水產研究 Journal of Taiwan Fisheries Research, 20 (1): 1-16, 2012

The Complete Mitochondrial DNA Sequence of the Deepwater Stingray *Plesiobatis daviesi* (Wallace, 1967): Unique Features in the Mitochondrial D-loop Region

Sheng-Tai Hsiao[1], Chi-Hsin Hsu[2], Don-Chung Liu[3] and I-Shiung Chen[4*]

[1]Marine Fisheries Division, Fisheries Research Institute
[2]Institute of Marine Resources, National Sun Yat-sen University
[3]Fisheries Research Institute
[4]Institute of Marine Biology, National Taiwan Ocean University

ABSTRACT

The 17514-nucleotide mitochondrial DNA (mtDNA) sequence of the deepwater stingray (*Plesiobatis daviesi*) was determined using long PCR and primer walking methods. The deepwater stingray genome contains 37 genes, including 2 ribosomal RNAs, 22 transfer RNAs (tRNAs), and 13 protein-coding genes, which are similar to other vertebrates. A comparison of the deepwater stingray mtDNA sequence with that of seven completely sequenced chondrichthyan mtDNAs revealed an identical gene order. However, the major noncoding region of the deepwater stingray genome, a D-loop sequence between the tRNA^{Pro} and tRNA^{Phe}, is the longest (1830 bp) among all known available sequences of chondrichthyan species. The components of D-loop sequences are two major copies of 47-bp tandem repeats and two characteristic conserved sequence blocks in this region. Phylogenetic analyses suggest that the deepwater stingray belongs to the basal group of Myliobatoidei that forms a sister group with Hexatrygonidae.

Key words: mitochondrial genome, Long PCR, deepwater stingray, *Plesiobatis daviesi*

INTRODUCTION

Mitochondrial genomic sequences have been used as genetic markers for molecular approaches in the phylogenetic and evolutionary studies of fish (Miya *et al*., 2001).

The order Myliobatiformes is a monophyletic group comprised of stingrays, round rays, butterfly rays, and eagle rays. The family Plesiobatidae belongs to the order Myliobatiformes and includes only one monotypic species, *Plesiobatis daviesi*, which was formerly recognized as *Urotrygon daviesi* and was placed under the family Urolophidae. Several different hypotheses have been proposed regarding the classification of Plesiobatidae (Fig. 1). The family was first established by Nishida (1990) based on several unique feature (i.e. nasal curtain being incompletely united and not reaching the mouth, with large size and soft disc, the average total length was ~ 200 cm and the average disc width ~ 100 cm). Later, McEachran *et al.* (1996) described musculature and skeletal structure of batoid fishes and provided a new topology on the relationship of myliobatoids. Based on the study by McEachran and Aschliman (2004) on the diversity of claspers in batoids, *P. daviesi* was placed under the family Urolophidae. However, recently Nelson (2006) placed *Plesiobatis* under the

---

*Correspondence: National Taiwan Ocean University, 2 Pei-Ning Rd., Keelung 202, Taiwan. TEL: 886-2-2462-2192 ext. 5320; FAX: 886-2-2463-3152; E-mail: isc@mail.ntou.edu.tw

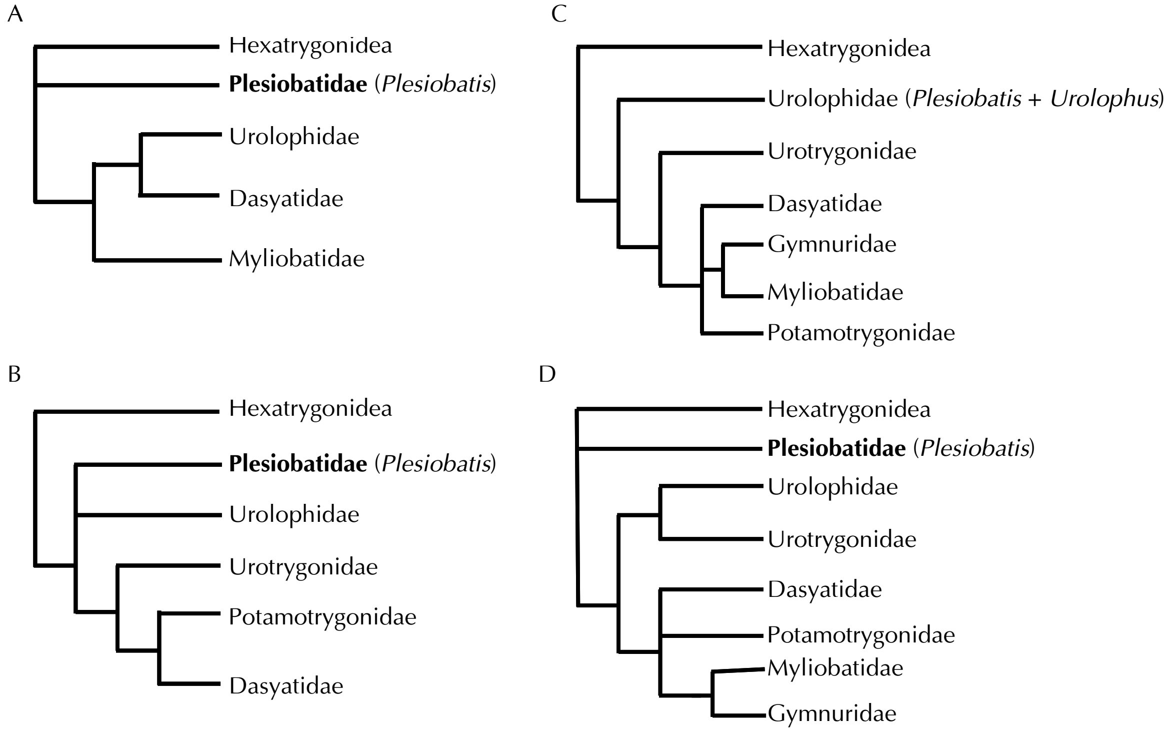

Fig. 1 The four most relevant hypotheses of phylogenetic relationship of myliobatoidei families. (A) Nishida's (1990) hypothesis based on morphological characters; (B) McEachran *et al.* (1996) proposed; (C) Phylogenetic relationship of myliobatoidei families, based on claspers characters, proposed by McEachran & Aschliman (2004); (D) The most recent hypothesis was made by Nelson (2006).

family Plesiobatidae. *Plesiobatis* is considered a primitive chondrichthyan. Therefore, in order to understand these taxa in terms of the evolutionary process among the myliobatoids, analysis of the complete mitochondrial genome sequence of the deepwater stingray would be beneficial.

Among chondrichthyans, complete mitochondrial DNA (mtDNA) sequences have been reported for only seven species: the small-spotted catshark *Scyliorhinus canicula* (Delarbre *et al.*, 1998), the starspotted smooth-hound *Mustelus manazo* (Cao *et al.*, 1998), the horn shark *Heterodontus francisci* (Arnason *et al.*, 2001), the spiny dogfish *Squalus acanthias* (Rasmussen and Arnason, 1999a), the thorny skate *Raja radiata* (Rasmussen and Arnason, 1999b), the ocellate spot skate *Raja porosa* (Kim *et al.*, 2005), and the rabbitfish *Chimaera monstrosa* (Arnason *et al.*, 2001). In contrast, there are no published mtDNA sequences on Myliobatiformes.

This study is to determine the complete mtDNA sequence of the deepwater stingray *P. daviesi* (Elasmobranchii: Myliobatiformes: Plesiobatidae) was determined using long polymerase chain reaction (PCR) and primer walking methods, which have been previously applied for teleost fish (Miya and Nishida, 1999; Miya and Nishida, 2000; Inoue *et al.*, 2000, 2001a, b, c; Kawaguchi *et al.*, 2001; Miya *et al.*, 2001). This is the first study to construct the complete mtDNA sequences of myliobatoid fish may provide an independent assessment of lineage evolution within Myliobatiformes to reconcile previous conflicting hypotheses of evolutionary relationships.

MATERIALS AND METHODS

1. Fish sample and DNA extraction

A deepwater stingray was captured in 2001 from Tungkang in southern Taiwan by bottom trawling at approximately 300 meters deep. Total

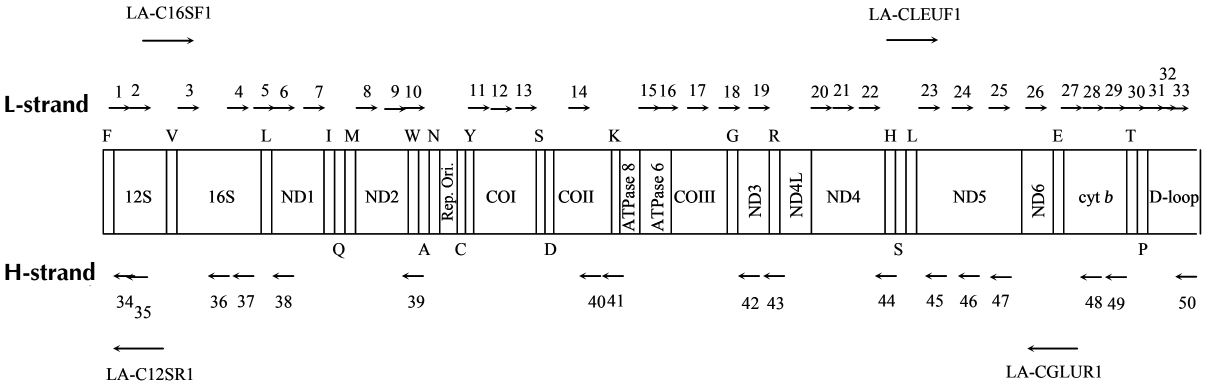

Fig. 2 Gene organization and sequencing strategy for the *Plesiobatis daviesi* mitochondrial genome. The circular genome is represented linearly.

DNA extraction was performed using the protocols of the High Pure PCR Template Preparation Kit (Roche Diagnostics, Indianapolis, Ind.). The voucher specimen was deposited in the collection of the National Museum of Marine Biology & Aquarium, Pingtung, Taiwan.

2. mtDNA amplification by long PCR

The complete mitochondrial genome of the deepwater stingray was amplified using a long PCR technique (Miya and Nishida, 1999, 2000). Two sets of primer pairs (first set, LA-C16SF1: GCCAACCC ACCTCTGTAGCAAAAGAGAGGGAAGACTCC; LA-CGLUR1: GCTAGGGCTAGTAATTTCTGCT GGGGTGGGTTGTGGTT; and second set, LA-CLEUF1: TTTCTCCCGCCCATGGTTCGAATCCC TGGCTCCCTTA; LA-C12SR1: CTCGTATAACC GCGGTGGCTGGCACGAGATTGACCAAC) were used to amplify the complete mitochondrial genome in two long PCR reactions.

Long PCR was performed in a Model 2400 or 9700 thermal cycler (Perkin-Elmer, Foster City, CA, USA), and reactions were performed with 30 cycles and a 25-μL reaction volume containing 15.25 μL of sterile distilled water, 2.5 μL of 10× LA PCR buffer II (Takara, Shiga, Japan), 4.0 μL of dNTPs (0.25 mM each), 1.0 μL of each primer (5 mM), 0.5 μL of 1.25-unit LA Taq (Takara), and 1.0 μL of template containing approximately 5 ng of DNA. The thermal cycle profile was that of "shuttle PCR," that is, denaturation at 98°C for 10 s and annealing and extension combined at the same temperature (68°C) for 16 min. Long PCR products were electrophoresed on a 1.0% agarose gel and then stained with ethidium bromide for band characterization via ultraviolet transillumination. The long PCR products were diluted with Tris-EDTA buffer for subsequent use as PCR templates.

3. Primer walking and sequencing

Fifty primers were used to amplify the contiguous and overlapping segments of the complete genome (Fig. 2). These primers were designed either from the obtained sequences of this species or from the aligned, complete nucleotide sequences of the mitochondrial genome of six species of cartilaginous fishes.

PCR was performed in a Model 2400 or 9700 thermal cycler, and reactions were performed with 30 cycles and a 25-μL reaction volume containing 14.4 μL of sterile distilled water, 2.5 μL of 10× PCR buffer (Perkin-Elmer), 2.0 μL of dNTPs (2.5 mM each), 2.5 μL of each primer (5 μM), 0.1 μL of 0.5-unit Ex Taq (Takara), and 1.0 μL of template. The thermal cycle profile was as follows: denaturation at 94°C for 15 s, annealing at 47–53°C for 15 s, and extension at 72°C for 30–60 s. PCR products were electrophoresed on a 1.0% agarose gel and then stained with ethidium bromide for band characterization via ultraviolet transillumination.

Double-stranded PCR products were purified using a High Pure PCR Product Purification Kit and were subsequently used for direct cycle sequencing with dye-labeled (BIG-DYE) terminators (ABI Perkin-Elmer). The primers used were the same as those for PCR. All sequencing reactions were performed according to the manufacturer's instructions. Labeled fragments were analyzed using a Model ABI 377 DNA Automated Sequencer (ABI Perkin-Elmer).

Table 1 Polymerase chain reaction (PCR) and sequencing primers in the analysis of the *Plesiobatis daviesi* mitochondrial genome

	Primers for L-strand		
	Name Long PCR primers	Position	Sequence (5′-3′)
	LA-C16SF1	L1417	GCCAA CCCAC CTCTG TAGCA AAAGA GAGGG AAGAC TCC
	LA-CLEUF1	L11856	TTTCT CCCGC CCATG GTTCG AATCC CTGGC TCCCT TA
	PCR and sequencing primers		
1	SK-12SF2	L199	TAGCC CAAGA CACCT TGCTA AG
2	P10AN	L501	CAAAC TGGGA TTAGA TACCC CA
3	SY164F	L1246	TACCG CAAGG GAAAG TTGAA A
4	SKH-16SF3	L2536	AAGGG TTTAC GACCT CGATG TT
5	SK-LEU-UURF1	L2815	CTGGT AADTG CAAAA GACCT AA
6	SK-ND1F1	L3285	GCMCT ACGNG CTGTA GCACA AAC
7	SK-ND1F2	L3793	TCCTA TGACA CATCA CCCTA CC
8	SKH-ND1F4	L4400	ACATT TCTGA CTACC AGAAG TACTT
9	SKH-ND2F2	L4900	TAGCC ACCAT TATAG CTCTA ACAAC
10	SKH-ND2F3	L5137	AAGCC TTCAA AGCTT TAAAC AA
11	SKH-ND2F4	L5659	ATCTA TAATG TGCTC GTAAC CGC
12	SKH-ND2F5	L6142	AATCT TAACA CAACT TTCTT TGACC
13	SKH-ND2F6	L6572	TCCTA GCCAA CTCAT CCCTC
14	SKH-CO2F3	L7670	GAAGA TGTCC TCCAC GCATG
15	SKH-CO2F4	L8134	ACTTC TTCAA TCAAT TCTTA AGCC
16	SKH-CO2F5	L8559	GAAAC TATTA GTCTA CTTAT TCGCC C
17	SKH-CO2F6	L9042	TGCGA TACGG AATAA TCTTG T
18	SK-GLYF1	L9632	AYTTA ATCTT GGTTA RAAYC CAAGG
19	SKH-ND3F1	L9963	TTCAG GCCTA ATTTA TGAAT GAC
20	SKH-ND3F2	L10572	ACTCA CATTA ACCTG TTGAC TTCTC C
21	SKH-ND4F1	L11088	ACTAG GTGGT TATGG AATAA TACG
22	SKH-ND4F2	L11640	CCACC TCACC TTAAT TAACC CCT
23	SKH-HISF1	L12063	CATCA ACTCA CATCA CAACT GTAA
24	SK-ND5F1	L13173	CYCAY YTAAA CGCCT GAGCC CT
25	SKH-ND5F2	L13512	ACCCA AACCT ATCCT ACCAT AAC
26	SKH-ND6F1	L13930	AACCA TGCTC ATTTC TATAG ACTCC
27	SKH-CYBMF6	L14797	AAATA TCATT CTGAG GCGCA AC
28	SK-LEU-CUNF1	L15196	CAGAG TGGTK ATGTG GTTGR CTTGA AA
29	SK-CYBMF1	L15213	TTYGC CTATG CCATY YTACG CTC
30	SKH-DL1	L15635	ACCCT TATCC TTGGC TCCCA A
31	SKH-CYBMF5	L15810	AATAC ATAAG ACAGT CTATG CTTAA TCC
32	SKH-DL3	L15972	CCACA CTAAC AGATT CTATA CCCTA AT
33	SKH-DL4	L16482	ACACC TTTGT TCGAT TTGTG G

Table 1 Continued

	Primers for H-strand		
	Name Long PCR primers	Position	Sequence (5′-3′)
	LA-C12SR1	H298	CTCGT ATAAC CGCGG TGGCT GGCAC GAGAT TGACC AAC
	LA-CGLUR1	H14184	GCTAG GGCTA GTAAT TTCTG CTGGG GTGGG TTGTG GTT
	PCR and sequencing primers		
34	SK-12SR1	H309	CGTAT AACCG CGGTG GCTGG CACGA
35	SK-16SR3	H574	TTTGA GCTAG CGCTT GTAGT ACTC
36	SKH-16SR2	H2059	GCAGG CGGGA CCTCT TATAC A
37	SY163R	H2620	CTCCG GTCTG AACTC AGATC ACGTA GG
38	SKH-16SR4	H2979	GACCT TTGCG GGATT GTATG TA
39	SK-TRPR2	H5132	AAAGC TTTGA AGGCT TTTGG T
40	SKH-CO2R4	H7457	GGTCA TTAAT TTCGT CCATT AA
41	SK-LYSR2	H7914	TTAGC TTAAA AGGCT AATGC T
42	SKH-ND3R5	H9774	GGTCA AAGCC GCACT CATAG
43	SK-ARGR2	H10042	ACCAT AATTT ACTAA GCCGA AAT
44	SKH-CO2R1	H11802	GCTAA TGCTG GACTG TTTAG CTT
45	SKH-ND5R4	H12978	ACCTG AGCAG AGGAA GAGTA TTG
46	SK-ND5R2	H13331	CTTCC GTAGG CGAGT CGTTT
47	SK-ND5R3	H13646	CTTTT WGGKC CWATT TTTTC
48	SKH-CYBMR3	H14605	CGGAT TATTC AGCCG TAGTT TA
49	SKH-CYBMR2	H14925	GGGAG GAGGA AGTGG AATGT
50	SKH-12SR2	H17368	ACATG TGTGT TTGAC ACTAT GTATG

L and H refer to the light and heavy strand, respectively.
The relative positions of primers in the mitochondrial genome are shown in Fig. 2.
Posotions with mixed bases are labeled with their IUB codes.

4. Sequence analysis

DNA sequences were aligned using the available complete mtDNA sequences of other cartilaginous fishes. The locations of the 13 protein-coding genes were determined by comparing the nucleotide or amino acid sequences of bony fish mitochondrial genomes. Twenty-two tRNA genes were identified by their proposed cloverleaf secondary structures and anticodon sequences, while two ribosomal RNA (rRNA) genes were identified by sequence homology and proposed secondary structures. Sequence data are available from GenBank under the accession number AY597334.

5. Phylogenetic analysis

Phylogenetic relationships among Myliobatiformes species were based on the nucleotide sequences of the mitochondrial 12S rRNA dataset from the 13 species of Myliobatiformes (e.g., Kao *et al.*, unpublished data) listed in the Appendix. DNA sequences were aligned using Clustal 1.83 as implemented in Mega 3.1 (Kumar *et al.*, 2004), and then the large gaps were manually deleted to optimize the alignment. Several approaches were applied to visualize 12S sequence diversity among the Myliobatiformes species investigated. Tree reconstruction was based on several approaches including neighbor-joining (NJ), maximum parsimony (MP), and Bayesian methods with bootstrap support where appropriate (Nei and Kumar, 2000). NJ and MP trees were constructed using Kimura's two-parameter method (Kimura, 1980) and a heuristic search with 1000 random sequence additions using PAUP* 4.0

b10 (Swofford, 2002). Bayesian analysis was performed using MrBayes version 3.1 (Ronquist and Huelsenbeck, 2003). The best-fit model of nucleotide substitution was the GTR model, which was selected using ModelTest version 3.7 (Posada and Crandall, 1998). The selected model was GTR + I + G (−lnL = 4905.52, K = 10, AIC = 9831.04) with base frequencies of A = 0.3538, C = 0.2235, G = 0.1751, and T = 0.2476; a proportion of invariable sites of I = 0.1777; and a gamma distribution shape parameter of variable sites of G = 0.3799. The analysis was run with the best-fit model for 1×10^6 generations, with a sampling frequency of 100 generations. The phylogenetic trees were visualized and edited with TreeView software (Page, 1996).

RESULTS AND DISCUSSION

The mitochondrial genomic content and the base composition of each genes are listed in Table 1. The genomic content of the deepwater stingray included 2 rRNA, 22 tRNA, and 13 protein-coding genes, an origin of replication, and a control region, as found in other cartilaginous fishes (Figs 2, 3; Table 1). Its gene order was also identical to that of the other cartilaginous fishes. The arrangements of most genes were encoded on the heavy (H)-strand, excluding ND6, origin of replication, and eight tRNA genes.

With a total length of 17514 bp, the deep water stingray genome is within the range of the known genomes of other cartilaginous fishes (16697–18580 bp; Cao *et al*., 1998; Delarbre *et al*., 1998; Rasmussen and Arnason, 1999a, b; Arnason *et al*., 2001). However, the total length of the deepwater stingray genome was approximately 730 bp longer than that of the thorny skate (16783 bp), which is under the order Rajiformes. By comparing these two sequences, it was found that the difference was apparently due to the length of the control region (1830 bp) of the deepwater stingray genome, which is longer than that of any other known sequences among cartilaginous fishes. However, the length of the control region of the deepwater stingray genome was shorter than that of the rabbitfish *Chimaera monstrosa*, which is split into two control regions: NC1 and NC2 (Arnason *et al*., 2001). Excluding the control region, the length of most genes in the mitochondrial genome of the deepwater stingray was pretty similar to that of other cartilaginous fishes (Table 2). These results indicated that the length of the coding region is somewhat conserved among teleost and cartilaginous fishes.

1. Protein-coding genes

Among the 13 protein-coding genes, there were four reading frame overlaps on the same strand (ATPase 8 and 6 shared 10 bp; ATPase 6 and COIII shared 1 bp; ND3 and tRNA^{Arg} shared 2 bp; ND4L and ND4 shared 7 bp) was observed. Conversely, there was one reading frame overlap on the opposite strand (ND5 and ND6 shared 4 bp). All mitochondrial protein-coding regions began with the ATG start codon excluding COI, which starts with GTG (Table 1). The open reading frame of the deepwater stingray ended with TAA (ND1, ND2, COI, ATPase 8, ATPase 6, COIII, ND4L, ND5, ND6, and *cyt* b), TAG (ND3), and the remaining genes had incomplete stop codons of T (COII and ND4).

2. tRNA genes

The mitochondrial genome of the deepwater stingray contained 22 tRNA genes dispersed between the RNA and protein-coding genes (Figs. 2, 3). Its size ranged from 67 to 75 nucleotides (Table 1), which is sufficient to permit the genes to fold into the four-arm cloverleaf secondary structure. However, the tRNA^{Trp} gene in *P. daviesi* exhibits unorthodox structures. The D-arms cannot form into a stable stem of 3–4 bp (replaced by a shorter stem of 2 bp), and the TΨC arm had 6 bp instead of 4–5 bp. There was conservation of the aminoacyl stem (7 bp) among the tRNA genes excluding tRNA^{Asp}, which forms a stem with only 6 bp. All postulated cloverleaf structures contained 4–5 bp in the anticodon stem. As shown in Table 3, most of the tRNA genes were located on the H-strand, with the exception of tRNA-Gln, -Ala, -Asn, -Cys, -Tyr, -Ser (UCN), -Glu, and –Pro, which are located on the light (L)-strand.

```
├ tRNA-Phe→                                                           ├ 12S rRNA→
GCTAGTGTAGCTTAAACCTAGAGCATGGCACTGAAGATGCCAAGATAAAAATTAATCTTTTTCACAAGCACGAAGGTTTGGTCCTAGCCTCAATATTAAT 100
TTTAACCTAATTTACACATGCAAGTCTCAGCACTCCGGTGAGAACGCCCTAATTATTACCCTCATTTTATTAGGAGCTGGTATCAGGCACATTCCCAATA 200
GCCCATGACACCTCGCCCAGCCACACCCACAAGGGAACTCAGCAGTGATAAACATTGTCCCATAAGCGTAAGCTTGACCCAATTAAAGTTATATAGTGTT 300
GGTCAATCTCGTGCCAGCCACCGCGGTTATACGAGCAACACAAATTAATATTTCACGGCATTAAGGGTGATTAGAAACATCTCTAATAAAATAAAGTTAA 400
AACCTTATTAAGCTGTCATACGCTTTTATATTTTAAAAACCCACTCACGAAAGTAACTTTAAATAAATACAGAACTTTTGACCTCACGACAGTTAAGACC 500
CAAACTAGGATTAGATACCCTACTATGCTTAACCATAAACATTGTTATAATAAACCCACCTTAATACCCGCCCGAGTACTACAAGCGCTAGCTCAAAACC 600
CAAAGGACTTGGCGGTGCTCCAAACCCCCCTAGAGGAGCCTGTTCTATAACCGATAATCCGCGTTCAATCCCACCACTTCTTGCCTTACCACCGCCTATA 700
TACCGCCGTCGTCAGCTCACCCTCTCAGGGCATAAAAGTAAGCAAAATGACCTTTCCCCTCAATACGTCAGGTCGAGGTGTAGCGAATGAAGTGGGAAGA 800
AATGGGCTACATTCCCTTTTCAGGGTATACGAACAGAAGCATGAAAATCTTCTTAAAGGTGGATTTAGCAGTAAGTAAATTTCAGGACATTATACTGAAA 900
CTGGCTCTGGAGCGCGCACACACCGCCCGTCACTCTCCTCAAAAATCATATTTAACTTTTATAAAAAAACTTTTTAGCAAGAGGAGGCAAGTCGTAACAT 1000
                             ┤      ├ tRNA-Val→
GGTAAGTGTACTGGAAAGTGCACTTGGATTAACAACCAAAGTGTAGCTAAATCAGTAAAGCACCTCCCTTACACCGAGAAGATTCCCGTGCAATCCGGGT 1100
       ┤ ├ 16S rRNA→
CACTTTGCTACCTCAAAGCTAGCCAAAATAAATTTATTAAGTTCCCCAATATTAACTAACACACACAACCTCTGTCCTCTAATTAAAACATTTTTCCCTT 1200
CCTAGTATGGGCGATAGAACAGAAACCTTTGAGCCATAGAAACAGTACCGCAAGGGAAAGCTGAAAAAGAAATGAAAAAACCATTAAAGTAAAAAGAAGC 1300
AGAGACCCGCCCTCGTACCTTTGCATCATGATTTAGCAAGAACAACTAGGCTAAAAAGCTTTTCCTAGCCTAGCCTCCCGAAACTAAACGAGCTACTTCG 1400
GAGCAGCTTACCCAGAGCCAACCCACCTCTGTAGCAAAAGAGAGGGAAGACTCCCAAGTAGCGGTGATAAGCCTACCGAGTTTAGTGATAGCTGGTCATC 1500
CAAAAAAAGAACTTAAATTCTGCATTAATTTTTCAACCAGCAACTAAAAACCTTTTACTAAGCTCACTTGTAAAAATTAAGAGTTATTCAAAAAAGGTAC 1600
AGCCTTTTTGAATCAAGAAACAACTTTATTAGGAGGGTAATGATTACATTCTTAAAGGGTTTCTCCTCAGTGGGCCTAAAAGCAGCCACCTGTTAAGCAA 1700
GCGTCATAGCTCAAGCCTCACCTGCCCCAACAAATTCCCATACTCATTCTCAACCCCCTACCTCACTATTGGATTATTTTATTAACCTAATTATAAAAGA 1800
AATTATGCTAAAATGAGTAATGAGGGAATAACCCCCTCCCCGACACCAGTGTATGTCAGAAAGAATTAAATCACTGACAATTAACCGATGCCATAGTTGA 1900
GGCTCTCATGACATAAACACTAAACACAAGAAAACCCCATACAAAACCTCGTTAACCCTACACAGGAGTGTCCCCGGGAAAGATTAAAAGAAAATAAAGG 2000
AACTCGGCAAACACAAACTCCGCCTGTTTACCAAAAACATCGCCTCTTGCCCCACTCATGTATAAGAGGTCCCGCCTGCCCTGTGATTTTTTAACGGCCG 2100
CGGTATTCTGACCGTGCGAAGGTAGCGTAATCACTTGTCTTTTAATTGAAGACCCGTATGAAAGGCATCACGAGAGTTTATCTGTCTCTATTTTCCAATC 2200
AATGAAATTGAACCTCTCGTGCAGAAGCGAGAATAAAAACATAAGACGAGAAGACCCTATGGAGCTTCAAACACTTAAGTTACTTTTAAAACATAAAATT 2300
CCTACCTTCGGGTATAAACTAAAAACTAATTTCTTAACTTAACCTGTTTTTGGTTGGGGCGACCAGGGGGAAAAACAAAACCCCCTTATCGAATGTGTTA 2400
AACACAAAAATTAGGACTACAGTCCTAATCAATAGAAAATCTAACGAACAATGACCCAGGGCCCGATCCCTGATCAATGAACCAAGTTACCCTAGGGATA 2500
ACAGCGCAATCCTTTCTTAGAGCCCCCATCGCCGAAAGGGTTTACGACCTCGATGTTGGATCAGGACATCCTAATGGTGTAGCCGCTATTAAGGGCTCGT 2600
TTGTTCAACGATTAACAGTCCTACGTGATCTGAGTTCAGACCGGAGTAATCCAGGTCAGTTTCTATCTATGATGGTATTTTCCCCAGTACGAAAGGACCG 2700
                                                                                                  ┤├
GGAAAATGAAGCCTATGCCCAAAGTACGCTTCCCCCCAACCTGCTGAAACCAACTCAAGCAGATAAAGGCAGCCACCTTTAAACCAAAGATAACGGTTGT 2800
tRNA-Leu(UUR)→                                                          ┤├ ND1→
TGGGGTGGCAGAGCCTGGTAAATGCGAAAGACCTAAGTTCTTTAATCCAGAGGTTCAACTCCTCTCCTTAACTATGTTAAACTTCACCCTTCCCTACATC 2900
                                                                          M  L  N  F  T  L  P  Y  I
ATCAACCCCCTAGCCTTCATCATCCCCATCCTTCTAGCCACAGCCTTCCTTACCTTAATTGAACGAAAAATTCTAGGTTACATACAATCCCGCAAAGGTC 3000
 I  N  P  L  A  F  I  I  P  I  L  L  A  T  A  F  L  T  L  I  E  R  K  I  L  G  Y  M  Q  S  R  K  G  P
CAAATGTAGTTGGCCCGTACGGTCTTCTACAACCCATCGCTGATGGACTTAAACTATTCACCAAAGAACCACTACGACCCTCCTTCTCCTCCCAATTCCT 3100
   N  V  V  G  P  Y  G  L  L  Q  P  I  A  D  G  L  K  L  F  T  K  E  P  L  R  P  S  F  S  S  Q  F  L
CTTTCTAACCGCCCCTGCCCTCGCCTTAACCCTCGCCCTATTAATATGAATACCCCTACCCCTCCCATACTCAATCCTAAACCTAAACCTAGGACTACTT 3200
  F  L  T  A  P  A  L  A  L  T  L  A  L  L  M  W  M  P  L  P  L  P  Y  S  I  L  N  L  N  L  G  L  L
TTTATCCTAGCTATCTCAAGCCTAACAGTCTACACAATTTTAGGCTCAGGCTGAGCATCAAACTCTAAATACGCTCTCATGGGAGCTCTCCGTGCAGTCG 3300
 F  I  L  A  I  S  S  L  T  V  Y  T  I  L  G  S  G  W  A  S  N  S  K  Y  A  L  M  G  A  L  R  A  V  A
CACAAACCATCTCTTACGAGGTAACACTCGCCCTAATTCTCTTAGCCCTCACCATCTTCGTAGGGGGGTTTACACTCCACACATTTAACATGAGTCAAGA 3400
   Q  T  I  S  Y  E  V  T  L  A  L  I  L  L  A  L  T  I  F  V  G  G  F  T  L  H  T  F  N  M  S  Q  E
AACCATTTGACTGATCGTCCCCACATGACCTCTAGCCATAATATGATATATCTGAACACTACCAAAAACAAACCGAGCCCCCTTTGACCTCACAGAGGGG 3500
  T  I  W  L  I  V  P  T  W  P  L  A  M  M  W  Y  I  W  T  L  P  K  T  N  R  A  P  F  D  L  T  E  G
GAATCCGAACTAGTCTCAGGTTTCAACACTGAGTACGCAGGGGGCTCTTTCGCCCTATTCTTCTTAGCCGAATACTCCAATATCTTACTAATAAATACCC 3600
 E  S  E  L  V  S  G  F  N  T  E  Y  A  G  G  S  F  A  L  F  F  L  A  E  Y  S  N  I  L  L  M  N  T  L
TATCCGCTATCCTATTCCTAGGCGCCTCCTACACCCCACACTTCCCACAACTAGCTACCCTTAACCTTATACTCAAAACAACCGCACTAACCCTACTATT 3700
   S  A  I  L  F  L  G  A  S  Y  T  P  H  F  P  Q  L  A  T  L  N  L  M  L  K  T  T  A  L  T  L  L  F
TCTATGAATTCGAGCCTCTTACCCCCGATTTCGATACGACCAACTCATACACCTTGTCTGAAAAAACTTTCTACCAATAACACTAGCCCTGATCCTATGA 3800
  L  W  I  R  A  S  Y  P  R  F  R  Y  D  Q  L  M  H  L  V  W  K  N  F  L  P  M  T  L  A  L  I  L  W
                                                  ┤  ├ tRNA-Ile→
CACATCACCCTACCAATCATTACAGCAAGCCTCCCACCAATAATAACTTAAACGGAAAAGTGCCTGAATAAAGGGCCACTTTGATAGAGTGGATAACGAA 3900
 H  I  T  L  P  I  I  T  A  S  L  P  P  M  M  T  *
                     ┤  ├ tRNA-Gln→                                                              ├
TGTTAAAACCATTCCCTTTCCTCACTAGAAAGACAGGATTTGAACCTATACCCTAGAGATCAAAACTCTTTGTACTCCCAACTATACTACTTCCTAAAGT 4000
tRNA-Met→                                                         ┤├ ND2→
AAAGTCAGCTAAATCAAGCTTTTGGGCCCATACCCCAAACATGTTGGTTAAAATCCTTCCTTTACTAATGAACCCATTAATTCTCACTATCTCAATCCTT 4100
                                                                    M  N  P  L  I  L  T  I  S  I  L
AGTCTAGGATTAGGCACCACAATAACACTCATTACCTCACACTGATTACTAATCTGACTTGGCCTCGAAATTAATACAATAGCCATTATCCCCCTCATAA 4200
 S  L  G  L  G  T  T  M  T  L  I  T  S  H  W  L  L  I  W  L  G  L  E  I  N  T  M  A  I  I  P  L  M  I
TCCATCAACACCATCCCCGTGCAACAGAAGCAACTACAAAATACTTCCTTACACAAGCCACTGCCTCAGCCCTCCTCCTATTTGCCAGCACCACCAACGC 4300
   H  Q  H  H  P  R  A  T  E  A  T  T  K  Y  F  L  T  Q  A  T  A  S  A  L  L  L  F  A  S  T  T  N  A
CTGGCTCACCGGCCAATGAAATATGACCGAAATAATTAATCCCCTCTCCACCACACTTCTCTCCATTGCTCTAGCACTAAAAATTGGCTTAGCCCCCTTA 4400
  W  L  T  G  Q  W  N  M  T  E  M  I  N  P  L  S  T  T  L  L  S  I  A  L  A  L  K  I  G  L  A  P  L
CATTTCTGACTACCAGAAGTACTTCAAGGACTTAACCTACTCACAGGACTCATCCTATCTACATGACAAAAACTTGCACCATTCGCAATTCTACTACAAC 4500
 H  F  W  L  P  E  V  L  Q  G  L  N  L  L  T  G  L  I  L  S  T  W  Q  K  L  A  P  F  A  I  L  L  Q  L
TCTACCCCCTCCTTAACCCAACATTACTTGTATCAATGGGCATACTATCAATCGTCATCGGTGGTTGAGGCGGCCTCAACCAAACACAACTACGAAAAAT 4600
   Y  P  L  L  N  P  T  L  L  V  S  M  G  M  L  S  I  V  I  G  G  W  G  G  L  N  Q  T  Q  L  R  K  I
TCTAGCATACTCCTCAATCGCGCACATCGGCTGAATAACTGTTATTCTCCATTACTCTCCTAACCTCACCCTCCTCAATCTAATCATTTACATTATCATA 4700
  L  A  Y  S  S  I  A  H  I  G  W  M  T  V  I  L  H  Y  S  P  N  L  T  L  L  N  L  I  I  Y  I  I  M
ACCTCCTCATTATTCCTTCTCCTCTACAACAACAACACCACAAAAATTAATTCAATCTCCACATCATCAACTAAATCCCCACTACTATCCATCATAACCA 4800
 T  S  S  L  F  L  L  L  Y  N  N  N  T  T  K  I  N  S  I  S  T  S  S  T  K  S  P  L  L  S  I  M  T  M
TATTAACACTCCTCTCTCTAGGGGGTCTACCCCCTCTCACAGGCTTCATGCCTAAGTGACTTATCCTCCAAGAAATAACCAAACAAAATCTCCTCATCCT 4900
   L  T  L  L  S  L  G  G  L  P  P  L  T  G  F  M  P  K  W  L  I  L  Q  E  M  T  K  Q  N  L  L  I  L
AGCCACCATTATAGCTCTAACAACACTACTTAGTCTATTTTTCTACCTACGTCTATGCTACATGATCTCCCTCACCCTCCCCCCAAACCCAATTTCTTCC 5000
  A  T  I  M  A  L  T  T  L  L  S  L  F  F  Y  L  R  L  C  Y  M  I  S  L  T  L  P  P  N  P  I  S  S
TCATCATCATGACGAACAAAAACTATCCAACCCAACTTAACCCTAATAACCTCCCTCTCCCTTCTTCTCCTCCCATTGACCCCCACAATACTTTCACTAA 5100
 S  S  S  W  R  T  K  T  I  Q  P  N  L  T  L  M  T  S  L  S  L  L  L  L  P  L  T  P  T  M  L  S  L  T
         ┤├ tRNA-Trp→                                                           ┤ ├ tRNA-Ala→
CTACTACATAAAGAAATTTAGGTTCAAATAAACCAAAAGCCTTCAAAGCTTTAAACAAGAGTGAAAACCCCTTAATTTCTGCTAAGGTTTGCAAAACTTT 5200
   T  T  *
                                                  ┤ ├ tRNA-Asn→
ACTTCACATCTCTTGAATGCAACCCAAGTGCTTTAATTAAGCTAAAACCTTACTAGATAAACAGGCCTCGATCCTATAAAATCTTAGTTAACAGCTAAGT 5300
                        ┤├ Rep-Origin→                   ┤├ tRNA-Cys→
```

```
GTCCAATCCAGCGAACTTTTATCTAGCTTCTCCCGCCGTTCAGGGCGGCGAGGCGGGAGAAGCCCCGGGAGAAGCCTATTCTCCTTCTCGGGACTTGCAA 5400
                                                 ┤     ├ tRNA-Tyr→
TCCCGTGTATTATAATACTACAGGGCTGAGTAGTAAGAGGAGGACTCGAACCTCCCTCTACGGGGCTACAATCCGCCGCTTTGGGTCTCAGCCATCTTAC 5500
┤ ├ COI→
CTGTGACAATTAATCGTTGATTATTTCTACCAATCACAAAGACATTGGCACCCTTTATTTAATCTTTGGTGCATGGGCAGGAATAGTGGGTACCGGCCT 5600
  V  T  I  N  R  W  L  F  S  T  N  H  K  D  I  G  T  L  Y  L  I  F  G  A  W  A  G  M  V  G  T  G  L
CAGCCTCTTAATTCGAACAGAGCTAAGCCAACCGGGGGCTTTATTAGGGGATGATCAAATCTATAATGTGCTCGTAACCGCTCATGCCTTTGTAATAATC 5700
 S  L  L  I  R  T  E  L  S  Q  P  G  A  L  L  G  D  D  Q  I  Y  N  V  L  V  T  A  H  A  F  V  M  I
TTCTTTATAGTTATACCAATTATAATCGGCGGGTTCGGCAATTGATTAGTTCCCTTAATAATCGGCGCCCCGGACATAGCCTTCCCGCGAATAAACAACA 5800
F  F  M  V  M  P  I  M  I  G  G  F  G  N  W  L  V  P  L  M  I  G  A  P  D  M  A  F  P  R  M  N  N  M
TAAGTTTCTGACTCCTTCCTCCCTCTTTTCTCTTATTACTAGCTTCCGCAGGGGTAGAAGCCGGAGCTGGCACAGGATGAACCGTTTACCCCCCATTAGC 5900
  S  F  W  L  L  P  P  S  F  L  L  L  L  A  S  A  G  V  E  A  G  A  G  T  G  W  T  V  Y  P  P  L  A
TGGAAATCTAGCACATGCCGGAGCTTCCGTGGACTTAGCCATTTTTTCCTTACATCTAGCAGGGGCCTCCTCTATTATAGCATCAATTAATTTTATTACT 6000
 G  N  L  A  H  A  G  A  S  V  D  L  A  I  F  S  L  H  L  A  G  A  S  S  I  M  A  S  I  N  F  I  T
ACGATCATTAATATGAAACCACCCGCGATCTCCCAATACCAAACGCCTCTTTTTGTTTGATCTATTCTCATTACAACTGTCCTTCTTTTATTGTCCCTTC 6100
T  I  I  N  M  K  P  P  A  I  S  Q  Y  Q  T  P  L  F  V  W  S  I  L  I  T  T  V  L  L  L  L  S  L  P
CTGTCCTAGCAGCAGGCATCACCATACTCCTCACAGACCGCAATCTTAACACAACTTTCTTTGACCCCGCTGGAGGAGGAGACCCCATTCTCTATCAACA 6200
  V  L  A  A  G  I  T  M  L  L  T  D  R  N  L  N  T  T  F  F  D  P  A  G  G  G  D  P  I  L  Y  Q  H
CCTCTTTTGATTTTTCGGACACCCAGAAGTATATATCCTAATCCTTCCTGGCTTTGGTATAATTTCCCATGTAGTCGCGTATTATTCTGGAAAAAAAGAA 6300
 L  F  W  F  F  G  H  P  E  V  Y  I  L  I  L  P  G  F  G  M  I  S  H  V  V  A  Y  Y  S  G  K  K  E
CCTTTTGGTTATATAGGTATAGTTTGAGCAATAATAGCTATCGGCCTTCTTGGCTTTATTGTTTGAGCCCATCATATATTCACAGTAGGTATAGACGTAG 6400
P  F  G  Y  M  G  M  V  W  A  M  M  A  I  G  L  L  G  F  I  V  W  A  H  H  M  F  T  V  G  M  D  V  D
ACACACGAGCCTACTTCACATCAGCAACTATAATTATCGCTATTCCAACAGGGGTAAAAGTATTTAGTTGACTAGCAACCCTTCATGGCGGCACCATTAA 6500
  T  R  A  Y  F  T  S  A  T  M  I  I  A  I  P  T  G  V  K  V  F  S  W  L  A  T  L  H  G  G  T  I  K
ATGAGAGACACCACTCCTTTGAGCCCTAGGGTTTATTTTCCTATTTACTGTCGGTGGCTTAACTGGTATTATCCTAGCCAACTCATCCCTCGATATTGTT 6600
 W  E  T  P  L  L  W  A  L  G  F  I  F  L  F  T  V  G  G  L  T  G  I  I  L  A  N  S  S  L  D  I  V
CTTCACGATACTTACTATGTTGTAGCCCACTTCCATTATGTTTTATCAATAGGGGCAGTATTCGCTATCATAGCTGGTTTCGTCCACTGATTTCCTTTAA 6700
L  H  D  T  Y  Y  V  V  A  H  F  H  Y  V  L  S  M  G  A  V  F  A  I  M  A  G  F  V  H  W  F  P  L  I
TCACAGGCTACACTCTTCACCCCACTTGAACTAAAGTACAATTCCTAGTAATATTCGTAGGAGTCAATATAACCTTCTTCCCTCAACACTTCTTGGGTCT 6800
  T  G  Y  T  L  H  P  T  W  T  K  V  Q  F  L  V  M  F  V  G  V  N  M  T  F  F  P  Q  H  F  L  G  L
AGCTGGAATACCACGCCGATATTCAGACTACCCGGACGCCTATACCTTTTGAAATGTAATTTCATCTATCGGTTCTTTAATCTCATTAGTAGCTGTAATC 6900
 A  G  M  P  R  R  Y  S  D  Y  P  D  A  Y  T  F  W  N  V  I  S  S  I  G  S  L  I  S  L  V  A  V  I
ATCATAATATTTATCCTTTGAGAAGCCTTCGCATCAAAACGTGAGATCCTTCACATTGAACTTCCCCATACAAACGTAGAATGACTTCACGGATGCCCCC 7000
I  M  M  F  I  L  W  E  A  F  A  S  K  R  E  I  L  H  I  E  L  P  H  T  N  V  E  W  L  H  G  C  P  P
                                                                             ┤      ├ tRNA-Ser→
CACCTTACCACACCTACGAAGAACCAGCATTCGTTCAAGTTCAACAACCTTAACTAAACAAGAAAAGAAGGAATTGAACCCCCATTAATTGGTTTCAAGC 7100
  P  Y  H  T  Y  E  E  P  A  F  V  Q  V  Q  Q  P  *
                             ┤ ├ tRNA-Asp→                                                   ┤
CAACCACATAACCACTCTGTCACTTTCTTGAGATTCTAGTAAAACTATTACATTTCCTTGTCAAGGCAAAATTGTGGGTTTAACCCCCACGAATCTTTAA 7200
├ COII→
TATGGCACATCCATCACAATTAGGTTTTCAAGACGCAGCCTCTCCAGTTATGGAAGAGCTCCTCCACTTTCACGATCATACCTTAATAATCGTATTTCTC 7300
 M  A  H  P  S  Q  L  G  F  Q  D  A  A  S  P  V  M  E  E  L  L  H  F  H  D  H  T  L  M  I  V  F  L
ATTAGCTCATTAGTCCTTTACGTTATTGTAGCAACAGTTTCAACCAAATTAACTAACAAATATATTCTAGACTCCCAAGAAATTGAAATCGTTTGAACTA 7400
I  S  S  L  V  L  Y  V  I  V  A  T  V  S  T  K  L  T  N  K  Y  I  L  D  S  Q  E  I  E  I  V  W  T  I
TCGTCCCAGCAATTATCTTAGTTATAATTGCCGTACCATCCCTGCGCATTCTCTATTTAATGGACGAAATTAATGACCCCCACATCACAATCAAAACCAT 7500
  V  P  A  I  I  L  V  M  I  A  V  P  S  L  R  I  L  Y  L  M  D  E  I  N  D  P  H  I  T  I  K  T  I
TGGCCACCAATGATATTGAAGTTATGAATACACAGACTACCAAAACCTGGAATTTGATTCTTACATAACTCAAACAGAAAATTTAACCCCAGGACAATTT 7600
 G  H  Q  W  Y  W  S  Y  E  Y  T  D  Y  Q  N  L  E  F  D  S  Y  M  T  Q  T  E  N  L  T  P  G  Q  F
CGCCTCCTAGAAACGGACCACCGCATAGTGGTCCCAATACAATCCCCCATTCGAGTCTTAGTAACAGCAGAAGATGTCCTCCACGCATGAACAGTCCCAG 7700
R  L  L  E  T  D  H  R  M  V  V  P  M  Q  S  P  I  R  V  L  V  T  A  E  D  V  L  H  A  W  T  V  P  A
CACTAGGAATTAAAATAGACGCAGTACCAGGGCGCTTAAACCAAACAGCCTTCATCATCTCTCGCCCAGGTATCTTCTATGGCCAATGCTCCGAAATTTG 7800
  L  G  I  K  M  D  A  V  P  G  R  L  N  Q  T  A  F  I  I  S  R  P  G  I  F  Y  G  Q  C  S  E  I  C
                                                                                             ┤├ tRNA-Lys→
TGGGGCTAACCATAGTTTTATGCCAATCGTAGTTGAAGCAGTTCCACTAGAACACTTCGAGAACTGATCCTCATTAATACTTGAAGAACTTTCATTAAGA 7900
 G  A  N  H  S  F  M  P  I  V  V  E  A  V  P  L  E  H  F  E  N  W  S  S  L  M  L  E  E  L  *
                                                                  ┤ ├ ATPase8→
AGCTAAACAGTCCAGCATTAGCCTTTTAAGCTAAAAATTGGTGACTCCTAACCACCCTTAATGACATGCCTCAACTTAACCCCGGTCCCTGATTTTTAAT 8000
                                                                   M  P  Q  L  N  P  G  P  W  F  L  I
TTTCCTATTTACATGACTTTTTTTCTTAGCTATTATACCAAACAAAGTAATATCTTACCTCCTCAATAATAATCCCACAACAAAAAATAACCAAAAGCCA 8100
 F  L  F  T  W  L  F  F  L  A  I  M  P  N  K  V  M  S  Y  L  L  N  N  N  P  T  T  K  N  N  Q  K  P
                          ├ ATPase6→  ┤
AAACCAAACCCCTGAAATTGACCATGGCCCTAAACTTCTTCAATCAATTCTTAAGCCCATCACTACTTGGACTCCCCCTCATTGCTCTAGCAATCATAAT 8200
K  P  N  P  W  N  W  P  W  P  *
                          M  A  L  N  F  F  N  Q  F  L  S  P  S  L  L  G  L  P  L  I  A  L  A  I  M  I
CCCTTGACTCATCTTTCCTCCCCCCTCTAAACAATGACTTACTAACCGCCTATTAACCCTTCAAACATGATTCATTAACCGATTCACCCATCAACTTATA 8300
 P  W  L  I  F  P  P  P  S  K  Q  W  L  T  N  R  L  L  T  L  Q  T  W  F  I  N  R  F  T  H  Q  L  I
CAACCACTAAGCCTCGGAGGCCATAAATGAGCCTCAATTCTTACCGCACTTATATTATTTTTAATCACTATTAACCTCTTGGGCTTACTCCCATATACAT 8400
Q  P  L  S  L  G  G  H  K  W  A  S  I  L  T  A  L  M  L  F  L  I  T  I  N  L  L  G  L  L  P  Y  T  F
TCACTCCCACAACCCAACTCTCCTTAAACATAGCATTCGCTCTACCCCTTTGACTAACTACAGTCCTAATTGGTATATTTAATCAACCAACAGTTGCTTT 8500
  T  P  T  T  Q  L  S  L  N  M  A  F  A  L  P  L  W  L  T  T  V  L  I  G  M  F  N  Q  P  T  V  A  L
AAGCCACTTCCTGCCAGAAGGAACACCCGCCCCCCTAATCCCTATTCTTATTATTATCGAAACTATTAGTCTACTTATTCGCCCCTTAGCCCTAGGAGTC 8600
 S  H  F  L  P  E  G  T  P  A  P  L  I  P  I  L  I  I  I  E  T  I  S  L  L  I  R  P  L  A  L  G  V
CGACTCACAGCTAACCTCACAGCAGGCCACCTTCTAATACAACTAATCGCAACTGCAGCATTTGCCCTAATCTCCATCATGCCTTCAATCGCCCTCCTCA 8700
R  L  T  A  N  L  T  A  G  H  L  L  M  Q  L  I  A  T  A  A  F  A  L  I  S  I  M  P  S  I  A  L  L  T
CCTCACTTATCCTCTTCCTCCTTACAATCCTAGAAGTTGCCGTAGCAATAATTCAAGCCTATGTGTTTGTTCTCCTCTTAAGTCTCTACTTACAAGAAAA 8800
  S  L  I  L  F  L  L  T  I  L  E  V  A  V  A  M  I  Q  A  Y  V  F  V  L  L  L  S  L  Y  L  Q  E  N
      ┼ COIII→
TGTTTAATGGCCCATCAAACACATGCATATCACATAGTTGACCCCAGCCCATGACCCTTAACAGGAGCAGTAGCAGCCCTTCTCATAACCTCAGGTCTCG 8900
  V  * M  A  H  Q  T  H  A  Y  H  M  V  D  P  S  P  W  P  L  T  G  A  V  A  A  L  L  M  T  S  G  L  V
TCATCTGATTCCATTTCCACTCCCTCCCTTTACTTTGCCTAGGCCTCCTCCTGCTAACCTTAACCATCATCCAATGATGACGAGACGTTATTCGAGAAGG 9000
  I  W  F  H  F  H  S  L  P  L  L  C  L  G  L  L  L  L  T  L  T  I  I  Q  W  W  R  D  V  I  R  E  G
AACATTCCAAGGTCACCACACCCTACCAGTCCAAAAAGGCCTGCGATACGGAATAATCTTGTTTATTACATCAGAAGTCTTCTTCTTTCTTGGCTTCTTC 9100
 T  F  Q  G  H  H  T  L  P  V  Q  K  G  L  R  Y  G  M  I  L  F  I  T  S  E  V  F  F  F  L  G  F  F
TGAGCCTTCTACCACTCAAGCCTCGCCCCAACCCCTGAACTTGGAGGCTGCTGACCACCTACAGGCATCCACCCCTTAGACCCCTTCGAAGTCCCCCTCC 9200
W  A  F  Y  H  S  S  L  A  P  T  P  E  L  G  G  C  W  P  P  T  G  I  H  P  L  D  P  F  E  V  P  L  L
TCAACACCGCCGTCCTCTTGGCCTCTGGAGTAACAGTCACCTGAGCCCACCATAGTGTCATAGAAGGTAACCGAAAAGAAGCCATTCAAGCACTCGCCCT 9300
  N  T  A  V  L  L  A  S  G  V  T  V  T  W  A  H  H  S  V  M  E  G  N  R  K  E  A  I  Q  A  L  A  L
```

```
TACAATCACTCTCGGCTTTTATTTCACCACACTCCAAGCCATAGAATACTATGAAGCCCCCTTCACCATAGCCGACAGCGTCTATGGAACAACCTTCTTC 9400
  T  I  T  L  G  F  Y  F  T  T  L  Q  A  M  E  Y  Y  E  A  P  F  T  M  A  D  S  V  Y  G  T  T  F  F
GTCGCCACAGGCTTCCACGGCCTCCATGTAATAATTGGCTCCTCATTTCTACTAGTTTGCCTTCTACGACAAATCCAATACCACTTCACATCCCAGCACC 9500
 V  A  T  G  F  H  G  L  H  V  M  I  G  S  S  F  L  L  V  C  L  L  R  Q  I  Q  Y  H  F  T  S  Q  H  H
                                                                                         ┤├ tRNA-Gly→
ATTTTGGCTTCGAAGCCGCTGCCTGATACTGACATTTTGTCGATGTAGTCTGATTATTCCTTTACGTCTCAATCTACTGATGAGGTTCATAAGCTTTTCT 9600
   F  G  F  E  A  A  A  W  Y  W  H  F  V  D  V  V  W  L  F  L  Y  V  S  I  Y  W  W  G  S  *
                                                              ┤├ ND3→
AGTATAAAACTAGTACAAATGACTTCCAATCATTTAATCTTGGTTTAAACCCAAGGAAAAGCAATGAACCTCATCACATTTGTTGTCGCCCTTACAGCCC 9700
                                                                M  N  L  I  T  F  V  V  A  L  T  A  L
TCATTTCCCTAATCTTAGCGATATTAGCTTTTTGACTGCCCACCCTTAGCCCAGATAACGAAAAAATATCCCCCTATGAGTGCGGCTTTGACCCACTAGG 9800
   I  S  L  I  L  A  M  L  A  F  W  L  P  T  L  S  P  D  N  E  K  M  S  P  Y  E  C  G  F  D  P  L  G
TACTGCACGCTTACCCTTCTCACTCCGCTTCTTCCTAGTTGCTATCCTCTTCCTCCTTTTTGACTTAGAAATCGCCCTACTTCTCCCACTCCCCTGAGCT 9900
  T  A  R  L  P  F  S  L  R  F  F  L  V  A  I  L  F  L  L  F  D  L  E  I  A  L  L  L  P  L  P  W  A
GTCCAACTCGACTCCCCCCTCATCACCTCTTTCTGAGCAACAACCATTTTACTCCTCCTAACTTCAGGCCTAATTTATGAATGACTCCAAGGGGGCCTAG 10000
 V  Q  L  D  S  P  L  I  T  S  F  W  A  T  T  I  L  L  L  L  T  S  G  L  I  Y  E  W  L  Q  G  G  L  E
           ├ ┤                                                       ← tRNA-Arg ┤├ ND4L
AATGAGCAGACTAGGTACTTAGTCCAAACTTAAGACTATTGATTTCGGCTTAGTAAATTATGGTGAAACCCCATAAGTATCTTCATGTCCCCAATCCACT 10100
   W  A  D  *                                                                        M  S  P  I  H  F
TCACCTTTTCCTCTGCTTTTGCCTTAAGCCTACTTGGCCTAGCCCTAAACCGCTCCCATCTTTTATCCGCCCTCATCTGCCTCGAAGGTATAATATTATC 10200
   T  F  S  S  A  F  A  L  S  L  L  G  L  A  L  N  R  S  H  L  L  S  A  L  I  C  L  E  G  M  M  L  S
CCTATTTACCGCCATCGCTCTTTGATCCACAACAATAACCACACCAACCTGCTCCCTCGCACCCATGATTCTTCTAACATTTTCAGCCTGTGAAGCAAGC 10300
  L  F  T  A  I  A  L  W  S  T  T  M  T  T  P  T  C  S  L  A  P  M  I  L  L  T  F  S  A  C  E  A  S
                                                                          ├ ND4→ ┤
GCAGGTTTAGCCCTCCTAGTAGCCACCACCCGCACCCACGGCTCAGATATACTAAAAAATCTTAATCTCCTACAATGTTAAAAATTATTATTCCTACAAT 10400
 A  G  L  A  L  L  V  A  T  T  R  T  H  G  S  D  M  L  K  N  L  N  L  L  Q  C  *
                                                                           M  L  K  I  I  I  P  T  I
TATACTTTACCCAATCTCATGGGCCACCCCCAAAAAATGACTATGAACTGCCTCCACATCCTACAGCCTTCTCATCGCATTTATCAGCTTATCTTGATTT 10500
  M  L  Y  P  I  S  W  A  T  P  K  K  W  L  W  T  A  S  T  S  Y  S  L  L  I  A  F  I  S  L  S  W  F
AAATGAGATACTGAAGTCAGCTGAGATTTCTCCAACCTCTACCTAGGAGTAGACCCCCTCTCATCCCCCCTACTCACATTAACCTGTTGACTTCTCCCAC 10600
 K  W  D  T  E  V  S  W  D  F  S  N  L  Y  L  G  V  D  P  L  S  S  P  L  L  T  L  T  C  W  L  L  P  L
TCATACTACTTGCTAGCCAAAATCACCTTAACAACGAACCACACACTCGACAACGCATTTATATTAACCTTCTTATCACTCTTCAACTCCTCCTAATCCT 10700
   M  L  L  A  S  Q  N  H  L  N  N  E  P  H  T  R  Q  R  I  Y  I  N  L  L  I  T  L  Q  L  L  L  I  L
AGCTTTCAGTGCTACTGAAATAATCTTATTTTATATCATATTTGAAGCAACCCTCATCCCAACACTCATCATCATTACCCGTTGAGGCAACCAAGCCGAA 10800
  A  F  S  A  T  E  M  I  L  F  Y  I  M  F  E  A  T  L  I  P  T  L  I  I  I  T  R  W  G  N  Q  A  E
CGCCTAAATGCAGGAATTTATTTTTATTTTATACACTAATAGCCTCCCTACCCCTCTTAATCGCCCTCCTCGCCTTACAAAATGACTTTGGCTCACTCT 10900
 R  L  N  A  G  I  Y  F  L  F  Y  T  L  M  A  S  L  P  L  L  I  A  L  L  A  L  Q  N  D  F  G  S  L  S
CAATGCTCACCTTTCAATTCCCCCAACTCCTAAACTCCAGCTCATGAACAAACAAGTTCTGATGAGCCGCATGTCTAATCGCCTTCTTAGTCAAAATACC 11000
   M  L  T  F  Q  F  P  Q  L  L  N  S  S  S  W  T  N  K  F  W  W  A  A  C  L  I  A  F  L  V  K  M  P
ACTCTACGGAACACACCTCTGACTACCCAAAGCCCATGTAGAAGCCCCTATCGCCGGATCCATGATCCTAGCCGCAATCCTACTAAAACTAGGTGGTTAT 11100
  L  Y  G  T  H  L  W  L  P  K  A  H  V  E  A  P  I  A  G  S  M  I  L  A  A  I  L  L  K  L  G  G  Y
GGAATAATACGAATTATCTCAATCCTTGATCCCCTCACAAAAGAAATGGCCTTCCCATTCCTAATTCTAGCCCTCTGAGGAATTATCATGACAAGTTCCA 11200
 G  M  M  R  I  I  S  I  L  D  P  L  T  K  E  M  A  F  P  F  L  I  L  A  L  W  G  I  I  M  T  S  S  T
CCTGTTTACGCCAAACAGACCTTAAATCCTTAATTGCCTACTCATCAGTAACCCACATAGGTCTTGTTGTAGCAGCCATCCTCATTCAAACACCATGAAG 11300
   C  L  R  Q  T  D  L  K  S  L  I  A  Y  S  S  V  T  H  M  G  L  V  V  A  A  I  L  I  Q  T  P  W  S
TTTCGCAGGGGCAACCGCCTTAATAATTGCCCATGGCCTCATTTCATCCGCCCTTTTCTGCTTAGCCAATACCAACTACGAACGAACACACACCCGAACT 11400
  F  A  G  A  T  A  L  M  I  A  H  G  L  I  S  S  A  L  F  C  L  A  N  T  N  Y  E  R  T  H  T  R  T
CTCCTCCTCACCCGAGGGATACAAATTGTTCTCCCATTAATAGCAACTTCGTGATTCTTAATTAACCTAGCAAACCTCGCCCTACCCCCAACCCCAAACC 11500
 L  L  L  T  R  G  M  Q  I  V  L  P  L  M  A  T  S  W  F  L  I  N  L  A  N  L  A  L  P  P  T  P  N  L
TAATGGGCGAACTCCTTATTATATCTTCTCTTTTCAAATGATCCCAATGAACCATCATCATAACTGGTTTAGGCGTCTTATTAACTGCCTCTTACTCCTT 11600
   M  G  E  L  L  I  M  S  S  L  F  K  W  S  Q  W  T  I  I  M  T  G  L  G  V  L  L  T  A  S  Y  S  L
ATATATATTCATCATTACACAACGAGGCCCTCTTCCTCCCCACCTCACCTTAATTAACCCCTCCCGCACACAAGAACATATACTAATTTACTTCCACCTA 11700
  Y  M  F  I  I  T  Q  R  G  P  L  P  P  H  L  T  L  I  N  P  S  R  T  Q  E  H  M  L  I  Y  F  H  L
                                                     ┤├ tRNA-His→
CTTCCCTCACTCCTCCTTATCTCCAAACCAGAACTAATTCTAGGCTGAACATCCTGTAATTATAGTTTAACAAAAACATTAGATTGTGGCTCTAAAAATA 11800
 L  P  S  L  L  L  I  S  K  P  E  L  I  L  G  W  T  S  *
                      ┤├ tRNA-Ser(AGY)→                                                    ┤ ├ tRNA-Leu(CUN)–
AAAGCTAAACCCTTTTTAACTACCAAGAGAGGTATGGAACAAGGAAAACTGCTAATTTCTCCCGCCCATGGTTCGAATCCCTGGCTCCCTTATGCCTCAG 11900
                                                              ┤├ ND5→
AAAGATAATAGACATCTATTGGTCTTAGGAACCAAAAACTCTTGGTGCAACTCCAAGCTAAGGCTATGAACCCCCTAATTTTCAACACATCCTTCACCCT 12000
                                                                  M  N  P  L  I  F  N  T  S  F  T  L
AATTTTCTTAATTCTCTCATACCCATTAATCCTCCCCTTATTCTCCATCAGCTCCCCCTCCCCATCAACTCACATCACAACTGTAAAAACCTCCTTCTTC 12100
  I  F  L  I  L  S  Y  P  L  I  L  P  L  F  S  I  S  S  P  S  P  S  T  H  I  T  T  V  K  T  S  F  F
ATTAGCCTACTCCCCCTCCTCATTCTCCTAGACCAGGGCTTAGAATCCATTACAACCAATTGAACTTGAATTAGCTTAGGAACCTTTGACATTCATATCA 12200
 I  S  L  L  P  L  L  I  L  L  D  Q  G  L  E  S  I  T  T  N  W  T  W  I  S  L  G  T  F  D  I  H  I  S
GTTTTAAATTTGACACCTATTCCATTATTTTCACTCCCATCGCCCTCTACGTAACCTGGTCTATCCTAGAATTTGCCCTATGATACATACACTCAGACCC 12300
   F  K  F  D  T  Y  S  I  I  F  T  P  I  A  L  Y  V  T  W  S  I  L  E  F  A  L  W  Y  M  H  S  D  P
CAACCTTAACAAATTCTTCAAATATCTCCTTCTATTCTTAATCACAATACTTATTCTCATCACAGCAAATAATCTTTTCCAACTATTCATCGGCTGAGAA 12400
  N  L  N  K  F  F  K  Y  L  L  L  F  L  I  T  M  L  I  L  I  T  A  N  N  L  F  Q  L  F  I  G  W  E
GGCGTAGGCATTATATCATTTCTTCTCATCGGCTGATGACTTAGTCGGGCCGACGCAAACACAGCTGCCCTTCAAGCTATCATCTACAACCGCATTGGCG 12500
 G  V  G  I  M  S  F  L  L  I  G  W  W  L  S  R  A  D  A  N  T  A  A  L  Q  A  I  I  Y  N  R  I  G  D
ATATCGGCCTGATCACAGCCATAGCATGACTAGCTATAAACCTCAACTCATGAGAAATTCAACAACTCTTCTTCCTCTCCAAAAATACAGATTTAACTCT 12600
   I  G  L  I  T  A  M  A  W  L  A  M  N  L  N  S  W  E  I  Q  Q  L  F  F  L  S  K  N  T  D  L  T  L
CCCACTTTTAGGTTTAGTTCTAGCAGCAGCAGGAAAATCCGCCCAATTTGGGTTACACCCATGACTTCCTGCCGCCATAGAAGGCCCCACACCAGTTTCC 12700
  P  L  L  G  L  V  L  A  A  A  G  K  S  A  Q  F  G  L  H  P  W  L  P  A  A  M  E  G  P  T  P  V  S
GCTTTACTCCACTCAAGCACAATAGTTGTCGCAGGGATTTTCTTACTTATCCGTCTTCACCCCCTAATTCAAGATAATCAACTAGTTCTATCCGCCTGCC 12800
 A  L  L  H  S  S  T  M  V  V  A  G  I  F  L  L  I  R  L  H  P  L  I  Q  D  N  Q  L  V  L  S  A  C  L
TATGTTTAGGTGCACTAACCACCCTTTTTACAGCAACCTGTGCACTTACCCAAAACGACATTAAAAAAATTATTGCTTTCTCTACATCCAGTCAACTAGG 12900
   C  L  G  A  L  T  T  L  F  T  A  T  C  A  L  T  Q  N  D  I  K  K  I  I  A  F  S  T  S  S  Q  L  G
ACTTATAATGGTCACTATCGGCCTTAACCAACCCCAATTAGCCTTCCTCCACATCTGCACACATGCCTTCTTTAAAGCAATACTCTTCCTCTGCTCAGGT 13000
  L  M  M  V  T  I  G  L  N  Q  P  Q  L  A  F  L  H  I  C  T  H  A  F  F  K  A  M  L  F  L  C  S  G
TCCATCATTCACAGCCTCAACGACGAACAAGACATTCGAAAAATAGGAGGCATACACAAACTTCTTCCCTTTACCTCCTCATCCCTAACAGTTGGTAGCC 13100
 S  I  I  H  S  L  N  D  E  Q  D  I  R  K  M  G  G  M  H  K  L  L  P  F  T  S  S  S  L  T  V  G  S  L
TAGCTCTCACTGGTATGCCCTTCTTATCCGGATATTTCTCTAAAGATGCCATCATTGAAGCAATAAACACATCACACCTTAACGCCTGAGCCCTAACTTT 13200
   A  L  T  G  M  P  F  L  S  G  Y  F  S  K  D  A  I  I  E  A  M  N  T  S  H  L  N  A  W  A  L  T  L
AACCCTTCTAGCTACGTCCTTCACCGCTATCTATAGCTTACGCCTAACATCCTTTTCACTCATGAACTACCCACGATTCACACCCCTTTCACCCATCAAC 13300
  T  L  L  A  T  S  F  T  A  I  Y  S  L  R  L  T  S  F  S  L  M  N  Y  P  R  F  T  P  L  S  P  I  N
```

```
GAAAATAACCCCCTACTCATTAACCCAATCAAACGACTCGCCTACGGAAGCATTCTAGCAGGCCTAATCATCACCTCTAACATACCACCAACTAAAACAC 13400
E  N  N  P  L  L  I  N  P  I  K  R  L  A  Y  G  S  I  L  A  G  L  I  I  T  S  N  M  P  P  T  K  T  Q
AAATCATAACAATAACACCTCTCCTTAAACTCTCCGCCCTCCTAGTAACAATCCTCGGCCTTATTTTAGCCCTAGAATTGACAAACCTAACCTCCACCCA 13500
  I  M  T  M  T  P  L  L  K  L  S  A  L  L  V  T  I  L  G  L  I  L  A  L  E  L  T  N  L  T  S  T  Q
ACTCAAAACCCACCCAAACCTATCCTACCATAACTTCTCTAACATACTAGGCTATTTCCCCTCAATTCTTCACCGACTCCCACCAAAACTCAGCTTATCT 13600
 L  K  T  H  P  N  L  S  Y  H  N  F  S  N  M  L  G  Y  F  P  S  I  L  H  R  L  P  P  K  L  S  L  S
TACGCCCAAACCGTTTCAACCCAAATAATTGATCTCTCATGAAACGAAAAAATCGGCCCAAAAAGCCCTATTATTCAACAAACATTCTTAATTAAATTAT 13700
Y  A  Q  T  V  S  T  Q  M  I  D  L  S  W  N  E  K  I  G  P  K  S  P  I  I  Q  Q  T  F  L  I  K  L  S
                                                                                     ├  ┤
CAACCTCCCCTCAACAAGGCTTAATCAAAACATACCTTCTCCTACTTCTCCTTACCCTTACCTTATCCTCAGCCCTCACCTTACTTTAAACCGTCCGCAA 13800
  T  S  P  Q  Q  G  L  I  K  T  Y  L  L  L  L  L  L  T  L  T  L  S  S  A  L  T  L  L  *
                                                                                       *  V  T  R  L
AGCCCCTCAAGATATCCCTCGCGTCAACTCCAACACCACAAACAAAGTCAAAAGCAACATTCAACCCCCTAAAACTAAAACTCCACTCCCTCAAGAATAC 13900
  A  G  W  S  M  G  R  T  L  E  L  V  V  F  L  T  L  L  L  M  W  G  G  L  V  L  V  G  S  G  W  S  Y
AATAACGCCACCCCACTGAAATCCCCACGAACCATGCTCATTTCTATAGACTCCTCACCTCCCACTCAACCCCCTCAACTTCAATCAACCACAAAATTTA 14000
 L  L  A  V  G  S  F  D  G  R  V  M  S  M  E  M  S  E  E  G  G  V  W  G  G  W  S  W  D  V  V  F  N
AACCCACTCCAAAAAGAATTCCAAAATACCCAACCACATAAGCTAAAACAGACCAATCACCCCACGACTCAGGATAAGGCTCAGCAGCTAACGCCGCCGT 14100
L  G  V  G  F  L  I  G  F  Y  G  V  V  Y  A  L  V  S  W  D  G  W  S  E  P  Y  P  E  A  A  L  A  A  T
ATAAGCAAACACAACTATTATCCCCCCCAAATAAATTAAAAACAAAATCAAAGACAAAAAAGAACTCCCAAATCCAACCAATAAACCACAACCCACCCCA 14200
  Y  A  F  V  V  M  M  G  G  L  Y  I  L  F  L  I  L  S  L  F  S  S  G  F  G  V  L  L  G  C  G  V  G
GCAGAAATTACTAGCCCTAGCGCAGCAAAATAAGGAGAAGGATTCGATGCCACAGCTACTAAGCCTAAAACAAAACTCACTATTAAAATCAACATTACAT 14300
 A  S  I  V  L  G  L  A  A  F  Y  P  S  P  N  S  A  V  A  V  L  G  L  V  F  S  V  M  L  I  L  M  V
 ← ND6 ┤├ tRNA-Glu→                                                        ┤     ├ cyt b→
ACATCATTATTCCTACTTAGACTCTAACTAAGACCTTTAATCCGAAAAACTACCGTTGTTATTCAACTATAGAAACCCTTAATGACTACTACAAACACCC 14400
Y  M  M                                                                          M  T  T  T  N  T  R
GCAAAACCCATCCCCTATTCAAAATTATCAATAACTCCGTAATTGACCTCCCAACTCCAACTAACATCTCTGCCTGATGAAACTTCGGCTCACTTGTAGG 14500
  K  T  H  P  L  F  K  I  I  N  N  S  V  I  D  L  P  T  P  T  N  I  S  A  W  W  N  F  G  S  L  V  G
CCTATGCTTAATTATCCAAATCCTGACAGGTCTCTTCCTAACTATACATTACACCGCAGACATCTCAACCGCATTCTCCTCAGTAGCACATATCTGCCGA 14600
 L  C  L  I  I  Q  I  L  T  G  L  F  L  T  M  H  Y  T  A  D  I  S  T  A  F  S  S  V  A  H  I  C  R
GACGTAAACTACGGCTGAATAATCCGCAACCTTCACGCCAATGGCGCCTCCCTCTTTTTCATCTGCGTTTACCTCCACATCGCCCGAGGACTTTACTACG 14700
D  V  N  Y  G  W  M  I  R  N  L  H  A  N  G  A  S  L  F  F  I  C  V  Y  L  H  I  A  R  G  L  Y  Y  G
GCTCCTACCTAAACAAAGAAACCTGAAATATTGGGGTAATTATCCTCTTACTCCTTATAGCCACTGCCTTCGTAGGATACGTCCTCCCATGAGGACAAAT 14800
  S  Y  L  N  K  E  T  W  N  I  G  V  I  I  L  L  L  L  M  A  T  A  F  V  G  Y  V  L  P  W  G  Q  M
ATCATTCTGAGGCGCAACCGTCATCACCAACTTACTATCCGCCCTCCCTTATATTGGAGACACACTCGTTCAATGACTCTGAGGTGGCTTCTCAATCGAC 14900
 S  F  W  G  A  T  V  I  T  N  L  L  S  A  L  P  Y  I  G  D  T  L  V  Q  W  L  W  G  G  F  S  I  D
AATGCAACACTAACCCGATTCTTTACATTCCACTTCCTCCTCCCCTTCCTAATCGCAGCCCTAACTATAGTCCACCTTCTCTTCCTCCACACCTCCGGTT 15000
N  A  T  L  T  R  F  F  T  F  H  F  L  L  P  F  L  I  A  A  L  T  M  V  H  L  L  F  L  H  T  S  G  S
CAAATAACCCAACTGGTCTTCCCTCTGACATAGACAAAATCCCATTTCACCCCTACTACTCATTTAAAGACCTCCTCGGCTTCTTTCTCCTCCTGCTCCT 15100
  N  N  P  T  G  L  P  S  D  M  D  K  I  P  F  H  P  Y  Y  S  F  K  D  L  L  G  F  F  L  L  L  L  L
ACTTACCCTCTTAGCCTTACTCACACCCAACCTCCTAACCGACACAGAAAACTTTATTCCAGCCAACCCTCTCGTCACACCCCCACACATCAAACCAGAG 15200
 L  T  L  L  A  L  L  T  P  N  L  L  T  D  T  E  N  F  I  P  A  N  P  L  V  T  P  P  H  I  K  P  E
TGGTACTTTCTATTCGCTTACGCCATTCTACGCTCTATCCCCAACAAACTAGGAGGCGTACTTGCCCTTGCCCTTTCAATCCTAATCCTCTTCTTAGTCC 15300
W  Y  F  L  F  A  Y  A  I  L  R  S  I  P  N  K  L  G  G  V  L  A  L  A  L  S  I  L  I  L  F  L  V  P
CCATTCTCCATACTTCTAAACTACGAAGCCTAACCTTTCGCCCGATCACACAGCTTCTATTCTGGCTCTTAGTGGCCAATACCCTTATCCTAACATGAAT 15400
  I  L  H  T  S  K  L  R  S  L  T  F  R  P  I  T  Q  L  L  F  W  L  L  V  A  N  T  L  I  L  T  W  I
CGGAGGACAACCCGTAGAACAACCATTTATTATCATTGGCCAAATCGCCTCCATCACCTATTTCTCCTTCTTTCTCCTCCTTTTTCCAATCGCCGGTTGA 15500
 G  G  Q  P  V  E  Q  P  F  I  I  I  G  Q  I  A  S  I  T  Y  F  S  F  F  L  L  L  F  P  I  A  G  W
                          ┤      ├ tRNA-Thr→
TGAGAAAACAAAATACTAAACCTTTAACCCAATGCTCTGGTGGCCCAGCTATTTAAGGCATCGGTCTTGTAAACCGAAAATCGGAGGTGAAATTCCTCCC 15600
W  E  N  K  M  L  N  L  *
      ┤         ├ tRNA-Pro→                                                        ┤├ Control region→
CAGTGCAAAAGCACTTTCAAGAAAAAGAGGACAAACCCTTATCCTTGGCTCCCAAAGCCAAGATTTTTATTAAACTATTTCCTGAACTCTATTAGAAAAT 15700
AGTCGACTAAATATGACGTTAGTTGACTAGATCACATATAATTATTAACCCCCATAATGCCTAAAATACATAATACATTACATAAATTAATCCCCATATA 15800
                                        Repeat 1                                 Repeat2
TTGTCATATAATACATAAGACAGTCTATGCTTAATCCTCATACATCTATATACCCCTATATCATAACATATCTATGCTTAATCCTCATACATCTATATAC 15900
                   Repeat 3 (partial)
CCCTATATCATAACATATCTATGCTTAATCCTCATTCATCTACATTCCCCTATTTCATTACCCACTCAACTCCACACTAACAGATTCTATACCCTAATCT 16000
ATTTATCCACCAACAGCCTCTCAATTATTATCTAACACGTCCCCATTAACAGATAACCTCCCAAACCTATATTAACCATACTGCTTAATCAACATATCAA 16100
                                                           TAS
GATATTCTATTATTTCATAGCATATATAACCTTAATCTACCATAACTTGACATTCACATACTTCATAACATATTCTCATGCTTAATCACCTATTATTGAT 16200
TAACTCCCACTACCGTACTGGTTAATTTCACTATTAATCATCTATTCACCTTCTGTCTTATCAGCTAGTGATTAATACTCCTTTACTTTCCAGTCCTCTC 16300
CAAATAATCTATAACTTATCAATGTTTAATCAATATTAATCGATAATATAACTATCTATGCTTAATCAGCATTAATCGACATTCCCCTATTTCATAACAT 16400
                                                                                           CSB-D
CAGATTCAAACATTAACACATTTCTATTATTCTCCGATAATTCTTCTGGTTGGCGTGAAATAATCAACGATCTAATAGAATACACCTTTGTTCGATTTGT 16500
GGCACAGAAATAGATTTATCCGCATAAACTGATCAGACCGACATTTGGCACCCTTAACACGCATACGTCTCTTAATCGCGTCAGAACTCACTTTCCTCTA 16600
GTTCCCTTGTAAGGCACAACTTGTCGACCCAATCCTGCTACGCAGCGGATCGGGTCCTACTAAGATTTCCAGTCCGCTAAGTCTTTTTAGTTTTCGTTGA 16700
CTTCATCCACTATTTCTTCGTTACGGTACTGTCTGATAGGATGGTAATGACCTCGACTACAAGTTAATTGATTAGCTACTTCTTTCATTCACTGATCACT 16800
CGGTCCAGAGGACTACCGTTGAAATTATCAAAAGGTTATAGTTGAGAATTTAATCAATATTCTATAATTAATTGAGCAATTAATCTAAAAAAGACATGTA 16900
AAAATTTTTCATAAGAATATTAATGATTAATCTTTGATAAAATAGGAATAACAATGAAGCCAAGTCTTAATTCTGAAGCTAAGATAGGTGCCCCCCCTGT 17000
                                                CSB II                                 CSB III
GCGGGCGTGCGTAAAAGATATTTATACAAGAAATTTTTTGGGAAAAACCCCCCCTCCCCCCAAAAAAACCCGGTTATCCTCGAAAAACCCCTAAAACGAG 17100
GGCCAAAGGAAACTTTTTTTATGCCATAATAATCCAAAAGAAATTTTTTATTTAAAACAAAAAACTTTTATATTATTACAACACTGGCAAAACGAAATTA 17200
ACTTGGGGAAAAAAATGAACTACTCATGATGATTTAAAAAAAAAAACAAAAAGATTACCGTGTCCAGGACATTATCTGGATCTATTATGTGAAAAATTAC 17300
TTTTTTTTGCCCTAATATACCTGTTATATTTAAATTTTTTTTACTGGTTGTGGGCGTGTGTGTGATGCATACATAGTGTCAAACACACATGTGTATTACC 17400
TGAATGTGCCTACATTATACACACAAATATGCATATTTTTGCACAAAAAACACACACATACACACATTTTCTCCACCATATCCTAACTTCACTTCAAATC 17500
CTTCCACTATGATC                                                                                       17514
```

Fig. 3 The complete L-strand nucleotide sequence of the deepwater stingray mitochondrial genome. Position 1 corresponds to the first nucleotide of the tRNA Phe gene. The direction of transcription for each gene is shown by arrows. The beginning and end of each are indicated by vertical bar (├ and ┤). Transfer RNA genes are boxed; corresponding anticodons are indicated in black boxes. Amino acid sequences presented below the nucleotide sequence were derived using mammalian mitochondrial genetic code (one latter amino acid abbreviations placed below the first nucleotide of each codon). Stop codons are overlined in gray boxes and indicated by asterisks. Non-coding sequences are underline with dots. TAS, putative termination-associated sequence; CSB2, 3, and D, conserved sequence blocks. Sequence data are available from DDBJ/EMBL/GenBank with accession number.

Table 2 Comparisons of lengths (bp) of 8 cartilaginous fish mitochondrial genomes

Subclass	Holocephali	Elasmobranchii	Elasmobranchii	Elasmobranchii	Elasmobranchii	Elasmobranchii	Elasmobranchii	Elasmobranchii
Order	Chimaeriformes	Carcharhiniformes	Carcharhiniformes	Heterodontiformes	Squaliformes	Myliobatiformes	Rajiformes	Rajiformes
Family	Chimaeridae	Scyliorhinidae	Triakidae	Heterodontidae	Squalidae	Plesiobatidae	Rajidae	Rajidae
Species name	*Chimaera monstrosa*	*Scyliorhinus canicula*	*Mustelus manazo*	*Heterodontus francisci*	*Squalus acanthias*	*Plesiobatis daviesi*	*Raja radiata*	*Raja porosa*
Common name	rabbitfish	small-spotted catshark	starspotted smooth-hound	horn shark	spiny dogfish	deepwater stingray	thorny skate	ocellate spot skate
Accession number	AJ_310140	Y_16067	AB_015962	AJ_310141	Y_18134	AY_597334	AF_106038	AY_525783
12S rRNA	949	957	952	952	951	960	967	965
16S rRNA	1665	1673	1669	1676	1676	1689	1678	1676
ND1	972	975	974	976	975	978	975	975
ND2	1044	1047	1046	1047	1047	1044	1047	1047
COI	1560	1554	1556	1557	1557	1551	1557	1557
COII	691	691	690	691	691	691	691	699
ATPase 8	168	168	167	168	168	168	168	168
ATPase 6	684	684	683	684	684	684	684	684
COIII	786	786	785	786	786	786	786	786
ND3	349	351	350	351	351	351	351	351
ND4L	297	297	296	297	297	297	297	297
ND4	1375	1381	1380	1381	1381	1387	1381	1387
ND5	1839	1830	1826	1830	1833	1824	1836	1836
ND6	522	522	521	510	522	522	519	519
cyt b	1144	1144	1145	1146	1146	1146	1143	1143
Control region	1602, 1503	1050	1067	1068	1080	1830	1064	1310
Total	18580	16697	16707	16708	16738	17514	16783	16972
Reference	Arnason *et al.* (2001)	Delarbre *et al.* (1998)	Cao et al. (1998)	Arnason *et al.* (2001)	Rasmussen *et al.* (1999)[1]	present study	Rasmussen *et al.* (1999)[2]	Kim *et al.* (2005)

3. Ribosomal RNA genes

The length of 12S and 16S rRNA genes of this deepwater stingray were 960 and 1689 bp, respectively (Table 1). They are located between the tRNA^{Phe} and tRNA$^{Leu(UUR)}$ genes and separated by the tRNA^{Val} gene (Figs. 2, 3).

4. Noncoding sequences

The L-strand replication origin in the deepwater stingray was located between the tRNA^{Asn} and tRNA^{Cys} genes, and it was 34 bp in length (Fig. 2), indicating that it has the potential to fold into a stem-loop secondary structure. This region has been observed in most cartilaginous fishes.

The major noncoding region found in the deepwater stingray was located between tRNA^{Pro} and tRNA^{Phe} and was up to 1830 bp in length, which was much longer than that of the sharks and other rays. Several unique characteristics were identified in the control region (Fig. 3), namely three conserved sequence blocks (CSB II, III, and D; Kawaguchi *et al.*, 2001), a termination-associated sequence, and the two copies of 47-bp repeat regions (repeat 1 and 2). Repeated sequences have also been observed in teleosts (Cecconi *et al.*, 1995; Lee *et al.*, 1995; Chen *et al.*, 1998; Chen *et al.*, 2002; Stefanni *et al.*, 2002). However, only the first 18 bp of the 3rd repeat region were identical to those of the 1st and 2nd repeat regions. Tandem repeats were located near the 5′ end of the control region. Such repeats may result from strand slippage of the third displaced DNA strand in the D-loop (Buroker *et al.*, 1990).

5. Phylogenetic relationships among other Myliobatoidei families

To determine the phylogenetic position among Myliobatoidei families, the 12S rRNA sequences of 13 fish species were retrieved from GenBank (Appendix). Sequence alignment resulted in 748 molecular characters. The same topology of

Table 3 Location of features in the mitochondrial genome of *Plesiobatis daviesi*

Gene	Position number		Size (bp)	Codon		Strand
	From	To		Start	Stop	
1 tRNA$^{\mathrm{Phe}}$	1	70	70			
2 12S rRNA	71	1030	960			
3 tRNA$^{\mathrm{Val}}$	1037	1108	72			
4 16S rRNA	1110	2798	1689			
5 tRNA$^{\mathrm{Leu(UUR)}}$	2799	2873	75			
6 ND1	2874	3851	978	ATG	TAA	
7 tRNA$^{\mathrm{Ile}}$	3854	3922	69			
8 tRNA$^{\mathrm{Gln}}$	3925	3996	72			L
9 tRNA$^{\mathrm{Met}}$	3998	4067	70			
10 ND2	4068	5111	1044	ATG	TAA	
11 tRNA$^{\mathrm{Trp}}$	5112	5181	70			
12 tRNA$^{\mathrm{Ala}}$	5183	5251	69			L
13 tRNA$^{\mathrm{Asn}}$	5253	5325	73			L
14 Rep-Origin	5326	5359	34			L
15 tRNA$^{\mathrm{Cys}}$	5360	5427	68			L
16 tRNA$^{\mathrm{Tyr}}$	5431	5501	71			L
17 CO Ⅰ	5503	7053	1551	GTG	TAA	
18 tRNA$^{\mathrm{Ser(UCN)}}$	7059	7129	71			L
19 tRNA$^{\mathrm{Asp}}$	7131	7197	67			
20 COⅡ	7202	7892	691	ATG	T--	
21 tRNA$^{\mathrm{Lys}}$	7893	7964	72			
22 ATPase 8	7966	8133	168	ATG	TAA	
23 ATPase 6	8124	8807	685	ATG	TAA	
24 COⅢ	8807	9592	786	ATG	TAA	
25 tRNA$^{\mathrm{Gly}}$	9593	9663	71			
26 ND3	9664	10014	351	ATG	TAG	
27 tRNA$^{\mathrm{Arg}}$	10013	10084	72			
28 ND4L	10085	10381	297	ATG	TAA	
29 ND4	10375	11755	1387	ATG	T--	
30 tRNA$^{\mathrm{His}}$	11756	11824	69			
31 tRNA$^{\mathrm{Ser(AGY)}}$	11825	11892	68			
32 tRNA$^{\mathrm{Leu(CUN)}}$	11894	11965	72			
33 ND5	11966	13789	1824	ATG	TAA	
34 ND6	13786	14307	522	ATG	TAA	L
35 tRNA$^{\mathrm{Glu}}$	14308	14376	69			L
36 cyt *b*	14382	15527	1146	ATG	TAA	
37 tRNA$^{\mathrm{Thr}}$	15534	15607	74			
38 tRNA$^{\mathrm{Pro}}$	15617	15684	68			L
39 Control Region	15685	17514	1830			

phylogenetic trees was obtained using the NJ, MP, and Bayesian methods with the rabbit fish *Chimaera monstrosa* as an outgroup. This topology is shown in Fig. 4 with bootstrap values. Phylogenetic analysis revealed that fish from the order Myliobatiformes were well separated according to taxonomic levels such as family, genus, and species. This appears to support the topology obtained in previous molecular studies (Dunn *et al*., 2003). Conflicting results regarding the phylogenetic relationships among Urolophidae, Plesiobatidae, and Hexatrygonidae have been reported by previous morphological studies.

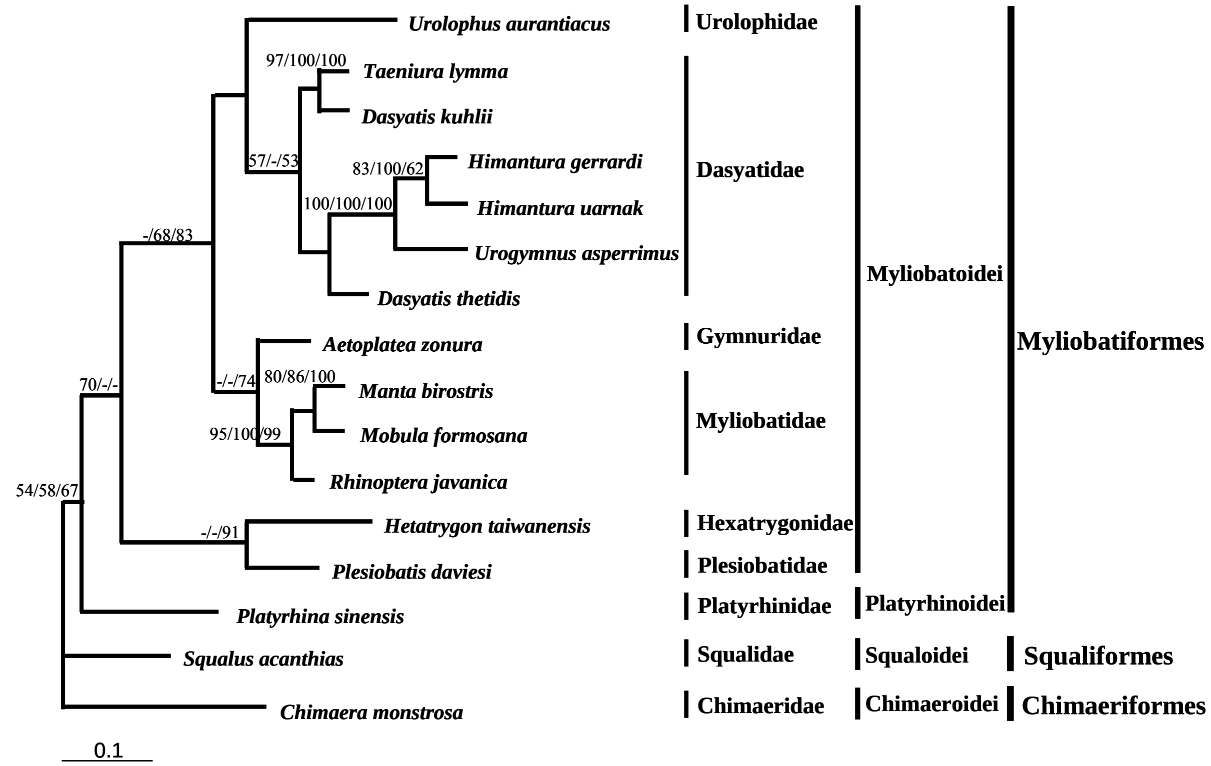

Fig. 4 Phylogenetic relationships among myliobatoidei families. The 50%-majority rule consensus of post-burn-in sampled trees from Bayesian inference analysis based on the 12S rRNA data set under GTR + I + G model is shown. Branch support values estimated by bootstrap pseudo-replicates in Bayesian inference, maximum parsimony and neighbor-joining, respectively, are shown above each branch. (-) indicates that either bootstrap value or Bayesian posterior probabilities below 50% in the analysis. Nodes with either BP or BPP below 50% are not numbered.

Nishida (1990) placed *Plesiobatis* and *Hexatrygon* as two basal lineages of Myliobatiformes, whereas McEachran *et al.* (1996) placed *Hexatrygon*, *Plesiobatis*, and *Urolophus* as a basal lineage. Based on the present study, *Plesiobatis* and *Hexatrygon* form a sister group, which supports the hypothesis proposed by Nishida (1990). Conversely, most morphological studies indicated a closer relationship between *Urolophus* and *Plesiobatis* (Nishida, 1990; McEachran *et al.*, 1996; McEachran and Aschliman, 2004). However, the sister relationship between *Urolophus* and *Plesiobatis* was not supported by the present study. Also in this study, the families Plesiobatidae and Urolophidae were separated from each other. The family Urolophidae was more closely related to Dasyatidae than to Plesiobatidae or Hexatrygonidae. This was in accordance with the hypothesis proposed by Nelson (2006). The molecular phylogeny of Myliobatoidei presented here is only tentative because several important genera were not included in our phylogenetic analyses. Moreover, it is important to clarify the placement of *Plesiobatis* within Myliobatiformes. Because no future resolution on the relative position of *Plesiobatis* was achieved, future work based on more complete mitochondrial genomes of Myliobatiformes will likely yield better resolved phylogenies.

ACKNOWLEDGEMENTS

The authors thank Mr. JH Wu for his kind help with data analysis and arrangements. We are very grateful to Dr. Lee-Shing Fang and Dr. Chi-Lun Wu for their encouragement and support. We are also grateful to the NSC (2002-2003) for partial support via research grants and to the National Museum of Marine Biology & Aquarium, Pingtung, Taiwan for biodiversity research funding.

APPENDIX

The 12 species of Myliobatoidei used for phylogenetic analysis in this study were as follows: *Urolophus aurantiacus* (AF448028); *Platyrhina sinensis* (AF448004); *Hexatrygon taiwanensis* (AF447995); *Aetoplatea zonura* (AF447986); *Rhinoptera javanica* (AF448019); *Manta birostris* (AF448000); *Mobula formosana* (AF448001); *Dasyatis kuhlii* (AF447991); *Dasyatis thetidis* (AF447993); *Urogymnus asperrimus* (AF448027); *Himantura gerrardi* (AF447996); *Himantura uarnak* (AF447997); *Squalus acanthias* (Y18134); *Chimaera monstrosa* (AJ310140).

REFERENCES

Arnason U., A. Gullberg and A. Janke (2001) Molecular phylogenetics of gnathostomous (Jawed) fishes: old bones, new cartilage. Zool. Scr., 30: 249-255.

Buroker, N. E., J. R. Brown, T. A. Gilbert, P. J. O'Hara, A. T. Beckenbach, W. K. Thomas and M. J. Smith (1990) Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. Genetics, 124: 157-163.

Cao, Y., P. Waddell, N. Okada and M. Hasegawa (1998) The complete mitochondrial DNA sequence of the shark *Mustelus manazo*: evaluating rooting contradictions to living bony vertebrates. Mol. Biol. Evol., 15: 1637-1646.

Cecconi, F., M. Giorgi and P. Mariottini (1995) Unique features in the mitochondrial D-loop region of the European seabass *Dicentrarchus labrax*. Gene, 165: 149-155.

Chen, I. S., C. H. Hsu, K. T. Shao, P. J. Miller and L. S. Fang (1998) Sequence length and variation in the mitochondrial control region of two freshwater gobiid fishes belonging to *Rhinogobius* (Teleostei: Gobioidei). J. Fish. Biol., 53: 179-191.

Chen, I. S., P. J. Miller, H. L. Wu and L. S. Fang (2002) Taxonomy and mitochondrial sequence evolution in non-diadromous species of *Rhinogobius* (Teleostei: Gobiidae) of Hainan Island, southern China. J. Mar. Freshwat. Res., 53: 259-273.

Delarbre, C., N. Spruyt, C. Delmarre, C. Gallut, V. Barriel, P. Janvier, V. Laudet and G. Gachelin (1998) The complete nucleotide sequence of the mitochondrial DNA of the dogfish, *Scyliorhinus canicula*. Genetics, 150: 331-344.

Dunn, K. A., J. D. McEachran and R. L. Honeycutt (2003) Molecular phylogenetics of myliobatiform fishes (Chondrichthyes: Myliobatiformes), with comments on the effects of missing data on parsimony and likelihood. Mol. Phylogenet. Evol., 27: 259-270.

Inoue, J. G., M. Miya, K. Tsukamoto and M. Nishida (2000) Complete mitochondrial DNA sequence of the Japanese sardine *Sardinops melanostictus*. Fish. Sci., 66: 924-932.

Inoue, J. G., M. Miya, J. Aoyama, S. Ishikawa, K. Tsukamoto and M. Nishida (2001a) Complete mitochondrial DNA sequence of the Japanese eel *Anguilla japonica*. Fish. Sci., 67: 118-125.

Inoue, J. G., M. Miya, K. Tsukamoto and M. Nishida (2001b). Complete mitochondrial DNA sequence of the Japanese anchovy *Engraulis japonicus*. Fish. Sci., 67: 828-835.

Inoue, J. G., M. Miya, K. Tsukamoto and M. Nishida (2001c). Complete mitochondrial DNA sequence of *Conger myriaster* (Teleostei: Anguilliformes): novel gene order for vertebrate mitochondrial genomes and the phylogenetic implications for Anguilliform families. J. Mol. Evol., 52: 311-320.

Kawaguchi, A., M. Miya and M. Nishida (2001) Complete mitochondrial DNA sequence of *Aulopus japonicus* (Teleostei: Aulopiformes), a basal Eurypterygii: longer DNA sequences and higher-level relationships. Ichthyol. Res., 48: 213-223.

Kim, I. C., S. O. Jung, Y. M. Lee, C. J. Lee, J. K. Park and J. S. Lee (2005) The complete mitochondrial genome of the rayfish *Raja porosa* (Chondrichthyes, Rajidae). DNA Seq., 16: 187-194.

Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol., 16: 111-120.

Kumar, S., K. Tamura and M. Nei (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform., 5: 150-163.

Lee, W. J., J. Conroy, W. H. Howell and T. D. Kocher (1995) Structure and evolution of teleost mitochondrial control region. J. Mol. Evol., 41: 54-66.

McEachran, J. D., K. A. Dunn and T. Miyake (1996) Interrelationships within the batoid fishes (Chondrichthyes: Batoidea). *In* Interrelationship of Fishes (M. L. J. Stiassny, L. R. Parenti and G. D. Johnson eds.), Academic Press, San Diego, CA, USA, 63-84.

McEachran, J. D. and N. Aschliman (2004) Phylogeny of Batoidea. *In* Biology of Sharks and their Relatives (J. C. Carrier, J. A. Musick and M. R. Heithaus eds.), CRC Press, Boca Raton, FL, USA, 79-114.

Miya, M. and M. Nishida (1999). Organization of the mitochondrial genome of a deep-sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): first example of transfer RNA gene rearrangements in bony fishes. Mar. Biotechnol., 1: 416-426.

Miya, M. and M. Nishida (2000) Use of mitogenomic information in teleostean molecular phylogenetics: a tree-based exploration under the maximum-parsimony optimality criterion. Mol. Phylogenet. Evol., 17: 437-455.

Miya, M., A. Kawaguchi A and M. Nishida (2001) Mitogenomic exploration of higher teleostean phylogenies: a case study for moderate-scale evolutionary genomics with 38 newly determined complete mitochondrial DNA sequences. Mol. Biol. Evol., 18: 1993-2009.

Nei, M. and S. Kumar (2000) Molecular Evolution and Phylogenetics. Oxford Univ. Press, New York, USA, 1-333.

Nelson, J. S. (2006) Fishes of the World (4th ed.) New Jersey, USA: John Wiley and Sons, 76–82.

Nishida, K. (1990) Phylogeny of the suborder Myliobatidoidei. Memoirs of the Faculty of Fisheries, Hokkaido Univ., 37: 259-298.

Page, R. D. M (1996) Tree View: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci., 12: 357-358.

Posada, D. and K. A. Crandall (1998) Modeltest: testing the model of DNA substitution. Bioinformatics, 14: 817-818.

Rasmussen, A. and U. Arnason (1999a) Molecular studies suggest that cartilaginous fishes have a terminal position in the piscine tree. P. Natl. Acad. Sci. USA, 96: 2177-2182.

Rasmussen, A. and U. Arnason (1999b). Phylogenetic studies of complete mitochondrial DNA molecules place cartilaginous fishes within the tree of bony fishes. J. Mol. Evol., 48: 118-128.

Ronquist, F. and J. P. Huelsenbeck (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19: 1572-1574.

Stefanni, S., I. S. Chen and P. J. Miller (2002) A very compact mt-DNA control region in the widely distributed goby *Pomatoschistus minutus* (Teleostei: Gobiidae). J. Mar. Biol. Assoc. UK, 82: 1001-1003.

Swofford, D. L. (2002) PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4.b10. Sinauer Associates, Sunderland, MA.

達氏深水尾魟（*Plesiobatis daviesi*）之完整粒線體 DNA 序列分析

蕭聖代[1]・徐基新[2]・劉燈城[3]・陳義雄[4*]

[1]行政院農業委員會水產試驗所海洋漁業組
[2]國立中山大學海洋生物科技暨資源學系
[3]行政院農業委員會水產試驗所
[4]國立台灣海洋大學海洋生物研究所

摘　要

本研究利用 Long PCR 及 primer walking 的方式將達氏深水尾魟完整粒線體 DNA 完成定序，並分析基因排序與組成。達氏深水尾魟完整粒線體 DNA 全長 17514 個鹼基對，共由 37 個基因所組成，包含 2 個核糖核酸基因 (rns、rnl)、13 個蛋白質基因 (atp6、atp8、cox1-3、cob、nad1-6、nad4L) 及 22 個傳遞者核酸基因 (tRNA)，與其他脊椎動物相似。達氏深水尾魟與其他 7 種軟骨魚類的完整粒線體基因組的排序上並無明顯不同，但是達氏深水尾魟其介於 $tRNA^{Pro}$ 和 $tRNA^{Phe}$ 之間的主要非編碼區，又稱 D-loop，是目前已知的軟骨魚序列中最長的，共 1830 個鹼基對。而這段序列包含了兩段 47 個鹼基對的重複性片段，以及兩段具有保守性特徵的片段。建構其親緣關係樹後發現，達氏深水魟在演化上是與六鰓魟科為姊妹群，並且屬於燕魟亞目的基群。

關鍵詞：粒線體基因組，Long PCR，深水尾魟（*Urotrygon daviesi*），達氏深水尾魟（*Plesiobatis daviesi*）

---

*通訊作者／國立海洋大學，基隆市 202 北寧路 2 號，TEL: 886-2-2462-2192 ext. 5320; Fax: 886-2-2463-3152; E-mail: isc@mail.ntou.edu.tw