

Complete mitochondrial genome profile of Randall's threadfin bream, *Nemipterus randalli* and its phylogenetic analysis

Neenu Raj^{1,2,*}, Sandhya Sukumaran¹, Lakshmi P. Mukundan¹, Anjaly Jose¹, Sujitha Mary¹, K. Nisha¹ and A. Gopalakrishnan¹

¹Marine Biotechnology Fish Nutrition and Health Division, ICAR-Central Marine Fisheries Research Institute, Ernakulam North P.O., Kochi 682 018, India

²Mangalore University, Mangalagangothri, Mangaluru 574 199, India

We characterized the complete mitogenome of *Nemipterus randalli* and performed comparative mitogenomic analysis with previously determined mitochondrial genomes belonging to the family Nemipteridae. Using Sanger sequencing, we identified a 16,642 bp mitogenome containing 37 genes (13 protein-coding, 2 rRNA, 22 tRNA) and a non-coding region. Its gene organization, nucleotide composition, tRNA secondary structure and codon usage are similar to other Nemipteridae mitogenomes. The phylogenetic analysis indicated that Nemipteridae is a monophyletic group. The present study forms the basis for further studies on the population genetics, evolution and phylogeny of *N. randalli* and its family.

Keywords: Mitogenome, *Nemipterus randalli*, nemipteridae, phylogeny, protein coding genes.

THE family Nemipteridae consists of economically important fishes in the tropical Indo-West Pacific region comprising five genera and 62 species. They are small to medium-sized and inhabit shallow mud and sandy bottoms¹. Nemipterid fish (threadfin breams) are commercially important as a food source² and account for 8.4% of the total landings³. Randall's threadfin bream, *Nemipterus randalli* formed the dominant species (54%) in demersal finfish landings and contributed significantly to the trawl fisheries of India³. It is widespread in Indian waters, Pakistan, the Persian Gulf, the Red Sea, the Gulf of Aden, the East African coast, the Seychelles, and Madagascar¹. The International Union for Conservation of Nature (IUCN) classifies the status of this species as 'Least concern'⁴.

Several studies have been conducted to describe the phylogenetic relationship among Nemipteridae, but they have largely depended on mitochondrial and nuclear markers⁵. The rapid development of high-throughput sequencing (HTS) technology has revolutionized the sequencing and characterization of whole mitogenomes. This allows for the efficient sequencing of large numbers of mitogenomes. In contrast to morphological traits, the mitogenome also plays a

crucial role in deciphering the phylogenetic relationship between closely related species⁶. Phylogenetic analysis based on the full mitogenome offers higher resolution than partial mitochondrial gene sequences⁷. Mitochondrial DNA (mt-DNA) is typically circular with a size of about 16–17 kb and encodes 13 protein coding genes (PCGs) as well as 22 tRNAs, 2 rRNAs and a control region⁸. Mitochondrial DNA plays a crucial role in evolutionary and population genetics⁹ due to its haploid nature, high mutation rate, and lack of genetic recombination¹⁰. Understanding the profile of the mitogenome is very important as it provides clues to the species taxonomy, phylogeny, adaptation mechanisms and evolutionary consequences in different habitats¹¹. Among 62 species belonging to the Nemipteridae family, only six complete mitogenomes are still available. Out of these six available mitogenomes, two are from the genus *Scolopsis* and the remaining from the genus *Nemipterus*.

Despite its commercial importance, there is a great paucity of genetic and phylogenetic information, especially for *N. randalli*¹². To bridge this resource gap, we sequenced the complete mitochondrial genome of *N. randalli* using Sanger sequencing (using universal primers + primer pairs designed for each region). We analysed the gene content, nucleotide composition, codon usage patterns of PCGs, the secondary structure of tRNAs, and control region (CR). We also performed a comparative mitogenomic analysis with other species of the family, Nemipteridae, to check the structure and variation within the family. In addition, divergence times was estimated to infer the evolutionary relationship of *N. randalli* compared to other closely related species.

Materials and methods

Sample collection and DNA extraction

The specimen of *N. randalli* was collected at Thoppumpady, Cochin (9.9356°N, 76.2583°E), India. A piece of muscle tissue from *N. randalli* stored in 90% ethanol was used for

*For correspondence. (e-mail: neenuraj993@gmail.com)

DNA extraction using the standard phenol-chloroform method¹³. The quality and quantity of DNA were measured with Nanodrop (Thermo Scientific, Waltham, MA, USA). The purified DNA was visualized on 1.0% agarose gel and stored at 4°C. The fish samples used in this study were handled according to the guidelines for the care and use of fish in research by DeTolla *et al.*¹⁴. The protocols were approved by the Ethics Committee of ICAR-Central Marine Research Institute, Kochi. These methods are also reported in accordance by the ARRIVE guidelines (<http://arrive-guidelines.org>).

PCR amplification and sequencing

Initial amplification was performed using universal primers for *ATPase 6/8* genes¹⁵, *16S rRNA*¹⁶, *COXI*¹⁷, *Cyt b*¹⁸ and displacement loop (D-loop) region¹⁹. Internal primers were designed using Primer3 (ref. 20) based on the aligned mitochondrial genome sequence of three closely related species: *Nemipterus virgatus*, *Nemipterus japonicus* and *Nemipterus bathybius* ([Supplementary Table 1](#)). Polymerase chain reaction (PCR) reactions were performed in a Biorad T100 thermal cycler (Biorad, USA) with a total reaction volume of 25 µl containing 10× PCR buffer (100 mM Tris-HCl, pH 8.3, 500 mM KCl, 15 mM MgCl₂ and 0.01% gelatin, Sigma Aldrich, USA), 1.5 KU of *Taq* DNA polymerase (Sigma Aldrich), 200 µM of each dNTPs (dATPs, dCTPs, dGTP, dTTPs) (Sigma Aldrich), 0.5 µmol of each primer and 20 ng of genomic DNA. The cycling profile was designed to have an initial denaturation of 94°C for 4 min, followed by 32 cycles of 94°C for 30 sec, annealing for 30 sec (at varying annealing temperatures), extension at 72°C (for varying time) and a final extension at 72°C for 10 min. The PCR amplified product was purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany), separated on 1.2% agarose gel, stained with ethidium bromide and visualized by a gel documentation system (Biorad Gel Doc XR). PCR products were sequenced from both directions using Sanger sequencing by an ABI 3730 automated sequencer. All sequences were checked manually against their chromatogram using the program Seqman from the software DNASTar Lasergene 6.0 (DNASTAR Inc., Madison, WI, USA).

Sequence analysis

The assembled mitochondrial genome was annotated using the MITOS web server²¹ and MitoAnnotator²². The circular map of the entire mitochondrial genome was drawn with OrganellarGenomeDRAW (OGDRAW)²³. Relative Synonymous Codon Usage (RSCU), nucleotide composition, and codon usage were analysed with MEGA X²⁴. Strand asymmetry (skewness) was calculated using the formula GC skew = $(G-C)/(G+C)$ and AT skew = $(A-T)/(A+T)$ ²⁵, where positive AT skew indicated the presence of more As than

Ts and negative GC skew indicated the presence of more Cs than Gs. The ratio of synonymous (Ks) to non-synonymous substitution (Ka) was estimated using DnaSP v6 (ref. 26). Non-synonymous (Ka) and synonymous (Ks) ratios were estimated to evaluate selective pressures and evolutionary relationships of species. $Ka/Ks > 1$ indicated the occurrence of positive (adaptive or diversifying) selection, $Ka/Ks < 1$ negative (purifying) selection and $Ka/Ks = 1$ neutral mutation²⁷. The overall p-genetic distance of 13 PCGs among species of the Nemipteridae family, including *N. randalli*, *N. bathybius*, *N. japonicus*, *N. virgatus*, *Nemipterus hexadon*, *Scolopsis ciliata* and *Scolopsis vosmeri* was assessed to evaluate the conservation of mitochondrial DNA genes. The secondary structure of tRNA genes was predicted by tRNA-scan-SE1.21 (ref. 28) and ARWEN²⁹ using the default search mode. The secondary structure of the origin of the light strand (O_L) was drawn with RNA structure³⁰. The presence of tandem repeats in the non-coding region was examined with Tandem Repeats Finder³¹.

Phylogenetic analysis

A total of 51 mitogenomes belonging to the 6 closely related families (Caesionidae, Haemulidae, Lethrinidae, Lutjanidae, Nemipteridae, Sparidae + Centranchidae) available in GenBank were used to investigate the phylogenetic relationships among these families. *Neotrygon kuhlii* (GenBank accession no: NC_021767.1) was selected as an outgroup. The 13 PCGs were aligned by Clustal W and concatenated without stop codon for constructing the phylogenetic tree via maximum likelihood (ML) and Bayesian inference (BI) methods using RaxML v 8.2.10 (ref. 32) and MrBayes v 3.2.4 (ref. 33) respectively. In the maximum likelihood analysis, the tree was constructed using the GTRGAMMA model with 1000 bootstrap replicates. Using MEGA X, GTR+G+I was selected as the best model for the BI method based on AIC (Akaike Information Criterion). For the Bayesian analysis, 1,000,000 generations were performed using the Markov chain Monte Carlo (MCMC) method. The Bayesian tree was run until the average standard deviation of the split frequency was less than 0.01. The first 25% of the sample data was discarded as burn-in, and the rest was used to generate a consensus tree. The resulting phylogenetic tree was viewed using FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>). Comparison of the *N. randalli* mitogenome with other mitogenomes of the Nemipteridae family was performed using the CG View comparison tool (CCT) to draw a circular genetic similarity map of the Nemipteridae family³⁴.

Divergence times

The divergence times of 51 species belonging to 6 closely related families were estimated with MEGA X (ref. 24) using the RelTime-ML method with GTR+G+I modelling.

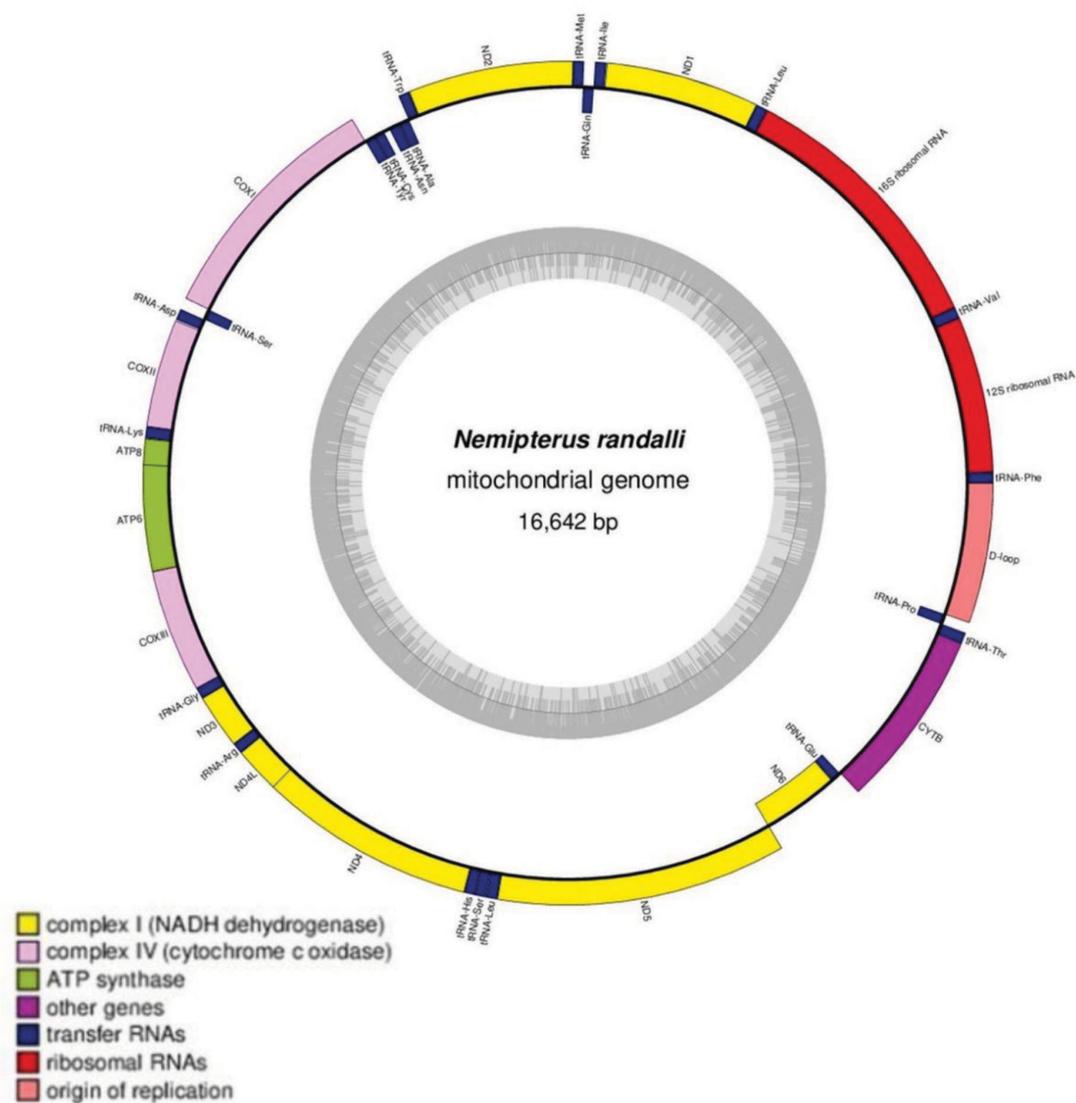


Figure 1. Circular map of Randall's Threadfin bream *Nemipterus randalli* mitochondrial genome. The inner and outer ring represents genes encoded on light and heavy strand respectively. The inner circle depicts the GC content.

The Bayesian phylogenetic tree obtained from the whole mitochondrial genome analysis was used to estimate the divergence times to ensure consistency between the divergence times and the phylogenetic tree. The calibration was obtained from the online Timetree database (<http://www.timetree.org>)³⁵.

Results and discussion

Mitochondrial genomic structure

The complete mitochondrial genome of *N. randalli* was 16,642 bp in length (GenBank accession number MT795184) and was comparable to the mitochondrial DNA of fishes belonging to the family Nemipteridae, ranging from

16,733 bp in *S. ciliata*³⁶ to 17,353 bp in *N. bathybius*³⁷. The genome of *N. randalli* consisted of 13 PCGs, 22 tRNAs, 2 rRNAs and a putative non-coding control region. The arrangement of the genes was similar to that of other Nemipteridae mitochondrial DNA. Among the 37 genes, 28 genes were encoded on the heavy strand (H strand) and all others (*ND6*, *tRNA^{Gln}*, *tRNA^{Ala}*, *tRNA^{Asn}*, *tRNA^{Cys}*, *tRNA^{Tyr}*, *tRNA^{Ser}*, *tRNA^{Glu}* and *tRNA^{Pro}*) on the light strand (L strand) (Figure 1).

The nucleotide composition of the *N. randalli* genome was 27.9% T, 26.8% C, 28.7% A and 16.5% G. It was biased towards A+T nucleotides (56.6%) like those of *N. bathybius* (57.3%), *N. japonicus* (57%), *N. virgatus* (57%), *N. hexadon* (57%), *S. ciliata* (54.4%), and *S. vosmeri* (56%) (Table 1). Most mitochondrial genomes exhibited a strand-specific asymmetry in the nucleotide composition

Table 1. Nucleotide composition of different regions in the mitochondrial genome of Nemipteridae family

Species	T%	C%	A%	G%	Size (bp)	A+T%	AT skewness	GC skewness
Whole mitogenome								
<i>Nemipterus randalli</i>	27.9	26.8	28.7	16.5	16,642	56.6	0.014	-0.237
<i>Nemipterus bathybius</i>	28.3	26.2	29	16.5	17,353	57.3	0.012	-0.227
<i>Nemipterus hexadon</i>	27.4	27	29.6	16.1	17,115	57	0.038	-0.252
<i>Nemipterus japonicus</i>	27.9	26.6	29.1	16.5	16,995	57	0.021	-0.234
<i>Nemipterus virgatus</i>	27.9	26.6	29.1	16.4	16,992	57	0.021	-0.237
<i>Scolopsis ciliata</i>	27.3	28.4	27.1	17.2	16,733	54.4	-0.003	-0.245
<i>Scolopsis vosmeri</i>	28.2	27	27.8	17	16,770	56	-0.007	-0.227
Protein coding genes								
<i>Nemipterus randalli</i>	29.9	27.9	26.2	16.1	11,400	56.1	-0.065	-0.268
<i>Nemipterus bathybius</i>	29.7	28.2	25.7	16.3	11,400	55.4	-0.072	-0.267
<i>Nemipterus hexadon</i>	29.5	28.1	26.7	15.7	11,400	56.2	-0.049	-0.283
<i>Nemipterus japonicus</i>	29.8	27.7	26.4	16.1	11,403	56.2	-0.060	-0.264
<i>Nemipterus virgatus</i>	29.8	27.8	26.3	16.1	11,396	56.1	-0.062	-0.266
<i>Scolopsis ciliata</i>	29.2	29.8	24.3	16.7	11,409	53.5	-0.091	-0.281
<i>Scolopsis vosmeri</i>	30.6	27.9	24.9	16.6	10,884	55.5	-0.102	-0.253
tRNA								
<i>Nemipterus randalli</i>	25.9	23.5	31.2	19.4	1,550	57.1	0.092	-0.095
<i>Nemipterus bathybius</i>	29	20.1	28.4	22.5	1,553	57.4	-0.010	0.056
<i>Nemipterus hexadon</i>	28.7	20.2	28.7	22.4	1,555	57.4	0	0.051
<i>Nemipterus japonicus</i>	28.3	20.6	28.4	22.7	1,552	56.7	0.001	0.048
<i>Nemipterus virgatus</i>	28.1	20.8	28.1	22.9	1,553	56.2	0	0.048
<i>Scolopsis ciliata</i>	28.5	20.8	28.1	22.6	1,570	56.6	-0.007	0.041
<i>Scolopsis vosmeri</i>	29	20.5	27.9	22.6	1,559	56.9	-0.019	0.048
rRNA								
<i>Nemipterus randalli</i>	23	24.6	32.1	20.3	2,728	55.1	0.165	-0.095
<i>Nemipterus bathybius</i>	23.4	23.8	32.5	20.4	2,727	55.9	0.162	-0.076
<i>Nemipterus hexadon</i>	22.8	24.7	32.5	20.1	2,726	55.3	0.175	-0.102
<i>Nemipterus japonicus</i>	22.8	24.7	32.3	20.2	2,652	55.1	0.172	-0.100
<i>Nemipterus virgatus</i>	22.9	24.4	32.4	20.3	2,712	55.3	0.171	-0.091
<i>Scolopsis ciliata</i>	22.7	25.1	31.3	20.9	2,722	54	0.159	-0.091
<i>Scolopsis vosmeri</i>	22.6	24.9	31.6	21	2,713	54.2	0.166	-0.084
Control region								
<i>Nemipterus randalli</i>	34	16.9	34	15.1	894	68	0	-0.056
<i>Nemipterus bathybius</i>	36.5	13.3	36.5	13.7	1,603	73	0	0.014
<i>Nemipterus hexadon</i>	29.1	19	36.9	15	1,357	66	0.118	-0.117
<i>Nemipterus japonicus</i>	32.8	17.3	35.2	14.7	1,260	68	0.035	-0.054
<i>Nemipterus virgatus</i>	33.2	16.9	35.7	14.2	1,342	68.9	0.036	-0.086
<i>Scolopsis ciliata</i>	29.9	19.9	32.7	17.5	987	62.6	0.044	-0.064
<i>Scolopsis vosmeri</i>	29.3	21	33.9	15.8	1,031	63.2	0.072	-0.141

between two strands³⁸, which can be calculated as GC (G–C/G+C) and AT (A–T/A+T)²⁵. The overall AT and GC skew for *N. randalli* were 0.014 and -0.237 respectively. Positive AT skew indicated the presence of more As than Ts and negative GC skew the presence of more Cs than Gs. Similar AT skew values were obtained for other species in the Nemipteridae family except for *S. ciliata* and *S. vosmeri*. The highest AT content was observed in the control region (68%), followed by tRNAs (57.1%), PCGs (56.1%) and rRNAs (55.1%) of *N. randalli*. A similar pattern was observed in *N. hexadon*, *N. japonicus*, *N. virgatus* and *S. vosmeri*, with the exception of *N. bathybius* and *S. ciliata* (Table 1).

The mitogenome of *N. randalli* exhibited both gene overlaps and intergenic spacers, typical attributes of vertebrate mitogenomes³⁹. The mitogenome of *N. randalli* contained a total of 24 bp overlap observed in six overlapping regions, comprising the *tRNA^{Ala}–tRNA^{Gln}* (1 bp), *tRNA^{Gln}–tRNA^{Met}* (1 bp), *ATP8–ATP6* (10 bp), *ND4L–ND4* (7 bp),

ND5–ND6 (4 bp), and *tRNA^{Thr}–tRNA^{Pro}* (1 bp) (Table 2). The intergenic spacers in *N. randalli* comprised 66 bp over 11 regions ranging from 1 bp to 42 bp. The longest spacer (42 bp) was located between *tRNA^{Asn}–tRNA^{Cys}*. The comparative analysis indicated a length variation in overlapping genes ranging from 24 bp (*N. randalli*, *N. hexadon*, *N. japonicus*) to 41 bp (*S. ciliata*). The length variation between intergenic spacers was highest in *N. japonicus* (127 bp) and lowest in *S. ciliata* (56 bp) ([Supplementary Table 2](#)).

The total nucleotide length of PCGs of *N. randalli* was 11,400 bp, representing 68.5% of the total mitochondrial genome similar to *N. bathybius* (65.7%), *N. japonicus* (67.09%), *N. virgatus* (67.06%), *N. hexadon* (66.6%), *S. ciliata* (68.1%), and *S. vosmeri* (64.9%). Among the 13 PCGs, 12 were encoded on the H strand, except for *ND6*, which was encoded on the L strand. The PCGs of *N. randalli* were biased towards A+T nucleotides (56.1%) with a

Table 2. Characteristics of the complete mitochondrial genome of *N. randalli*

Gene	Strand	Start	Stop	Size	Anticodon	Start codon	Stop codon	Intergenic spacer (+)/overlap (-)
<i>tRNA^{Phe}</i>	H	1	68	68	GAA			0
<i>12S rRNA</i>	H	69	1,065	997				0
<i>tRNA^{Val}</i>	H	1,066	1,137	72	TAC			0
<i>16S rRNA</i>	H	1,138	2,868	1,731				0
<i>tRNA^{Leu}</i>	H	2,869	2,942	74	TAA			0
<i>ND1</i>	H	2,943	3,917	975		ATG	TAG	4
<i>tRNA^{Ile}</i>	H	3,922	3,991	70	GAT			-1
<i>tRNA^{Gln}</i>	L	3,991	4,061	71	TTG			-1
<i>tRNA^{Met}</i>	H	4,061	4,130	70	CAT			0
<i>ND2</i>	H	4,131	5,176	1,046		ATG	TA	0
<i>tRNA^{Trp}</i>	H	5,177	5,247	71	TCA			0
<i>tRNA^{Ala}</i>	L	5,248	5,316	69	TGC			1
<i>tRNA^{Asn}</i>	L	5,318	5,390	73	GTT			42
<i>tRNA^{Cys}</i>	L	5,433	5,498	66	GCA			0
<i>tRNA^{Tyr}</i>	L	5,499	5,569	71	GTA			1
<i>COX1</i>	H	5,571	7,121	1,551		GTG	TAA	1
<i>tRNA^{Ser}</i>	L	7,123	7,193	71	TGA			2
<i>tRNA^{Asp}</i>	H	7,196	7,266	71	GTC			7
<i>COX2</i>	H	7,274	7,964	691		ATG	T	0
<i>tRNA^{Lys}</i>	H	7,965	8,039	75	TTT			1
<i>ATP8</i>	H	8,041	8,208	168		ATG	TAA	-10
<i>ATP6</i>	H	8,199	8,881	683		ATG	TA	0
<i>COX3</i>	H	8,882	9,666	785		ATG	TA	0
<i>tRNA^{Gly}</i>	H	9,667	9,737	71	TCC			0
<i>ND3</i>	H	9,738	10,086	349		ATG	T	0
<i>tRNA^{Arg}</i>	H	10,087	10,155	68	TCG			0
<i>ND4L</i>	H	10,156	10,452	297		ATG	TAA	-7
<i>ND4</i>	H	10,446	11,826	1,381		ATG	T	0
<i>tRNA^{His}</i>	H	11,827	11,895	69	GTG			0
<i>tRNA^{Ser}</i>	H	11,896	11,962	67	GCT			2
<i>tRNA^{Leu}</i>	H	11,965	12,037	73	TAG			0
<i>ND5</i>	H	12,038	13,876	1,839		ATG	TAA	-4
<i>ND6</i>	L	13,873	14,394	522		ATG	TAG	1
<i>tRNA^{Glu}</i>	L	14,396	14,464	69	TTC			4
<i>Cyt b</i>	H	14,469	15,609	1,141		ATG	T	0
<i>tRNA^{Thr}</i>	H	15,610	15,680	71	TGT			-1
<i>tRNA^{Pro}</i>	L	15,680	15,748	69	TGG			0
Control region	H	15,749	16,642	894				0

base composition of 29.9% T, 27.9% C, 26.2% A and 16.1% G, similar to other species in the Nemipteridae family (Supplementary Table 3). All PCGs in *N. randalli* started with a canonical codon ATG except *COX1* which began with GTG. Six PCGs in *N. randalli* terminated with the complete stop codons TAA (*COX1*, *ATP8*, *ND4L*, *ND5*) and TAG (*ND1*, *ND6*). The remaining genes terminated with incomplete stop codons TA (*ND2*, *ATP6*, *COX3*) and T (*COX2*, *ND3*, *Cyt b*), which were completed by adding 3' A residues to the mRNA through post-transcriptional polyadenylation⁴⁰. Although members of the family Nemipteridae shared similarities in a stop codon, a slight difference was observed in *N. randalli* (Supplementary Table 4). The AT and GC skew for PCGs of *N. randalli* is shown in Supplementary Figure 1. All PCGs exhibited negative AT skews except *COX2*, *ATP8* and *ND5* which had positive AT skews. The GC skew was negative except for *ND6*, consistent with all other species in the family Nemipteridae (Supplementary Figure 2). An

anti-G bias (8.2%) was observed at the third codon position of PCGs of *N. randalli*, similar to all other vertebrate mitogenomes³⁹ (Supplementary Table 5).

Protein coding genes

A total of 3800 codons (excluding stop codons for 20 amino acids) were identified in all PCGs of *N. randalli*. The most frequent amino acid in the PCGs was Leucine (16.2%), followed by Alanine (8.97%). The least used amino acid was Cysteine (0.81%). Leucine was encoded by 6 different codons, while all the other amino acids were encoded by 2–4 different codons (Figure 2 a). This situation was consistent with all other species in the Nemipteridae family (Supplementary Figure 3). The RSCU of the 13 PCGs of *N. randalli* indicated that the most frequently used codons were CGA (Arg), CAA (Gln), AAA (Lys), TGA (Trp), and the least used were ACG (Thr), CAG

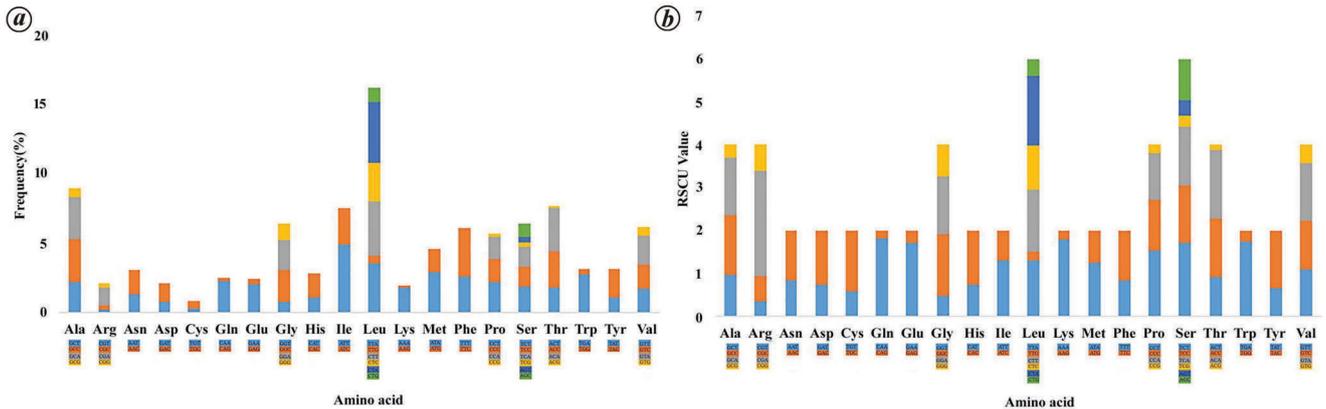


Figure 2. Codon usage of the mitochondrial protein coding genes of *N. randalli*. (a) Codon usage frequency and (b) relative synonymous codon usage.

(Gln), CCG (Pro) and AAG (Lys) (Figure 2 b). The RSCU values indicated that the codons with A or C in the third position were used more frequently than G or T. Codons ending in A were most common for an amino acid with six-fold and four-fold degenerate codons. In contrast, C and T were the most frequent two-fold degenerate codons. A comparative analysis indicated that the RSCU values of *N. randalli* were identical with other species of the Nemipteridae family (Supplementary Figure 3).

The mean Ka/Ks substitution ratio of 13 PCGs from seven species of the Nemipteridae family was 0.065858, ranging from 0.00939 (*COX1*) to 0.1514 (*ND6*). The Ka/Ks value was less than 1 for all PCGs, indicating strong purifying selection (Supplementary Figure 4). The highest Ka/Ks value was associated with the *ND6* gene and the lowest with the *COX1* gene, implying that the highest evolutionary rates were associated with the *ND6* gene and the lowest with the *COX1* gene. In addition, *ND6* and *ATP8* had the highest ratios, indicating that the selection pressure was independent of the position of the gene on the strand.

The p-genetic distance at the third position was higher than at the other two positions, implying that the third position evolved faster than the first and second positions. Based on the first and second codon and full-length PCGs, the highest overall p-genetic distance was in *ND6* (0.186 and 0.274 respectively) and the lowest in *COX1* (0.023 and 0.139 respectively), corroborating the previous findings that *ND6* had comparatively higher evolutionary rates. When full-length PCGs were considered, the highest overall p-genetic distance was in *ND6* (0.274) and the lowest in *COX1* (0.139) (Supplementary Figure 5). The two results showed that the data agree with the values. The overall p-genetic distance again proved that *ND6* is the least conserved gene with higher evolutionary rates.

Ribosomal RNA

Nemipterus randalli had two ribosomal rRNAs (*12S rRNA* and *16S rRNA*) with a total length of 2728 bp located on

the H strand. It was separated by *tRNA^{Val}*, which was the same as other species in the Nemipteridae family. The *12S rRNA* was located between *tRNA^{Phe}* and *tRNA^{Val}* with a length of 997 bp, whereas the *16S rRNA* was located between *tRNA^{Val}* and *tRNA^{Leu}* with a length of 1731 bp. The base composition of rRNA genes in *N. randalli* was 23% T, 24.6% C, 32.1% A and 20.3% G and biased towards A+T nucleotides (55.1%). The overall AT and GC skews were 0.165 and -0.095 respectively, indicating more As and Cs than Ts and Gs. The AT and GC skews of other species of the family Nemipteridae exhibited a similar pattern to that of *N. randalli*, with AT skew of 0.159 to 0.175 and GC skew of -0.102 to -0.076 (Table 1).

Transfer RNA

The mitochondrial genome of *N. randalli* encoded 22 tRNAs ranging in size from 66 to 75 bp with a total concatenated length of 1550 bp. The overall nucleotide composition of tRNA genes was 25.9% T, 23.5% C, 31.2% A and 19.4% G (Table 1). The A+T content was 57.1% with AT skew (0.092) and GC skew (-0.095). On the other hand, the size of the tRNA of other species in the Nemipteridae family ranged from 66 to 76 bp. The highest A+T content of tRNAs was observed in *N. bathybius* and *N. hexadon* and the lowest in *N. virgatus*. Among 22 tRNAs, 14 were encoded on the H strand and the remaining 8 on the L strand, identical to a typical vertebrate mitogenome. Except for *tRNA^{Ser}* that lacks a DHU (dihydrouracil) loop, all other tRNA genes could be folded into canonical cloverleaf secondary structures. However, some non-complementary bases exist in the stem region. *tRNA^{Ser}* considered an aberrant tRNA lacking the DHU arm, also fits the ribosome by adjusting its structural conformation like other tRNAs⁴¹. In *N. bathybius*, *tRNA^{Cys}* also lacked a DHU loop. In *N. randalli*, the 7 bp amino acid acceptor stem was conserved in all tRNAs except *tRNA^{Met}* (8 bp), *tRNA^{Cys}* (8 bp) and *tRNA^{Ser}* (6 bp). Variations of three to four nucleotides in the DHU stem and variations of four to six

nucleotides in the TΨC stem were observed. The anticodon stem in *N. randalli* was 5 bp except for *tRNA^{Ser}* (6 bp).

Among four loops in the tRNA of *N. randalli*, the anticodon loop with a length of 7 bp was highly conserved, and the other three loops showed significant variation. Cross comparison of *N. randalli* tRNA with other species of the Nemipteridae family based on the size of each region is shown in [Supplementary Table 6](#). Three tRNA clusters (IQM, WANCY and HSL) were well conserved in *N. randalli*, typical of vertebrate mitogenomes. In *N. randalli*, *tRNA^{Ser}* was determined by two anticodons (GCT and TGA) and *tRNA^{Leu}* by TAA and TAG. The secondary structure of tRNAs is shown in [Supplementary Figure 6](#). The stem region of the tRNA contained non-Watson–Crick pairs (G–T mismatches and other mismatches), a common phenomenon for mitochondrial tRNA genes. A total of 56 non-canonical base pairs were found in *N. randalli*, including A–A, A–C, G–T, C–T, T–T and A–G mismatches. In a comparative analysis with other species of the Nemipteridae family, most mismatches were G–T pairs (239) ([Supplementary Table 7](#)). These mismatches in the tRNA stem were probably modified by post-transcriptional editing processes⁴².

Non-coding regions

The mitochondrial control region also called the displacement loop (D-loop) is the most variable region in the mitochondrial DNA. It is essential for the replication of the mitochondrial DNA and as in other teleosts, was characterized by several conserved sequence blocks. The non-coding region in *N. randalli* comprised the control region (D-loop) and the O_L region. The control region was 894 bp in length and was located between the *tRNA^{Pro}* and *tRNA^{Phe}* genes in *N. randalli* similar to other species in the Nemipteridae family. The length of the control region for other species of the family Nemipteridae ranged from 987 bp (*S. ciliata*) to 1603 bp (*N. bathybius*) (Table 1). The difference in length occurred due to the presence of tandem repeats in this region. The control region of *N. randalli* was A+T rich (68%) with an overall base composition of 34% T, 16.9% C, 34% A and 15.1% G. The general composition of the control region of *N. randalli* was similar for other species in the family Nemipteridae (Table 1). In a comparative analysis with seven species belonging to the Nemipteridae family, it showed the tripartite structure with a termination-associated sequence domain (TAS), a central conserved domain (CCD), and a conserved sequence block domain (CSB) ([Supplementary Figure 7](#)). The TAS domain included ETAS (extended termination associated sequence), TAS motif and complementary TAS motif and was located at the 5' end of the control region similar to other teleost fishes. The key sequence of the TAS domain in *N. randalli* included ETAS (TACATTCATATGTATTATCACCATAATATATATTAACATT), TAS (TACAT),

cTAS (ATGTA). The TAS domain might be associated with the termination of H-strand synthesis⁴³.

Six conserved sequence blocks CSB-F, E, D, C, B and A have been identified in the CCD of species belonging to the Nemipteridae family. Out of six conserved sequence blocks, five have been identified in the control region of mammals⁴⁴. The central conserved sequence of *N. randalli* was CSB-F (ATGTAGTAAGAACCGACCATC) followed by CSB-E (GGGACAATAATTGTGGGGG). The G box (GTGGGGG), the most conserved region in CSB-E, was present in *N. randalli*. The CSB-D (TATTCCTGGCATT-TGGTTCCTATTTTCAGG) was located downstream of CSB-E. CSB-C (GCATAAGTT), CSB-B (ATGGCG) and CSB-A (CCATGCCGA) were also present in *N. randalli*.

CSB, located at the 3' end of the control region, included CSB 1, CSB 2 and CSB 3. All three CSBs were present, along with the conserved sequence blocks of species belonging to Nemipteridae except *S. vosmeri* ([Supplementary Table 8](#)). Each block consisted of CSB1 (GATTTCAAGTGCATAG), CSB2 (TAAACCCCCCTACCCCC) and CSB3 (AAAACCCCCCGGAAACAGGAA). The conserved sequence block contained elements significant for priming replication and positioning RNA polymerase for transcription of the mitochondrial DNA^{44,45}.

The putative O_L region was located in a cluster of five tRNA genes (WANCY region) between *tRNA^{Asn}* and *tRNA^{Cys}* genes, which is similar in other species of the Nemipteridae family. The length of the O_L region in *N. randalli* was 42 bp, with the potential to fold into a stable stem-loop secondary structure containing 13 paired nucleotides in the stem and 12 bp in the loop (Figure 3 a). This stem-loop region plays a role in initiating light strand replication⁴⁶. A single polypyrimidine tract, characteristic of fish O_L loop, has been observed in *N. randalli*⁴⁷. The secondary stem-loop structure similar to *N. randalli* was also present in other species of the Nemipteridae family ([Supplementary Figure 8](#)). A conserved motif 5' GCCGG 3' was found within the stem of the *tRNA^{Cys}* associated with the transition from RNA to DNA synthesis (Figure 3 b). This conserved sequence plays a role in DNA replication at O_L (ref. 46). A total of three tandem sequence repeats were detected in the control region of *N. randalli* with three copies of a 29 bp repeat (TTATACATTTATATGTATTATCACCATTA), two copies of a 59 bp repeat (TATATGTATTATCACCATTATATATATATATATATGCATATATATGTAGTACTTATACATT), and four copies of a 17 bp repeat (TTATATGTATTATACCA). Tandem repeats in the control region were also observed in *N. hexadon*, *N. virgatus*, *N. japonicus* and *N. bathybius*, except *S. ciliata* and *S. vosmeri*.

CG view comparison tool

Using CCT, six previously reported mitogenomes from the Nemipteridae family were compared to *N. randalli* as

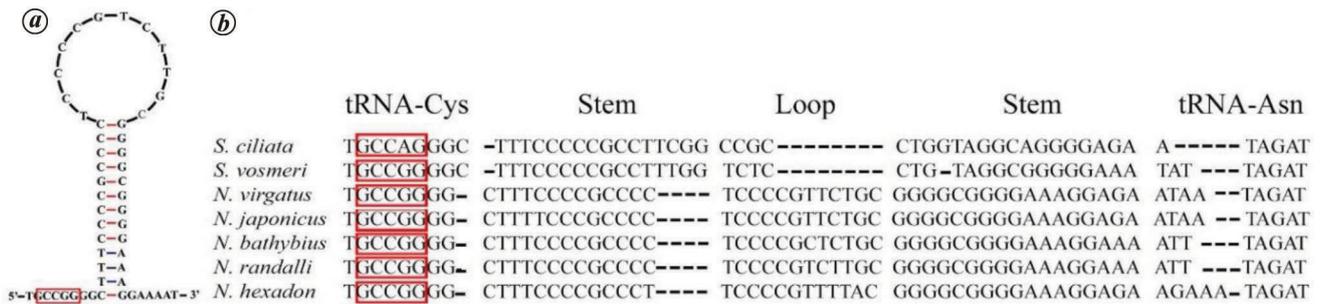


Figure 3. (a) The secondary structure of putative origin of light strand replication in *N. randalli*. The box represents the conserved sequence motif in tRNA-Cys (b) alignment of origin of light strand sequences of Nemipteridae family.

the reference genome. The comparative analysis indicated that the nucleotide composition, number and arrangement of genes were similar in the mitogenomes of all other species belonging to the family Nemipteridae. The results indicated that the highest similarity was between the mitogenomes of *N. randalli* and *N. bathybius*. The nucleotide identity was lower than coding DNA sequences (Supplementary Figures 9 and 10). Of 37 genes, *COX1* and *Cyt b* were the most conserved genes, and *ATP8* was the least conserved among all the species. The cluster of orthologous groups (COGs) is used to predict the function of homologous proteins and track the evolutionary divergence from a common ancestor⁴⁸. Twelve COGs were identified in eleven genes (except *ATP8* and *ND6*) compared to the other mitogenomes. Of these eleven genes, the *ND5* gene carries two COGs (c COG and p COG), and the remaining genes carry a single COG.

Phylogenetic analysis

The phylogenetic position of *N. randalli*, based on 13 concatenated PCGs from 51 species belonging to 6 families was analysed with *N. kuhlii* as the outgroup. Phylogenetic analyses were based on two commonly used methods: Maximum likelihood and Bayesian methods. The phylogenetic tree included 51 species representing six families (Caesionidae, Haemulidae, Lethrinidae, Lutjanidae, Nemipteridae and Sparidae + Centracanthidae) and 20 genera. The results generated analogous tree topologies with similar branch lengths, strong bootstraps and posterior probability values. Therefore, only the phylogenetic tree generated by Bayesian inference is shown in the figure with both the values, including the bootstrap values for ML analysis and the posterior probability value for Bayesian analysis (Figure 4).

The phylogenetic tree constructed by RaxML and MrBayes supported the monophyly of the family Nemipteridae. Since *N. randalli*, along with all other species of *Nemipterus* and two species of *Scolopsis* formed a single clade and was well supported by a posterior probability value of 1, Akazaki⁴⁹ placed Nemipteridae together with

Lethrinidae and Sparidae, referred to as a 'spariform' fish⁵⁰. Johnson⁵¹ later included the family Centracanthidae along with spariform fish based on similarities it shared with Sparidae and erected the superfamily Sparoidea and supported the monophyly of superfamily Sparoidea⁴⁹. In our analysis, Nemipteridae, which was previously placed as the sister group of Lethrinidae and Sparidae + Centracanthidae, was not placed with Lethrinidae and Sparidae + Centracanthidae. Instead, they formed a separate clade in the basal portion of the phylogenetic tree with high support value. Sparidae + Centracanthidae formed a separate clade and was sister to Lethrinidae along with Haemulidae, Lutjanidae and Caesionidae. In the phylogenetic tree, *N. randalli* was most closely related to *N. bathybius*, forming a sub-branch and a sister-species relationship with *N. virgatus* with a high support value. The family Haemulidae was found to have a sister relationship with Lutjanidae and Caesionidae. *Lutjanus erythropterus* (Lutjanidae) clustered with species belonging to the family Caesionidae and was placed as a sister to the Caesionidae species. The phylogenetic tree showed that Nemipteridae species clustered into one clade in the basal region, and other associated families (Lethrinidae and Sparidae + Centracanthidae) formed a separate clade with remaining closely related families (Haemulidae, Lutjanidae, Caesionidae) as that of previous study⁵².

In the present study, the monophyly of the superfamily Sparoidea was not supported by the mitogenomic data. Therefore, the superfamily Sparoidea cannot be considered a monophyletic group.

Divergence times

The divergence times of Nemipteridae and five closely related families (Caesionidae, Haemulidae, Lethrinidae, Lutjanidae, Sparidae + Centracanthidae) were estimated and marked at each node of the tree (Supplementary Figure 11). The divergence between Nemipteridae and closely related families occurred 121.43 million years ago (MYA) during the Lower Cretaceous period of the Mesozoic era. The first divergence within the Nemipteridae has been

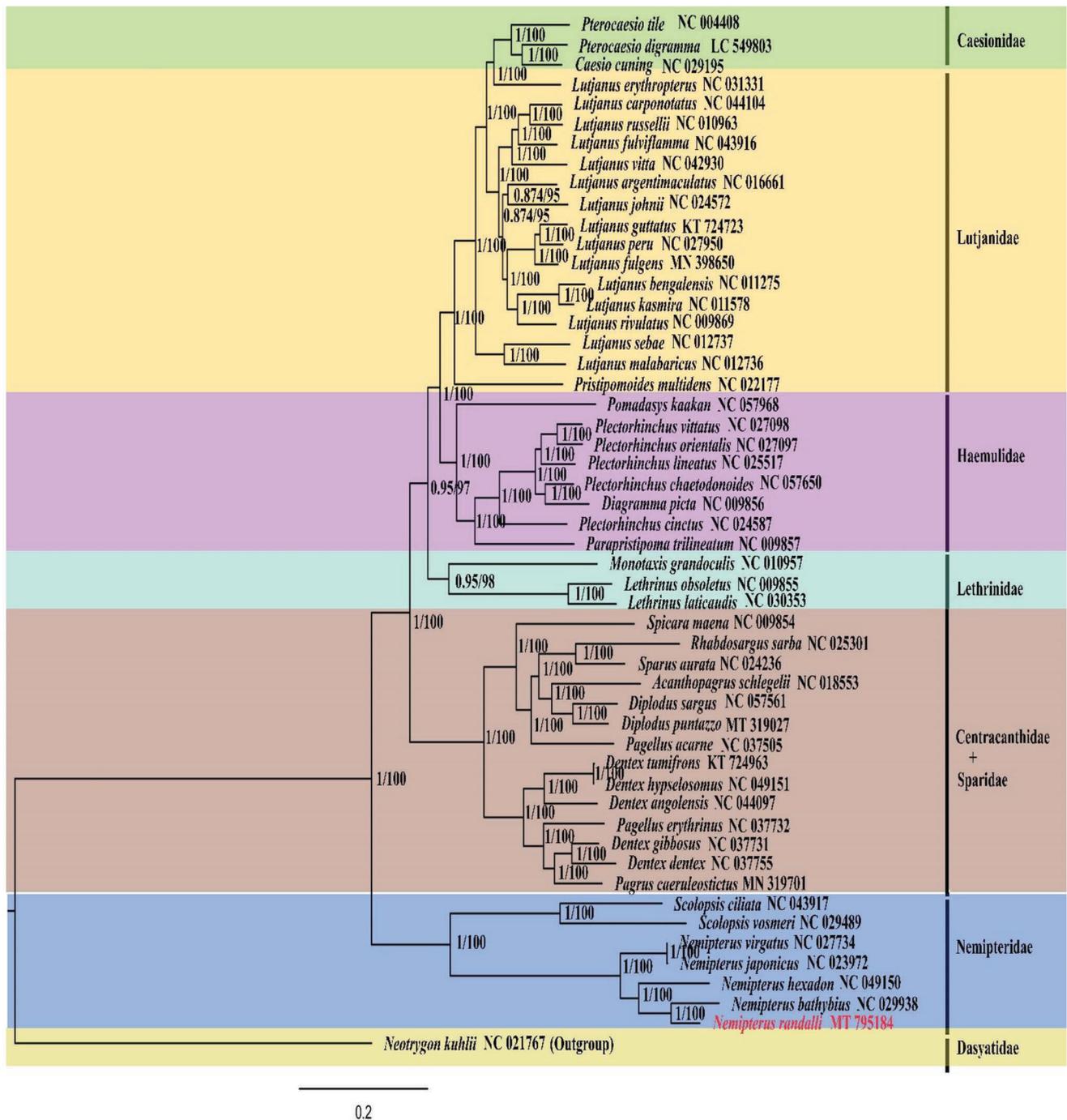


Figure 4. Phylogenetic relationships of *N. randalli* with other Percoidae species constructed using Bayesian inference analysis and maximum likelihood method. The numbers at the nodes are bootstrap value and Bayesian posterior probabilities.

estimated to have occurred during the Upper Cretaceous period of the Mesozoic era at 86.82 MYA with a split between the genus *Nemipterus* and *Scolopsis*. In the genus *Nemipterus*, the divergence time of *N. japonicus* and *N. virgatus* from other *Nemipterus* species was the earliest (22.04 MYA) during the Neogene period of the Cenozoic era. *N. hexadon* differentiated at 18.20 MYA, and *N. bathybius* and *N. randalli* were differentiated at 9.28 MYA

during the Neogene period of the Cenozoic era. The separation of Centracanthidae + Sparidae from Lethrinidae + Haemulidae + Lutjanidae + Caesionidae occurred around 114.90 MYA in the lower Cretaceous period of the Mesozoic era. Centracanthidae diverged from Sparidae around 63 MYA during the Paleogene period of the Cenozoic era. During the lower Cretaceous period of the Mesozoic era, the split between the families Lethrinidae,

Haemulidae and Lutjanidae + Caesionidae occurred at 103.49 MYA.

Conclusion

This study is the first report of the complete mitochondrial genome of *N. randalli*. The mitogenome sequence is 16,642 bp and comprises 13 PCGs, 22 tRNAs, 2 rRNAs and 2 non-coding control regions. Our results shows that structure, gene order, nucleotide composition, protein coding genes, codon usage and non-coding regions of *N. randalli* is consistent with other species of Nemipteridae but shows some deviation in usage of stop codon among protein coding genes. *N. randalli* has a shorter mitogenome than other Nemipteridae species since the length of control region is shorter. Phylogenetic analysis based on protein coding genes shows well supported monophyly of family Nemipteridae. None of our phylogenetic results supported the monophyly of superfamily Sparoidea. The complete mitogenome of *N. randalli* could provide information for further studies on population genetics and elucidation of phylogenetic relationships. The complete mitochondrial genome is a valuable tool for clearing taxonomic ambiguities in the family Nemipteridae. Fisheries management aims at ensuring conservation of biodiversity and biocomplexity to ensure climatic resilience and long-term sustainability of marine fishery resources. Accurate identification and conservation is the key to ensure that the biological complexity is not lost in the long run. Mitogenomic data ensures accurate characterization and conservation of genetic and biological diversity. It is also a great resource to understand evolutionary changes in the OXPHOS gene repertoire which contributes to the metabolic performance of the organism. We suggest that more mitogenomic data are necessary for better resolving the phylogeny of Nemipteridae.

- Russell, B. C., FAO species catalogue. Nemipterid fishes of the world (Threadfin breams, Whiptail breams, Monocle breams, Dwarf monocle breams and Coral breams) family Nemipteridae. An annotated and illustrated catalogue of Nemipterid species known to date, FAO Fisheries Synopsis, FAO Rome, 1990, vol. 12, pp. 1–149.
- Karuppasamy, K., Kingston, S. D., Jawahar, P. and Vidhya, V., Spatio-temporal variation in the diversity of threadfin breams (Family: Nemipteridae) from Wadge Bank, South India. *J. Entomol. Zool. Stud.*, 2018, **6**, 450–454.
- CMFRI, Annual Report 2018–19, 2018.
- IUCN, The IUCN Red List of Threatened Species. IUCN Red List of Threatened Species, 2019; <https://www.iucnredlist.org/en>
- Hung, K. W., Russell, B. C. and Chen, W. J., Molecular systematics of threadfin breams and relatives (Teleostei, Nemipteridae). *Zool. Scr.*, 2017, **46**, 536–551; <https://doi.org/10.1111/zsc.12237>.
- Ballard, J. W. O. and Whitlock, M. C., The incomplete natural history of mitochondria. *Mol. Ecol.*, 2004, **13**, 729–744; <https://doi.org/10.1046/j.1365-294X.2003.02063.x>.
- Sun, Y., Liu, D., Xiao, B. and Jiang, G., The comparative mitogenomics and phylogenetics of the two grouse-grasshoppers (Insecta, Orthoptera, Tetrigoidea). *Biol. Res.*, 2017, **50**, 1–8; <https://doi.org/10.1186/s40659-017-0132-9>.
- Boore, J. L., Animal mitochondrial genomes. *Nucleic Acids Res.*, 1999, **27**, 1767–1780; <https://doi.org/10.1093/nar/27.8.1767>.
- Gupta, A., Bhardwaj, A., Supriya, Sharma, P., Pal, Y., Mamta and Kumar, A., Mitochondrial DNA – a tool for phylogenetic and biodiversity search in equines. *J. Biodivers. Endanger. Species*, 2013, **01**, 1–8; <https://doi.org/10.4172/2332-2543.s1-006>.
- Wolstenholme, D. R., Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.*, 1992, **141**, 173–216; [https://doi.org/10.1016/S0074-7696\(08\)62066-5](https://doi.org/10.1016/S0074-7696(08)62066-5).
- Sebastian, W., Sukumaran, S., Zacharia, P. U., Muraliedharan, K. R., Dinesh Kumar, P. K. and Gopalakrishnan, A., Signals of selection in the mitogenome provide insights into adaptation mechanisms in heterogeneous habitats in a widely distributed pelagic fish. *Sci. Rep.*, 2020, **10**, 1–14; <https://doi.org/10.1038/s41598-020-65905-1>.
- Kalhor, M. A., Tang, D., Ye, H., Morozov, E., Liu, Q., Memom, K. H. and Kalhor, M. T., Population dynamics of Randall's threadfin bream *Nemipterus randalli* from Pakistani waters, Northern Arabian Sea. *Indian J. Mar. Sci.*, 2017, **46**, 551–561.
- Russell, D. W. and Sambrook, J., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, 2001.
- DeTolla, L. J., Srinivas, S., Whitaker, B. R., Andrews, C., Hecker, B., Kane, A. S. and Reimschuessel, R., Guidelines for the care and use of fish in research. *ILAR*, 1995, **37**, 159–173; <https://doi.org/10.1093/ilar.37.4.159>.
- Taylor, M. I., Fox, C., Rico, I. and Rico, C., Species-specific Taq-Man probes for simultaneous identification of (*Gadus morhua* L.), haddock (*Melanogrammus aeglefinus* L.) and whiting (*Merlangius merlangus* L.). *Mol. Ecol. Notes*, 2002, **2**, 599–601; <https://doi.org/10.1046/j.1471-8286.2002.00269.x>.
- Palumbi, S., Martin, A., Romano, S., Mcmillan, W. O., Stice, L. and Grabowski, G., *The Simple Fool's Guide to PCR*, University of Hawaii Press, Honolulu, 1991.
- Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. and Hebert, P. D. N., DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. B.*, 2005, **360**, 1847–1857; <https://doi.org/10.1098/rstb.2005.1716>.
- Bartlett, S. E. and Davidson, W. S., Identification of *Thunnus* tuna species by the polymerase chain reaction and direct sequence analysis of their mitochondrial cytochrome b genes. *Can. J. Fish. Aquat. Sci.*, 1991, **48**, 309–317; <https://doi.org/10.1139/f91-043>.
- Cheng, Y. Z. *et al.*, Universal primers for amplification of the complete mitochondrial control region in marine fish species. *Mol. Biol.*, 2012, **46**, 727–730; <https://doi.org/10.1134/S0026893312040024>.
- Untergasser, A., Cutcutache, L., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. and Rozen, S. G., Primer3-new capabilities and interfaces. *Nucleic Acids Res.*, 2012, **40**, 1–12; <https://doi.org/10.1093/nar/gks596>.
- Bernt, M. *et al.*, MITOS: Improved *de novo* metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.*, 2013, **69**, 313–319; <https://doi.org/10.1016/j.ympev.2012.08.023>.
- Iwasaki, W. *et al.*, Mitofish and mitoannotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. *Mol. Biol. Evol.*, 2013, **30**, 2531–2540; <https://doi.org/10.1093/molbev/mst141>.
- Greiner, S., Lehwark, P. and Bock, R., OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.*, 2019, **47**, W59–W64; <https://doi.org/10.1093/nar/gkz238>.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K., MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.*, 2018, **35**, 1547–1549; <https://doi.org/10.1093/molbev/msy096>.
- Perna, N. T. and Kocher, T. D., Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.*, 1995, **41**, 353–358; <https://doi.org/10.1007/BF00186547>.
- Rozas, J., Ferrer-Mata, F., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E. and Sanchez-Gracia, A.,

- DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.*, 2017, **34**, 3299–3302; <https://doi.org/10.1093/molbev/msx248>.
27. Yang, Z. and Bielawski, J. P., Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.*, 2000, **15**, 496–503; [https://doi.org/10.1016/S0169-5347\(00\)01994-7](https://doi.org/10.1016/S0169-5347(00)01994-7).
 28. Lowe, T. M. and Eddy, S. R., TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, 1996, **25**, 955–964; <https://doi.org/10.1093/nar/25.5.0955>.
 29. Laslett, D. and Canbäck, B., ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics*, 2008, **24**, 172–175; <https://doi.org/10.1093/bioinformatics/btm573>.
 30. Reuter, J. S. and Mathews, D. H., RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinform.*, 2010, **11**, 1–9; <https://doi.org/10.1186/1471-2105-11-129>.
 31. Benson, G., Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.*, 1999, **27**, 573–580; <https://doi.org/10.1093/nar/27.2.573>.
 32. Stamatakis, A., RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 2014, **30**, 1312–1313; <https://doi.org/10.1093/bioinformatics/btu033>.
 33. Ronquist, F. *et al.*, MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.*, 2012, **61**, 539–542; <https://doi.org/10.1093/sysbio/sys029>.
 34. Grant, J. R., Arantes, A. S. and Stothard, P., Comparing thousands of circular genomes using the CGView comparison tool. *BMC Genom.*, 2012, **13**, 1–8; <https://doi.org/10.1186/1471-2164-13-202>.
 35. Kumar, S., Stecher, G., Suleski, M. and Hedges, S. B., TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.*, 2017, **34**, 1812–1819; <https://doi.org/10.1093/molbev/msx116>.
 36. Andriyono, S., Sektiana, S. P., Alam, M. J. and Kim, H. W., Complete mitochondrial genome of saw-jawed monacle bream *Scolopsis ciliata* and its phylogenetic relationship in genus *Scolopsis*. *Mitochondrial DNA B: Resour.*, 2019, **4**, 393–394; <https://doi.org/10.1080/23802-359.2018.1547149>.
 37. Wu, R., Niu, S., Qin, C. and Liu, J., Complete mitochondrial genome of the yellowbelly threadfin bream, *Nemipterus bathybius* (Perciformes, Nemipteridae). *Mitochondrial DNA A: DNA Mapp. Seq. Anal.*, 2016, **27**, 4624–4626; <https://doi.org/10.3109/19401736.2015.1101587>.
 38. Reyes, A., Gissi, C., Pesole, G. and Saccone, C., Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Mol. Biol. Evol.*, 1998, **15**, 957–966; <https://doi.org/10.1093/oxfordjournals.molbev.a026011>.
 39. Satoh, T. P., Miya, M., Mabuchi, K. and Nishida, M., Structure and variation of the mitochondrial genome of fishes. *BMC Genom.*, 2016, **17**, 1–20; <https://doi.org/10.1186/s12864-016-3054-y>.
 40. Ojala, D., Montoya, J. and Attardi, G., TRNA punctuation model of RNA processing in human mitochondria. *Nature*, 1981, **290**, 470–474; <https://doi.org/10.1038/290470a0>.
 41. Ohtsuki, T., Kawai, G. and Watanabe, K., The minimal tRNA: unique structure of *Ascaris suum* mitochondrial tRNA^{Ser}UCU having a short T arm and lacking the entire D arm. *FEBS Lett.*, 2002, **514**, 37–43; [https://doi.org/10.1016/S0014-5793\(02\)02328-1](https://doi.org/10.1016/S0014-5793(02)02328-1).
 42. Sharma, A., Siva, C., Ali, S., Sahoo, P. K., Nath, R., Laskar, M. A. and Sarma, D., The complete mitochondrial genome of the medicinal fish, *Cyprinion semiplotum*: insight into its structural features and phylogenetic implications. *Int. J. Biol. Macromol.*, 2020, **164**, 939–948; <https://doi.org/10.1016/j.ijbiomac.2020.07.142>.
 43. Xu, T. J., Cheng, Y. Z., Sun, Y. N., Shi, G. and Wang, R. X., The complete mitochondrial genome of bighead croaker, *Collichthys niveatus* (Perciformes, Sciaenidae): Structure of control region and phylogenetic considerations. *Mol. Biol. Rep.*, 2011, **38**, 4673–4685; <https://doi.org/10.1007/s11033-010-0602-4>.
 44. Southern, Š. O., Southern, P. J. and Dizon, A. E., Molecular characterization of a cloned dolphin mitochondrial genome. *J. Mol. Evol.*, 1988, **28**, 32–42; <https://doi.org/10.1007/BF02143495>.
 45. Clayton, D. A., Nuclear gadgets in mitochondrial DNA replication and transcription. *Trends Biochem. Sci.*, 1991, **16**, 107–111; [https://doi.org/10.1016/0968-0004\(91\)90043-U](https://doi.org/10.1016/0968-0004(91)90043-U).
 46. Shadel, G. S. and Clayton, D. A., Mitochondrial DNA maintenance in vertebrates. *Annu. Rev. Biochem.*, 1997, **66**, 409–436; <https://doi.org/10.1146/annurev.biochem.66.1.409>.
 47. Hixson, J. E., Wong, T. W. and Clayton, D. A., Both the conserved stem-loop and divergent 5'-flanking sequences are required for initiation at the human mitochondrial origin of light-strand DNA replication. *J. Biol. Chem.*, 1986, **261**, 2384–2390; [https://doi.org/10.1016/S0021-9258\(17\)35948-3](https://doi.org/10.1016/S0021-9258(17)35948-3).
 48. Hurst, C. D., Bartlett, S. E., Davidson, W. S. and Bruce, I. J., The complete mitochondrial DNA sequence of the Atlantic salmon, *Salmo salar*. *Gene*, 1999, **239**, 237–242; [https://doi.org/10.1016/S0378-1119\(99\)00425-4](https://doi.org/10.1016/S0378-1119(99)00425-4).
 49. Akazaki, M., Studies on the spariform fishes anatomy, phylogeny, ecology and taxonomy. *Misaki Mar. Biol. Inst. Kyoto Univ. Spec. Rep.*, 1962, **1**, 1–368.
 50. Galperin, M. Y., Kristensen, D. M., Makarova, K. S., Wolf, Y. I. and Koonin, E. V., Microbial genome analysis: The COG approach. *Brief. Bioinform.*, 2018, **20**, 1063–1070; <https://doi.org/10.1093/bib/bbx117>.
 51. Johnson, G. D., *The Limits and Relationships of the Lutjanidae and Associated Families*, University of California Press, Berkeley, Los Angeles and London, 1980, pp. 1–114.
 52. Wu, R. X., Zhai, Y., Miao, B. B., Niu, S. F., Zhang, H. R., Liu, F. and Ou, C. X., The second complete mitogenome of *Nemipterus virgatus* to dissect control region structure and phylogenetic problem of the superfamily Sparoidea (Teleostei, Perciformes). *Mitochondrial DNA B: Resour.*, 2019, **4**, 3409–3411; <https://doi.org/10.1080/23802359.2019.1674708>.

ACKNOWLEDGMENTS. The authors thank Dr P. Vijayagopal, Head-in-Charge, Marine Biotechnology Division (MBTD), and the Director, ICAR-Central Marine Fisheries Research Institute (CMFRI), Kochi, for providing necessary facilities to carry out the research. Neenu Raj thanks the University Grants Commission (UGC), New Delhi, India for the fellowship provided to her for Ph.D. research.

Received 12 September 2023; revised accepted 5 March 2024

doi: 10.18520/cs/v127/i3/322-332