

Fish Aquat Sci. 2022;25(3):158-166 https://doi.org/10.47853/FAS.2022.e14



The mitochondrial genome of *Tremoctopus violaceus* (Octopoda, Tremoctopodidae) and its phylogenetic consideration

Dae-Ju Oh, Jong-Chul Lee, Yong-Hwan Jung*

Biodiversity Research Institute, Jeju Technopark, Jeju 63208, Korea

Abstract

The complete mitochondrial genome of *Tremoctopus violaceus* was sequenced to analyze its organization and phylogenetic status within the order Octopoda. The mitochondrial genome of *T. violaceus* had a structure and organization similar to that of other Octopoda. The content of the nucleotides A, C, G, and T was 31.68 %, 7.71 %, 20.02 %, and 40.58 %, respectively. All protein-coding genes (PCG) began with the ATG codon, excluding *ND4* and *ATP6*, which began with ATC and ATT, respectively, and terminated with TAG, TAA, TA, or T. Codons for isoleucine were the most used codons, whereas those for arginine were used the least. Two extra tRNAs, trnN and trnL, were found in the control region. These tRNAs have a D-armless structure. The control region had excess A + T content (83.16 %) and a stem-loop structure with two elements, which is reported for the first time in Octopoda by our study. Bayesian inference using 13 PCG revealed that *Octopus* and Octopodidae were polyphyletic, and that Tremoctopodidae diverged relatively earlier within Octopoda. The mitochondrial genome of *T. violaceus* and its characteristics may help to understand the evolutionary history of Octopoda and establish a marine biodiversity conservation strategy.

Keywords: Mitochondrial genome, Octopoda, Phylogeny, *Tremoctopus*

Introduction

The order Octopoda has two main distinguished suborders, Cirrata and Incirrata. Cirrata possess rows of cirri, while Incirrata are devoid of cirri. The Incirrata suborder consists of two superfamilies: Argonautoidea and Octopodoidea. *Tremoctopus* is a genus of Tremoctopodidae, which belongs to the Argonautoidea superfamily. *Tremoctopus* is commonly known as the

blanket *octopus* and contains four species (Finn, 2016): *Tremoctopus violaceus*, *Tremoctopus gracilis*, *Tremoctopus robsoni*, and *Tremoctopus gelatus*. *T. violaceus* was previously classified into two subspecies: *Tremoctopus violaceus violaceus* and *Tremoctopus violaceus gracilis* (Thomas, 1977). However, the two subspecies are now treated as distinct species, *T. violaceus* and *T. gracilis* (Finn, 2016).

In East Asia, the presence of *T. violaceus* has been recorded

Received: Nov 29, 2021 Revised: Jan 17, 2022 Accepted: Jan 19, 2022

*Corresponding author: Yong-Hwan Jung

Biodiversity Research Institute, Jeju Technopark, Jeju 63208, Korea Tel: +82-64-720-2802, Fax: +82-64-720-2801, E-mail: yhjung@jejutp.or.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. Copyright © 2022 The Korean Society of Fisheries and Aquatic Science

158 https://www.e-fas.org

in the coastal waters of Japan and the nearby South China Sea (Georeferenced records in GBIF, https://www.gbif.org), as well as on the shore of Samcheok-si, Kangwon-do, South Korea in 2019. Specimens of *T. violaceus* are very rarely collected in East Asia; therefore, its distribution status, dispersing history, and phylogenetic relationships with populations found in other regions are unclear.

Ever since the mitochondrial (mt) genome of Octopus vulgaris was first reported (Yokobori et al., 2004), complete mt genomes of over 20 Octopoda species have been reported. Most animal mt genomes generally have a compact size (14-19 kb), circular form, and comprise 13 protein-coding genes (PCGs), 22 tRNAs, two rRNAs, and a noncoding control region (Boore, 1999). The mt genome has been used for identifying species, tracking evolutionary history, and making inferences related to comparative genomics from nucleotide composition, structural features, and gene rearrangement, etc. (Boore, 1999; Kumazawa et al., 1996; Li et al., 2016; Yu et al., 2019). Molecular phylogeny performed using mt genomes is useful for resolving controversial evolutionary relationships among various animal groups. Uribe & Zardoya (2017) revisited the phylogeny of Cephalopoda using complete mt genomes, and suggested that Octopus and Octopodidae are polyphyletic, as reported in previous studies (Cheng et al., 2013; Magallón-Gayón et al., 2020).

Recently, the importance of biodiversity has led to world-wide attention because biodiversity affects ecosystem functions (Baert et al., 2018; Isbell et al., 2018; O'Connor et al. 2017). Biodiversity is highly threatened by climate change, therefore, many advanced countries are making a lot of efforts to preserve global biodiversity through the Nagoya protocol, etc. Research on biodiversity begins with identifying species, and genetic analysis methods are widely used as a tool to identify species and analyze phylogenetic relationships (Chen & Wang, 2021; Yi et al., 2021).

This study presents the sequencing of the complete mt genome of *T. violaceus* and the description of the *T. violaceus* mt genome organization and phylogenetic analysis using the mt 13 PCG sequences that are generally found in species of the Octopoda.

Materials and Methods

DNA extraction and mitochondrial genome sequencing

Our study was performed using a *T. violaceus* specimen caught by a fisherman in the sea near the Jeju east pier, and donated for

research. Genomic DNA was extracted from a small piece of muscle using a Nucleospin Tissue Kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's protocol. The complete mt genome was amplified and sequenced using newly designed primers, including nested primers for primer walking sequencing. All sequencing reactions were conducted at Bionics (Seoul, Korea).

Gene annotation and sequence analysis

The complete mt genome was annotated using MITOS (Bernt et al., 2013) and ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder) using the invertebrate mitochondrial code. The complete mt genome map was drawn using SnapGene 5 software. The codon usage of 13 PCGs was analyzed using MEGA X (Kumar et al., 2018). The tRNA genes were identified using webbased tRNAscan-SE software (Lowe & Eddy, 1997) and ARW-EN (Laslett & Canbäck, 2008).

Phylogenetic analysis

To reconstruct the phylogenetic tree within the Octopoda, previously reported mt genomes of species in Octopoda were collected from the GenBank database and 13 PCG sequences were aligned using Clustal X (Thompson et al., 1997) with default settings. Bayesian inference (BI) performed using MrBayes 3.2 software (Ronquist et al., 2012) was used for phylogenetic tree reconstruction. The GTR + G + I model was selected as the best evolutionary model using jmodeltest2 (Darriba et al., 2012). Four Markov chains were run for 100,000 generations and sampled every 100 generations to obtain a posterior probability (PP) distribution of 1,000 trees. *Vampyroteuthis infernalis* (AB266515) was used as the outgroup (Uribe & Zardoya, 2017).

Results

Mitochondrial genome organization

The complete mt genome sequence of T. violaceus was 16,191 bp long (GenBank accession No. MZ043857). The content of A, C, G, and T was 31.68 % (5,130/16,191), 7.71 % (1,248/16,191), 20.02 % (3,242/16,191), and 40.58 % (6,571/16,191), respectively. The mt gene order was identical to that of other Octopoda species.

As observed in other animals, the mt genome of *T. violaceus* has 13 PCGs encoded on the heavy and the light strand. NADH dehydrogenase subunit 1 (*ND1*), *ND6*, cytochrome *b*, *ND4L*, *ND4*, and *ND5* were encoded on the heavy strand and



the other genes were encoded on the light strand (Fig. 1). On the heavy strand, *ND6* and *CytB* genes overlapped by 8 nucleotides, while ATPase subunit 8 and 6 genes overlapped by eight nucleotides. On the light strand, cytochrome c oxidase subunit (CO) 1 and *ND2* genes overlapped by 23 nucleotides. All PCGs

were initiated with the ATG codon, excluding *ND4* and *ATP6*, that were initiated with ATC and ATT, respectively, and terminated with TAG, TAA, TA, or T (Table 1). Among the 13 PCGs, only two PCGs, *ND1*, and *ND5*, terminated with the incomplete stop codons, T and TA, respectively.

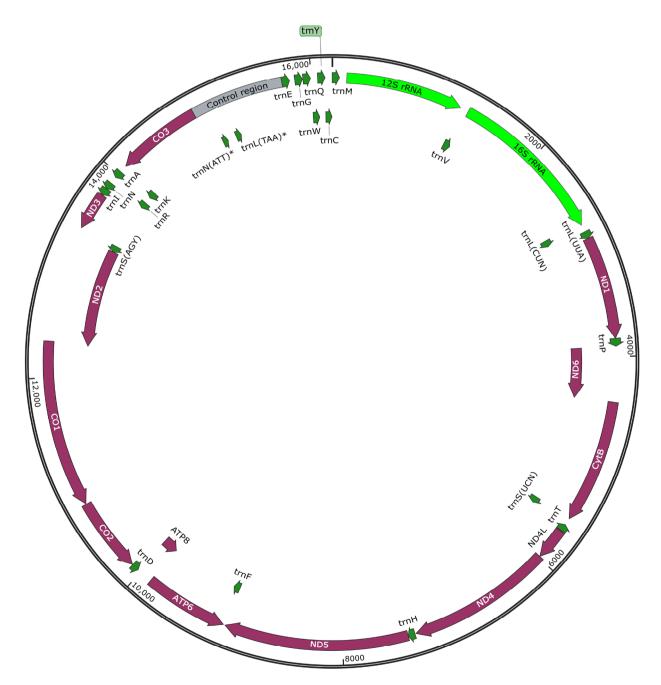


Fig. 1. The complete mitochondrial genome of *Tremoctopus violaceus***.** ND1-6, NADH dehydrogenase subunits 1-6; CO1-3, cytochrome *c* oxidase subunits 1-3; ATP8 and 6, ATPase subunit 8 and 6; CytB, cytochrome *b*; 12S and 16S, 12S and 16S ribosomal RNA.

Table 1. Location of feature of Tremoctopus violaceus mt genome

Features Location Length Start Stop trnM 1 67 67	genome					
Noncoding region-1 68 133 66 I	Features	Location		Length	Start	Stop
125 rRNA 134 1,210 1,077 trnV 1,220 1,288 69 respect to the control of the control region -2 1,287 2,758 1,472 trnL(UN) 2,271 2,785 65 trnL(UN) 2,824 39 trnL(UN) 2,825 2,890 66 trnL(UN) 2,825 2,890 66 trnL(UN) 1,070 4,823 339 949 ATG T 7 1,000 1,000 7 1,000 <td>trnM</td> <td>1</td> <td>67</td> <td>67</td> <td>,</td> <td></td>	trnM	1	67	67	,	
trnV 1,220 1,288 69 Head Head 165 rRNA 1,287 2,758 1,472 Head	Noncoding region-1	68	133	66		
16S rRNA 1,287 2,758 1,472 trnL(CUN) 2,721 2,785 65 Noncoding region-2 2,786 2,824 39 trnL(UUA) 2,825 2,890 66 ND1 2,891 3,839 949 ATG T trnP 3,840 3,909 70 TAG TAG ND6 3,911 4,423 513 ATG TAG CytB 4,416 5,555 1,140 ATG TAA trnS(UCN) 5,554 5,618 65 TAG trnT 5,688 5,984 297 ATG TAG ND4L 5,987 7,324 1,338 ATC TAG trnH 7,335 7,398 64 TAG TAG trnF 9,092 9,156 65 ATG TAA trnF 9,092 9,156 65 ATG TAA trnD 10,032 10,099 68	12S rRNA	134	1,210	1,077		
trnL(CUN) 2,721 2,786 65 Noncoding region-2 2,786 2,824 39 trnL(UUA) 2,825 2,890 66 ND1 2,891 3,839 949 ATG T tmP 3,840 3,909 70 TAG TAG ND6 3,911 4,423 513 ATG TAG CytB 4,416 5,555 1,140 ATG TAA trnS(UCN) 5,554 5,618 65 TAG TAG ND4L 5,688 5,984 297 ATG TAG ND4L 5,987 7,324 1,338 ATC TAG trnH 7,335 7,398 64 TAG TAG ND5 7,399 9,092 1,694 ATG TAG ATP6 9,109 9,882 774 ATG TAA CO2 10,104 10,790 68 ATG TAA ND2 12,30	trnV	1,220	1,288	69		
Noncoding region-2 2,786 2,824 39 trnL(UUA) 2,825 2,890 66 ND1 2,891 3,839 949 ATG T tmP 3,840 3,909 70 TAG TAG ND6 3,911 4,423 513 ATG TAG CytB 4,416 5,555 1,140 ATG TAA trnS(UCN) 5,554 5,618 65 TAG TAG ND4L 5,688 5,984 297 ATG TAG ND4L 5,987 7,324 1,338 ATC TAG ND5 7,399 9,092 1,694 ATG TA ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA CO2 10,104 10,790 687 ATG TAA TMD2 12,306 13,340 1,035 ATG TAA	16S rRNA	1,287	2,758	1,472		
trnL(UUA) 2,825 2,890 66 ND1 2,891 3,839 949 ATG T trnP 3,840 3,909 70 TAG TAG ND6 3,911 4,423 513 ATG TAG CytB 4,416 5,555 1,140 ATG TAG trnS(UCN) 5,554 5,618 65 TTAG TAG trnT 5,619 5,683 65 TTAG TAG ND4L 5,987 7,324 1,338 ATC TAG ND5 7,399 9,092 1,694 ATG TA trnF 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA CO2 10,104 10,790 687 ATG TAA CO3 11,794 13,340 1,333 ATG TAA trnS(AGY) 13,341 13,409 69 TAA <td>trnL(CUN)</td> <td>2,721</td> <td>2,785</td> <td>65</td> <td></td> <td></td>	trnL(CUN)	2,721	2,785	65		
ND1 2,891 3,839 949 ATG T tmP 3,840 3,909 70 TAG ND6 3,911 4,423 513 ATG TAG CytB 4,416 5,555 1,140 ATG TAA tmT 5,619 5,683 65 TAG ND4L 5,688 5,984 297 ATG TAG ND4L 5,688 5,984 297 ATG TAG ND4L 5,987 7,324 1,338 ATC TAG ND5 7,399 9,092 1,694 ATG TA tmF 9,092 9,156 65 TA ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA CO2 10,104 10,790 687 ATG TAA ND2 12,336 13,340 1,035 ATG TAA tmS(AGY	Noncoding region-2	2,786	2,824	39		
trnP 3,840 3,909 70 ND6 3,911 4,423 513 ATG TAG CytB 4,416 5,555 1,140 ATG TAA trnS(UCN) 5,554 5,618 65 TAG TAG ND4L 5,688 5,984 297 ATG TAG ND4L 5,987 7,324 1,338 ATC TAG ND4L 5,987 7,324 1,338 ATG TAG ND5 7,399 9,092 1,694 ATG TAG trnH 7,335 7,398 64 TAG TAG trnF 9,092 9,156 65 TAG TAG ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA CO2 10,104 10,790 687 ATG TAA ND2 12,306 13,340 1,035 ATG	trnL(UUA)	2,825	2,890	66		
ND6 3,911 4,423 513 ATG TAA CytB 4,416 5,555 1,140 ATG TAA trnS(UCN) 5,554 5,618 65 TTA trnT 5,619 5,683 65 TAG ND4L 5,987 7,324 1,338 ATC TAG ND5 7,399 9,092 1,694 ATG TA ND5 7,399 9,092 1,694 ATG TA ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA CO2 10,104 10,790 68 TAA CO1 10,796 12,328 1,533 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA trnI 13,894 13,958 65 TAA trnI <td>ND1</td> <td>2,891</td> <td>3,839</td> <td>949</td> <td>ATG</td> <td>T</td>	ND1	2,891	3,839	949	ATG	T
CytB 4,416 5,555 1,140 ATG TAA trmS(UCN) 5,554 5,618 65 TTA trmT 5,619 5,683 65 TAG ND4L 5,688 5,984 297 ATG TAG ND4L 5,987 7,324 1,338 ATC TAG trnH 7,335 7,398 64 TAG trnH 7,339 9,092 1,694 ATG TA trnF 9,092 9,156 65 ATG TAA ATP6 9,109 9,882 774 ATG TAA ATP8 9,875 10,030 156 ATG TAA CO2 10,104 10,790 687 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA trnI 13,380 13,894 65 TAA trnR	trnP	3,840	3,909	70		
trnS(UCN) 5,554 5,618 65 trnT 5,619 5,683 65 ND4L 5,688 5,984 297 ATG TAG ND4L 5,987 7,324 1,338 ATC TAG trnH 7,335 7,398 64 ATG TA ND5 7,399 9,092 1,694 ATG TA trnF 9,092 9,156 65 ATG TA ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA trnD 10,032 10,099 68 TAA CO1 10,796 12,328 1,533 ATG TAA trnS(AGY) 13,341 13,409 69 TAA trnS(AGY) 13,348 13,758 351 ATG TAA trnN 13,894 13,894 65 TAA trnR 13,894 13,9	ND6	3,911	4,423	513	ATG	TAG
trnT 5,619 5,683 65 ND4L 5,688 5,984 297 ATG TAG ND4L 5,987 7,324 1,338 ATC TAG trnH 7,335 7,398 64 TA ND5 7,399 9,092 1,694 ATG TA trnF 9,092 9,156 65 ATT TAG ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA trnD 10,032 10,099 68 TAA CO2 10,104 10,790 687 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA trnI 13,759 13,829 71 TAA trnR 13,894 13,958 65 TAA trnA 14,028 14,029	CytB	4,416	5,555	1,140	ATG	TAA
ND4L 5,688 5,984 297 ATG TAG ND4L 5,987 7,324 1,338 ATC TAG trnH 7,335 7,398 64 TA ND5 7,399 9,092 1,694 ATG TA trnF 9,092 9,156 65 ATT TAG ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA trnD 10,032 10,099 68 TAG TAA CO2 10,104 10,790 687 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA trnI 13,759 13,829 71 TAA trnR 13,894 13,958 65 TAA trnR 14,028 14,029 70 TAA CO3	trnS(UCN)	5,554	5,618	65		
ND4L 5,987 7,324 1,338 ATC TAG trnH 7,335 7,398 64 ATG TA ND5 7,399 9,092 1,694 ATG TA trnF 9,092 9,156 65 ATT TAG ATP8 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA trnD 10,032 10,099 68 ATG TAA CO2 10,104 10,790 687 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 69 ATG TAA trnI 13,759 13,829 71 71 TAA TAA TAA TAA TAM	trnT	5,619	5,683	65		
trnH 7,335 7,398 64 ND5 7,399 9,092 1,694 ATG TA trnF 9,092 9,156 65	ND4L	5,688	5,984	297	ATG	TAG
ND5 7,399 9,092 1,694 ATG TA trnF 9,092 9,156 65 ATT TAG ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA trnD 10,032 10,099 68 ATG TAA CO2 10,104 10,790 687 ATG TAA CO1 10,796 12,328 1,533 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA ND3 13,408 13,758 351 ATG TAA trnI 13,759 13,829 71 TTAA trnR 13,894 13,958 65 TTAA trnA 14,028 14,097 70 trnK 14,098 14,877 780 ATG TAA CO3 <td>ND4L</td> <td>5,987</td> <td>7,324</td> <td>1,338</td> <td>ATC</td> <td>TAG</td>	ND4L	5,987	7,324	1,338	ATC	TAG
trnF 9,092 9,156 65 ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA trnD 10,032 10,099 68 ATG TAA CO2 10,104 10,790 687 ATG TAA ND2 12,306 13,340 1,035 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 ATG TAA trnI 13,759 13,829 71 TAA trnN 13,830 13,894 65 TAA trnR 13,894 13,958 65 TAA trnK 14,028 14,029 70 TAA CO3 14,878 15,732 855 TAA Control region 14,878 15,732 855 TAA Noncoding region-3 15,804 15,851 48 trnQ 15,927 15,994 68 <tr< td=""><td>trnH</td><td>7,335</td><td>7,398</td><td>64</td><td></td><td></td></tr<>	trnH	7,335	7,398	64		
ATP6 9,109 9,882 774 ATT TAG ATP8 9,875 10,030 156 ATG TAA trnD 10,032 10,099 68 ATG TAA CO2 10,104 10,790 687 ATG TAA CO1 10,796 12,328 1,533 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA ND3 13,408 13,758 351 ATG TAA trnI 13,8759 13,829 71 TAA trnR 13,894 13,958 65 5 trnA 13,960 14,029 70 70 trnK 14,028 14,097 70 70 CO3 14,098 14,877 780 ATG TAA Noncoding region-3 15,804 15,851 48 trnG 15,927 15,994 68 trnQ 15,995 16,061 67	ND5	7,399	9,092	1,694	ATG	TA
ATP8 9,875 10,030 156 ATG TAA trnD 10,032 10,099 68 TAA CO2 10,104 10,790 687 ATG TAA CO1 10,796 12,328 1,533 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA ND3 13,408 13,758 351 ATG TAA trnI 13,759 13,829 71 TAA TAA trnR 13,894 13,958 65 TAA TAA trnA 13,960 14,029 70 TAA TAA CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 TAA trnE 15,733 15,804 15,851 48 trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnW 16,060 16,127	trnF	9,092	9,156	65		
trnD 10,032 10,099 68 CO2 10,104 10,790 687 ATG TAA CO1 10,796 12,328 1,533 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA ND3 13,408 13,758 351 ATG TAA trnI 13,759 13,829 71 TAA trnR 13,894 13,958 65 TAA trnA 13,960 14,029 70 TAA CO3 14,098 14,097 70 TAA Control region 14,878 15,732 855 TAA Noncoding region-3 15,804 15,851 48 48 trnQ 15,927 15,994 68 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68 68	ATP6	9,109	9,882	774	ATT	TAG
CO2 10,104 10,790 687 ATG TAA CO1 10,796 12,328 1,533 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA ND3 13,408 13,758 351 ATG TAA trnI 13,759 13,829 71 71 TTAA trnR 13,894 13,958 65 5 5 trnA 13,960 14,029 70 70 70 70 CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 71 trnE 15,733 15,803 71 71 Noncoding region-3 15,804 15,851 48 48 trnQ 15,927 15,994 68 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	ATP8	9,875	10,030	156	ATG	TAA
CO1 10,796 12,328 1,533 ATG TAA ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA ND3 13,408 13,758 351 ATG TAA trnI 13,759 13,829 71 TAA trnN 13,830 13,894 65 TAA trnA 13,960 14,029 70 TAA trnK 14,028 14,097 70 TAA CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 TAA Noncoding region-3 15,804 15,851 48 TAA trnQ 15,927 15,994 68 TAA trnW 15,995 16,061 67 trnY 16,060 16,127 68	trnD	10,032	10,099	68		
ND2 12,306 13,340 1,035 ATG TAA trnS(AGY) 13,341 13,409 69 TAA ND3 13,408 13,758 351 ATG TAA trnI 13,759 13,829 71 70 <t< td=""><td>CO2</td><td>10,104</td><td>10,790</td><td>687</td><td>ATG</td><td>TAA</td></t<>	CO2	10,104	10,790	687	ATG	TAA
trnS(AGY) 13,341 13,409 69 ND3 13,408 13,758 351 ATG TAA trnI 13,759 13,829 71 trnN 13,830 13,894 65 trnR 13,960 14,029 70 trnK 14,028 14,097 70 CO3 14,878 15,732 855 trnE 15,733 15,803 71 Noncoding region-3 15,804 15,851 48 trnQ 15,927 15,921 70 trnQ 15,995 16,061 67 trnY 16,060 16,127 68	CO1	10,796	12,328	1,533	ATG	TAA
ND3 13,408 13,758 351 ATG TAA trnI 13,759 13,829 71 71 trnN 13,830 13,894 65 65 trnA 13,960 14,029 70 70 trnK 14,028 14,097 70 70 CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 48 trnE 15,831 15,803 71 70 70 70 trnG 15,852 15,921 70	ND2	12,306	13,340	1,035	ATG	TAA
trnl 13,759 13,829 71 trnN 13,830 13,894 65 trnR 13,894 13,958 65 trnA 13,960 14,029 70 trnK 14,028 14,097 70 CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 trnE 15,733 15,803 71 Noncoding region-3 15,854 15,951 48 trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	trnS(AGY)	13,341	13,409	69		
trnN 13,830 13,894 65 trnR 13,894 13,958 65 trnA 13,960 14,029 70 trnK 14,028 14,097 70 CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 trnE 15,733 15,803 71 Noncoding region-3 15,804 15,851 48 trnG 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	ND3	13,408	13,758	351	ATG	TAA
trnR 13,894 13,958 65 trnA 13,960 14,029 70 trnK 14,028 14,097 70 CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 trnE 15,733 15,803 71 Noncoding region-3 15,804 15,851 48 trnG 15,927 15,991 70 trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	trnl	13,759	13,829	71		
trnA 13,960 14,029 70 trnK 14,028 14,097 70 CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 trnE 15,733 15,803 71 Noncoding region-3 15,854 15,851 48 trnG 15,852 15,921 70 trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	trnN	13,830	13,894	65		
trnK 14,028 14,097 70 CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 trnE 15,733 15,803 71 Noncoding region-3 15,804 15,851 48 trnG 15,852 15,921 70 trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	trnR	13,894	13,958	65		
CO3 14,098 14,877 780 ATG TAA Control region 14,878 15,732 855 TAA trnE 15,733 15,803 71 71 Noncoding region-3 15,804 15,851 48 48 trnG 15,852 15,921 70 70 trnQ 15,927 15,994 68 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	trnA	13,960	14,029	70		
Control region 14,878 15,732 855 trnE 15,733 15,803 71 Noncoding region-3 15,804 15,851 48 trnG 15,852 15,921 70 trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	trnK	14,028	14,097	70		
trnE 15,733 15,803 71 Noncoding region-3 15,804 15,851 48 trnG 15,852 15,921 70 trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	CO3	14,098	14,877	780	ATG	TAA
Noncoding region-3 15,804 15,851 48 trnG 15,852 15,921 70 trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	Control region	14,878	15,732	855		
trnG 15,852 15,921 70 trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	trnE	15,733	15,803	71		
trnQ 15,927 15,994 68 trnW 15,995 16,061 67 trnY 16,060 16,127 68	Noncoding region-3	15,804	15,851	48		
trnW 15,995 16,061 67 trnY 16,060 16,127 68	trnG	15,852	15,921	70		
trnY 16,060 16,127 68	trnQ	15,927	15,994	68		
	trnW	15,995	16,061	67		
trnC 16,126 16,190 65	trnY	16,060	16,127	68		
	trnC	16,126	16,190	65		

The total number of codons used in the 13 PCGs was 3,750, wherein codons corresponding to trnI were the most used (257 times), while codons for trnR were used the least (three times) (Table 2). Among the 3,750 codons, the proportion of codons for trnL was 15.2 % (571/3,750, the most used codon), whereas that for trnR was 1.3 % (50/3,750, the least used codon), trnL and trnS were coded by six and eight different codons, respectively. trnA, trnR, trnG, trnP, trnT, and trnV were coded by four different codons, while the other sequences were coded by two different codons.

The 12S and 16S rRNA genes of the T. violaceus mt genome were 1,077 bp and 1,472 bp in length, respectively (Table 1). As observed in other Octopus genomes, these two rRNA genes were located between trnM and trnL, and were separated by trnV.

The mt genome of *T. violaceus* was seen to have 24 tRNAs, whereas the mt genomes of most animals have 22 tRNAs. The two extra tRNAs, found located in the control region, were trnN and trnL (Fig. 1). The secondary structure of both additional tRNAs showed an abnormal cloverleaf structure without the D-arm (Fig. 2).

The mt genome of *T. violaceus* had four noncoding regions, including the control region known as the D-loop. The control region, a major noncoding region, was 855 bp long and located

Table 2. Codon usage of 13 protein-coding genes of Tremoctopus violaceus mt genome

Codon	Count	Codon	Count	Codon	Count	Codon	Count
UUU(F)	287	UCU(S)	92	UAU(Y)	128	UGU(C)	57
UUC(F)	53	UCC(S)	46	UAC(Y)	38	UGC(C)	8
UUA(L)	301	UCA(S)	81	UAA(*)	9	UGA(W)	80
UUG(L)	100	UCG(S)	11	UAG(*)	4	UGG(W)	20
CUU(L)	51	CCU(P)	72	CAU(H)	47	CGU(R)	19
CUC(L)	28	CCC(P)	23	CAC(H)	36	CGC(R)	3
CUA(L)	80	CCA(P)	28	CAA(Q)	42	CGA(R)	25
CUG(L)	11	CCG(P)	6	CAG(Q)	15	CGG(R)	3
AUU(I)	257	ACU(T)	57	AAU(N)	123	AGU(S)	42
AUC(I)	77	ACC(T)	29	AAC(N)	51	AGC(S)	9
AUA(M)	191	ACA(T)	58	AAA(K)	64	AGA(S)	56
AUG(M)	76	ACG(T)	5	AAG(K)	33	AGG(S)	50
GUU(V)	110	GCU(A)	55	GAU(D)	59	GGU(G)	73
GUC(V)	11	GCC(A)	19	GAC(D)	14	GGC(G)	15
GUA(V)	78	GCA(A)	36	GAA(E)	53	GGA(G)	99
GUG(V)	56	GCG(A)	10	GAG(E)	25	GGG(G)	55

^{*} Asterisk indicates termination codon.



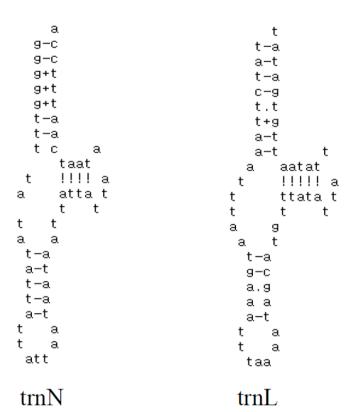


Fig. 2. Secondary structures of two extra putative tRNAs. trnN and trnL have anticodons corresponding to the codons AAU and UUA, respectively. Both extra tRNAs have abnormal structures in the D-arm.

between CO3 and trnE, while other noncoding regions — 1, 2, and 3 — were 66 bp, 39 bp, and 48 bp, respectively (Table 1). Similar to the mt genomes in other animals, the control region of the *T. violaceus* mt genome showed high A + T content (83.16 %), and had a stem-loop structure flanked by two elements (TATA motif and $G(A)_nT$ motif) (Fig. 3).

Phylogenetic analysis

In this study, nine genera from four families registered in the GenBank database were analyzed for obtaining Octopoda phylogeny. BI using 13 PCGs showed that the Tremoctopodidae family was supported as a sister group to Argonautidae (Fig. 4). The branches of Tremoctopodidae and Argonautidae were located in the basal position of the Incirrata clade. The phylogenetic tree showed *Octopus* and Octopodidae to be polyphyletic. *Octopus* species were seen to be dispersed around the phylogenetic tree and the Octopodidae family was divided into two clades, Octopodidae-I and Octopodidae-II. Octopodidae-II was

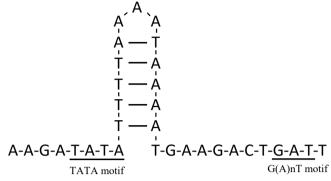


Fig. 3. Predicted stem-loop structure with the TATA motif and $G(A)_nT$ motif in the control region.

the first to branch from the Incirrata clade, and Tremoctopodidae and Argonautidae branched later.

Discussion

Mitochondrial genome of T. violaceus

The general features of the mt genome were identical to those of other Octopus mt genomes. ND6 and CytB genes, as well as the ATPase subunit 8 and 6 genes had an overlap of eight nucleotides, while CO1 and ND2 genes had an overlap of 23 nucleotides. These overlapping features are common and have been found in the mt genomes of other octopuses (Cheng et al., 2013; Chiu et al., 2018; Tang et al., 2018). The alternative initiation codons, AUA, AUU, ATC, GUG, and UUG have been reported in invertebrate mt genomes (Donath et al., 2019). Among these, ATC and ATT were used in ND4 and ATP6, respectively, in the T. violaceus mt genome. Meanwhile, incomplete termination codons, T and TA, were used in ND1 and ND5, respectively. Incomplete termination codons are very common in various animal mt genes, including those of mammals, fishes, reptiles, amphibians, and invertebrates (Cui et al., 2007; Kumazawa et al., 1996; Oh & Jung, 2019; Oh et al., 2019; Sumida et al., 2001). Incomplete termination codons are said to be completed via posttranscriptional polyadenylation (Ojala et al., 1981).

Transfer RNA genes

Although an animal mt genome normally has 22 tRNAs, the *T. violaceus* mt genome was seen to have 24 tRNAs. The additional tRNAs, trnN and trnL, were located in the noncoding control region (Fig. 1). Additional tRNAs in the Octopoda mt genome have been reported in *Cistopus chinensis* (Cheng et al., 2013).

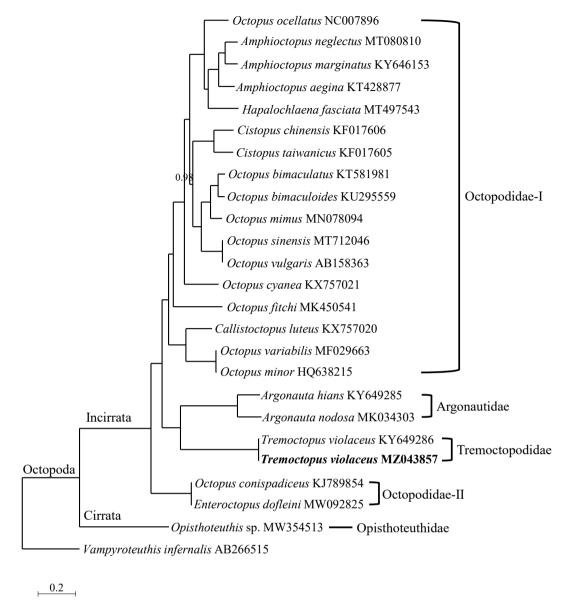


Fig. 4. Bayesian inference using 13 mitochondrial protein-coding genes in Octopoda. Number at node is the statistical support value for the BI analysis (1.0 support value was not shown). BI, Bayesian inference.

Both extra tRNAs have D-armless structures; abnormal secondary structures of some tRNAs in the mt genome have often been identified in various animals (Danic-Tchaleu et al., 2011; Jarošová et al., 2016; Morrison, 2010; Pons et al., 2019; Yokobori et al., 2005). Although the occurrence of extra tRNAs, and their functional mechanisms, in the animal mt genome remain unclear, they might share the same amino acid with conventional tRNAs, as observed in the ascidian mt genome (Kondow et al., 1999).

Control region

Four noncoding regions, including the control region, were identified in the mt genome of *T. violaceus*. Meaningful structural features were absent in the three noncoding regions other than the control region. Generally, the control region is believed to be involved in DNA replication and transcriptional regulation (Clayton, 1982). The control region of the *T. violaceus* mt genome has a stem-loop structure and two elements that are presumed to be associated with the processes of replication



and transcriptional regulation (Wei et al., 2010). Although the TATA motif and $G(A)_nT$ motif have been reported in other invertebrates (Kuhn et al., 2008; Schultheis et al., 2002), this study is the first to report these elements in Octopoda.

Phylogenetic status of Octopoda

The Tremoctopodidae, which was well supported as a sister group to the Argonautidae, was seen to consist only one genus, *Tremoctopus*, and the branches of both families were located at the basal position in Incirrata (Fig. 4). Consequently, *Tremoctopus* was considered to have diverged relatively earlier within the Octopoda.

Our BI tree showed the polyphyly of *Octopus* and Octopodidae, which has been supported by molecular data from previous studies (Carlini et al., 2001; Cheng et al., 2013; Chiu et al., 2018; Magallón-Gayón et al., 2020; Uribe & Zardoya, 2017). Our BI tree showed Octopodidae to be divided into two clades (Fig. 4). Octopodidae-I consisted of all species in Octopodidae, excluding *Octopus conispadiceus* and *Enteroctopus dofleini*, which were found in Octopodidae-II. *Octopus ocellatus*, *Octopus variablilis*, *Octopus minor*, and *Octopus conispadiceus* require taxonomic reconsideration (Cheng et al., 2013; Uribe & Zardoya, 2017).

In this study, we described the complete mt genome of *T. violaceus* and showed its phylogenetic status. In this study, the two putative extra tRNAs that were found represented a rare feature observed in Octopoda, and, for the first time, a stemloop structure with two elements was reported in Octopoda. Phylogenetic analysis showed the polyphyly of *Octopus* and Octopodidae, and the relatively early divergence of the *Tremoctopus* genus. Our data provides useful information for studying comparative mt genomics and phylogenomics of *Tremoctopus*, and further facilitates evolutionary studies of Octopoda.

Genetic analysis is very useful as a tool to measure biodiversity levels in each ecosystem, and sufficient genetic data can help determine the priorities of biodiversity conservation organisms. The mt genome sequence is also used in many studies for speciation, gene flow, population dynamics, etc. Therefore, the mt genome will provide useful information for establishing biodiversity conservation strategies.

Competing interests

No potential conflict of interest relevant to this article was reported.

Funding sources

All authors thank the research funding of the Jeju Special Self-Governing Province.

Acknowledgements

Upon reasonable request, the datasets of this study can be available from the corresponding author.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

This article does not require IRB/IACUC approval because there are no human and animal participants.

ORCID

 Dae-Ju Oh
 https://orcid.org/0000-0003-2110-6341

 Jong-Chul Lee
 https://orcid.org/0000-0001-6344-3878

 Yong-Hwan Jung
 https://orcid.org/0000-0003-2606-8663

References

Baert JM, Eisenhauer N, Janssen CR, De Laender F. Biodiversity effects on ecosystem functioning respond unimodally to environmental stress. Ecol Lett. 2018;21:1191-9.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: improved *de novo* metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013;69:313-9.

Boore JL. Animal mitochondrial genomes. Nucleic Acids Res. 1999;27:1767-80.

Carlini DB, Young RE, Vecchione M. A molecular phylogeny of the Octopoda (Mollusca: Cephalopoda) evaluated in light of morphological evidence. Mol Phylogenet Evol. 2001;21:388-97.

Chen J, Wang W. Genetic diversity and genetic differentiation of *Megalobrama* populations inferred by mitochondrial markers. Genes Genomics. 2021;43:1119-32.

Cheng R, Zheng X, Ma Y, Li Q. The complete mitochondrial genomes of two octopods *Cistopus chinensis* and *Cistopus taiwanicus*: revealing the phylogenetic position of the genus *Cistopus* within the order Octopoda. PLOS ONE. 2013;8:e84216.

Chiu YW, Chang CW, Lin HD, Shen KN. The complete mitogenome of the winged argonaut *Argonauta hians* and its

- phylogenetic relationships in Octopoda. Conserv Genet Resour. 2018;10:359-62.
- Clayton DA. Replication of animal mitochondrial DNA. Cell. 1982;28:693-705.
- Cui P, Ji R, Ding F, Qi D, Gao H, Meng H, et al. A complete mitochondrial genome sequence of the wild two-humped camel (*Camelus bactrianus ferus*): an evolutionary history of camelidae. BMC Genomics. 2007;8:241.
- Danic-Tchaleu G, Heurtebise S, Morga B, Lapègue S. Complete mitochondrial DNA sequence of the European flat oyster *Ostrea edulis* confirms Ostreidae classification. BMC Res Notes. 2011;4:400.
- Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012;9:772.
- Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, et al. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. Nucleic Acids Res. 2019;47:10543-52.
- Finn JK. Family Tremoctopodidae. In: Jereb P, Roper CFE, Norman MD, Finn JK, editors. Cephalopods of the world: an annotated and illustrated catalogue of Cephalopod species known to date. Rome: FAO; 2016. p. 240-3.
- Isbell F, Cowles J, Dee LE, Loreau M, Reich PB, Gonzalez A, et al. Quantifying effects of biodiversity on ecosystem functioning across times and places. Ecol Lett. 2018;21:763-78.
- Jarošová A, Půža V, Žurovcová M. The complete mitochondrial genome of the facultative entomopathogenic nematode *Oscheius chongmingensis* (Rhabditida: Rhabditidae). Mitochondrial DNA A. 2016;27:3109-10.
- Kondow A, Suzuki T, Yokobori S, Ueda T, Watanabe K. An extra tRNA^{Gly}(U*CU) found in ascidian mitochondria responsible for decoding non-universal codons AGA/AGG as glycine. Nucleic Acids Res. 1999;27:2554-9.
- Kuhn K, Streit B, Schwenk K. Conservation of structural elements in the mitochondrial control region of *Daphnia*. Gene. 2008;420:107-12.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35:1547-9.
- Kumazawa Y, Ota H, Nishida M, Ozawa T. Gene rearrangements in snake mitochondrial genomes: highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster. Mol Biol Evol. 1996;13:1242-54.

- Laslett D, Canbäck B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics. 2008;24:172-5.
- Li T, Yang J, Li Y, Cui Y, Xie Q, Bu W, et al. A mitochondrial genome of Rhyparochromidae (Hemiptera: Heteroptera) and a comparative analysis of related mitochondrial genomes. Sci Rep. 2016;6:35175.
- Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25:955-64.
- Magallón-Gayón E, del Río-Portilla MÁ, de los Angeles Barriga-Sosa I. The complete mitochondrial genomes of two octopods of the eastern Pacific Ocean: *Octopus mimus* and '*Octopus' fitchi* (Cephalopoda: Octopodidae) and their phylogenetic position within Octopoda. Mol Biol Rep. 2020;47:943-52.
- Morrison DA. How and where to look for tRNAs in Metazoan mitochondrial genomes, and what you might find when you get there. 2010. http://arxiv.org/abs/1001.3813
- O'Connor MI, Gonzalez A, Byrnes JEK, Cardinale BJ, Duffy JE, Gamfeldt L, et al. A general biodiversity–function relationship is mediated by trophic level. Oikos. 2017;126:18-31.
- Oh DJ, Jung YH. Mitochondrial genome of Japanese amberjack, Seriola quinqueradiata, and yellowtail amberjack, Seriola lalandi. Mitochondrial DNA B Resour. 2019;4:826-7.
- Oh DJ, Yang KS, Jung YH. The mitochondrial genome of the Jeju ground beetle *Carabus smaragdinus monilifer* (Coleoptera, Carabidae). Mitochondrial DNA B Resour. 2019;5:39-40.
- Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria. Nature. 1981;290:470-4.
- Pons J, Bover P, Bidegaray-Batista L, Arnedo MA. Arm-less mitochondrial tRNAs conserved for over 30 millions of years in spiders. BMC Genomics. 2019;20:665.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61:539-42.
- Schultheis AS, Weigt LA, Hendricks AC. Arrangement and structural conservation of the mitochondrial control region of two species of Plecoptera: utility of tandem repeat-containing regions in studies of population genetics and evolutionary history. Insect Mol Biol. 2002;11:605-10.
- Sumida M, Kanamori Y, Kaneda H, Kato Y, Nishioka M, Hasegawa M, et al. Complete nucleotide sequence and gene rearrangement of the mitochondrial genome of the Jap-



- anese pond frog *Rana nigromaculata*. Genes Genet Syst. 2001:76:311-25.
- Tang Y, Zheng X, Ma Y, Cheng R, Li Q. The complete mitochondrial genome of *Amphioctopus marginatus* (Cephalopoda: Octopodidae) and the exploration for the optimal DNA barcoding in Octopodidae. Conserv Genet Resour. 2018;10:115-8.
- Thomas RF. Systematics, distribution, and biology of cephalopods of the genus *Tremoctopus* (Octopoda: Tremoctopodidae). Bull Mar Sci. 1977;27:353-92.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 1997;25:4876-82.
- Uribe JE, Zardoya R. Revisiting the phylogeny of Cephalopoda using complete mitochondrial genomes. J Molluscan Stud. 2017;83:133-44.
- Wei SJ, Shi M, Chen XX, Sharkey MJ, van Achterberg C, Ye GY, et al. New views on strand asymmetry in insect mitochondrial genomes. PLOS ONE. 2010;5:e12708.
- Yi CH, Yoon M, Kim JM, Kim IH, Cho IY, An HS. Genetic analysis and population genetic structure of hard-shelled mussel, *Mytilus coruscus* Gould 1861 (Mytiloida: Mytilidae) from the coasts of South Korea based on mitochondrial cytochrome oxidase (COI) gene sequences. Genes Genomics. 2021;43:577-85.
- Yokobori S, Fukuda N, Nakamura M, Aoyama T, Oshima T. Long-term conservation of six duplicated structural genes in cephalopod mitochondrial genomes. Mol Biol Evol. 2004;21:2034-46.
- Yokobori S, Oshima T, Wada H. Complete nucleotide sequence of the mitochondrial genome of *Doliolum nationalis* with implications for evolution of urochordates. Mol Phylogenet Evol. 2005;34:273-83.
- Yu X, Tan W, Zhang H, Jiang W, Gao H, Wang W, et al. Characterization of the complete mitochondrial genome of *Harpalus sinicus* and its implications for phylogenetic analyses. Genes. 2019;10:724.