MJFAS
Malaysian Journal of Fundamental and Applied Sciences

Bakri *et al.* | Malaysian Journal of Fundamental and Applied Sciences, Vol. 20 (2024) 385-400

RESEARCH ARTICLE

Comparative Mitogenome Characterization of Cryptic Asian Seabass (*Lates calcarifer*) in Captivity within Peninsular Malaysia

Athirah Mohd Bakri[a]*, Yuzine Esa[a,b]

[a]Department of Aquaculture, Faculty of Agriculture, University Putra Malaysia, 43400, Selangor, Malaysia; [b]International Institute of Aquaculture and Aquatic Science, University Putra Malaysia, Lot 960 Jalan Kemang 6, 71050 Port Dickson, Negeri Sembilan, Malaysia

Abstract *Lates calcarifer* is one of the most commercial species in the world. Yet, phylogenetic relationships among this cryptic species remain unclear despite its significance in the fishery sector. The first report on bifurcation of captive *Lates calcarifer* in Peninsular Malaysia (PM) inferred from complete mitochondrial genome was discussed in the present study. Two mitochondrial genomes (mitogenomes) of *Lates* species were sequenced and assembled. The mitogenome length was 16,627 bp and 16,515 bp for K4 and S15, respectively. The majority of PCGs in both species exhibited the common initiation marker ATG codon and ended with termination marker TAA codon. Also, the incomplete termination codon T/TA was found in both species. Most AT-skew and GC-skew values observed in the protein-coding genes (PCGs) across the two samples were negative. The phylogenetic analysis based on 13 protein-coding genes by Maximum-Likelihood tree displayed two lineages from different regions which were from Myanmar/Indian waters (K4) and Southeast Asia/ Australian waters (S15). The striking genetic distance values between both specimens indicated 13 PCGs were suffering purifying selection. This study offers significant information for future analyses of evolutionary relationships among the numerous and taxonomically puzzling perciform fishes.

Keywords: *Lates calcarifer*, cryptic, phylogenetic analysis, mitogenome.

Introduction

Ever since 1990s, *L. calcarifer* or commonly known as 'Siakap' (Figure 1) emerged as a prominent species in Malaysia, constituting 50-90% of the overall annual cultured production. However, scientific research pertaining to this species remained insufficient and constrained [1]. The absence of an organized breeding program compelled local hatcheries to rely exclusively on wild and imported broodstock that could potentially influence the population composition of *L. calcarifer* within the country. *Lates calcarifer* or Asian seabass was locally known as 'Siakap Putih' in Malaysia. Since 90s, *Lates calcarifer* had conquered 50-90% of the cultured marine fishes' production in Malaysia.

Lates calcarifer was recognized as a notable species due to high demands for aquaculture products [2,3] caused by the declining of natural fishery stocks which mainly came from human-induced environmental pollutions, frequent and extreme climate change [4]. By being euryhaline species (can tolerate large range of salinity) where culture condition can differ from tank, pond, or cage-farmed base [5], rapid growth rates (2 kg weight in only 12 months on farm), high fecundity (females spawn an average of 300,000 eggs per kg of body weight) [6] and worthy market acceptance [7], it had merits in the aquaculture sector.

Genomic resources for *L. calcarifer* was indeed available [8,9] compare to other species (*Lates niloticus, Lates japonicus* and *Psammoperca waigiensis*) in the family, but the relationships of the family to other Perciform groups was still ambiguous or vague, especially in the perspective of evolutionary and

***For correspondence:**
athirahbakri17@gmail.com

Received: 10 Sept. 2023
Accepted: 19 Dec. 2023

taxonomic [10]. *Lates calcarifer* also had a wide geographical range [11] that allow emergence of new or cryptic species to occur. There was also the presence of considerable genetic variation between populations of *L. calcarifer* [12].

Figure 1 Photograph of *L. calcarifer*

Based on the latest finding using partial analysis of COI gene, two distinct mitochondrial lineages of *Lates* sp. were found representing the Australian/Southeast Asia and Indian/Myanmar groups [13]. Thus, subsequent analysis of the two newly identified groups using the complete mitochondrial genome would provide better information on their genomic characteristics and organizations as well as inferring their phylogenetic relationships using highly longer sequences. This could be attributed to the highly conserved and compacted nature of the encoded gene repertoire within the entire mitochondrial genome. However, the data presented by mitogenome was idiosyncratic between different species [14]. Research/Previous studies of fish mitochondrial genomes were not deep enough because most of them just simply defined the gene structure of single species without comprehensive comparisons [15,16]. So, this study aims to compare and differentiate the two *Lates* specimens from two different clades using complete mitochondrial genome.

Materials and Methods

Sampling Collection and Sequencing

Specimens of *L. calcarifer* specimens in this study were collected from commercial hatcheries located in the east coast (Kelantan; K4: 5.789504, 102.598868) and west coast (Selangor; S15: 2.809926, 101.417420) of PM as described elsewhere [13]. DNA extraction was extracted from the fish tissue (muscle) following the standard extraction protocol of ReliaPrep gDNA Tissue Miniprep System (Promega Corp, Madison, USA). The specimen was kept in the Depository Museum, Department of Aquaculture, Faculty of Agriculture, University Putra Malaysia (Dr. Zafri Hassan, +6039769432, mzafri@upm.edu.my). The specimen has a voucher number of JAQ/LC/00002. The relevant mitogenome sequence was submitted to the NCBI database (https://www.ncbi.nlm.nih.gov/) (OP359071; K4 and OP373725; S15). A DNA library was prepared using the NEB Ultra II library preparation (NEB, Ipswich, MA) according to the manufacturer's instructions. The extracted DNA was sequenced on a NovaSEQ6000 (Illumina, San Diego, CA) with 150 paired-end modes.

Mitogenome Assembly, Annotation and Sequence Analysis

Low-quality bases and Illumina adapter sequences were removed by trimming raw reads using fastp v0.21 [17] followed by de novo assembly in MegaHIT [18]. Identification and circularization of the mitochondrial-derived contings were performed with MitoZ. The assembled mitogenome sequences were then annotated and re-orientated using MitoAnnotator [19]. MEGA 11 was applied to determine the nucleotide composition of the mitogenome [20]. Nucleotide bias of AT/GC skew were computed using the formulas: AT-skew = (A – T)/(A + T) and GC-skew = (G – C)/(G + C).

Codon frequency denotes the frequency at which codons appear in all protein-coding genes (PCGs), while relative synonymous codon usage (RSCU) signifies the probability that a specific codon is used in the synonymous codon for a corresponding amino acid. The frequencies of codons and RSCU values were computed using MEGA 11, and both metrics reflect the utilization of codons [17]. The RSCU values were unaffected by the usage of amino acids, while the abundance of codons directly indicated the level of preference for codon usage [18]. In the absence of any bias towards codon usage, the RSCU value

of a codon equaled one. A codon with an RSCU value greater than one indicated its relatively frequent usage, while a value lower than one indicated the opposite.

Phylogenetic Analysis

The phylogenetic tree was constructed using concatenated sequences of 13 PCGs. The MASCE [21] algorithm in PhyloSuite 1.2.2 [22] was utilized to align the PCGs with vertebrate mitochondrial genetic code. The alignments of each individual gene were concatenated as different dataset with five mitogenomes retrieved from the GenBank database. The multiple alignments of the concatenated nucleotide sequences of the 13 PCGs were conducted using MEGA v. 11 program [20]. The bootstrap confidence of 1000 replicates were applied to evaluate the result in Maximum-Likelihood tree [23]. To determine the evolutionary connections between the specimen, the entire mitochondrial genome sequences of other closely related species were obtained from the GenBank database. The tree was rooted by using the GenBank entry of other Perciforms (*Lates niloticus, Istiompax indica and Istiophorus platypterus*) as outgroup.

Results and Discussion

Genome Organization and Base Composition

Two sequences from different origin were sent for mitogenome sequencing namely K4 and S15 with GenBank accession number OP359071 and OP373725 respectively. The mitochondrial genome of K4 and S15 were 16,627 bp and 16,515 bp in length, correspondingly (Figure 2 and Figure 3). The 37 genes followed the typical vertebrate canonical organization with 13 protein-coding genes (PCGs), two ribosomal RNA (rRNA) genes, 22 transfer RNA (tRNA) genes, and a control region or D-loop. Table 1 showed that genes were mostly encoded on the heavy(H) strand, excluding ND6 genes and eight tRNA genes (Gln, Ala, Asn, Cys, Tyr, Ser, Glu, Pro) which were encoded by light (L) strand. Like other vertebrates such as *Acanthopagrus schlegelii* [24] and *Seriola lalandi* [25], the D-loop was situated between tRNA- Pro and tRNA-Phe gene.

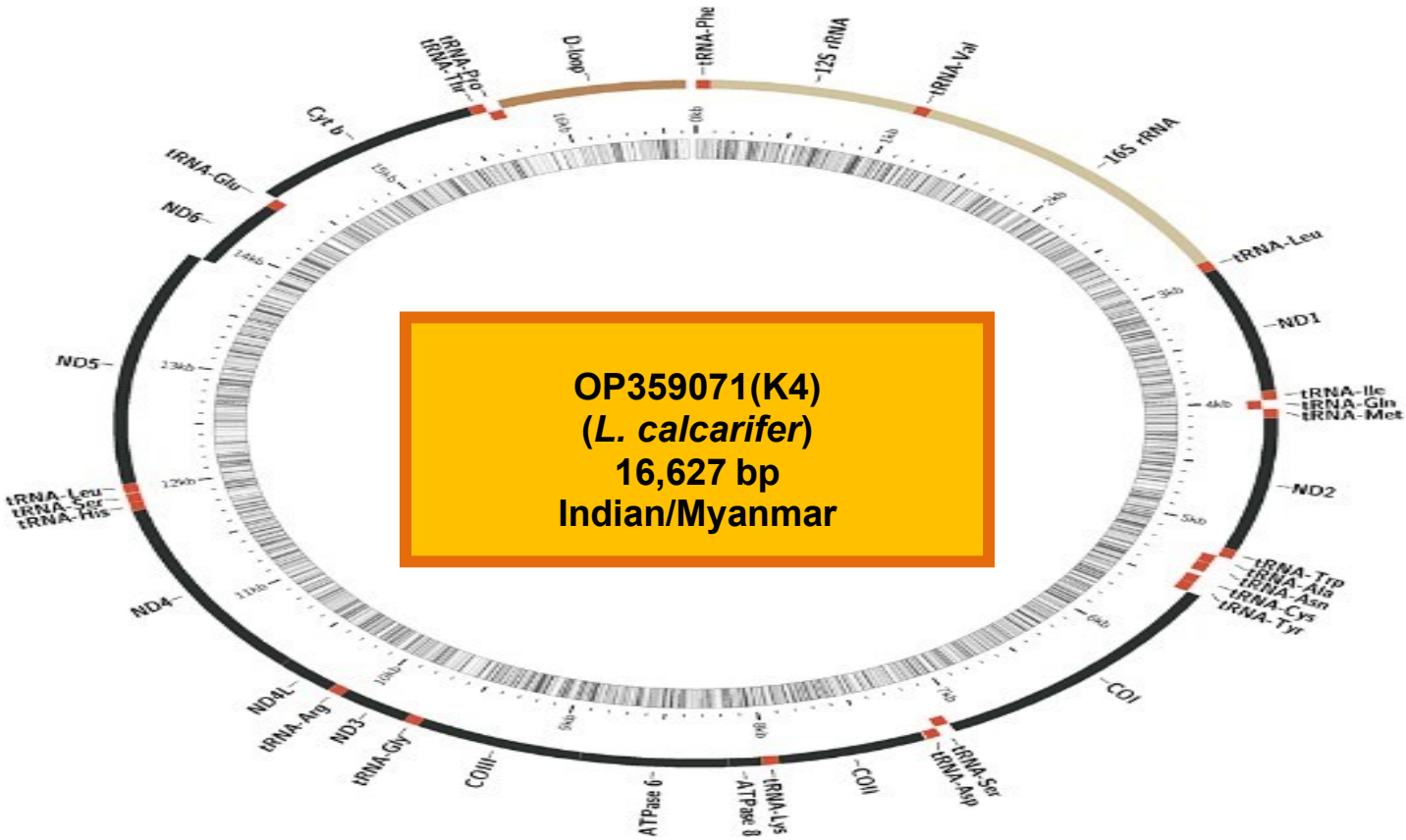


Figure 2. The mitogenome map structure of *Lates calcarifer,* S15 collected from Selangor's hatchery. The inner genes of the circular map were L-strand while the outer genes were H-strand

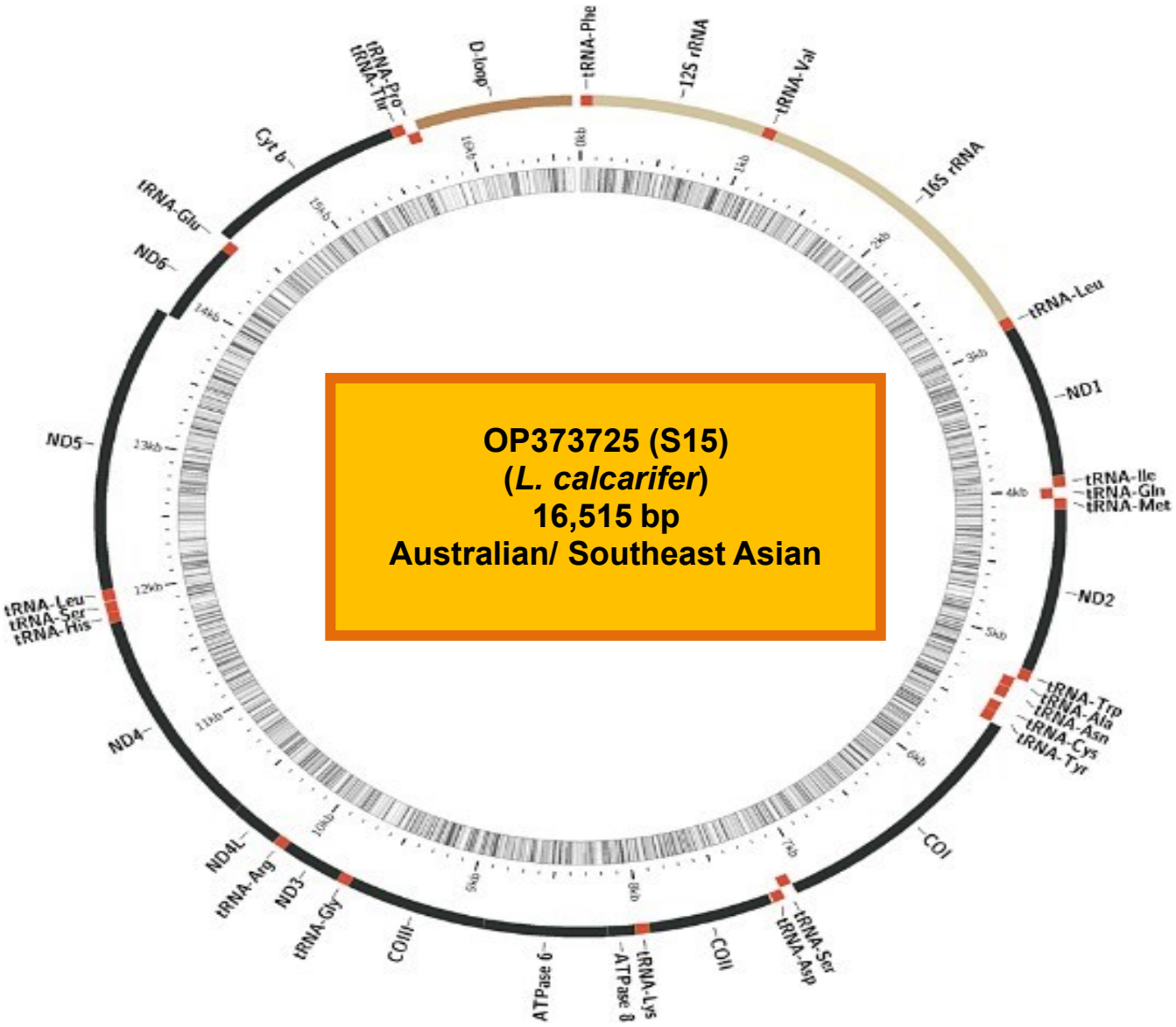


Figure 3. The mitogenome map structure of *Lates calcarifer,* K4 collected from Kelantan's hatchery. The inner genes of the circular map were L-strand while the outer genes were H-strand

The lowest GC content was found in tRNA-Arg (30.43 %) and the highest in tRNA-Cys (58.21%) of S15. Meanwhile, tRNA- Gly in K4 and tRNA-Arg in S15 were detected as the richest AT content regions in each mitogenome (Table 2). In the same table, slight differences of the entire AT composition were detected for both K4 (54.2%) and S15 (54.0%). Besides, low G content represents the obvious antiguanine bias in PCGs just like as discovered in another teleost [26,27]. At the third codon position, high values of A and C content show the preference of codon usage for A and C nucleotides at this position. The highest T content was in the second codon position.

Table 1. Gene features of the mitochondrial genome of *Lates calcarifer*

Genes	Strand	Position Start/End		Size (bp)		Codon Start/End		GC (%)		Intergenic Nucleotide (bp)	
		K4	S15	K4	S15	K4	S15	K4	S15	K4	S15
tRNA-Phe	+	1/68	1/68	68	68			41.18	44.12	0	0
12SrRNA	+	69/1035	69/1035	967	967			48.29	47.77	0	0
tRNA-Val	+	1036/1108	1034/1106	73	71			49.32	43.84	0	0
16SrRNA	+	1109/2857	1107/2860	1749	1754			44.65	44.87	0	0
tRNA-Leu	+	2858/2932	2861/2935	75	75			49.33	50.67	0	0
ND1	+	2933/3907	2936/3910	975	975	ATG/TAA	ATG/TAA	48.21	47.69	3	3
tRNA-Ile	+	3911/3981	3914/3984	74	74			49.3	47.89	1	1
tRNA-Gln	-	3983/4053	3986/4056	72	72			47.89	47.89	-1	-1
tRNA-Met	+	4053/4122	4056/4125	69	69			40	41.43	0	0
ND2	+	4123/5169	4126/5172	1047	1047	ATG/TAA	ATG/TAA	47.47	48.42	0	0
tRNA-Trp	+	5170/5241	5173/5244	72	72			51.39	51.39	2	1
tRNA-Ala	-	5244/5312	5246/5314	71	70			42.03	43.48	1	1
tRNA-Asn	-	5314/5386	5316/5388	74	74			47.95	47.95	34	34
tRNA-Cys	-	5421/5487	5423/5489	67	67			58.21	49.25	0	0
tRNA-Tyr	-	5488/5557	5490/5559	70	70			48.57	50	1	1
COI	+	5559/7109	5561/7111	1552	1552	GTG/TAA	GTG/TAA	48.03	48.1	0	0
tRNA-Ser	-	7110/7180	7112/7182	71	71			46.48	47.89	2	2
tRNA-Asp	+	7183/7251	7185/7253	71	71			40.58	42.03	8	8
COII	+	7260/7953	7262/7955	702	702	ATG/T	ATG/T	44.24	46.25	0	0
tRNA-Lys	+	7954/8029	7956/8031	76	76			47.37	48.68	2	2
ATPase-8	+	8032/8199	8034/8201	170	170	ATG/TAA	ATG/TAA	50	47.62	-14	-14
ATPase 6	+	8184/8872	8186/8874	673	673	GTG/TA	GTG/TA	46.3	46.3	0	0
COIII	+	8873/9657	8875/9659	785	785	ATG/TA	ATG/TA	48.41	48.66	0	0
tRNA-Gly	+	9658/9729	9660/9731	72	72			33.33	31.94	0	0
ND3	+	9730/10078	9732/10080	349	349	ATG/T	ATG/T	44.99	46.99	0	0
tRNA-Arg	+	10079/10147	10081/10149	69	69			36.23	30.43	1	1
ND4L	+	10149/10445	10151/10447	298	298	ATG/TAA	GTG/TAA	51.52	51.18	-5	-5
ND4	+	10439/11819	10441/11821	1374	1374	ATG/T	ATG/T	46.05	46.92	0	0

Genes	Strand	Position	Start/End	Size	(bp)	Codon Start/End		GC (%)		Intergenic Nucleotide (bp)	
tRNA-His	+	11820/11888	11822/11890	69	69			36.23	36.23	0	0
tRNA-Ser	+	11889/11956	11891/11958	68	68			54.41	58.82	4	4
tRNA-Leu	+	11961/12033	11963/12035	77	77			41.1	42.47	0	0
ND5	+	12034/13869	12036/13871	1836	1836	ATG/TAA	ATG/TAA	45.37	45.26	-2	-2
ND6	-	13866/14387	13868/14389	518	518	ATG/TAG	ATG/TAG	46.55	45.4	0	0
tRNA-Glu	-	14388/14456	14390/14458	69	69			39.13	37.68	5	4
Cyt b	+	14462/15602	14463/15603	1146	1145	ATG/T	ATG/T	44.61	45.14	0	0
tRNA-Thr	+	15603/15674	15604/15675	72	72			54.17	54.17	0	0
tRNA-Pro	-	15675/15745	15676/15746	71	71			33.8	33.8	0	0
D-loop		15746/16627	15747/16515	882	769			34.7	35	0	0

Table 2. Base composition for protein-coding, tRNA, rRNA genes, and D-loop region of the mitogenomes of K4 and S15

	K4					Total number	S15					Total number
	A	T	G	C	A+T		A	T	G	C	A+T	
Complete genome	28.8	25.4	15.7	30.1	54.2	16627	28.7	25.3	16.1	30.0	54.0	16515
Protein-coding genes												
First	26.1	20.7	25.0	28.2	46.8	3815	26.0	40.0	13.4	27.9	66.0	3815
Second	18.2	40.6	13.4	27.8	58.8	3811	18.3	40.0	13.4	27.9	58.3	3811
Third	33.6	20.9	7.8	37.8	54.5	3809	33.2	21.0	8.4	37.8	54.2	3809
Total	25.96	27.40	15.39	31.25	53.36	11435	25.83	27.23	15.65	31.28	53.06	11435
tRNA	28.3	26.8	24.0	20.9	55.1	1558	27.9	26.8	24.2	21.1	54.7	1558
rRNA	33.0	21.0	19.6	26.4	54.0	2716	32.5	21.6	19.9	26.0	54.1	2719
D-loop	34.5	30.8	13.8	20.9	65.3	882	34.9	30.2	16.3	18.7	65.1	769

The average base composition for K4 and S15 whole mitochondrial genome in ascending order were G%,25.36 T%,28.75 A%,30.01 C% (Table 2). This depicted the higher content of C and A than G and T. This discovery was consistent with the previous result reported for the same species which also had negative AT-skew and positive GC-skew values [8], displaying significant chain asymmetry bias or chain-specific. Asymmetrical directional mutation stress was often related to the chain bias mechanism [28]. C nucleotide was most often found in the first and third codon position of PCGs while the second position was dominated by T like in other bony fishes [29]. The tRNA- Ser (58.82-54.41%) and tRNA-Thr (54.17%) were observed as the genes with highest GC content (Table 1).

The length of 13 PCGs was 11435 bp for both samples, but account differently in the whole mitogenome where K4 and S15 were 16,627 bp (68.78%) and 16,515 bp (69.24%) respectively (Table 2). In the H-strand of K4 and S15, G and T bases were dominating while A and C bases were discovered more in L-strand. Other than the first codon position of K4, all PCGs expressed obviously low G content same as other bony fish like *Acheilognathus barbatulus* and *Oliotius oligolepis* [30]. ND6 harbored the highest GC-skew value with the lowest AT-skew, and the only gene that had positive GC-skew, which was consistent to other studies [31,32]. Out of the two initiation codons (ATG, GTG) detected, ATG was the frequent initiation codon of both specimens. COI (K4 and S15), ATP6 (K4 and S15), and ND4L (S15) were the only genes which started with GTG codon. K4 and S15 both got seven termination codons and five incomplete termination codons (Table 1). For TA-, the genes include COIII and ATP6 while COII, ND3, ND4 and Cytb were ended by T--.

The lengths of PCGs, tRNAs, rRNAs, and control regions for both mitogenomes were compared in Figure 4. D-loop was detected as the most diversified which responsible for variation in length of the complete genome [15]. The differences were almost 100 bp in length between both samples where K4 and S15 were 882 bp and 769bp, respectively. Further research was needed to discover information about the rapid variation of D-loop in relation to species revolution. Like other typical mtDNA vertebrates, 22 tRNA genes of *Lates calcarifer* were scattered between PCGs and rRNA that correspond to the previous record of this species [8]. Most of the tRNA genes ranged from 67 to 75 in size that enough to fold into cloverleaf secondary structures (Table 1). Meanwhile, both K4 and S15 produced 967 bp for small subunit of rRNA (12S rRNA) whereas 1749 and 1754 bp for large subunit of rRNA (16S rRNA) in K4 and S15, respectively. Both rRNA's positions in the genome were the same as those of other teleost; they were situated between the tRNA-Phe and tRNA-Leu genes, separated by the tRNA-Val gene.

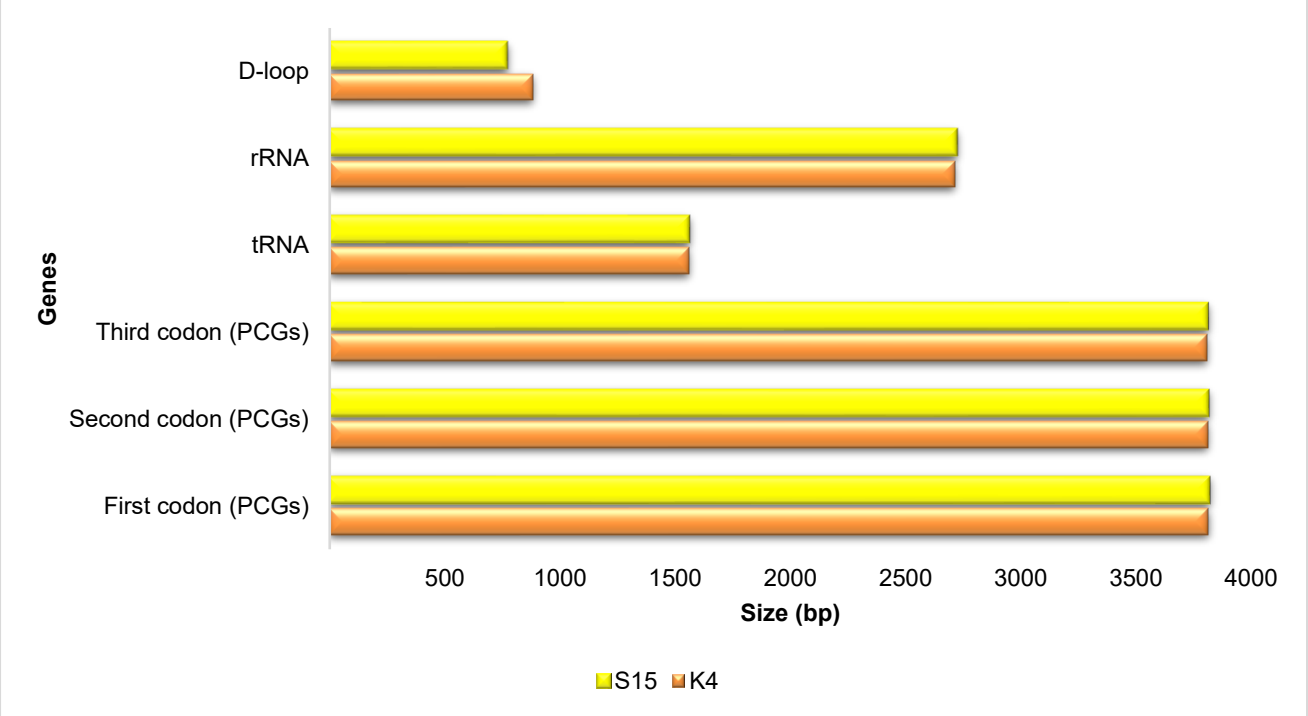


Figure 4. The length of protein-coding genes (PCGs), tRNAs, rRNAs, and control regions among two *Lates* mitogenome

Intergenic Spacer and Overlapping

Compared to gene regions, short spacer with rapidly changing ratio of animal mitochondrial genes were more effective for evolutionary studies. Ten small intergenic spacers (IGS) regions were identified in K4 and S15 with 64 and 62 bp in length, respectively (Table 1). The length variation in mitogenome was also influenced by the size and number of IGSs [33]. The IGS between tRNA-Asn and tRNA-Cys (WANCY region) in the length of 34 bp was the putative origin of L-strand replication (OL). This non-coding region was often found between two conserved genes [34] where it was discovered in both samples. Other than IGS, 4 overlap sites were found in these mitogenomes with 20 bp in length for both K4 and S15. Overlapped PCGs were usually related to the partial sharing of transcript between adjacent regions. Overlap of reading-frames were detected in the ATP8-ATP6 region, followed by ATP6- COXIII, ND4L-ND4, and ND5-ND6.2.3.

Strand Specific Bias

Bias of strand-specific nucleotide composition was the notable characteristic of complete mitochondrial genome for metazoan [35,36]. It was measured using GC-skew (G%-C%)/(G%+C%) and AT-skew (A%-T%)/(A%+T%), respectively. The patterns of nucleotide composition and strand asymmetry in DNA sequences could be determined by analyzing AT-skew and GC-skew [31]. GC-skew of ND6 gene were drastically different from the other H-strand PCGs, but it was still common in vertebrates. Overall, GC-skew and AT-skew of the H-strand were -0.31 and 0.06 respectively, showing a compositional bias on the strand depicted from the excess nucleotide preference of C over G and slight excess of A over T nucleotides [37].

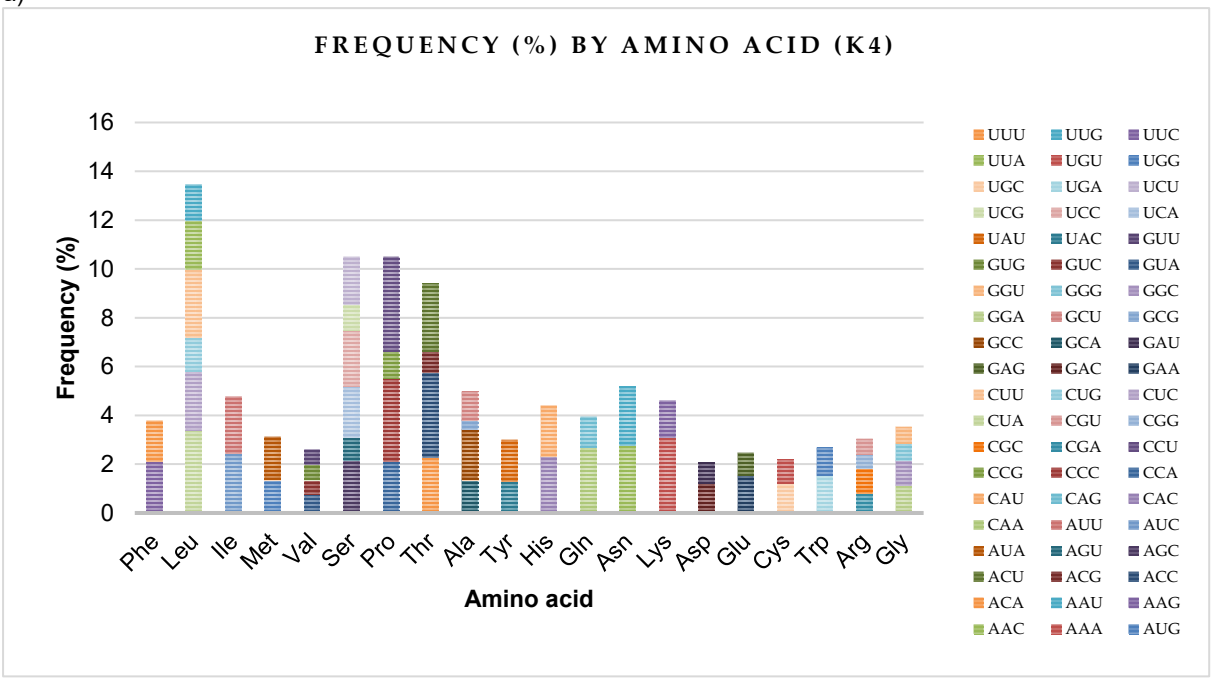
Table 4. The composition and skewness between *Lates calcarifer* from Australian clade (S15) and Indian clade (K4)

Feature	S15		K4	
	AT-skew	GC-skew	AT-skew	GC-skew
Whole mitogenome	**0.06**	**-0.31**	**0.06**	**-0.31**
PCGs	-0.03	-0.33	-0.03	-0.34
ND1	0.01	-0.39	0.03	-0.42
ND2	0.09	-0.54	0.09	-0.55
COI	-0.09	-0.22	-0.08	-0.23
COII	0.05	-0.27	0.03	-0.26
COIII	0.01	-0.29	0.00	-0.30
ATPase-8	0.05	-0.55	0.10	-0.57
ATPase 6	-0.03	-0.41	-0.03	-0.43
ND3	-0.19	-0.31	-0.18	-0.35
ND4L	-0.24	-0.41	-0.21	-0.41
ND4	0.03	-0.40	0.01	-0.39
ND5	0.07	-0.42	0.07	-0.43
ND6	-0.47	0.48	-0.48	0.46
Cytb	-0.07	-0.37	-0.08	-0.37
tRNA	0.03	-0.06	0.03	0.07
rRNA	0.2	-0.13	0.22	-0.15

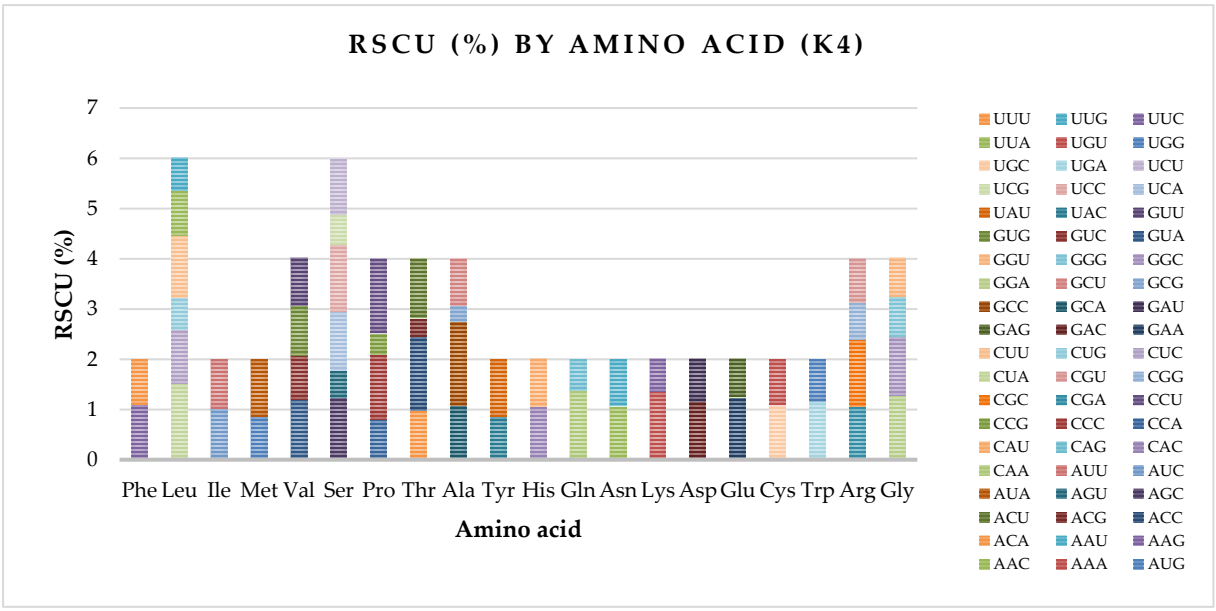
Usage of Mitogenome Codon

The amino acid encoded by six codons each were Serine and Leucine and the rest were encoded by 2 to 4 codons. The highest used amino acids for both mitogenomes was Leucine while Aspartic was the least used in PCGs (Figure 5a). Synonymous codon usage bias was usually assessed by calculating relative synonymous codon usage (RSCU) values [38]. RSCU was defined as the observed frequency of codons to the expected frequency ratio where all the synonymous codons for the same amino acids were used equally. Like other fish mitogenome (*Neoepinnula minetomai* and *Thyrsites atun*) [39], the RSCU for the present study also shows preference towards A and C-ending compared to G and T-ending codon. The frequency of codon used was like other degenerate codons when the RSCU value=1 and codon was often used when RSCU value > 1. Only Arginie (CGU) had RSU value = 1, indicating the variety levels of codon usage bias in each amino acid (Figure 5b). Both samples got different codon fondness but contained the same type of codon encoded for the highest (CCU) and lowest (GCG) codon usage.

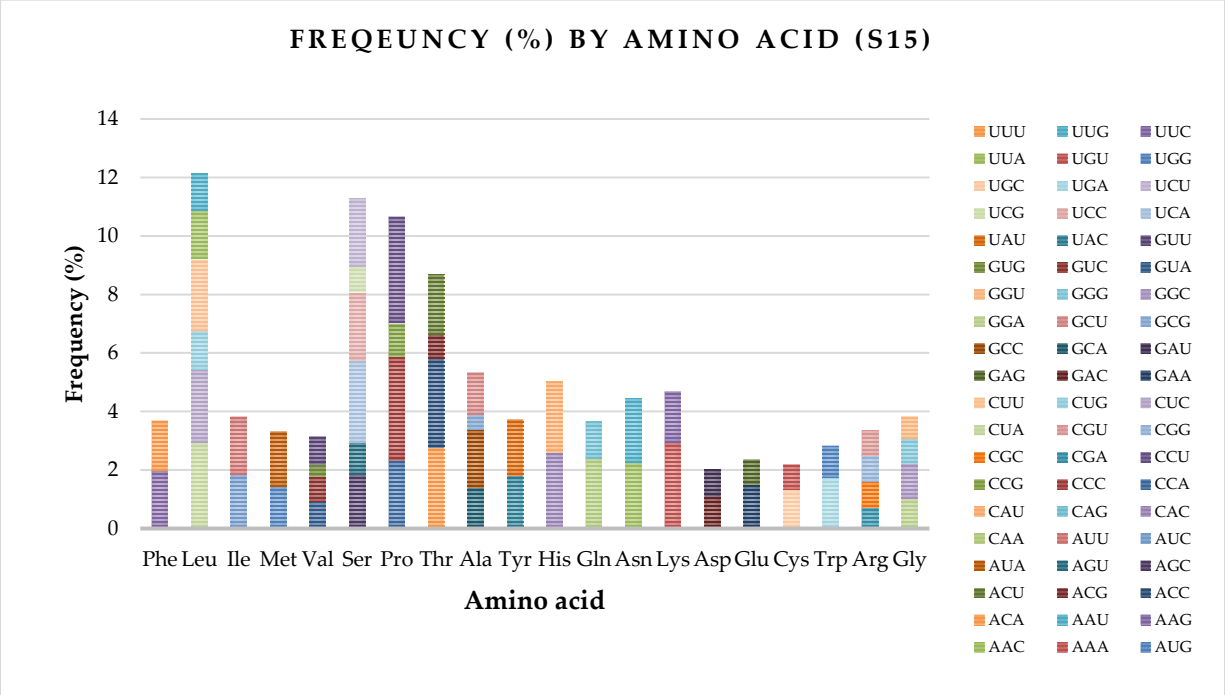
a)



b)



c)



d)

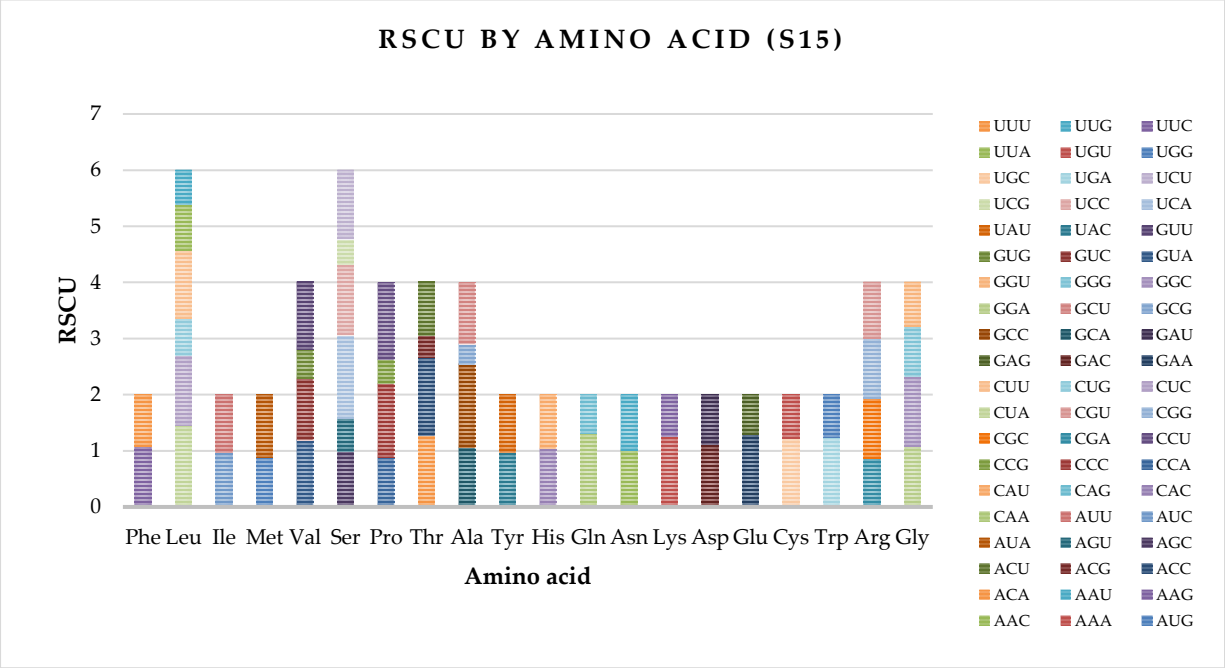


Figure 5. Codon frequency (a and c) and Relative Synonymous Codon Usage (b and d) of mitochondrial genome for *Lates calcarifer* (S4 and S15)

Phylogenetic Tree and Genetic Distance

Evolutionary relationships among *Lates* and related species were established by adding other Perciformes' complete mitogenomes sequences from GenBank entry. GenBank entry of closely related family, *Istiompax indica* (KJ510417) and *Istiophorus platypterus* (KU315124) were used as outgroups to

root the phylogenetic tree. As displayed in Figure 6, *Lates calcarifer* were divided into two parts where K4 was excluded from the Australian lineage even though it was from the same species. GenBank exposed that the closest match (> 99% similarity) was between the newly sequenced mitogenomes of S15 (OP373725) from Australian clade and the mitogenomes of *Lates calcarifer* (KY213962) from Australia. Meanwhile, the K4 mitogenome was the first complete mitochondrial genome deposited to GenBank as the representative of Indian clade when compare to a single gene sequences (not showed in the figure). The tree based on the single gene can be referred elsewhere [13].

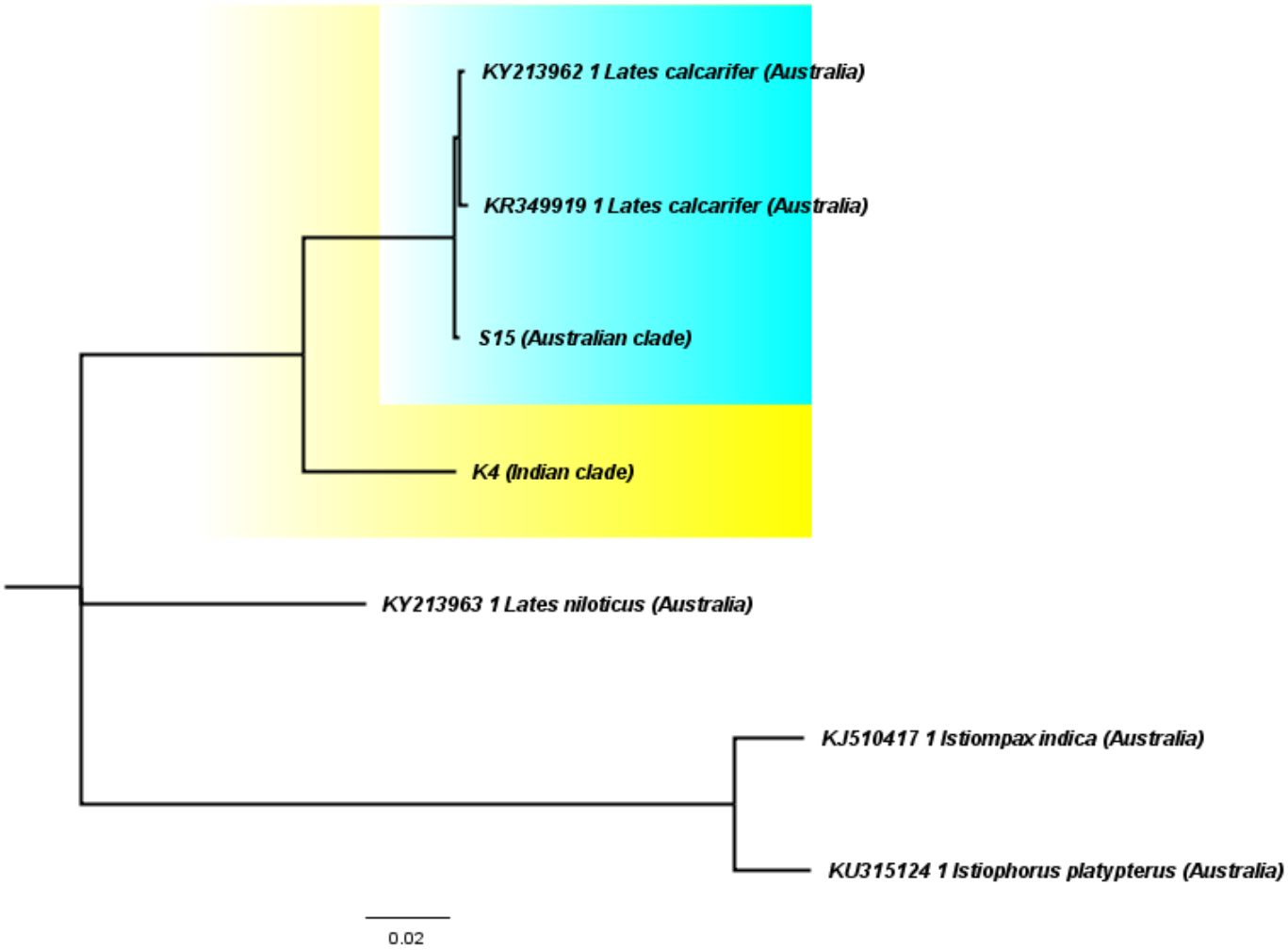


Figure 6. A Maximum likelihood tree of two captive *Lates calcarifer* (K4 and S15) in Malaysia with outgroup species and subspecies based on 13 protein-coding genes were created. Scale bars indicate the relative evolutionary distance

The sequences of PCGs' divergence and conservation among *Lates* sp. were shown by the pairwise genetic distances, p-distance (Figure 7). The obvious difference in genetic distance between the third and sum of the first and second position depicted the occurrence of rapid evolution in third position compared to the first and second. In the third nucleotide position, ND4L was found to have the longest p-distance while ND4 was on the first and second position. The lowest p-distance value was ATP8 for the third and CO1 for first + second analysis. Among the two samples, ND4L and ND4 might contain high evolutionary rates which was in contrast with ATP8 and COI.

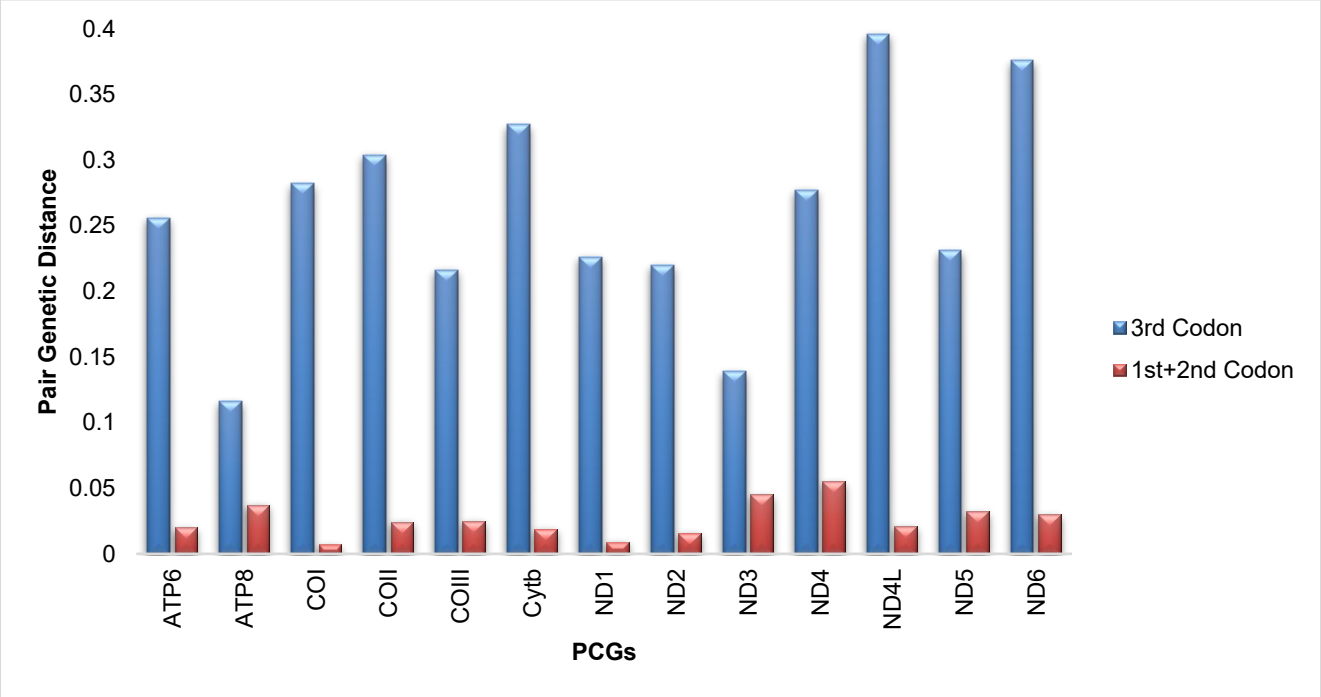


Figure 7. The pair genetic distance of 13 PCGs between K4 and S15. The calculated values were based on the first and second nucleotide position, and on third nucleotide position, respectively

All genes excluding D-loop for the recent sequenced mitogenome samples were nearly similar in size, but both samples were grouped in different clade when focused directly on phylogenetic kinship and genetic distance. Besides, the codons usage also distributed differently which was related to nucleotide composition event [40]. The 37 genes were following the same genes arrangement of another teleost [41]. The K4 mitogenome was longer than S15 and had a lower AT content (first codon position) than those in S15, AT percentage played a vital role to indicate how far the two closely related species were. Because the GC pairs have three hydrogen bonds and the AT pairs only have two, the low AT content was more stable than the high AT concentration. Hence, a structure with a high AT composition was less stable and might have a stronger evolutionary adaptation. Regarding mitochondrial DNA variations and evolutionary adaptation, the decreasing of AT content might be a significant adaptive trait [42]. Phylogenetic, phylogeographic, and evolutionary research frequently involved mitochondrial genes [43,44]. Constructed trees from complete mitogenomes got the best phylogenetic branch supports with the strongest taxonomic resolution [44]. Evolutionary among the two samples were portrayed through phylogenetic analysis.

In this study, the monophyly for the former Australian group and polyphyly for the newly established Indian group were well supported through Maximum Likelihood tree. *Lates calcarifer* were parted into two parts where K4 was excluded from the Australian lineage although it was the same species. The previous study on COI also depicted the same phylogenetic tree which indicated that these newly published mitogenome sequences could help in resolving some of the evolution issues. The partition of all phylogenetic trees into two clades were supported by the Pairwise Genetic distance that show big differences between the nucleotide in third position. Even samples between Australia and Southeast Asia got some differences when compared using deeper mechanisms such as analysis of B-related fragments [45], re- sequenced genomes [46] and SNP analysis [47]. It was common to find the truncated stop codon in metazoans' mitochondrial gene where mitochondrial gene translation or transcription was not affected due to the possibility that post-transcriptional polyadenylation produces the entire stop codon [48]. Incomplete termination codons would also occur during RNA processing when poly A tail was added [49].

The RSCU value of most amino acids in both species were not equal to one, implying that the usage of each amino acid had varying degrees of bias. The gene function in Latidae family might share similar

traits as the RSCU values in both samples were almost the same. Codon usage bias in the mitochondrial genomes of various species was significant, and it might predict the gene's function when distinct selection was pressured onto each gene [50]. Quantity of NNA and NNC were displayed more which correspond to the preference towards A and C at the third nucleotide position. Codon bias was influenced by natural selection, genetic drift, and mutation pressure. In addition, codon bias also influenced GC contents at the third codon position, gene expression levels and gene length. In animals, high content of either A + T or G + C was the primary evolutionary force, and third codon position would be easier to methylate and caused mutations when the GC content was high [51]. The gene overlap that was found in the same area might indicate recent common ancestry and a pattern that was specific to a particular taxon [52]. Mitochondrial genome compact could occur due to the gene overlap and genome info often passed down to offspring by the smaller mitochondrial [53]. But genome size variation selection always involved the adaptation ability to the new location [54]. The *Lates* sp. (K4 and S15) showed similar overlapping zones, suggesting they might have a unique defense mechanism against the environment [55].

Conclusions

In conclusion, sequencing mitogenome research gave a wider information and picture about the bifurcation of *Lates* resources in Malaysia. The mitogenome length of *Lates* sp. (K4) was longer compared to other *Lates* sp. mainly due to variations in the D-loop region Upcoming research on *L. calcarifer* should also include the specimens from Sabah and Sarawak as well as samples solely from Indian/Myanmar to have a better comparation with our specimens. Other than that, hypervariable markers, either microsatellites or single nucleotide polymorphisms (SNPs) should be applied for the better findings at the broader geographical region.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgment

This research was funded by Higher Education, Malaysia, FRGS research grant (FRGS/1/2019/WAB01/UPM/02/26)—Genetic diversity and relatedness estimates for selective breeding. We also thank the Fish Genetic and Breeding Laboratory, Faculty of Agriculture, University Putra Malaysia, for laboratory facilities and everyone involved in sampling collection.

References

[1] Idris, S. M., Noordin, W. N. M., Manah, F. O., & Hamzah. (2022). A. toward systematic breeding of Asian Sea Bass, *Lates calcarifer* (Bloch, 1790), in Malaysia: Status, challenges and prospects for future development. *Asian Fisheries Science*, *35*(1). Doi: 10.33997/j.afs.2022.35.1.001.

[2] Gjedrem, T., Robinson, N., & Rye, M. (2012). The importance of selective breeding in aquaculture to meet future demands for animal protein: A review. *Aquaculture*, *350*, 117-129. Doi: 10.1016/j.aquaculture.2012.04.008.

[3] Yue, G. H. (2014). Recent advances of genome mapping and marker-assisted selection in aquaculture. *Fish and Fisheries*, *15*(3), 376-396. Doi: org/10.1111/faf.12020.

[4] Barange, M., Merino, G., Blanchard, J. L., Scholtens, J., Harle, J., Allison, E. H., Allen, J. I., Holt, J., & Jennings, S. (2014). Impacts of climate change on marine ecosystem production in societies dependent on fisheries. *Nature Climate Change*, *4*(3), 211-216. DOI:10.1038/nclimate2119.

[5] Wang, L., Bai, B., Huang, S., Liu, P., Wan, Z. Y., Ye, B., Wu, J., & Yue, G. H. (2017). QTL Mapping for resistance to iridovirus in Asian Seabass using genotyping-by-sequencing. *Marine Biotechnology*, *19*(5), 517-527. Doi: 10.1007/s10126-017-9770-8.

[6] Khang, P. V., Phuong, T. H., Dat, N, K., Knibb, W., & Nguyen, N. H. (2018). An 8-year breeding program for Asian seabass *Lates calcarifer*: Genetic evaluation, experiences, and challenges. *Frontiers in Genetics*, *9*, 191. Doi: 10.3389/fgene.2018.00191.

[7] Robinson, N. A., Schipp, G., Bosmans, J., & Jerry, D. R. (2018). Modelling selective breeding in protandrous, batch-reared Asian sea bass (*Lates calcarifer*, Bloch) using walkback selection. *Aquaculture Research*, *41*(10), e643-e655. Doi: 10.1111/j.1365-2109.2010.02584.x.

[8] Lin, G., Lo, L. C., Zhu, Z. Y., Feng, F., Chou, R., & Yue, G. H. (2006). The complete mitochondrial genome sequence and characterization of single-nucleotide polymorphisms in the control region of the Asian seabass (*Lates calcarifer*). *Marine Biotechnology*, *8*, 71-79. Doi: 10.1007/s10126-005-5051-z.

[9] Vij, S., Purushothaman, K., Gopikrishna, G., Lau, D., Saju, J. M., Shamsudheen, K. V., Vinaya Kumar, K.,

Basheer, V. S., Gopalakrishnan, A., Hossain, M. S., Sivasubbu, S., Scaria, V., Jena, J. K., Ponniah, A. G., & Orbán. L. (2014). Barcoding of Asian seabass across its geographic range provides evidence for its bifurcation into two distinct species. *Frontiers in Marine Science*, *1*, 1-13. Doi: 10.3389/fmars.2014.00030.

[10] Harrington, R. C., Faircloth, B. C., Eytan, R. I., Smith, W. L., Near, T. J., Alfaro, M. E., & Friedman, M. (2016). Phylogenomic analysis of carangimorph fishes reveals flatfish asymmetry arose in a blink of the evolutionary eye. *BMC Evolutionary Biology*, *16*(1), 1-14. Doi: 10.1186/s12862-016-0786-x.

[11] Haque, M. A., Hossain, M. I., Uddin, S. A., & Dey, P. K. (2019). Review on distribution, culture practices, food and feeding, brood development and artificial breeding of Seabass, *Lates calcarifer* (BLOCH 1790): Bangladesh perspective. *Research in Agriculture Livestock and Fisheries*, *6*(3), 405-414. Doi: 10.3329/ralf.v6i3.44806.

[12] Pethiyagoda, R., & Gill, A. C. (2012). Description of two new species of sea bass (Teleostei: Latidae: Lates) from Myanmar and Sri Lanka. *Zootaxa*, *3314*(1), 1-16. Doi: 10.11646/zootaxa.3314.1.1.

[13] Bakri, A. M., & Esa, Y. (2021, October). Analysis of genetic diversity in five captive population of Asian Seabass (*Lates calcarifer*) for selective breeding in Malaysia. *1st Postgraduate Seminar on Agriculture and Forestry 2021 (PSAF 2021)* (p. 71).

[14] Yu, P., Zhou, L., Zhou, X. Y., Yang, W. T., Zhang, J., Zhang, X. J., Wang, Y., & Gui, J. F. (2019). Unusual AT-skew of Sinorhodeus microlepis mitogenome provides new insights into mitogenome features and phylogenetic implications of bitterling fishes. *International Journal of Biological Macromolecules*, *129*, 339-350. Doi: 10.1016/j.ijbiomac.2019.01.200.

[15] Nevatte, R. J., Clark, J. A., Williamson, J. E., & Gillings, M. R. (2019). The complete mitochondrial genome of the Epaulette Shark, Hemiscyllium ocellatum (Bonnaterre, 1788). *Mitochondrial DNA Part B*, *4*(1), 534-535. Doi: 10.1080/23802359.2018.1553511.

[16] Satoh, T. P., Miya, M., Mabuchi, K., & Nishida, M. (2016). Structure and variation of the mitochondrial genome of fishes. *BMC Genomics*, *17*, 1-20. Doi: 10.1186/s12864-016-3054-y.

[17] Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884-i890.

[18] Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674-1676. Doi: 10.1093/bioinformatics/btv033.

[19] Iwasaki, W., Fukunaga, T., Isagozawa, R., Yamada, K., Maeda, Y., Satoh, T. P., Sado, T., Mabuchi, K., Takeshima, H., Miya, M., & Nishida, M. (2013). MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. *Molecular Biology and Evolution*, *30*(11), 2531-2540. Doi: 10.1093/molbev/mst141.

[20] Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, *38*(7), 3022-3027. Doi: 10.1093/molbev/mst197.

[21] Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. J. (2011). MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. *PloS One*, *6*(9), e22594. Doi: 10.1371/journal.pone.0022594.

[22] Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., & Wang, G. T. (2020). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources*, *20*(1), 348-355. Doi: 10.1111/1755-0998.13096.

[23] Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, *10*(3), 512-526. Doi: 10.1093/oxfordjournals.molbev.a040023.

[24] Ma, X., Xie, Z., Zhou, L., & Chen, Y. (2016). The complete mitochondrial genome of Acanthopagrus schlegelii (Perciformes: Sparidae) with phylogenetic consideration. *Mitochondrial DNA Part B*, *1*(1), 348-349. Doi: 10.1080/23802359.2016.1144105.

[25] Oh, D. J., & Jung, Y. H. (2019). Mitochondrial genome of Japanese amberjack, Seriola quinqueradiata, and yellowtail amberjack, Seriola lalandi. *Mitochondrial DNA Part B*, *4*(1), 826-827. Doi: 10.1080/23802359.2019.1567281.

[26] Norfatimah, M. Y., Teh, L. K., Salleh, M. Z., Isa, M. M., & Siti Azizah, M. N. (2014). Complete mitochondrial genome of Malaysian Mahseer (Tor tambroides). *Gene*, *548*(2), 263-269. Doi: 10.1016/j.gene.2014.07.044.

[27] Xie, X. Y., Huang, G. F., Li, Y. T., Zhang, Y. T., & Chen, S. X. (2016). Complete mitochondrial genome of *Acrossocheilus parallens* (Cypriniformes, Barbinae). *Mitochondrial DNA Part A*, *27(5),* 3339-3340. Doi: 10.3109/19401736.2015.1018212.

[28] Tomkova, M., Tomek, J., Kriaucionis, S., & Schuster-Böckler, B. (2018). Mutational signature distribution varies with DNA replication timing and strand asymmetry. *Genome Biology*, *19*(1), 1-12. Doi: 10.1186/s13059-018-1509-y.

[29] Oh, D. J., Kim, J. Y., Lee, J. A., Yoon, W. J., Park, S. Y., & Jung, Y. H. (2007). Complete mitochondrial genome of the rock bream Oplegnathus fasciatus (Perciformes, Oplegnathidae) with phylogenetic considerations. *Gene*, *392*(1-2), 174-180. Doi: 10.1016/j.gene.2006.12.007.

[30] Miya, M., Friedman, M., Satoh, T. P., Takeshima, H., Sado, T., Iwasaki, W., Yamanoue, Y., Nakatani, M., Mabuchi, K., Inoue, J. G., Poulsen, J. Y., Fukunaga, T., Sato, Y., & Nishida, M. (2013). Evolutionary origin of the Scombridae (tunas and mackerels): members of a Paleogene adaptive radiation with 14 other pelagic fish families. *PloS One*, *8*(9), e73535. Doi: 10.1371/journal.pone.0073535.

[31] Lü, Z., Zhu, K., Jiang, H., Lu, X., Liu, B., Ye, Y., Jiang, L., Liu, L., & Gong, L. (2019). Complete mitochondrial genome of *Ophichthus brevicaudatus* reveals novel gene order and phylogenetic relationships of Anguilliformes. *International Journal of Biological Macromolecules*, *135*, 609-618. Doi: 10.1016/j.ijbiomac.2019.05.139.

[32] Mu, X., Liu, Y., Lai, M., Song, H., Wang, X., Hu, Y., & Luo, J. (2015). Characterization of the *Macropodus opercularis* complete mitochondrial genome and family Channidae taxonomy using Illumina-based de novo transcriptome sequencing. *Gene*, *559*(2), 189-195. Doi: 10.1016/j.gene.2015.01.056.

[33] Li, Q., Wang, Q., Jin, X., Chen, Z., Xiong, C., Li, P., Zhao, J., & Huang, W. (2019). Characterization and comparison of the mitochondrial genomes from two Lyophyllum fungal species and insights into phylogeny of Agaricomycetes. *International Journal of Biological Macromolecules*, *121*, 364-372. Doi: 10.1016/j.ijbiomac.2018.10.037.

[34] Brockman, S. A., & McFadden, C. S. (2012). The mitochondrial genome of *Paraminabea aldersladei* (Cnidaria: Anthozoa: Octocorallia) supports intramolecular recombination as the primary mechanism of gene rearrangement in octocoral mitochondrial genomes. *Genome Biology and Evolution*, *4*(9), 994-1006. Doi: 10.1093/gbe/evs074.

[35] Hassanin, A., Leger, N., & Deutsch, J. (2005). Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa, and consequences for phylogenetic inferences. *Systematic Biology*, *54*(2), 277-298. Doi: 10.1080/10635150590947843.

[36] Reyes, A., Gissi, C., Pesole, G., & Saccone, C. (1998). Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Molecular Biology and Evolution*, *15*(8), 957-966. Doi: 10.1093/oxfordjournals.molbev.a026011.

[37] Zhao, L., Gao, T., & Lu, W. (2015). Complete mitochondrial DNA sequence of the endangered fish (*Bahaba taipingensis*): Mitogenome characterization and phylogenetic implications. *ZooKeys*, *546*, 181. Doi: 10.3897/ZOOKEYS.546.5964.

[38] Liu, X. Y., Li, Y., Ji, K. K., Zhu, J., Ling, P., Zhou, T., Fan, L.Y., & Xie, S. Q. (2020). Genome-wide codon usage pattern analysis reveals the correlation between codon usage bias and gene expression in Cuscuta australis. *Genomics*, *112*(4), 2695-2702. Doi: 10.1016/j.ygeno.2020.03.002.

[39] Mthethwa, S., Bester-van der Merwe, A. E., & Roodt-Wilding, R. (2023). Addressing the complex phylogenetic relationship of the Gempylidae fishes using mitogenome data. *Ecology and Evolution*, *13*(6), e10217. Doi: 10.1002/ece3.10217.

[40] Zahdeh, F., & Carmel, L. (2019). Nucleotide composition affects codon usage toward the 3'-end. *Plos One, 14*(12), e0225633. Doi: 10.1371/journal.pone.0225633.

[41] Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research*, *27*(8), 1767-1780. Doi: 10.1093/nar/27.8.1767.

[42] Zhang, L., Sun, K., Csorba, G., Hughes, A. C., Jin, L., Xiao, Y., & Feng, J. (2021). Complete mitochondrial genomes reveal robust phylogenetic signals and evidence of positive selection in horseshoe bats. *BMC Ecology and Evolution*, *21*, 1-15. Doi: 10.1186/s12862-021-01926-2.

[43] Wang, Y., Shen, Y., Feng, C., Zhao, K., Song, Z., Zhang, Y., Yang, L., & He, S. (2016). Mitogenomic perspectives on the origin of Tibetan loaches and their adaptation to high altitude. *Scientific Reports*, *6*(1), 1-10. DOI: 10.1038/srep29690.

[44] Yuan, M. L., Zhang, Q. L., Zhang, L. I., Jia, C. L., Li, X. P., Yang, X. Z., & Feng, R. Q. (2018). Mitochondrial phylogeny, divergence history and high-altitude adaptation of grassland caterpillars (Lepidoptera: Lymantriinae: Gynaephora) inhabiting the Tibetan Plateau. *Molecular Phylogenetics and Evolution*, *122*, 116-124.

[45] Komissarov, A., Vij, S., Yurchenko, A., Trifonov, V., Thevasagayam, N., Saju, J., Sridatta, P. S. R., Purushothaman, K., Graphodatsky, A., Orbán, L., & Kuznetsova, I. B. (2018). Chromosomes of the Asian seabass (*Lates calcarifer*) contribute to genome variations at the level of individuals and populations. *Genes*, *9*(10), 464.

[46] Vij, S., Kuhl, H., Kuznetsova, I. S., Komissarov, A., Yurchenko, A. A., Van Heusden, P., ... & Orbán, L. (2016). Chromosomal-level assembly of the Asian seabass genome using long sequence reads and multi-layered scaffolding. *PLoS Genetics*, *12*(4), e1005954.

[47] Wang, L., Wan, Z. Y., Lim, H. S., & Yue, G. H. (2016). Genetic variability, local selection and demographic history: genomic evidence of evolving towards allopatric speciation in Asian seabass. *Molecular Ecology*, *25*(15), 3605-3621.

[48] Masta, S. E., & Boore, J. L. (2014). The complete mitochondrial genome sequence of the spider *Habronattus oregonensis* reveals rearranged and extremely truncated tRNAs. *Molecular Biology and Evolution*, *21*(5), 893-902.

[49] Ojala, D., Montoya, J., & Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. *Nature*, *290*(5806), 470-474.

[50] Sharp, P. M., & Matassi, G. (1994). Codon usage and genome evolution. *Current Opinion in Genetics & Development*, *4*(6), 851-860.

[51] Tatarinova, T. V., Alexandrov, N. N., Bouck, J. B., & Feldmann, K. A. (2010). GC 3 biology in corn, rice, sorghum and other grasses. *BMC Genomics*, *11*, 1-18.

[52] Montoya, J., Gaines, G. L., & Attardi, G. (1983). The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. *Cell*, *34*(1), 151-159.

[53] Hogan, R. I., Hopkins, K., Wheeler, A. J., Allcock, A. L., & Yesson, C. (2019). Novel diversity in mitochondrial genomes of deep-sea Pennatulacea (Cnidaria: Anthozoa: Octocorallia). *Mitochondrial DNA Part A*, *30*(6), 764-777.

[54] Aguilar, C., Miller, M. J., Loaiza, J. R., Krahe, R., & De León, L. F. (2018). Mitochondrial genomes and phylogenetic analysis of Central American weakly-electric fishes: *Apteronotus rostratus*, *Brachyhypopomus occidentalis* and *Sternopygus dariensis*. *bioRxiv*, 353565.

[55] Ruan, H., Li, M., Li, Z., Huang, J., Chen, W., Sun, J., Liu, L., & Zou, K. (2020). Comparative analysis of complete mitochondrial genomes of three Gerres fishes (Perciformes: Gerreidae) and primary exploration of their evolution history. *International Journal of Molecular Sciences*, *21*(5), 1874.

[56] Zhu, K. C., Liang, Y. Y., Wu, N., Guo, H. Y., Zhang, N., Jiang, S. G., & Zhang, D. C. (2017). Sequencing and characterization of the complete mitochondrial genome of Japanese Swellshark (*Cephalloscyllium umbratile*). *Scientific Reports*, *7*(1), 15299.