]
5	3961	4134	57	atp8	Heavy	Sense	ARO34960.1 ATP synthase F0 subunit 8 [Syrista parreyssii]
65	4113	4802	229	atp6	Heavy	Sense	ARO34961.1 ATP synthase F0 subunit 6 [Syrista parreyssii]
40	4799	5617	272	cox3	Heavy	Sense	ARO34962.1 cytochrome c oxidase subunit III [Syrista parreyssii]
41	5696	6052	118	nd3	Heavy	Sense	ARO34963.1 NADH dehydrogenase subunit 3 [Syrista parreyssii]
164	8212	6518	564	nd5	Light	Sense	ARO34964.1 NADH dehydrogenase subunit 5 [Syrista parreyssii]
108	9639	8284	451	nd4	Light	Sense	ARO34965.1 NADH dehydrogenase subunit 4 [Syrista parreyssii]
133	9935	9633	100	nd4L	Light	Sense	ARO34966.1 NADH dehydrogenase subunit 4L [Syrista parreyssii]
49	10,001	10,591	196	nd6	Heavy	Sense	ARO34967.1 NADH dehydrogenase subunit 6 [Syrista parreyssii]
50	10,595	11,770	391	cytB	Heavy	Sense	ARO34968.1 cytochrome b [Syrista parreyssii]
96	12,969	11,959	336	nd1	Light	Sense	ARO34969.1 NADH dehydrogenase subunit 1 [Syrista parreyssii]
2	2116	2208	30	cox1	Heavy	Sense	AAO50289.1 cytochrome oxidase subunit I [Walterianella biarcuata]
33	2279	2434	51	cox1	Heavy	Sense	AEI26129.1 cytochrome c oxidase subunit I [Samia ricini]
77	7593	7709	38	nd5	Light	Antisense	AGA54148.1 NADH dehydrogenase subunit 5 [Aleurodicus dispersus]
45	8939	9139	66	nd4	Light	Antisense	AGA54149.1 NADH dehydrogenase subunit 4 [Aleurodicus dispersus]
23	11,317	11,490	57	cytB	Heavy	Sense	AJE26419.1 cytochrome b [Cosmoscarta bispecularis]
21	10,969	11,145	58	cytB	Heavy	Sense	ALO76560.1 cytochrome b [Urodontus sp. URO01]
32	2018	2260	80	cox1	Heavy	Sense	BBI28832.1 cytochrome c oxidase subunit 1, partial [Glossina morsitans morsitans]
103	10,791	10,513	92	nd6-IG-cytB	Heavy	Antisense	CAC80825.1 cytochrome-b protein [Sabanejewia bulgarica]

(Continues)

TABLE 1 (Continued)

ORF	Start	Stop	aa lenght	Location	Strand	Sense or antisense	Blastp search
102	10,938	10,795	47	cytB	Heavy	Antisense	CAC80825.1 cytochrome-b protein [Sabanejewia bulgarica]
34	2483	2704	73	cox1	Heavy	Sense	CAD44502.1 cytochrome oxidase I [Chaoborus flavicans]
35	2708	2923	71	cox1	Heavy	Sense	CAD44502.1 cytochrome oxidase I [Chaoborus flavicans]
95	13,593	13,483	36	rrnL	Light	Sense	EJY57844.1 AAEL017413-PA [*Aedes aegypti*]
171	4798	4490	102	atp6	Heavy	Antisense	KAF2897351.1 hypothetical protein ILUMI_08823 [Ignelater luminosus]
76	7353	7550	65	nd5	Light	Antisense	KOX68003.1 hypothetical protein WN51_07944 [Melipona quadrifasciata]
175	3763	3347	138	cox2	Heavy	Antisense	KRZ64503.1 hypothetical protein T10_1603 [Trichinella papuae]
101	11,091	10,942	49	cytB	Heavy	Antisense	MCC8467664.1 hypothetical protein [Rickettsia endosymbiont of Eriopis connexa]
51	12,167	12,511	114	nd1	light	antisense	QED22670.1 NADH dehydrogenase subunit 1 [Amynthas sp. HU201607-04]
182	1627	1409	72	cox1	Heavy	Antisense	SBT57303.1 hypothetical protein POVWA2_078610 [Plasmodium ovale wallikeri]
75	7176	7349	57	nd5	light	antisense	SGA03430.1 Uncharacterized protein [Chlamydia abortus]
126	13,673	13,578	31	rrnL	Light	Sense	SGA22881.1 Uncharacterized protein [Chlamydia abortus]
180	2155	1853	100	cox1	Heavy	Antisense	SHE23799.1 cytochrome oxidase subunit 1, partial [Paraphaenops breuilianus espanoli]
181	1822	1628	64	cox1	Heavy	Antisense	SSC84599.1 cytochrome oxidase subunit 1 [Folsomia candida]
27	13,600	13,692	30	rrnL	Light	Antisense	UGN61594.1 cytochrome c oxidase subunit III [Janus megamaculatus]
173	4384	4100	94	atp6-atp8	Heavy	Antisense	WP_129335196.1 hypothetical protein [Enterobacter cloacae complex sp. 2DZ2F20B]
113	5187	4942	81	cox3	Heavy	Antisense	WP_146043407.1 hypothetical protein [Vibrio vulnificus]
177	3004	2795	69	cox1	Heavy	Antisense	WP_219622595.1 hypothetical protein [Enterococcus faecium]
179	2578	2378	66	cox1	Heavy	Antisense	WP_228494194.1 hypothetical protein [Bacillus thuringiensis]
100	11,436	11,167	89	cytB	Heavy	Antisense	WP_257974106.1 hypothetical protein [Enterococcus faecium]
111	5592	5401	63	cox3	Heavy	Antisense	WP_261308635.1 hypothetical protein [Chlamydia abortus]

(Mears et al., 2006). Additionally, the fact that rRNA genes in mitogenomes encode MDPs may contribute to their high expression profiles (Lee et al., 2013; Lee et al., 2016). The Cq value of the most distant *nd5* gene region from the *gapdh* reporter gene was measured to be 30, and this high expression level may be related to the fact that very few genes of the nd protein complex are encoded by the mitochondrial genome. Thus, the high expression of the nuclear genome-encoded congeners may compensate for the low expression level of the *nd* genes. This may be due to deviations in the *nd* gene products being relaxed by nuclear coevolution models (Castellana et al., 2011; Meiklejohn et al., 2007).

The antisense transcription of the *cox1* gene was remarkable (Figure 2), and it has been discovered that 3′ UTR sequences are present in the antisense transcripts in humans (Temperley et al., 2010). UTR sequences have important functions in post-transcriptional gene expression and in the termination of translation (Lucy et al., 2013), which suggests that antisense conjugates of genes involved in the control of gene expression can be utilized. Another possible explanation for the expression of the antisense partner could be that certain genes have snoRNA and/or miRNA functions that control mitochondrial gene expression levels (Mercer et al., 2011). The regulation of mitochondrial gene expression may be controlled by these genes, which are the most conserved in terms of mitochondrial nucleotide data and typically exhibit species-specific barcode characteristics (Lynn & Strüder-Kypke, 2006; Meiklejohn et al., 2007; Santamaria et al., 2009). Alternatively, the expression levels of the *cox1* gene may be controlled by the highly differentiated fragmented cox1 gene located in the broad antisense chain of the organism from its own *cox1* gene. This may be due to important functional limitations of the self-replicating replacement gene fragment, which is controlled by both sequential (Meiklejohn et al., 2007; Pesole et al., 1999) and functional (Decoster et al., 1990) constraints. These limitations may be specific to the organism or may require the *cox1* transcript. Thus, the mitochondria may protect the *cox1* gene to avoid disrupting cellular processes, which may be the easiest way to keep copies in antisense conjugate.

An additional explanation for the high abundance of *nd5* and *cox1* antisense transcripts may be the existence of group 1 introns in these genes, particularly in certain Cnidarians (Beagley et al., 1996; Fukami et al., 2007). These genes and transcripts may be shielded by the mitogenome due to the function of this intron group in the splicing mechanism. Alternatively, these potential sequences could serve as a template for lncRNAs that regulate the expression of conjugate or other mitochondrial genes. For example, lncRNA has been found to be expressed for the light chain gene *nd6* in human mitogenomes. In insect mitogenomes, nearly half of the genes are encoded by the light chain (Figure 1), and in *S. parreyssi*, 14 out of 37 genes were encoded by the light chain, suggesting a high likelihood for the production of lncRNAs (Figure 2).

The majority of data obtained from the ORF Finder program was derived from transcriptomic studies, which may suggest that the relevant ORFs are indeed expressed or have copies in the nuclear genome. Although it is challenging to detect antisense transcripts by real-time PCR, if these ORFs are expressed, we would expect that they are expressed at least to some degree in each region of the genome. The fact that 54.94% of the ORFs detected may be encoded from the antisense chain could also explain the production of antisense polycistronic messages (Supporting Information: Table S3). Further research, including pure mitochondrial transcriptomics studies, should be conducted for products that cannot be obtained in sufficient numbers by real-time PCR analysis. The existence of nuclear equivalents of these ORFs is a second possibility. Current knowledge suggests that this possibility is more likely due to the ongoing gene transfers between the mitochondrial and nuclear genomes (De Gray, 2005; Müller & Martin, 1999).

Localization of the highly conserved duplicated fragment of the *cox3* gene to a conserved region, such as rrnL domain 5-domain 6, may increase the accuracy of the *cox3* gene by DNA repair mechanisms at the replication level. If the antisense pair of the rrnL gene also has high expression levels, we could suggest that it is also controlled at the transcription and/or translation level, but we could not detect antisense expression. Since mitochondrial duplication commonly results in the random loss of one of the duplicated genes (TDRL) (Bernt et al., 2013), the mitochondria may have copied this duplex control gene into the rrnL gene. Such a mechanism is rare, but it involves intramolecular duplexing, usually with short but highly similar sequences (Yu et al., 2007). This duplication is located

not only in the stem-loop structure but also in the region that does not form the secondary structure, so as not to disrupt the rrnL topology of the related species.

AUTHOR CONTRIBUTIONS

The project development, data management, and manuscript writing were conducted by Habeş Bilal Aydemir. Data collection and analysis were performed by Merve Nur Aydemir. Ertan Mahir Korkmaz reviewed the manuscript. All authors have read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

This study did not receive any dedicated financial support from public, commercial, or nonprofit organizations. The authors acknowledge the valuable laboratory assistance provided by Berşan Seçil Durel Avcıoğlu and Sevcan Yangın.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Habeş Bilal Aydemir http://orcid.org/0000-0002-0806-8531
Ertan Mahir Korkmaz http://orcid.org/0000-0003-0699-1354

REFERENCES

Alexeyev, M.F. (2017) Mitochondrial transcription in mammalian cells. *Frontiers in Bioscience*, 22(5), 835–853.

Asin-Cayuela, J. & Gustafsson, C.M. (2007) Mitochondrial transcription and its regulation in mammalian cells. *Trends in Biochemical Sciences*, 32(3), 111–7. http://www.ncbi.nlm.nih.gov/pubmed/17291767

Aydemir, M.N., Aydemir, H.B., Budak, M., Kızıltepe, B., Çelebi, M.Ş., Korkmaz, E.M. et al. (2023) A novel, conserved and possibly functional motif "WHWGHTW" in mitochondrial transcription across Bilateria. *Mitochondrion*, 68, 72–80. https://linkinghub.elsevier.com/retrieve/pii/S1567724922000939

Beagley, C.T., Okada, N.A. & Wolstenholme, D.R. (1996) Two mitochondrial group I introns in a metazoan, the sea anemone *Metridium senile*: one intron contains genes for subunits 1 and 3 of NADH dehydrogenase. *Proceedings of the National Academy of Sciences*, 93(11), 5619–5623.

Bernt, M., Braband, A., Schierwater, B. & Stadler, P.F. (2013) Genetic aspects of mitochondrial genome evolution. *Molecular Phylogenetics and Evolution*, 69(2), 328–338. https://linkinghub.elsevier.com/retrieve/pii/S1055790312004265

Berthier, F., Renaud, M., Alziari, S. & Durand, R. (1986) RNA mapping on drosophila mitochondrial DNA: precursors and template strands. *Nucleic Acids Research*, 14(11), 4519–4533.

Brzezniak, L.K., Bijata, M., Szczesny, R.J. & Stepien, P.P. (2011) Involvement of human ELAC2 gene product in 3′ end processing of mitochondrial tRNAs. *RNA Biology*, 8(4), 616–626.

Burzio, V.A., Villota, C., Villegas, J., Landerer, E., Boccardo, E., Villa, L.L. et al. (2009) Expression of a family of noncoding mitochondrial RNAs distinguishes normal from cancer cells. *Proceedings of the National Academy of Sciences*, 106(23), 9430–9434.

Cameron, S.L. (2014) How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. *Systematic Entomology*, 39(3), 400–411. https://doi.org/10.1111/syen.12071

Cantatore, P., Roberti, M., Polosa, P.L., Mustich, A. & Gadaleta, M.N. (1990) Mapping and characterization of *Paracentrotus lividus* mitochondrial transcripts: multiple and overlapping transcription units. *Current Genetics*, 17(3), 235–245.

Castellana, S., Vicario, S. & Saccone, C. (2011) Evolutionary patterns of the mitochondrial genome in metazoa: exploring the role of mutation and selection in mitochondrial protein-coding genes. *Genome Biology and Evolution*, 3(1), 1067–1079 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3229188&tool=pmcentrez&rendertype=abstract [Accessed November 25, 2014].

Clayton, D.A., Hughes, H. & Chase, C. (2000) Transcription and replication of mitochondrial DNA., 15(Figure 1), pp. 11–17.
Decoster, E., Simon, M., Hatat, D. & Faye, G. (1990) The MSS51 gene product is required for the translation of the COX1 mRNA in yeast mitochondria. *Molecular & General Genetics: MGG*, 224(1), 111–118.
De Grey, A.D.N.J. (2005) Forces maintaining organellar genomes: is any as strong as genetic code disparity or hydrophobicity? *BioEssays*, 27(4), 436–446.
Fukami, H., Chen, C.A., Chiou, C.Y. & Knowlton, N. (2007) Novel group I introns encoding a putative homing endonuclease in the mitochondrial cox1 gene of scleractinian corals. *Journal of Molecular Evolution*, 64(5), 591–600.
Gray, M.W. (1989) Origin and evolution of mitochondrial DNA. *Annual Review of Cell Biology*, 5, 25–50.
Lee, C., Kim, K.H. & Cohen, P. (2016) MOTS-c: a novel mitochondrial-derived peptide regulating muscle and fat metabolism. *Free Radical Biology and Medicine*, 100, 182–187. Available at https://doi.org/10.1016/j.freeradbiomed.2016.05.015
Lee, C., Yen, K. & Cohen, P. (2013) Humanin: a harbinger of mitochondrial-derived peptides? *Trends in Endocrinology & Metabolism*, 24(5), 222–228. https://linkinghub.elsevier.com/retrieve/pii/S1043276013000179
Lee, C., Zeng, J., Drew, B.G., Sallam, T., Martin-Montalvo, A., Wan, J. et al. (2015) The mitochondrial-derived peptide MOTS-c promotes metabolic homeostasis and reduces obesity and insulin resistance. *Cell Metabolism*, 21(3), 443–454. https://linkinghub.elsevier.com/retrieve/pii/S1550413115000613
Lopez Sanchez, M.I.G., Mercer, T.R., Davies, S.M.K., Shearwood, A.M.J., Nygård, K.K.A., Richman, T.R. et al. (2011) RNA processing in human mitochondria. *Cell Cycle*, 10(17), 2904–2916. http://www.tandfonline.com/doi/abs/10.4161/cc.10.17.17060
Lucy, W.B., Sue, F. & Wilton, S.D. (2013) *Untranslated gene regions and other non-coding elements*.
Lung, B., Zemann, A., Madej, M.J., Schuelke, M., Techritz, S., Ruf, S. et al. (2006) Identification of small non-coding RNAs from mitochondria and chloroplasts. *Nucleic Acids Research*, 34(14), 3842–3852.
Lynn, D.H. & Strüder-Kypke, M.C. (2006) Species of tetrahymena identical by small subunit rRNA gene sequences are discriminated by mitochondrial cytochrome c oxidase I gene sequences. *The Journal of Eukaryotic Microbiology*, 53(5), 385–7.
Mears, J.A., Sharma, M.R., Gutell, R.R., McCook, A.S., Richardson, P.E., Caulfield, T.R. et al. (2006) A structural model for the large subunit of the mammalian mitochondrial ribosome. *Journal of Molecular Biology*, 358(1), 193–212.
Meiklejohn, C.D., Montooth, K.L. & Rand, D.M. (2007) Positive and negative selection on the mitochondrial genome. *Trends in Genetics*, 23(6), 259–263. https://doi.org/10.1016/j.tig.2007.03.008
Mercer, T.R., Neph, S., Dinger, M.E., Crawford, J., Smith, M.A., Shearwood, A.M.J. et al. (2011) The human mitochondrial transcriptome. *Cell*, 146(4), 645–658. https://linkinghub.elsevier.com/retrieve/pii/S0092867411007677
Müller, M. & Martin, W. (1999) The genome of rickettsia prowazekii and some thoughts on the origin of mitochondria and hydrogenosomes. *BioEssays*, 21(5), 377–81.
Osigus, H. et al., 2013. Molecular Phylogenetics and Evolution Mitogenomics at the base of Metazoa., 69, pp.339–351.
Pesole, G., Gissi, C., De Chirico, A. & Saccone, C. (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. *Journal of Molecular Evolution*, 48, 427–434.
Polosa, P.L., Deceglie, S., Falkenberg, M., Roberti, M., Di Ponzio, B., Gadaleta, M.N. et al. (2007) Cloning of the sea urchin mitochondrial RNA polymerase and reconstitution of the transcription termination system. *Nucleic Acids Research*, 35(7), 2413–2427.
Rackham, O., Shearwood, A.M.J., Mercer, T.R., Davies, S.M.K., Mattick, J.S. & Filipovska, A. (2011) Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. *RNA*, 17(12), 2085–2093. Available at: http://rnajournal.cshlp.org/cgi/doi/10.1261/rna.029405.111
Santamaria, M., Vicario, S., Pappadà, G., Scioscia, G., Scazzocchio, C. & Saccone, C. (2009) Towards barcode markers in fungi: an intron map of ascomycota mitochondria. *BMC Bioinformatics*, 10(Suppl 6), S15.
Scarpulla, R.C. (2008) Transcriptional paradigms in mammalian mitochondrial biogenesis and function.
Schmittgen, T.D. & Livak, K.J. (2008) Analyzing real-time PCR data by the comparative CT method. *Nature Protocols*, 3(6), 1101–1108.
Stewart, J.B. & Beckenbach, A.T. (2009) Characterization of mature mitochondrial transcripts in drosophila, and the implications for the tRNA punctuation model in arthropods. *Gene*, 445(1–2), 49–57. http://www.ncbi.nlm.nih.gov/pubmed/19540318
Temperley, R.J., Wydro, M., Lightowlers, R.N. & Chrzanowska-Lightowlers, Z.M. (2010) Human mitochondrial mRNAs-like members of all families, similar but different. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1797(6–7), 1081–1085.
Torres, T.T., Dolezal, M., Schlötterer, C. & Ottenwälder, B. (2009) Expression profiling of drosophila mitochondrial genes via deep mRNA sequencing. *Nucleic Acids Research*, 37(22), 7509–7518.
Yu, D.J., Xu, L., Nardi, F., Li, J.G. & Zhang, R.J. (2007) The complete nucleotide sequence of the mitochondrial genome of the oriental fruit Fly, *Bactrocera dorsalis* (Diptera: Tephritidae). *Gene*, 396(1), 66–74.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Aydemir, H.B., Aydemir, M.N. & Korkmaz, E.M. (2023) Differential sense and antisense expression profiles of *Syrista parreyssi* (Hymenoptera: Cephidae) mitochondrial transcripts. *Archives of Insect Biochemistry and Physiology*, 113, e22026. https://doi.org/10.1002/arch.22026