Polyadenylation-based mitotranscriptomics of Apis mellifera (Insecta: Hymenoptera)

Merve Aydemir[1], Habeş Bilal AYDEMİR[2], and Ertan Korkmaz[3]

[1]Gaziosmanpasa University
[2]Tokat Gaziosmanpaşa University
[3]Sivas Cumhuriyet University Faculty of Science

April 12, 2023

Abstract

Mitochondrial transcription is being studied with the increasing interest to better understand the coding capacity of mitochondrial DNA (mtDNA). Based on studies showing that mtDNA encodes additional genes that are not functional in oxidative phosphorylation (OXPHOS), it has been understood that mitochondrial transcription is more complex than previously thought. In this study, mitochondrial transcription was investigated in terms of polyadenylation patterns in Apis mellifera (Insecta: Hymenoptera). We found that both sense and antisense transcripts were captured and polyadenylated, and gene expression levels and polyadenylation length were highly variable between genes in Apis. Mitochondrial RNAs of A. mellifera were found to be polyadenylated with an average size of 31 adenines. Nevertheless, the highest transcript representatives of each gene generally appeared to have polyadenylation of size 5-9 bases. Generally, transcripts containing longer and/or shorter forms of each gene were thought to indicate the existence of putative UTR-containing genes, functional non-coding RNA genes and variant gene forms. We showed that some transcripts contain non-template dependent additional nucleotides before polyadenyl tail. Transcripts containing at least some bases of the intergenic regions downstream of the gene had the highest transcript representation after the monocistronic gene units. These findings support the possible functional role of the intergenic regions.

1. INTRODUCTION

Animal mitogenome is a highly encoded genome (>90%) since it contains no introns. It typically encodes 37 genes (22 tRNA, 13 PCG and 2 rRNA) functional in oxidative phosphorylation (OXPHOS) process which effects all mitochondrial function either directly or indirectly. Thus, mitochondrial transcription is key to maintain mitochondrial function, both in terms of the expression of oxidative phosphorylation subunits encoded from mitochondria and mtDNA replication. Although advances have been made with more than 40 years of work to elucidate mitochondrial transcription, there are still challenges and contradictions that need to be addressed (Shokolenko and Alexeyev, 2017).

In insect mitogenomes, these "formal" 37 genes are distributed almost equally in heavy (GC-rich) and light (GC-poor) chains, and the entire chains are transcribed as polycistronic units. Accordingly, almost half of each polycistronic unit is not converted into known RNA molecules (mRNA, tRNA or rRNA) after transcripts are processed post- transcriptionally (Figure 1).

Interestingly, novel protein coding genes which are encoded in the mitogenome and not functional in OXPHOS have recently been identified (Caricasole *et al.* , 2002; Lee, Yen and Cohen, 2013; Lee, Kim and Cohen, 2016). Not only protein coding genes but also novel non-coding RNAs have been found especially in human mitogenomes including lncRNAs (Mercer *et al.* , 2011; Rackham *et al.* , 2011; Antonicka*et al.* , 2013). Even, some non-coding intergenic regions have also been found to be present in the transcripts,

especially in mRNAs of*Drosophila* (Stewart and Beckenbach, 2009). Consequently, it has been hypothesized that the mitochondrial transcription capacity may be more complex and intense than previously assumed.

Polyadenylation of mitochondrial transcripts can be functionally different when compared to their nuclear counterparts. While it is known that polyadenylation ensures mRNA stability in nuclear transcription, mitochondrial polyadenylation can lead to the degradation of mRNAs (Bouda, Stapon and Garcia-Diaz, 2019). Especially, non-polyadenylated forms of some mRNAs can induce relatively higher expression and be more stable than the polyadenylated forms, as in the case of human*cox1* (Mercer *et al.* , 2011).

Even processes of the mitochondrial transcription and post-transcriptionally RNA maturation have been studied in*Drosophila* and sea-urchin, it also needs to be studied in different invertebrates due to the variation of transcription profiles (Cantatore *et al.* , 1990; Polosa *et al.* , 2007; Torres, Dolezal, Schlötterer and Ottenwälder, 2009). In this study, we predicted the polyadenylation pattern of the mitogenome of *Apis mellifera*bioinformatically by conducting high-throughput RNA sequencing data. Additionally, we determined the transcription patterns of mitochondrial intergenic regions.

2. MATERIAL and METHOD

2. 1. Sample selection

Apis mellifera RNA-seq data (SRX5798518) and *Apis mellifera sahariensis* data (NC_035883) as reference mitogenomes were downloaded from NCBI database. The RNA-seq data was generated using the brain tissues of *Apis mellifera* infected with deformed wing virus type A, and produced by Traniello et al. (2020) in Illumina HiSeq 4000 platform with Chromium Next GEM Single Cell 3' library preparation. This RNA-seq data was 34.2G bases in length and 18.4Gb in size. Mitogenome reference was a verified RefSeq data and 16,569 bp in length with non-coding A+T- rich region.

2. 2. Determination of mitochondrial transcription pattern

PolyA (polyT in reverse direction) selection was applied to determine the polyadenylation patterns and relative abundances of mitochondrial monocistronic units. Firstly, reads from the RNA-seq data were reduced to only those containing polyA sequences with a minimum five bases in size at any end, using the manually written script in PERL (Appendix 1). Then, selected sequences with polyadenylation were aligned to the reference mitogenome with "map to reference" option and using the bowtie2 algorithm in Geneious R9.1.5 (*https://www.geneious.com*) (Kearse *et al.* , 2012) and their patterns were checked manually. Whole RNA-seq data without polyadenylation selection was also aligned to mitogenome, and RPKM (Reads Per Kilobase of transcript per Million mapped reads) values were calculated for each gene.

2.3. Prediction of potential coding capacity of mitogenome

Mitogenome data of *Apis mellifera* was analyzed in ORFFinder (Open Reading Frame Finder, NCBI) in order to predict additional gene regions which are able to encode proteins different than known 13 OXPHOS proteins. Search parameters were adjusted as "invertebrate mitochondrial genetic code" and "ATG" and "alternative initiation codons". Potential protein encoding sequences were then verified using regular BLASTP.

3. RESULTS and DISCUSSION

102.775 sequences were aligned to reference mitogenome from polyadenylation selected RNA-seq data. 76.175 of those mapped to any end of any mitochondrial gene. 33 of 37 mitochondrial genes were represented with at least one mapped polyadenylated sequence. It has been determined that not only protein-coding genes but also rRNA and tRNA genes are polyadenylated in the *Apis mellifera* mitochondrial transcriptome. In general, polyadenyls of five to 124 bases long were added to mitochondrial genes. The highest representation for all genes were polyadenylated forms with 5-9 adenines in length (19%). Polyadenylated transcripts with 10-14 and 15-19 adenines were represented by 16% and 13%, respectively. The length of polyadenylation and the number of representative transcripts seemed to be inversely proportional in logarithmic scale (Figure 2, R^2= 0,9578). Specifically, polyadenylation length and transcript counts of each gene were given in Figure 3.

Results obtained from polyadenylation selection and related properties were given as a table in Appendix -2.

trnQ, trnS1, trnE and trnL2 tRNA genes were not mapped to any sense or antisense polyadenylated end-read. trnQ, trnS1 and trnE genes are located quite close to the noncoding A+T rich region; thus, it may be due the reduced transcript abundance that any end-read were not mapped to these genes. On the other hand, it is remarkable that there was no mapped read to trnL2 gene, which is located between the most expressed genes *cox1* and *cox2* . *cox2* gene was represented with 60%, *cox1* gene with 25% and rrnL rRNA gene with 9% of the total end- reads; while others have less than one percent representation (Figure 4c). It is determined that the rRNA genes were more under-expressed than expected, probably because of the sequencing platform used with polyA selection.

While 99.19% of the mapped reads represent sense orientated transcripts, antisense transcripts were also found (Figure 4d). Sense transcripts were polyadenylated longer than antisense counterparts for each gene group (Figure 5). The presence of both sense and antisense transcripts indicated that functional RNA molecules were generated in insects, similar to 7S RNA and *nd6* lncRNA in humans (Mercer*et al.* , 2011). Furthermore, insect mitochondrial gene content was distributed equally in both heavy and light chains unlike human mitogenome which codes only *nd6* and a few tRNAs in light chain. Thus, it can be speculated that the potential of producing noncoding- functional RNAs from insect mitogenomes can be much more likely compared to that in human mitogenomes.

rrnL, *nd4* , rrnS, *nd3* and trnV genes had only sense, and trnN, trnW and trnM genes had only antisense transcripts (Figure 4b). Antisense transcripts had 22.85 adenines and sense transcripts had 31.00 adenines on average in their polyadenyl tails. Our data showed variable polyadenylation length for sense and antisense transcripts of each gene, and patterns of polyadenylation were generally independent of the expression level (Figure 3, 4e and Appendix-2).

We found that some of mitochondrial genes predominantly appear to form longer or shorter transcripts than annotated formal gene length. 21 of 33 mitochondrial genes tended to be found in longer transcripts. Nine of these genes add whole sequence following the intergenic region (*atp6* , *cytB* , *nd4* , *nd5* , *nd6* , rrnL, trnN, trnP, trnY) and other eight of 21 add at least one base following the intergenic region (*cox3* , *nd1* , *nd3* , *nd4L* , trnG, trnR, trnS2, trnT). The remaining four contain, if they have any intergenic region, add at least one base following mitochondrial gene in the transcript (*atp8* , *nd2* , trnC, trnW) (Figure 6).

Intergenic regions that are estimated as gene rearrangement residues, seem to be conserved despite the reduced nature of mitogenome in*Apis* . Mitochondrial regulator elements of transcription can spread into whole genome, even into gene regions (Barshad *et al.* , 2018).Thus, it was suggested that intergenic regions may be considered as transcription control regions (Barshad *et al.* , 2018; Aydemir and Korkmaz, 2020).

Long RNA molecules were evaluated as a result of a mis-processing of transcription termination in human mitochondria (Temperley *et al.* , 2010). Although this assumption is acceptable for protein-coding genes, the results of the rrnL gene is remarkable: 86.25% and 94.02% of all rrnL transcripts contain whole sequence and half of whole sequence of neighboring intergenic region, respectively. These high rates may also suggest the possibility of incorrect annotation of this noncoding gene since RNA gene annotations were performed using PCG boundaries and there is no accepted, standard, consensus system for annotation (Stewart and Beckenbach, 2009). *cox1* , *cox2* , rrnS, trnG and trnK genes tend to form a transcript that contain only annotated sequences. Seven tRNAs seem to be transcribed mostly in truncated forms (trnA, trnD, trnH, trnI, trnL1, trnM, trnV). Truncated mRNA molecules are determined in human mitochondria, potentially produced by secondary structures that may occur within the genes and evaluated as residues of RNA surveillance pathway based on their low frequencies with 1/161 (Szczesny *et al.* , 2009; Mercer *et al.* , 2011). These polyadenylated truncated transcripts have been evaluated as a clue for the occurrence of polyadenylation-dependent RNA degradation processes in human mitochondria (Slomovic *et al.* , 2005). Nevertheless, it is determined that 9,80% of all mitochondrial transcripts of *Apis* represent truncated forms. In recent studies, it is supposed that truncated mitochondrial RNA molecules may be precursors for circRNA molecules and are not translated (Mance *et al.* , 2020). Unlike in humans, truncated forms were found not only in mRNAs

but also in rRNA and tRNA transcripts. Truncated mitochondrial tRNA molecules were detected and evaluated as t-elements in a freshwater alga *Cyanophora paradoxa* (Salinas-Giegé, Ubrig and Drouard, 2021). It has been proposed that these truncated tRNA molecules potentially hybridize with neighboring mRNAs and thereby protect and stabilize mRNAs from exonuclease digestion (Salinas-Giegé, Ubrig and Drouard, 2021).

Accordingly, 75.93% of mitochondrial transcripts containing polyadenyl residues at their ends consist of only the annotated gene (gene + polyA/T), 11.42% of them consist of the gene and the entire nucleotides of following intergenic region (gene + IG + polyA/T), 9.78% consist of truncated genes (gen- X base + polyA/T), 2.40% consist of the gene and a part of the following intergenic region (gene + X base IG + polyA/T) and 0.47% consist of gene, and if any intergenic region, and a part of the neighbor gene (gene(+ IG)+ gene+ polyA/T) (Figure 4f).

Genes which transcribed in the same polycistronic unit did not show similar expression abundance in RPKM analysis. *cox2* was the most expressed, while trnE is the least expressed mitochondrial gene, and general pattern in expression levels was*cox2* >*cox1* >> trnD> rrnL> trnL2> trnC> trnY> trnL1> trnW> *nd6* > rrnS> trnI> *nd2* > trnV>*atp6* > *nd1* > trnR>*cytB* > trnN> trnF> trnK> trnG> trnS2> trnM> *cox3* > *nd3* > trnS1> trnH> trnP>*nd5* > trnA> *atp8* >*nd4* > trnQ> trnT>*nd4L* > trnE. RPKM value of each gene was given in Table 1 and Figure 4a. Genes with high RPKM values were found to exhibit variation in their polyadenylation profiles (Figure 2 and Appendix 2). While rRNA expression levels are expected to be the highest, the overexpression of *cox2* and *cox1* is remarkable (Gelfand and Attardi, 1981). *cox* genes, especially *cox2* mRNA accumulation can be the result of DWV (deformed wing virus) infection of*Apis* in this study. It was found that the expression of*cox2* increases due to the modulation of prostaglandin E2 (PGE_2) synthesis during viral infection (Steer and Corbett, 2003).

According to the results of ORF Finder analyses, 205 potential ORFs were found to have equal or longer than 75 nucleotides (25 amino acids) (Appendix- 3). 32 of these were verified with BLAST searches, and 13 of those were known OXPHOS genes. 13 and one of them were aligned with bacterial and insect hypothetic proteins, respectively, and need further validation to understand whether they are actually translated into proteins. Four and one of 32 seemed to be variant forms of *cox1*and *atp8* genes, respectively.

3. 1. Polyadenylation patterns of gene groups

In particular, the following findings were obtained when each mitochondrial gene cluster was examined individually.

Polyadenyl length varied from 15.41 (*atp8*) to 41.88 (*nd6*), and 23.83 adenines for all protein-coding genes were found in sense transcripts. Antisense transcripts of protein- coding genes had minimum of 12 (*atp6*), mean of 20.27 and maximum of 56 (*cox1*) bases polyadenylated (Figure 3, 4e). 11 of 13 protein coding genes seemed to be transcribed in longer units than monocistrons which include only ORF. The excessive presence of intergenic regions in both sense and antisense transcripts indicates that intergenic regions are also transcribed and therefore may be functional in transcription or translation, like recommended as a transcription termination signal in Aydemir et al. 2022 and Roberti et al. 2003. We hypothesized that these extra-added sequences can act as UTRs for mitochondrial mRNAs. Both 5‘ and 3‘ UTRs are known to play critical roles in the post-transcriptional regulation of gene expression by modulating the transport of mRNAs. It is known that 3‘ UTRs regulate translation efficiency and are important in the intracellular localization and the stability of mRNAs (Van Der Velden and Thomas, 1999; Bashirullah, Cooperstock and Lipshitz, 2001; Jansen, 2001). Although it is assumed that mitochondrial RNAs do not contain UTRs, some genes may contain 3‘ or 5‘ UTR sequences (Gagliardi *et al.* , 2004; Nagaike *et al.* , 2005). In general, it has been suggested that these UTR-containing sequences are the products of non-canonical processes and may be associated with RNA secondary structure clusters (Mercer *et al.* , 2011). 88.03% of antisense transcripts of *Apis* have additional genomic sequences in monocistron, and these can be considered as UTRs for insects, such as in humans (Temperley *et al.* , 2010). tRNA sequences can act as 3‘ UTR through the tRNA punctuation model in human mitogenomes (Temperley*et al.* , 2010). Intergenic regions surrounding almost all mRNAs including tRNAs may also act as UTRs for mRNAs in animal mitogenomes, unlike that

in human. It has been suggested that 3‘ UTR sequences in sense transcripts may be regulated by post-transcriptional processes and control the excessive expression of mRNAs in *Drosophila* and human (Torres, Dolezal, Schlötterer, Ottenwälder, *et al.* , 2009; Mercer*et al.* , 2011). We propose that premature stop codons which are commonly exist in mitochondrial protein coding genes may lead to the incorporation of additional amino acids to the protein and may produce canonically longer transcripts, such has been previously detected in*cox2* gene of *D. melanogaster* with additional 11 amino acids (Stewart and Beckenbach, 2009).

Polyadenyl length was varied from 17,49 (rrnS) to 19,35 (rrnL) with an average of 18.42 bases for sense rRNAs. The shortest polyadenylated gene group was rRNA transcripts (Figure 4e). This may be the result of the tendency to economize the mitogenome due to the proportional abundance of rRNA transcripts.

Polyadenyl length was varied from 11.64 (trnA) to 85 (trnF), and 45.08 adenines for all tRNA genes on average were found in sense transcripts. Antisense transcripts of tRNAs have minimum of 13.25 (trnT), mean of 28.84 and maximum of 57 (trnF) bases polyadenylated (Figure 4e). Antisense transcripts of each tRNA were much more intensely detected than sense transcripts (10 tRNAs of 17 total). Antisense tRNA pool could generate tRFs (tRNA- derived RNA fragments), which are a group of small RNAs detected in human and yeast mitochondria (Lee *et al.* , 2009; Mercer *et al.* , 2011; Ro *et al.* , 2013; Shang *et al.* , 2018). Although the biogenesis of these RNA molecules are not known, it was predicted that they play a role in the regulation of mtDNA gene expression (Ro *et al.* , 2013). Since tRNAs are known to be extensively transported from the nuclear genome into the mitochondria, the scarcity of sense tRNA molecules does not have the adverse effect on mitochondrial transcription such as inhibition or reduction of gene expression (Mercer *et al.* , 2011).

3.2. Non-templated additional nucleotides before polyadenylation

Interestingly, we found that some genes tend to add some nucleotides just before the polyadenylation sequence. Remarkably, this bias was determined in both sense and antisense transcripts of *Apis* . We found that 18.70% of all mapped transcripts (14,248 sequences) contain at least one and at most seven additional nucleotides before polyadenyls. 42.25% of the additional nucleotides were C, 30.28% G, 18.28% T and 9.19% A nucleotide (Figure 4g). It is suggested that additional nucleotides before polyadenyls may be functional in the processing of 3‘ ends in maize mitochondrial *rps12* , *cox2*and *atp9* genes (Williams, Johzuka and Mulligan, 2000). Non-templated additional nucleotides were also found to be present in human U2 snRNA and in some of *Arabidopsis* nuclear mRNAs in other studies (Cho *et al.* , 2002; Jin and Bian, 2004). The additional nucleotide detected in these three species mentioned above was nucleotide C, and it was assumed that this nucleotide determines the heterogeneity in transcript abundance (Jin and Bian, 2004). RNA-editing mechanism is prevalent and mainly sustained with C to U editing in plant mitochondrial transcription (Wu *et al.* , 2015). While it is assumed that animal mitochondrial transcription system shorn of RNA-editing mechanism, this additional nucleotide sites may be templates of post-transcriptional regulations such as RNA-editing (Lynch, Koskella and Schaack, 2006).

4. CONCLUSION

In this study, the entire mitotranscriptomics of *Apis mellifera*were characterised for the first time. The following important findings were also obtained in the present study: (i) mitochondrial genes had polyadenyl sequences of 5 to 124 bases long, but the highest representation for all genes was 5-9 base polyadenylated forms, (ii) antisense transcripts were found as polyadenylated forms in transcript pool, (iii) polyadenylation of annotated, monocistronic gene units was the most common pattern, (iv) some of mitochondrial genes predominantly appeared to form longer or shorter transcripts than annotated formal gene length, (v) the overexpression of *cox2* gene was potentially due to the viral infection of selected sample, (vi) *Apis*mitogenome had the coding capacity of 32 alternative ORFs, (vii) 18.70% of studied transcripts contain at least one and at most seven additional nucleotides before polyadenyls, especially the cytosine. Apart from, comparisons from different tissues sequenced with different approaches may provide further testing of mitochondrial gene expression profiles due to the nature of transcriptomic analyzes.

REFERENCES

Antonicka, H., Sasarman, F., Nishimura, T., Paupe, V. and Shoubridge, E. A. (2013) 'The Mitochondrial RNA-Binding Protein GRSF1 Localizes to RNA Granules and Is Required for Posttranscriptional Mitochondrial Gene Expression', *Cell Metabolism* , 17(3), pp. 386–398. doi: 10.1016/j.cmet.2013.02.006.

Aydemir, M. N. and Korkmaz, E. M. (2020) 'Comparative mitogenomics of Hymenoptera reveals evolutionary differences in structure and composition', *International Journal of Biological Macromolecules* . doi: 10.1016/j.ijbiomac.2019.12.135.

Barshad, G., Marom, S., Cohen, T. and Mishmar, D. (2018) 'Mitochondrial DNA Transcription and Its Regulation: An Evolutionary Perspective',*Trends in Genetics* , 34(9), pp. 682–692. doi: 10.1016/j.tig.2018.05.009.

Bashirullah, A., Cooperstock, R. L. and Lipshitz, H. D. (2001) 'Spatial and temporal control of RNA stability', *Proceedings of the National Academy of Sciences* , 98(13), pp. 7025–7028. doi: 10.1073/pnas.111145698.

Bouda, E., Stapon, A. and Garcia-Diaz, M. (2019) 'Mechanisms of mammalian mitochondrial transcription', *Protein Science* , 28(9), p. pro.3688. doi: 10.1002/pro.3688.

Cantatore, P., Roberti, M., Polosa, P. L., Mustich, A. and Gadaleta, M. N. (1990) 'Mapping and characterization of Paracentrotus lividus mitochondrial transcripts: multiple and overlapping transcription units', *Current Genetics* , 17(3), pp. 235–245. doi: 10.1007/BF00312615.

Caricasole, A., Bruno, V., CAPPUCCIO, I., Melchiorri, D., Copani, A. and Nicoletti, F. (2002) 'A novel rat gene encoding a Humanin-like peptide endowed with broad neuroprotective activity', *The FASEB Journal* , 16(10), pp. 1331–1333. doi: 10.1096/fj.02-0018fje.

Cho, H. D., Tomita, K., Suzuki, T. and Weiner, A. M. (2002) 'U2 Small Nuclear RNA Is a Substrate for the CCA-adding Enzyme (tRNA Nucleotidyltransferase)', *Journal of Biological Chemistry* , 277(5), pp. 3447–3455. doi: 10.1074/jbc.M109559200.

Gagliardi, D., Stepien, P. P., Temperley, R. J., Lightowlers, R. N. and Chrzanowska-Lightowlers, Z. M. A. A. (2004) *Messenger RNA stability in mitochondria: Different means to an end* , *Trends in Genetics* . doi: 10.1016/j.tig.2004.04.006.

Gelfand, R. and Attardi, G. (1981) 'Synthesis and turnover of mitochondrial ribonucleic acid in HeLa cells: the mature ribosomal and messenger ribonucleic acid species are metabolically unstable.',*Molecular and Cellular Biology* , 1(6), pp. 497–511. doi: 10.1128/mcb.1.6.497.

Jansen, R. P. (2001) 'mRNA localization: Message on the move',*Nature Reviews Molecular Cell Biology* , pp. 247–256. doi: 10.1038/35067016.

Jin, Y. and Bian, T. (2004) 'Nontemplated nucleotide addition prior to polyadenylation: A comparison of Arabidopsis cDNA and genomic sequences', *RNA* , 10(11), pp. 1695–1697. doi: 10.1261/rna.7610404.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. and Drummond, A. (2012) 'Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data', *Bioinformatics* , 28(12), pp. 1647–1649. doi: 10.1093/bioinformatics/bts199.

Lee, C., Kim, K. H. and Cohen, P. (2016) 'MOTS-c: A novel mitochondrial-derived peptide regulating muscle and fat metabolism',*Free Radical Biology and Medicine* . Elsevier, 100, pp. 182–187. doi: 10.1016/j.freeradbiomed.2016.05.015.

Lee, C., Yen, K. and Cohen, P. (2013) 'Humanin: A harbinger of mitochondrial-derived peptides?', *Trends in Endocrinology and Metabolism* , 24(5), pp. 222–228. doi: 10.1016/j.tem.2013.01.005.

Lee, Y. S., Shibata, Y., Malhotra, A. and Dutta, A. (2009) 'A novel class of small RNAs: tRNA-derived RNA fragments (tRFs)', *Genes and Development* , 23(22), pp. 2639–2649. doi: 10.1101/gad.1837609.

Lynch, M., Koskella, B. and Schaack, S. (2006) 'Mutation pressure and the evolution of organelle genomic architecture', *Science* , pp. 1727–1730. doi: 10.1126/science.1118884.

Mance, L. G., Mawla, I., Shell, S. M. and Cahoon, A. B. (2020) 'Mitochondrial mRNA fragments are circularized in a human HEK cell line', *Mitochondrion* , 51, pp. 1–6. doi: 10.1016/j.mito.2019.11.002.

Mercer, T. R., Neph, S., Dinger, M. E., Crawford, J., Smith, M. A., Shearwood, A.-M. J. M. J., Haugen, E., Bracken, C. P., Rackham, O., Stamatoyannopoulos, J. A., Filipovska, A. and Mattick, J. S. (2011) 'The Human Mitochondrial Transcriptome', *Cell* , 146(4), pp. 645–658. doi: 10.1016/j.cell.2011.06.051.

Nagaike, T., Suzuki, T., Katoh, T. and Ueda, T. (2005) 'Human mitochondrial mRNAs are stabilized with polyadenylation regulated by mitochondria-specific poly(A) polymerase and polynucleotide phosphorylase', *Journal of Biological Chemistry* , 280(20), pp. 19721–19727. doi: 10.1074/jbc.M500804200.

Polosa, P. L., Deceglie, S., Falkenberg, M., Roberti, M., Di Ponzio, B., Gadaleta, M. N. and Cantatore, P. (2007) 'Cloning of the sea urchin mitochondrial RNA polymerase and reconstitution of the transcription termination system', *Nucleic Acids Research* , 35(7), pp. 2413–2427. doi: 10.1093/nar/gkm159.

Rackham, O., Shearwood, A.-M. J., Mercer, T. R., Davies, S. M. K., Mattick, J. S. and Filipovska, A. (2011) 'Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins', *RNA* , 17(12), pp. 2085–2093. doi: 10.1261/rna.029405.111.

Ro, S., Ma, H. Y., Park, C., Ortogero, N., Song, R., Hennig, G. W., Zheng, H., Lin, Y. M., Moro, L., Hsieh, J. T. and Yan, W. (2013) 'The mitochondrial genome encodes abundant small noncoding RNAs', *Cell Research* , 23(6), pp. 759–774. doi: 10.1038/cr.2013.37.

Roberti, M., Loguercio Polosa, P., Bruni, F., Musicco, C., Gadaleta, M. N. and Cantatore, P. (2003) 'DmTTF, a novel mitochondrial transcription termination factor that recognises two sequences of Drosophila melanogaster mitochondrial DNA', *Nucleic Acids Research* , 31(6), pp. 1597–1604. doi: 10.1093/nar/gkg272.

Salinas-Giege, T., Ubrig, E. and Drouard, L. (2021) 'Cyanophora paradoxa mitochondrial tRNAs play a double game', *Plant Journal* . doi: 10.1111/tpj.15222.

Shang, J., Yang, Y., Wu, L., Zou, M. and Huang, Y. (2018) 'The S. pombe mitochondrial transcriptome', *RNA* , 24(9), pp. 1241–1254. doi: 10.1261/rna.064477.117.

Shokolenko, I. N. and Alexeyev, M. F. (2017) 'Mitochondrial transcription in mammalian cells', *Frontiers in Bioscience - Landmark* , pp. 835–853. doi: 10.2741/4520.

Slomovic, S., Laufer, D., Geiger, D. and Schuster, G. (2005) 'Polyadenylation and Degradation of Human Mitochondrial RNA: the Prokaryotic Past Leaves Its Mark', *Molecular and Cellular Biology* , 25(15), pp. 6427–6435. doi: 10.1128/MCB.25.15.6427-6435.2005.

Steer, S. A. and Corbett, J. A. (2003) 'The Role and Regulation of COX-2 during Viral Infection', *Viral Immunology* , 16(4), pp. 447–460. doi: 10.1089/088282403771926283.

Stewart, J. B. and Beckenbach, A. T. (2009) 'Characterization of mature mitochondrial transcripts in Drosophila, and the implications for the tRNA punctuation model in arthropods', *Gene* . Elsevier B.V., 445(1–2), pp. 49–57. doi: 10.1016/j.gene.2009.06.006.

Szczesny, R. J., Borowski, L. S., Brzezniak, L. K., Dmochowska, A., Gewartowski, K., Bartnik, E. and Stepien, P. P. (2009) 'Human mitochondrial RNA turnover caught in flagranti: Involvement of hSuv3p helicase in RNA surveillance', *Nucleic Acids Research* , 38(1), pp. 279–298. doi: 10.1093/nar/gkp903.

Temperley, R. J., Wydro, M., Lightowlers, R. N. and 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), pp. 1081–1085. doi: 10.1016/j.bbabio.2010.02.036.

Torres, T. T., Dolezal, M., Schlotterer, C. and Ottenwalder, B. (2009) 'Expression profiling of Drosophila mitochondrial genes via deep mRNA sequencing', *Nucleic Acids Research* . doi: 10.1093/nar/gkp856.

Torres, T. T., Dolezal, M., Schlotterer, C., Ottenwalder, B., Ottenwa, B., Torres, T. T., Dolezal, M., Schlo, C., Schlotterer, C. and Ottenwalder, B. (2009) 'Expression profiling of Drosophila mitochondrial genes via deep mRNA sequencing', *Nucleic Acids Research* , 37(22), pp. 7509–7518. doi: 10.1093/nar/gkp856.

Traniello, I. M., Bukhari, S. A., Kevill, J., Ahmed, A. C., Hamilton, A. R., Naeger, N. L., Schroeder, D. C. and Robinson, G. E. (2020) 'Meta-analysis of honey bee neurogenomic response links Deformed wing virus type A to precocious behavioral maturation', *Scientific Reports* , 10(1), p. 3101. doi: 10.1038/s41598-020-59808-4.

Van Der Velden, A. W. and Thomas, A. A. M. (1999) 'The role of the 5' untranslated region of an mRNA in translation regulation during development', *International Journal of Biochemistry and Cell Biology* , pp. 87–106. doi: 10.1016/S1357-2725(98)00134-4.

Williams, M. A., Johzuka, Y. and Mulligan, R. M. (2000) 'Addition of non-genomically encoded nucleotides to the 3'-terminus of maize mitochondrial mRNAs: Truncated rps12 mRNAs frequently terminate with CCA', *Nucleic Acids Research* , 28(22), pp. 4444–4451. doi: 10.1093/nar/28.22.4444.

Wu, Z., Stone, J. D., Štorchová, H. and Sloan, D. B. (2015) 'High transcript abundance, RNA editing, and small RNAs in intergenic regions within the massive mitochondrial genome of the angiosperm Silene noctiflora', *BMC Genomics* , 16(1), p. 938. doi: 10.1186/s12864-015-2155-3.

LEGENDS

Figure 1: Mitochondrial RNA pool with sense and antisense units

Figure 2: Preferred polyadenylation length among whole mitochondrial transcripts

Figure 3: Transcript frequencies of mitochondrial genes with different polyadenylation length

Figure 4: The summary of mitotranscriptomics of *Apis mellifera*

Figure 5: Polyadenylation length of sense and antisense transcripts

Figure 6: Relatively amounts of preferred transcript patterns of mitochondrial genes

Data Accessibilty Statement

Not applicable.

Competing Interests Statement

Not applicable

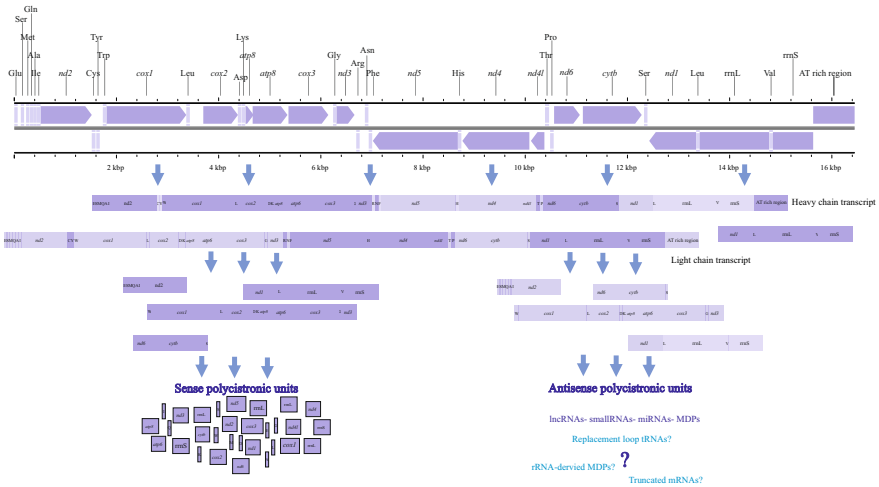
Author Contributions Section

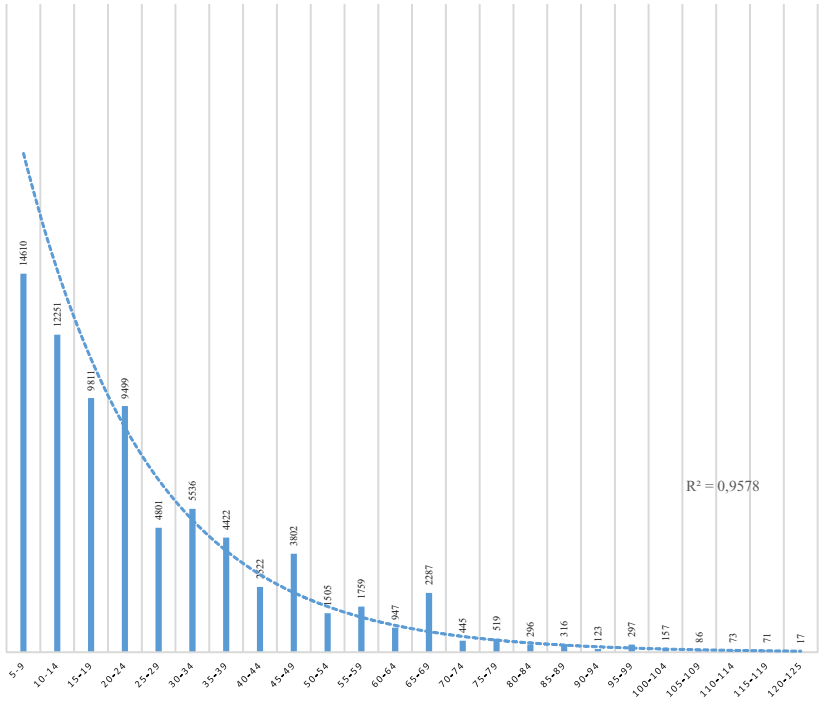
Data was collected and analysed by H. Bilal Aydemir. Project development, data management and manuscript writing was conducted by Merve N. Aydemir. The manuscript was reviewed by E. Mahir Korkmaz.

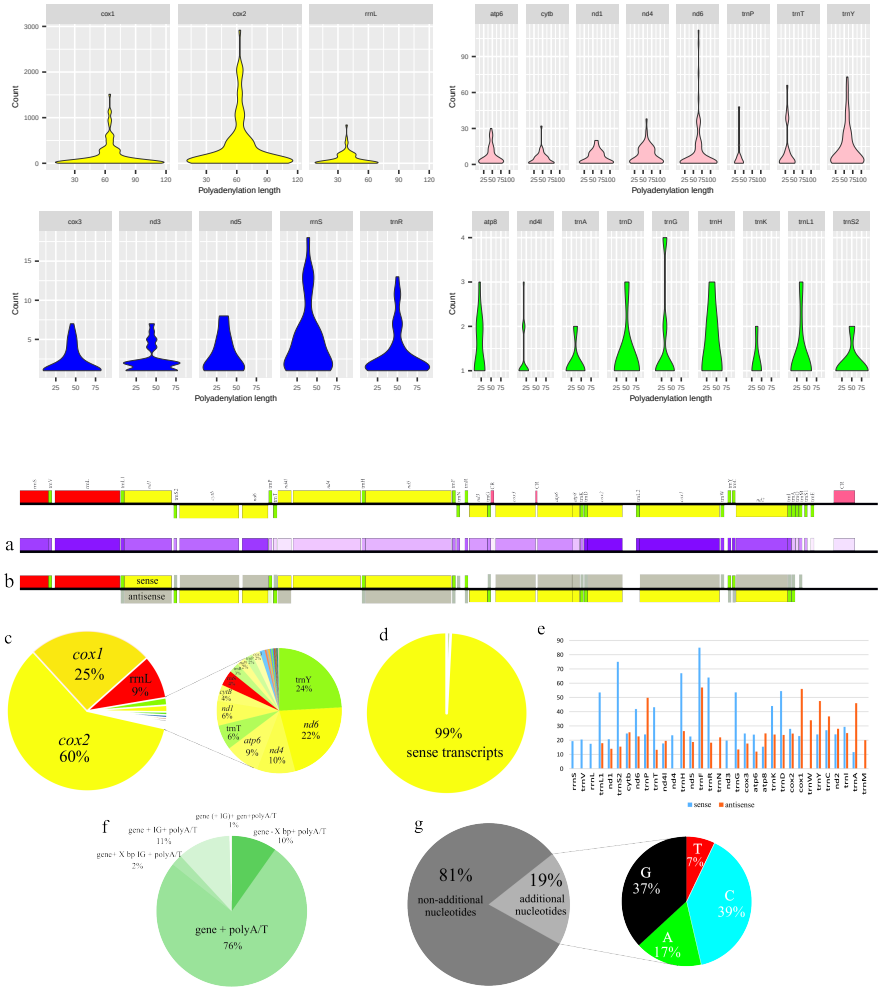
All authors have read and approved the manuscript. A part of this project was presented as a PhD thesis of the first author (M.N.A.).

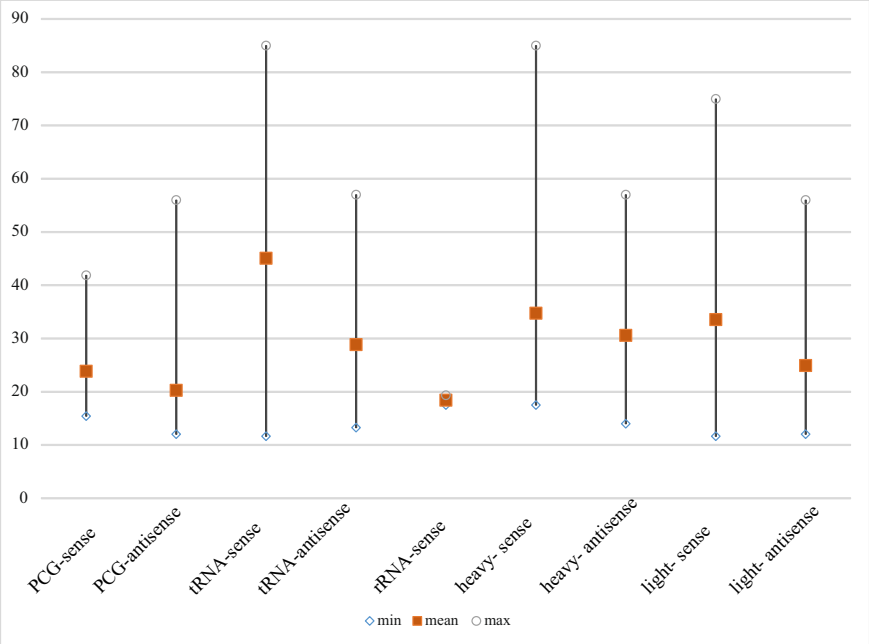
Acknowledgements

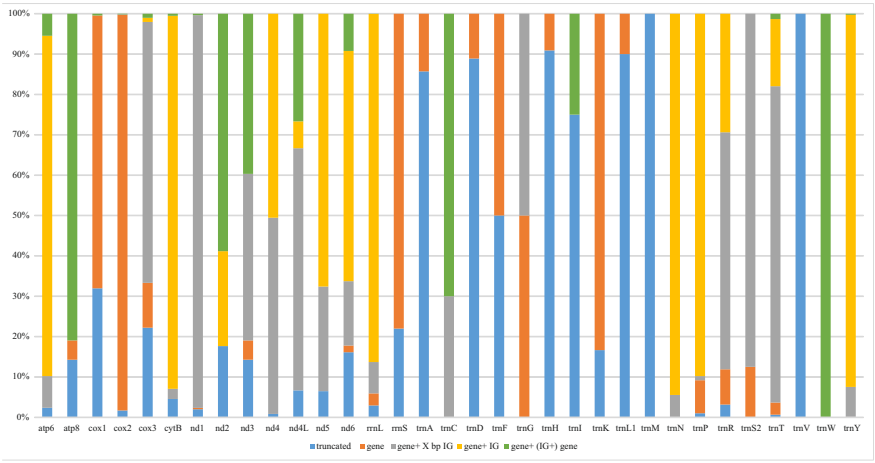
We thank Dr. Mahir Budak for his valuable help in every stages of the study. This study was financially supported by Sivas Cumhuriyet University (CÜBAP F-573).











Hosted file

`Table 1.xlsx` available at https://authorea.com/users/605653/articles/635051-polyadenylation-based-mitotranscriptomics-of-apis-mellifera-insecta-hymenoptera