



Comparative synthesis of tandem repeats in the control region of *Epinephelus* mitogenomes (Peciformes: Epinephelidae)

Carla Bessa-Brito¹ , Rodrigo Petry Corrêa de Sousa¹ , Wanny Gomes de Lima¹,
Renata Furtado do Rosário¹, Luan Pinto Rabelo^{1,2} , Yrlan Sousa Oliveira¹, Iracilda Sampaio¹,
Grazielle Fernanda Evangelista Gomes³ and Marcelo Vallinoto^{1,4}

¹Universidade Federal do Pará, Instituto de Estudos Costeiros, Laboratório de Evolução, Bragança, PA, Brazil.

²Instituto Tecnológico Vale (ITV), Belém, PA, Brazil.

³Universidade Federal do Pará, Instituto de Estudos Costeiros (IECOS), Laboratório de Genética Aplicada (LAGA), Bragança, PA, Brazil.

⁴Universidade do Porto, Centro de Investigação em Biodiversidade e Recursos Genéticos, Vairão, Portugal.

Abstract

In this study, we analyzed the evolutionary patterns of tandem repeats in the mitochondrial Control Region (CR) of groupers of the genus *Epinephelus*. This research was based on a comprehensive survey of *Epinephelus* mitogenomes available in public databases, aiming to recover complete CR sequences. A total of 97 specimens were recovered, including 31 species that contain tandem repeat sequences. The repeated sequences were classified into three categories according to their length: short, medium, and long. The long repeats were restricted to a monophyletic group, while short and medium repeats were more widely distributed across different clades, whereas some species lacked repeats entirely. Given the similarities found among the repeat sequence alignments and the phylogenetic arrangements, it was possible to infer that these segments evolved in synchrony over time, with a process of concerted evolution being the most likely explanation for the evolutionary dynamics of these repetitive DNAs. Correlation analyses further revealed that mitogenome length covaries with CR length and that CR length increases with the full repeat region length, whereas motif length is inversely related to repeat copy number. These patterns underscore how repeat architecture evolves through time in the mitochondrial CR of epinephelid mitogenomes.

Keywords: Mitogenomes, VNTR, concerted evolution, phylogeny.

Received: November 07, 2025; Accepted: March 29, 2026.

Introduction

The groupers of the family Epinephelidae are reef-dwelling marine fish that include 249 species distributed in 32 genera (Fricke *et al.*, 2025). The species of this family are highly valued as ornamental fish, particularly as top-quality seafood (Sadovy *et al.*, 2013). *Epinephelus* is the most diverse genus of Epinephelidae, with more than 90 recognized species (Nelson *et al.*, 2016; Ma and Craig, 2018; Frable *et al.*, 2019). Many of these species, such as *Epinephelus itajara* and *Epinephelus akaara*, are listed as either vulnerable or endangered as a result of population decline caused by overfishing and the degradation of habitats (Silva-Oliveira *et al.*, 2008; Sadovy *et al.*, 2013; Bertocini *et al.*, 2018).

Given the importance of these fishes, several studies have used sequences of mitochondrial DNA (mtDNA) to resolve phylogenetic, phylogeographic or population-level questions, with a conservation focus (Pinhal *et al.*, 2012; Sodré *et al.*, 2012; Guimarães-Costa *et al.*, 2016; Cárcamo-Tejer

et al., 2021; Sun *et al.*, 2021; Davidović *et al.*, 2022). The mtDNA has been an important tool for determining indices of genetic diversity in the grouper, making some important contributions to studies in both evolution and conservation, as well as providing molecular tools for the identification of species (Damasceno *et al.*, 2015; Qu *et al.*, 2018; Fadli *et al.*, 2021), metabarcoding (Li *et al.*, 2023; Qing *et al.*, 2024) or phylogenomics (Song *et al.*, 2024; Winn *et al.*, 2024).

Notably, the mtDNA is a potentially important marker for molecular studies because of its unique set of characteristics, such as haploidy, easy extraction, and, in particular, its high mutation rates (Lawless *et al.*, 2020). The arrangement of this molecule is also highly conserved in vertebrates, with approximately 16–17 kb, no introns and only 37 genes (13 encoding proteins, two for rDNA, and 22 tRNAs), in addition to a noncoding region, known as the Control Region (CR) (Zhuang *et al.*, 2013).

The mitochondrial CR plays a central role in regulating mitochondrial DNA replication and transcription, and it is commonly described as comprising three domains (Fonseca *et al.*, 2014; Bernacki and Kilpatrick, 2020). The 5' domain contains termination-associated sequences of the heavy strand, in addition to a varying number of tandem repeats. The

central domain contains conserved sequence blocks (CSBs), whereas the 3' domain typically includes three short conserved sequence blocks (CSB-1, CSB-2, and CSB-3). These elements are associated with the initiation of H-strand replication and with transcriptional regulation via the light-strand promoter (LSP) and heavy-strand promoter (HSP) (Clayton, 1992), and the 3' domain may also harbor tandem repeats. This interesting feature may involve considerable variation in the number of tandem repeats or VNTR (Morris-Pocock *et al.*, 2010; Akiyama *et al.*, 2017; Shu *et al.*, 2018; Kornienko *et al.*, 2019). These repeats typically contain between 10 and 100 base pairs (bps) and thus have a considerable influence on some aspects of the mitochondrial genome, including its size, structure, replication, and recombination (Røyrvik and Johnston, 2020; Lin *et al.*, 2021; Thapana *et al.*, 2022).

Several studies have shown that VNTR configurations are often similar among closely related species and even among populations within the same species, reflecting concerted evolution that homogenizes repeats within a genome (Feliner and Rosselló, 2012; Gomes *et al.*, 2016; Belyayev *et al.*, 2019). This mechanism is a biological process that results from the recombination of the DNA through the unequal exchange of the units of gene repetition, repair, and conversion, ultimately resulting in a similar set of repetitive sequences (Calonje *et al.*, 2009; Gomes *et al.*, 2016; Akiyama *et al.*, 2017).

Recent mitogenomic sequencing efforts in the genus *Epinephelus* have identified numerous tandem repeats, although these studies have focused primarily on genome characterization and CR description (Han J *et al.*, 2011; Han S *et al.*, 2014; Cheng *et al.*, 2015; He *et al.*, 2024; Padgett and Baeza, 2025). Given these limitations, the application of an evolutionary perspective to the investigation of the origin, structure, and dynamics of these segments may provide important insights into the genomic diversity and evolution of these repeats in the genus *Epinephelus*. For example, what is the relationship between the different types of repeats in phylogenetically closely related species? Do these phylogenetic relationships reveal how the repetition motif expanded? Additionally, is it possible to infer when a repetition motif increased in size? In this context, this review examines the evolutionary trends of tandem repeats in the mitochondrial CR of *Epinephelus* species, exploring their importance for comprehending phylogenetic connections and the dynamics of repeat size fluctuations.

Material and Methods

Data acquisition and database compilation

Complete mitochondrial genomes from the Epinephelidae family, of the genus *Epinephelus*, were obtained through a comprehensive search of the GenBank public online DNA sequence database, which is available on the platform of the National Center for Biotechnology Information (NCBI). Only complete and annotated mitogenomes were analyzed, and no raw sequences were used. Thus, all analyses used assembled genomes from public repositories.

The search strategy was based on the current taxonomic structure proposed by Ma and Craig (2018). The accession numbers, species names, and taxonomic information are

provided in Table 1 and Table S1. Species from genera other than *Epinephelus* were included for comparative purposes and were also used as outgroups in phylogenetic analyses.

Extraction and characterization of the mitochondrial control region

The mitochondrial CR was identified on the basis of existing genomic annotations available in GenBank. When necessary, CR boundaries were manually refined by locating the region between the tRNA-Pro and tRNA-Phe genes, following the canonical organization of vertebrate mitogenomes. The total mitogenome length and CR length were recorded, and only CR sequences derived from complete mitochondrial genomes were included in subsequent analyses.

Identification of tandem repeats

Tandem repeats within the mitochondrial CR were identified via Tandem Repeats Finder (TRF) version 4.07b (Benson, 1999). Analyses were performed via the software's default settings, which enable detection of a set of repeat descriptors, such as the consensus motif sequence, including its nucleotide composition (the Motif Length of the repeat unit that is copied in tandem within the array; that is, the consensus repeat unit reported by TRF), and Repeat Copy Number (the number of repeat units, that is, how many times the motif is consecutively repeated within the array) and the Full Repeat Region Length (start–end of the repeat) within the CR.

When more than one repeat array was detected in the same CR, each array was treated as an independent record. For descriptive purposes, the motifs were grouped into three practical classes the basis of motif length: short, medium, and long. The medium category was used for intermediate cases that did not fit naturally into the short or long classes.

All repeat arrays detected by TRF are summarized in Table 1 and Table S1. To compare repeats within and between species, we focused on the motif itself. The consensus motif sequences reported by TRF, that can be found in the supplementary material (Table S1), were aligned via MAFFT v7.313, and similarity was assessed via direct comparison of the aligned motifs, together with motif length and repeat copy number. Intraspecific comparisons were performed when more than one specimen was available for a species, and variation was recorded primarily as differences in repeat copy number. We refer to VNTR-like variation when specimens of the same species share the same or highly similar motif sequence but differ in the number of repeat units. Interspecific comparisons were based on motif sequence similarity across taxa, allowing us to identify shared motif types and to evaluate how these patterns were distributed across the phylogenetic framework.

Phylogenetic analyses

Phylogenetic analyses were conducted on the basis of the sequences of 13 protein-coding genes from 41 species of *Epinephelus*, along with three species of the genus *Hyporthodus*. Additionally, *Cephalopholis sonnerati*, *Variola louti*, *Plectropomus areolatus* and *Grammistes sexlineatus* were included as outgroup taxa (Table 1; Table S1). The sequences were aligned via MAFFT v7.313 software (Katoh and Standley, 2013) and then concatenated via SequenceMatrix

Table 1 – Results of the analyses of the genus *Epinephelus* run using Tandem Repeats Finder. The asterisks (*) indicate the species that have more than one repeat. The numbers highlighted in bold represent the largest tandem repeats identified within each species, evidencing intraspecific variability in repeat length, which is further characterized in detail in Table S1.

Species	Number of genomes analyzed	Size of the genome (bps)	Size of the control region (bps)	Full Repeat Region Length	Motif Length (bps)	Type	Number of Repeat Units	VNTR
<i>Epinephelus lanceolatus</i>	5	16714 – 16743	880 – 1041	1 – 134	17	Short	7.8 – 17.8	Yes
<i>Epinephelus fuscoguttatus</i>	3	16373 – 16648	673 – 948	1 – 190	17	Short	10.9	No
<i>Epinephelus coioides</i>	3	16418 – 16458	720 – 760	1 – 153	20	Short	5.8 – 7.8	Yes
<i>Epinephelus trimaculatus</i>	3	16761 – 16777	1056 – 1074	166 – 366	18	Short	10.2 – 11.2	Yes
<i>Epinephelus moara</i>	3	16696	998 – 999	1 – 373	17	Short	22.1	No
<i>Epinephelus bruneus</i>	3	16686 – 16692	991 – 995	1 – 372	17	Short	22 – 22.1	Yes
<i>Epinephelus epistictus</i>	2	16920	1217	7 – 553	33	Short	16.4	No
<i>Epinephelus corallicola</i>	2	16647	950	101 – 249	18	Short	8.3	No
<i>Epinephelus malabaricus</i>	3	16423	720	1 – 113	20	Short	5.8	No
<i>Epinephelus tukula</i>	2	16503	805	8 – 170	21	Short	7.7	No
<i>Epinephelus septemfasciatus</i>	2	16558	850	10 – 230	18	Short	12.3	No
<i>Epinephelus tauvina</i>	2	16787	1080	214 – 307	22	Short	4.2	No
<i>Epinephelus quoyanus</i>	2	16797	1093	181 – 382	18 or 36	Short	11.2 or 5.6	No
<i>Epinephelus bontoides</i>	2	16903	1200	173 – 497	37	Short	8.8	No
<i>Epinephelus chlorostigma</i> *	2	16894	1191	505 – 556	23	Short	2.3	No
<i>Epinephelus chlorostigma</i> *				52 – 503	76	Medium	5.9	
<i>Epinephelus areolatus</i>	3	16893 – 16965	1191 – 1264	52 – 579	76	Medium	5.9 – 6.7	Yes
<i>Epinephelus akaara</i>	5	16795	1093	20 – 435	133	Long	3.1	No
<i>Epinephelus awoara</i>	3	16798 – 16802	1098	21 – 427	134	Long	3 – 3.2	Yes
<i>Epinephelus fasciatomaculosus</i>	2	16682	980	6 – 349	141	Long	2.4	No
<i>Epinephelus sexfasciatus</i>	2	16786	1090	14 – 422	132	Long	3.1	No
<i>Anypserodon leucogrammicus</i>	1	16616	916	3 – 254	17	Short	14.8	No
<i>Epinephelus amblycephalus</i>	2	16869	1169	47 – 419	50	Medium	7.5	No
<i>Epinephelus adscensionis</i>	2	16963	881	34 – 227	19	Short	10.2	No

Table 1 – Cont.

Species	Number of genomes analyzed	Size of the genome (bps)	Size of the control region (bps)	Full Repeat Region Length	Motif Length (bps)	Type	Number of Repeat Units	VNTR
<i>Epinephelus hexagonatus</i>	2	16872	1172	197 – 447	17	Short	14.8	No
<i>Epinephelus itajara</i>	1	16561	857	1 – 205	17	Short	12.1	No
<i>Epinephelus marginatus</i>	1	16984	1286	46 – 541	57	Medium	8.7	No
<i>Epinephelus morio</i>	2	16756	1057	1 – 462	18	Short	25.7	No
<i>Epinephelus polyphemadion</i>	2	16619	915	27 – 140	18	Short	6.3	No
<i>Epinephelus undulosus</i>	2	17126	954	29 – 189	10	Short	16.1	No
<i>Epinephelus bilobatus*</i>	1	17354	981	169 – 209	2	Short	20.5	No
<i>Epinephelus bilobatus*</i>				838 – 874	8		4.6 – 4.8	Yes
<i>Epinephelus maculatus</i>	3	17066 – 16746	924 – 937	726 – 785	29	Short	2.1	No
<i>Epinephelus merra</i>	2	17017	988	–	–	–	–	–
<i>Epinephelus bleekeri</i>	2	17227	900	–	–	–	–	–
<i>Epinephelus latifasciatus</i>	2	16389	684	–	–	–	–	–
<i>Epinephelus stictus</i>	2	16524	824	–	–	–	–	–
<i>Epinephelus aeneus</i>	1	16578	891	–	–	–	–	–
<i>Epinephelus multinotatus</i>	2	17057	766	–	–	–	–	–
<i>Epinephelus cyanopodus</i>	2	16649	768	–	–	–	–	–
<i>Epinephelus flavocaeruleus</i>	2	16758	767	–	–	–	–	–
<i>Epinephelus japonicus</i>	1	17084	1091	–	–	–	–	–
<i>Epinephelus longispinis</i>	1	17221	954	–	–	–	–	–
<i>Epinephelus altivelis</i>	4	16503 – 16497	799 – 805	–	–	–	–	–
<i>Hyporthodus haifensis</i>	1	16525	798	–	–	–	–	–
<i>Hyporthodus octofasciatus</i>	2	16545	840	–	–	–	–	–

The asterisks (*) indicate the species that have more than one repeat. The numbers highlighted in bold represent the largest tandem repeats identified within each species, evidencing intraspecific variability in repeat length, which is further characterized in detail in Table S1.

v1.10 (Vaidya *et al.*, 2011). The database was used for the production of a maximum likelihood tree in IQTREE v3.0.1 software (Nguyen *et al.*, 2015), and ultrafast bootstrap analysis, with 1,000 pseudoreplicates, was used to determine the statistical support for each node. The IQTREE automatic substitution model options were used to select the evolutionary model for each partition.

Statistical analyses

Statistical analyses were performed to evaluate the contribution of tandem repeats to length variation in mitochondrial genomes, with a focus on the CR. We applied both Pearson's (r) and Spearman's rank (ρ) correlation tests to examine a specific set of structural hypotheses. The choice of these methods followed normality assessments via the Shapiro-Wilk test. While mitogenome and CR lengths followed a normal distribution ($p > 0.05$), repeat-related variables (motif length, repeat copy number and full repeat region length) deviated significantly ($p < 0.05$). Consequently, Spearman's ρ was prioritized as a robust estimator for nonparametric associations, whereas Pearson's r provided a measure of linear expansion.

First, we tested whether the overall mitogenome length covaries with CR length. Because repeat arrays can vary through two components, motif length and copy number, we evaluated whether CR length or mitogenome length are better explained by the full repeat region length. Finally, because this region occupies a limited segment of the CR, we tested whether longer motifs tend to occur in fewer copies or shorter motifs tend to occur in more copies.

Results

The CR sequences were selected from the complete genomes of 97 specimens belonging to seven genera (Table S1). Tandem repeats were detected in 31 species (Table 1), and the alignment of the sequences revealed that the repeats were concentrated primarily in the 5' domain. In 13 of the species that had repeats, these motifs began between bases 1 and 10 of the CR, including 9 species in which they began between bases 1 and 3 (Table 1 and Figure 1).

Many sequences are unique to a given taxon, which precludes the conclusive confirmation of the existence of VNTRs for all species. However, in the species for which at least two sequences were available, only seven had VNTRs (Table 1). The motifs of the repetitive sequences vary in size from 17-141 bps, which permitted the classification of the repeats into three categories: short (2-37 bps), medium (57-76 bps), and long (132-141 bps) (Table 1). Some of the repeated sequences are highly homogeneous in terms of motif sequence composition, such that both their alignments and distribution in the phylogeny confirm the similarities between the repetitive sequences (Figure 1).

Two unique species, *Epinephelus chlorostigma* (KR872887.1, and NC_032086.1) and *Epinephelus bilobatus* (ON321831.1 and NC_069198.1), presented two types of repeats. In *E. chlorostigma*, one medium repeat of 76 bps followed by one short repeat of 23 bps were identified; these repeats were arranged *in tandem* and located between bases 52-503 and 505-556, respectively. In *E. bilobatus*, two short

repeats were found, positioned between bases 169-209 and 838-879 (Table 1; Figure 1).

In the phylogenetic analysis, long repeats were found exclusively in a single monophyletic group of species. In contrast, medium and short repeats were broadly distributed across the phylogeny, occurring in multiple clades and even in distinct genera (Figure 1). As mentioned previously, all the repetitions exhibited phylogenetic congruence. The alignment of the motif sequences can be easily performed, and these are well conserved among closely related species. For example, *Epinephelus tukula* (NC_024039.1 and KJ414470.1), *E. coioides* (NC_011111.1, EU043376.1, KM377093.1 and MW752082.1), and *E. malabaricus* (NC_028406.1 and KM873711.1) presented highly similar repeat sequences, reinforcing their close evolutionary relationships (Figure 1). Similarly, *E. bruneus* (NC_013820.1, JQ518289.1 and FJ594964.1) and *E. moara* (JQ518290.1, KP009977.1 and NC_017891.1) shared an identical repeat sequence, as did *Epinephelus itajara* (OP056827.1) and *Epinephelus lanceolatus* (NC_011715.1, OP980559.1, FJ472837.1, HQ660062.1, KM386619.1 and KJ451389.1), further supporting their recognition as sister species. In contrast, *E. tauvina* (NC_056898.1 and MW194890.1) and a set of closely related congeners presented a greater degree of motif sequence divergence, with sequence similarity decreasing in parallel with increasing phylogenetic distance.

In some species, the TRF identified more than one motif length for the same repeat sequence in the same region. For example, analysis of repeats in *Epinephelus quoyanus* revealed that the motif sequence occurs in two forms, differing in size, with repeats of 18 and 36 bps detected in the same region, between bases 181 and 382 (Table 1). The closely related species *E. trimaculatus* (NC_021612.1, KC847086.1, and KC593372.1) and *E. bontoides* (NC_028428.1 and KT619054.1) also contained short repeats, with 18 bps in *E. trimaculatus* and 39 bps in *E. bontoides*. In other words, TRF was unable to distinguish the actual length of the motif, resulting in alternative computational interpretations that suggest different sizes for the same repeating motif sequence.

Finally, to test hypotheses linking tandem repeats to variations in the mitogenome and mitochondrial CR length, we performed correlation analyses. The results are summarized in Table S2 and all tests are supported by either Spearman's rank or Pearson's correlation analyses. First, CR length showed a strong positive association with overall mitogenome length (Test 1). Second, mitogenome length also increased with increasing length of the full repeat region (Test 2), indicating that repeat-rich segments can contribute to broader mitogenome length variation. Third, and most directly relevant to our main question, the mitochondrial CR length increased with the length of the full repeat region (Test 3), which is consistent with the idea that the physical footprint of the array translates into measurable differences in CR length.

In addition, motif length was strongly and inversely associated with the repeat copy number (Test 4). In practical terms, longer motifs tend to occur in fewer copies, whereas shorter motifs can be maintained in more copies, which is consistent with a limited segment of the CR being filled by different motif architectures over evolutionary time.

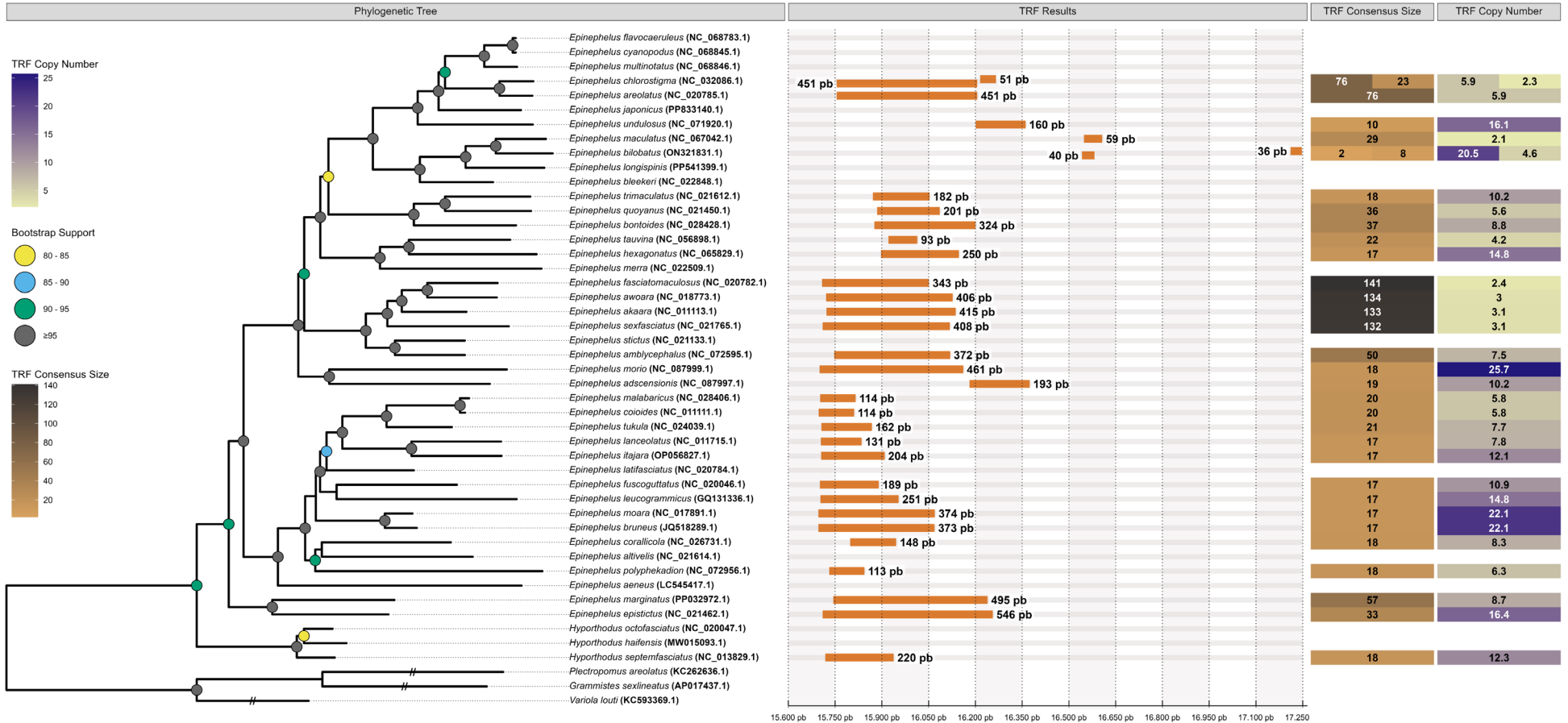


Figure 1 – Phylogeny of groupers based on 13 protein-coding genes. The TRF Consensus Size panel indicates the lengths of the tandem repeat cores, with different colors representing short, medium, and long repeat motifs. The TRF Number of Copies panel shows the number of times the core is repeated. The TRF Results panel displays the distribution of repeats along the control region, predominantly in the 5' domain, and their correspondence within the phylogeny.

Discussion

The present study revealed that Epinephelidae groupers share highly similar tandem repeat motif sequences both within species and among species. Belyayev *et al.* (2019) proposed that family-level satDNA mutations may initially spread gradually through mechanisms of homogenization to eventually become fixed in the population, resulting in the concerted evolution of these segments, a process observed in many different vertebrates (Lorite *et al.*, 2017; Utsunomia *et al.*, 2017; de Lima *et al.*, 2020; Jiang *et al.*, 2024; Dos Santos *et al.*, 2024).

Other mechanisms, such as species hybridization, could also explain the similarities in repetitive sequences among some species. For example, alleles from one species can introgress into the gene pool of a second species through hybrids and, in some cases, the complete mitochondrial genome of another species can be incorporated into that of the first species (Harrison and Larson, 2014; Read *et al.*, 2025). Hybridization events such as these are very common in the groupers (Kiriya *et al.*, 2011; Gao *et al.*, 2017; Murata *et al.*, 2017; Chen *et al.*, 2018; Triastuti *et al.*, 2018; Shapawi *et al.*, 2019; Fan *et al.*, 2020; Li *et al.*, 2020a; Zhang *et al.*, 2024; Aoki *et al.*, 2025).

Nevertheless, hybridization alone cannot explain the full evolutionary history of the tandem repeats observed across *Epinephelus*, because it is limited to a subset of species known to hybridize; even in those cases, shared repeats would be expected only when introgression involves the same repeat type. However, although the species *E. tukula* and *E. moara* can hybridize naturally (Li *et al.*, 2020b), they present different mitogenomes, as well as different repeats (Table 1; Table S1). In this case, the maintenance of the identity of these different groups of tandem repeats (Figure 1 and Table 1) could be better explained by the effects of concerted evolution on the CRs of these organisms, which are common observed both in this region in other species and multigene families (Hänske *et al.*, 2020; Shi *et al.*, 2020; Yano *et al.*, 2020; Thapana *et al.*, 2022; Jiang *et al.*, 2024).

These tandem repeats result from a series of varied and complex mechanisms, two of which are cited most often in published studies: gene conversion linked to slippage and the concerted evolution of regions of the mitochondrial CR (Morris-Pocock *et al.*, 2010; Akiyama *et al.*, 2017; Lorite *et al.*, 2017; Xiao *et al.*, 2017; Wynn and Christensen, 2019; Jiang *et al.*, 2024; Pham *et al.*, 2024). Point mutation is another plausible mechanism. Following the appearance of this type of variation (i.e. base insertion or deletion), the processes mentioned above are fundamental to the variation in the size (increase or decrease) of the motif sequences of the repeats (Xiao *et al.*, 2017; Wynn and Christensen, 2019; Lin *et al.*, 2021).

Maintenance vs. elimination of repeats in Epinephelidae.

The groups of repeats identified in the groupers suggest an ancient origin, with an evolutionary dynamic that is best understood from a phylogenetic perspective. The homogeneity found both within and among the groups of repeats, as mentioned above, is the result of concerted evolution, which

means that, the closer the species are in genetic terms, the more similar their repetitive motif sequences are (Ellingsen *et al.*, 2007; Schirrmeister *et al.*, 2012; De Luca *et al.*, 2021; Zhang *et al.*, 2021). This, together with their arrangement in the tree, implies that these types of repeats share a common ancestry (Belyayev *et al.*, 2019; Jiang *et al.*, 2024).

However, interestingly, some species lack repeats, suggesting that repeats may have been lost during the evolutionary process from an ancestor in which this feature was present. For example, *Epinephelus latifasciatus* and *Epinephelus aeneus* are the only species in their clade that do not have an *in tandem* repeats (Table 1; Figure 1). This suggests that the absence of repeats in some species is likely due to random loss through genetic drift (Xiao *et al.*, 2017; Butenko *et al.*, 2024) or, possibly, differential selection pressures.

Several studies have reported that sequence features associated with replication and transcription control can be conserved within mitochondrial genomes (Fernández-Silva *et al.*, 2003; San Mauro *et al.*, 2006; Zhuang *et al.*, 2013; Xu *et al.*, 2024). This may help explain, at least in part, the maintenance of certain repeats, given that the mitochondrial CR functions as the origin of H-strand replication and contains promoters involved in transcription from both the light and heavy strands (Kurabayashi and Sumida, 2013; Zhuang *et al.*, 2013). In addition, expansions or contractions of the full repeat region can change the local sequence context and spacing near the CR boundaries. In that sense, such changes may be associated with variation in adjacent regions, including nearby coding genes, through shifts in genome organization rather than by directly affecting coding function (Zardoya *et al.*, 1995; Dadkhah *et al.*, 2023).

Consistent with this interpretation, repeat-driven changes in CR architecture have also been linked to length variation in the mitochondrial CR across diverse taxa. For example, Liu *et al.* (2013) demonstrated that the mitochondrial genomes of the clam *Scapharca broughtonii* ranged in length from approximately 47 kb to approximately 50 kb due to the variation in the number of repeat copies. In this case, the number of tandem repeats found in this segment is correlated with the size of the CR. Lin *et al.* (2021) also reported tandem repeat motifs rich in GCs in the CR of the parasitoid wasp *Nasonia vitripennis* and concluded that these repeats are responsible for the expansion of the size of this region. This is because the repeat motifs found in this species may be folded into large palindromic structures, similar to loops, which contribute to the expansion of these motifs and, in turn, the size of the CR.

A strong relationship between motif length and the number of times these units were repeated was also observed in the *Epinephelus* species analyzed in the present study (Table 1 and Table S2). This finding is corroborated and supported by the correlation test used in our study, which revealed that motif length is strongly and inversely associated with repeat copy number. This is a key result for the biological question we raised, because it directly supports the hypothesis that tandem repeat architecture is shaped by a trade-off between motif length and repeat copy number within a constrained array.

Moreover, CR length was strongly positively associated with overall mitogenome length (Table S2). In addition, mitogenome length increased with the total size of the full

repeat region within the CR (Table S2), suggesting that repeat-rich segments can contribute, at least in part, to broader variation in mitogenome length. However, the most direct signal remains within the CR itself. In this case, the mitochondrial CR length increased with the length of the full repeat region. In other words, because the full tandem repeat region occupies a finite portion of the CR, variation in how much sequence it spans translates into measurable differences in mitochondrial CR length (Wang *et al.*, 2015; Zhao *et al.*, 2015).

Our data do not support the view that repeat variation alone explains mitochondrial CR or mitogenome length. Instead, repeats appear to represent one component within a broader set of factors shaping length variation (e.g., variation in intergenic spacers, localized insertions or deletions, and other lineage-specific structural features), and their relative contribution is likely to differ among lineages and/or species. Even so, the strong associations observed here, both between CR length and full repeat region length and between mitogenome length and repeat region length, indicate that repeats can represent a measurable component of length variation in this group.

These findings reinforce the notion that tandem repeat architecture contributes to length variation in the mitochondrial CR, and they are compatible with the idea that the CR of these organisms may be under constraints that limit its overall size (Hassanin *et al.*, 2009).

Concerted evolution and implications of *tandem* repeats for the phylogeny of Epinephelidae species

While the short and medium repeats are distributed throughout the phylogenetic tree, the long repeats are limited to a specific monophyletic clade (Figure 1). Short repeats are found all over the phylogenetic tree, and sometimes, closely related species (sister groups) can share the same basic repeat motif sequence, such as *E. moara* and *E. bruneus*, which have the same repeat size (17 bps) within the same region (between bases 1-372/373). Other groups possess different sizes of repeats, such as the monophyletic group composed of *E. quoyanus* (repeat motif with 18 or 36 bps), *E. trimaculatus* (18 bps) and *E. bontoides* (37 bps). These kinds of groups could be the keys to understanding how the repeat size evolves over time (Figure 2a). Here, we observed a clear alignment between the first 18 bases of the *E. quoyanus* repeat sequence and the 18 latter bases (Figure 2b), with only four differences. The same phenomenon was observed in the *E. bontoides* repeat sequence alignment (Figure 2b).

As noted above, with respect to the repeats of *E. quoyanus*, the analysis revealed two types of repeats in the same region (Figures S1a and S1b). This reflects a limitation of Tandem Repeats Finder analysis due to the characteristics observed in the repeat motif length of this region, but in fact, *E. quoyanus* has only a single type of repeat size. Initially, a more systematic analysis revealed that the 18 bps repeats of *E. quoyanus* are not homogeneous, but rather, are repeated in a nonconsecutive manner (Figure S1a), whereas the 36 bps motif

is repeated regularly and conserved *in tandem* (Figure S1b). Analysis of the short motifs revealed that odd repeats are homogeneous but differ in four bases from even repeats, which are also homogeneous (Figure 2b and S1a).

In other words, in the case of *E. quoyanus*, this appears to be a repeat of 36 bps (Figure 3a) that has been changed by mutations (Figure 3b) and the concerted homogenization of a motif that now contains two different repeat motif lengths in one, one odd and one even (Figure 3b-c). That is, it can be recognized as a single repeat motif length, now of 36 bps (Figures 2a and 3c).

If this is true, we could expect that other species of the same clade, with similar motif length patterns, would show the same pattern, that is, an increase in the size of the motif. First, the repeat of *E. bontoides* has only a single repeat motif sequence of 37 bps, which is highly similar to the 36 bps repeat sequence observed in *E. quoyanus* (Figure 2a). Second, the comparison of the 5' (16 bps) and 3' (17 bps) regions of the repeat of *E. bontoides* (Figure 2b) revealed a similar pattern to that observed in *E. quoyanus* (Figures S1c and S1d), with five differences in the bases between the two regions. In other words, both species present the same pattern of motif expansion, with the motifs of the different repeats in both species being formed by random mutations and concerted evolution independently.

This pattern is also observed in *E. trimaculatus*, which has a short repeat of 18 bps (Figure 3a), likely reflecting a subsequent reduction in size. This is possible because the different sizes of the repeat motifs compete to occupy the 5' region of the CR and may become shorter again through a new round of concerted evolution (Figure 3d).

One last example supports the hypothesis of the growth of the repeat motif. As mentioned above, two repeats of different sizes are found *in tandem* in *E. chlorostigma* (Table 1). A comparison of the two motifs revealed that the medium (76 bps) and short (24 bps) repeats can be easily aligned, as observed in the species mentioned above (Figure 2). However, this medium repeat is not an exact copy of three different motifs of ancestral tandem repeats, which may reflect the long period of time since the medium repeat was formed. This would increase the chance of new mutations and increase the motif length, which would ultimately result in greater differentiation within the repeat. In any case, these two repeats are arranged side by side and represent a clear example of sequences that compete to accumulate the largest possible number of copies.

This means that these sequences behave selfishly and, by genetic drift, may be able to increase either their frequency or size without necessarily being advantageous or disadvantageous in terms of the fitness of the organism while also guaranteeing their transmission (Hurst and Werren, 2001; Gardner and Úbeda, 2017; McLaughlin and Malik, 2017). A similar pattern is found in many groups of animals; in the present case, there is a selfish dispute among sequences with different motif lengths, which can ensure their dispersal and fixation (Klucnika and Ma 2019; Dubie *et al.*, 2020; Wagner *et al.*, 2020; Butenko *et al.*, 2024; Sequeira *et al.*, 2024).

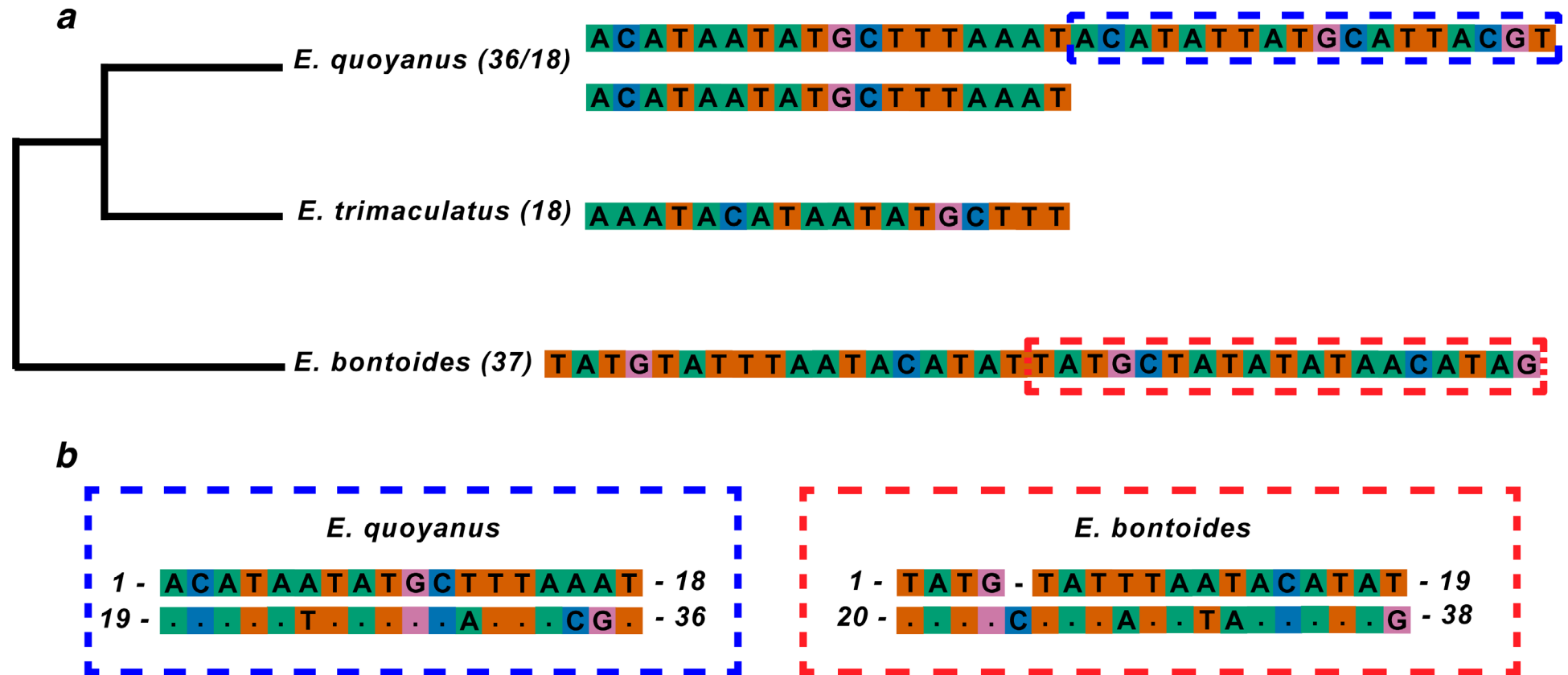


Figure 2 – (a) Phylogenetic relationships among *E. quoyanus*, *E. trimaculatus*, and *E. bontoides*. The tandem repeats found in these species are shown beside them. The dotted rectangles indicate the sequences highlighted in (b). In *E. quoyanus*, the same repeat may have two distinct forms (18 and 36 bps). The alignment of the halves of the sequences of 36 bps and 38 bps in *E. quoyanus* and *E. bontoides* (b) reveals some differences, which are clearly the result of mutations in a single ancestral repeat.

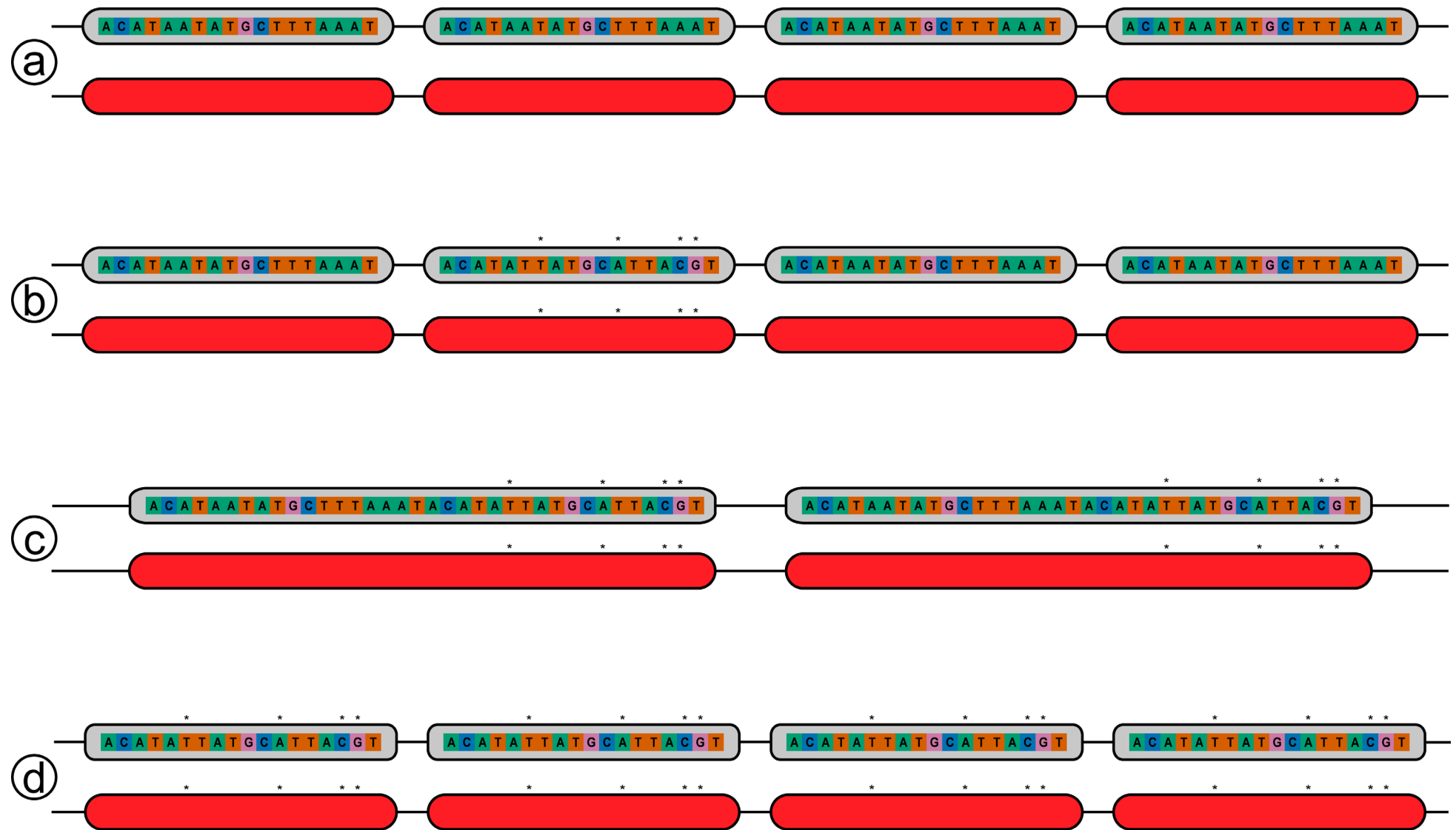


Figure 3 – Effects of in-concert evolution on the Control Region: (a) the core of a repeat, repeated four times; (b) four hypothetical mutations (*) in the second repeat; (c) homogenization derived from the in-concert evolution of a core made up of repeats 1 and 2, with a mean repeat of 36 bps; (d) potential competition between the short (ancestral) and medium (derived) repeats, resulting in homogenization of the short repeat, which is now different from the ancestral short repeat (a).

Acknowledgements

The authors are thankful to the Federal University of Pará for providing all the necessary research facilities to carry out this research. This study was financed by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) through a research project (Universal 422744/2018-2) and research fellowships to MV (303889/2022-5), and, in part, by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance Code 001.

Conflicts of Interest

All the authors declare that they have no conflicts of interest.

Author Contributions

CBB investigated, analyzed the data, conducted the analysis and writing – original draft preparation; WGL, RFR, LPR and YSO writing – review and editing; IS funding acquisition and supervision; RPCS and GFEG methodology, supervision, writing – review and editing; MV conceived the study, supervision, writing – review and editing. All the authors have read and approved the final version.

Data Availability

All the data used to support the findings of this study are included within the article.

References

- Akiyama T, Nishida C, Momose K, Onuma M, Takami K and Masuda R (2017) Gene duplication and concerted evolution of mitochondrial DNA in crane species. *Mol Phylogenet Evol* 106:158–163.
- Aoki R, Matsumasa T, Kumada T, Jin N and Masuma S (2025) Hatchability and growth performance of F1, F2, and backcross progenies of *Epinephelus bruneus* and *Epinephelus lanceolatus*. *Aquac Int* 33:291.
- Belyayev A, Josefiová J, Jandová M, Kalendar R, Krak K and Mandák B (2019) Natural history of a satellite DNA family: From the ancestral genome component to species-specific sequences, concerted and non-concerted evolution. *Int J Mol Sci* 20:1201.
- Benson G (1999) Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* 27:573–580.
- Bernacki LE and Kilpatrick CW (2020) Structural variation of the turtle mitochondrial control region. *J Mol Evol* 88:618–640.
- Bertoncini AA, Aguilar-Perera A, Barreiros J, Craig MT, Ferreira B and Koenig C (2018) *Epinephelus itajara*. IUCN 2018: T195409A46957794.
- Butenko A, Lukeš J, Speijer D and Wideman JG (2024) Mitochondrial genomes revisited: Why do different lineages retain different genes? *BMC Biol* 22:15.
- Calonje M, Martín-Bravo S, Dobeš C, Gong W, Jordon-Thaden I, Kiefer C and Koch MA (2009) Non-coding nuclear DNA markers in phylogenetic reconstruction. *Plant Syst Evol* 282:257–280.
- Cárcamo-Tejer V, Vila I, Llanquín-Rosas F, Sáez-Arteaga A and Guerrero-Jiménez C (2021) Phylogeography of high Andean killifishes *Orestias* (Teleostei: Cyprinodontidae) in Caquena and Lauca sub-basins of the Altiplano (Chile): Mitochondrial and nuclear analysis of an endangered fish. *PeerJ* 9:e11917.
- Chen ZF, Tian YS, Wang PF, Tang J, Liu JC, Ma WH and Zhai JM (2018) Embryonic and larval development of a hybrid between kelp grouper *Epinephelus moara* ♀ × giant grouper *E. lanceolatus* ♂ using cryopreserved sperm. *Aquac Res* 49:1407–1413.
- Cheng SS, Senoo S, Siddiquee S and Rodrigues KF (2015) Genetic variation in the mitochondrial genome of the giant grouper *Epinephelus lanceolatus* (Bloch, 1790) and its application for the identification of broodstock. *Aquac Rep* 2:139–143.
- Dadkhah K, Mianji GR, Barzegar A and Farhadi A (2023) Characterization of the mitochondrial *Huso huso* genome and new aspects of its organization in the presence of tandem repeats in 12S rRNA. *BMC Ecol Evol* 23:55.
- Damasceno JS, Siccha-Ramirez R, Morales MJ, Oliveira C, Torres RA, Costa EN and Hostim-Silva M (2015) Mitochondrial DNA evidences reflect an incipient population structure in Atlantic goliath grouper (*Epinephelus itajara*, Epinephelidae) in Brazil. *Sci Mar* 79:419–429.
- Davidović S, Marinković S, Kukobat M, Mihajlović M, Tanasić V, Hribšek I and Stamenković-Radak M (2022) Genetic diversity analysis of mitochondrial cytb gene, phylogeny and phylogeography of protected griffon vulture (*Gyps fulvus*) from serbia. *Life* 12:164.
- de Lima LG, Hanlon SL and Gerton JL (2020) Origins and evolutionary patterns of the 1.688 satellite DNA family in *Drosophila* phylogeny. *G3 (Bethesda)* 10:4129–4146.
- De Luca D, Kooistra WH, Sarno D, Biffali E and Piredda R (2021) Empirical evidence for concerted evolution in the 18S rDNA region of the planktonic diatom genus *Chaetoceros*. *Sci Rep* 11:807.
- Dos Santos GE, Crepaldi C, Da Silva MJ and Parise-Maltempi PP (2024) Revealing the satellite DNA content in *Ancistrus* sp. (Siluriformes: Loricariidae) by genomic and bioinformatic analysis. *Cytogenet Genome Res* 164:52–59.
- Dubie JJ, Caraway AR, Stout MM, Katju V and Berghthorsson U (2020) The conflict within: Origin, proliferation and persistence of a spontaneously arising selfish mitochondrial genome. *Philos Trans R Soc B* 375:20190174.
- Ellingsen A, Slamovits CH and Rossi MS (2007) Sequence evolution of the major satellite DNA of the genus *Ctenomys* (Octodontidae, Rodentia). *Gene* 392:283–290.
- Fadli N, Muchlisin ZA and Siti-Azizah MN (2021) DNA barcoding of commercially important groupers (Epinephelidae) in Aceh, Indonesia. *Fish Res* 234:105796.
- Fan B, Yang S, Wang L, Chen X, Liu X, Zhang Y and Lin H (2020) Hybridization of tiger grouper (*Epinephelus fuscoguttatus* ♀) × giant grouper (*Epinephelus lanceolatus* ♂) using cryopreserved sperm. *Cryobiology* 95:84–89.
- Feliner GN and Rosselló JA (2012) Concerted evolution of multigene families and homoeologous recombination. In: Wendel J, Greilhuber J, Dolezel J, Leitch I (eds) *Plant genome diversity*. Springer, Vienna, vol 1, pp 171–193.
- Fernández-Silva P, Enriquez JA and Montoya J (2003) Replication and transcription of mammalian mitochondrial DNA. *Exp Physiol* 88:41–56.
- Frale BW, Tucker SJ and Walker Jr HJ (2019) A new species of grouper, *Epinephelus craigi* (Perciformes: Epinephelidae), from the South China Sea. *Ichthyol Res* 66:215–224.
- Gao F, Wei M, Zhu Y, Guo H, Chen S and Yang G (2017) Characterization of the complete mitochondrial genome of the hybrid *Epinephelus moara* ♀ × *Epinephelus lanceolatus* ♂, and phylogenetic analysis in subfamily Epinephelinae. *J Ocean Univ China* 16:555–563.
- Gardner A and Úbeda F (2017) The meaning of intragenomic conflict. *Nat Ecol Evol* 1:1807–1815.
- Gomes C, Rodrigues-Filho LF, Sodré D, Neckel-Oliveira S, Gordo M, Gallati U and Vallinoto M (2016) Concerted evolution in the mitochondrial control region of the Amazon small-bodied frog *Pseudopaludicola canga* (Anura, Leiuperidae). *Mitochondrial DNA A DNA Mapp Seq Anal* 27:4270–4273.

- Guimarães-Costa A, Vallinoto M, Giarrizzo T, Pezold F, Schneider H and Sampaio I (2016) Molecular evidence of two new species of *Eleotris* (Gobiiformes: Eleotridae) in the western Atlantic. *Mol Phylogenet Evol* 98:52-56.
- Han J, Lv F and Cai H (2011) Detection of species-specific long VNTRs in mitochondrial control region and their application to identifying sympatric Hong Kong grouper (*Epinephelus akaara*) and yellow grouper (*Epinephelus awoara*). *Mol Ecol Resour* 11:215-218.
- Han SH, Lee YD, Baek HJ, Oh HS and Noh CH (2014) Genetic structure and phylogenetic relationship of red spotted grouper (*Epinephelus akaara*) based on the haplotypes and polymorphisms of mitochondrial COI gene sequences. *J Life Sci* 24:626-632.
- Hänske J, Hammacher T, Grenkowitz F, Mansfeld M, Dau TH, Maksimov P and Kammerer R (2020) Natural selection supports escape from concerted evolution of a recently duplicated CEACAM1 paralog in the ruminant CEA gene family. *Sci Rep* 10:3404.
- Harrison RG and Larson EL (2014) Hybridization, introgression, and the nature of species boundaries. *J Hered* 105:795-809.
- Hassanin A, Ropiquet A, Couloux A and Cruaud C (2009) Evolution of the mitochondrial genome in mammals living at high altitude: New insights from a study of the tribe Caprini (Bovidae, Antilopinae). *J Mol Evol* 68:293-310.
- He H, Gao Z, Hu Z, Cai L, Huang Y, Zhou M and Liang R (2024) Comparative characterization and phylogenetic analysis of complete mitogenome of three taxonomic confused groupers and the insight to the novel gene tandem duplication in *Epinephelus*. *Front Mar Sci* 11:1450003.
- Hurst GD and Werren JH (2001) The role of selfish genetic elements in eukaryotic evolution. *Nat Rev Genet* 2:597-606.
- Jiang C, Kang H, Zhou Y, Zhu W, Zhao X, Mohamed N and Li B (2024) Selected lark mitochondrial genomes provide insights into the evolution of second control region with Tandem Repeats in *Alaudidae* (Aves, Passeriformes). *Life* 14:881.
- Kiriyakit A, Gallardo WG and Bart AN (2011) Successful hybridization of groupers (*Epinephelus coioides* x *Epinephelus lanceolatus*) using cryopreserved sperm. *Aquaculture* 320:106-112.
- Klucnika A and Ma H (2019) A battle for transmission: The cooperative and selfish animal mitochondrial genomes. *Open Biol* 9:180267.
- Kornienko IV, Faleeva TG, Oreshkova NV, Grigoriev SE, Grigorieva LV, Putintseva YA and Krutovsky KV (2019) Structural and functional organization of the mitochondrial DNA control region in the woolly mammoth (*Mammuthus primigenius*). *Mol Biol* 53:560-570.
- Katoh K and Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 30:772-780.
- Kurabayashi A and Sumida M (2013) Afrobatrachian mitochondrial genomes: Genome reorganization, gene rearrangement mechanisms, and evolutionary trends of duplicated and rearranged genes. *BMC Genom* 14:633.
- Lawless C, Graves L, Reeve AK, Turnbull DM and Vincent AE (2020) The rise and rise of mitochondrial DNA mutations. *Open Biol* 10:200061.
- Li Q, He X, Ren Y, Xiong C, Jin X, Peng L and Huang W (2020a) Comparative mitogenome analysis reveals mitochondrial genome differentiation in ectomycorrhizal and asymbiotic *Amanita* species. *Front Microbiol* 11:1382.
- Li Z, Tian Y, Cheng M, Wang L, Zhang J, Wu Y and Zhai J (2020b) The complete mitochondrial genome of the hybrid grouper *Epinephelus moara* (♀) x *Epinephelus tukula* (♂), and phylogenetic analysis in subfamily Epinephelinae. *Acta Oceanol Sin* 39:65-75.
- Lin ZJ, Wang X, Wang J, Tan Y, Tang X, Werren JH and Wang X (2021) Comparative analysis reveals the expansion of mitochondrial DNA control region containing unusually high GC tandem repeat arrays in *Nasonia vitripennis*. *Int J Biol Macromol* 166:1246-1257.
- Li J, Seeber P, Axtner J, Crouthers R, Groenenberg M, Koehncke A and Greenwood AD (2023) Monitoring terrestrial wildlife by combining hybridization capture and metabarcoding data from waterhole environmental DNA. *Biol Conserv* 284:110168.
- Liu YG, Kurokawa T, Sekino M, Tanabe T and Watanabe K (2013) Complete mitochondrial DNA sequence of the ark shell *Scapharca broughtonii*: An ultra-large metazoan mitochondrial genome. *Comp Biochem Physiol D Genom Proteom* 8:72-81.
- Lorite P, Muñoz-López M, Carrillo JA, Sanllorente O, Vela J, Mora P and Palomeque T (2017) Concerted evolution, a slow process for ant satellite DNA: Study of the satellite DNA in the *Aphaenogaster* genus (Hymenoptera, Formicidae). *Org Divers Evol* 17:595-606.
- Ma KY and Craig MT (2018) An inconvenient monophyly: An update on the taxonomy of the groupers (Epinephelidae). *Ichthyol Herpetol* 106:443-456.
- McLaughlin Jr RN and Malik HS (2017) Genetic conflicts: The usual suspects and beyond. *J Exp Biol* 220:6-17.
- Clayton DA (1992) Structure and function of the mitochondrial genome. *J Inherit Metab Dis* 15:439-447.
- Morris-Pocock JA, Taylor SA, Birt TP and Friesen VL (2010) Concerted evolution of duplicated mitochondrial control regions in three related seabird species. *BMC Evol Biol* 10:14.
- Murata O, Itakura S, Yamamoto S, Hattori N, Kurata M, Ohta H and Masuma S (2017) Interspecific hybridization of longtooth grouper *Epinephelus bruneus* x giant grouper *E. lanceolatus* and growth of hybrid larvae and juveniles. *Aquacult Sci* 65:93-95.
- Nelson JS, Grande TC and Wilson MV. (2016) *Fishes of the World*. John Wiley & Sons, New Jersey.
- Nguyen LT, Schmidt HA, Von Haeseler A and Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268-274.
- Padgett K and Baeza JA (2025) The mitochondrial genome of the imperiled Goliath Grouper *Epinephelus itajara*: Selective pressures in protein coding genes, secondary structure of tRNA genes, and phylogenetic placement. *Ecol Evol* 15:e171795.
- Pham LTK, Van Quyen D, Saijuntha W, Doan HTT, Le TH and Lawton SP (2024) Mitogenomics of the zoonotic parasite *Echinostoma miyagawai* and insights into the evolution of tandem repeat regions within the mitochondrial non-coding control region. *Parasitology* 151:543-1554.
- Pinhal D, Shivji MS, Vallinoto M, Chapman DD, Gadig OBF and Martins C (2012) Cryptic hammerhead shark lineage occurrence in the western South Atlantic revealed by DNA analysis. *Mar Biol* 159:829-836.
- Qing X, Karlicki M, Guo F, Karnkowska A and Li H (2024) Soil nematode community profiling using reference-free mitogenomics. *Soil Biol Biochem* 199:109613.
- Qu M, Tang W, Liu Q, Wang D and Ding S (2018) Genetic diversity within grouper species and a method for interspecific hybrid identification using DNA barcoding and RYR3 marker. *Mol Phylogenet Evol* 121:46-51.
- Read WJ, Laver RJ, Lau CC, Moritz C and Zozaya SM (2025) Repeated mitochondrial capture with limited genomic introgression in a lizard group. *Mol Ecol* 34:e17766.
- Røyrvik EC and Johnston IG (2020) MtDNA sequence features associated with 'selfish genomes' predict tissue-specific segregation and reversion. *Nucleic Acids Res* 48:8290-8301.
- Fonseca MM, Harris DJ and Posada D (2014) The inversion of the control region in three mitogenomes provides further evidence for an asymmetric model of vertebrate mtDNA replication. *PLoS One* 9:9.

- Sadovy de Mitcheson Y, Craig MT, Bertoni AA, Carpenter KE, Cheung WW, Choat JH and Sanciangco J (2013) Fishing groupers towards extinction: A global assessment of threats and extinction risks in a billion dollar fishery. *Fish Fish* 14:119-136.
- San Mauro D, Gower DJ, Zardoya R and Wilkinson M (2006) A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. *Mol Biol Evol* 23: 227-234.
- Schirrmeister BE, Dalquen DA, Anisimova M and Bagheri HC (2012) Gene copy number variation and its significance in cyanobacterial phylogeny. *BMC Microbiol* 12:177.
- Sequeira AN, O'Keefe IP, Katju V and Berghthorsson U (2024) Friend turned foe: Selfish behavior of a spontaneously arising mitochondrial deletion in an experimentally evolved *Caenorhabditis elegans* population. *G3* 14jkae018.
- Shapawi R, Abdullah FC, Senoo S and Mustafa S (2019) Nutrition, growth and resilience of tiger grouper (*Epinephelus fuscoguttatus*) × giant Grouper (*Epinephelus lanceolatus*) hybrid- a review. *Rev Aquac* 11:1285-1296.
- Shi W, Gong L and Yu H (2020) Double control regions of some flatfish mitogenomes evolve in a concerted manner. *Int J Biol Macromol* 142:11-17.
- Shu L, Wanghe K, Wang J, Niu Y, Zhao K and Peng Z (2018) The complete mitochondrial genomes of two schizothoracine fishes (Teleostei, Cypriniformes): A novel minisatellite in fish mitochondrial genomes. *J Appl Ichthyol* 34:1028-1034.
- Silva-Oliveira GC, Rêgo PSD, Schneider H, Sampaio I and Vallinoto M (2008) Genetic characterisation of populations of the critically endangered Goliath grouper (*Epinephelus itajara*, Serranidae) from the Northern Brazilian coast through analyses of mtDNA. *Genet Mol Biol* 31:988-995.
- Sodré D, Rodrigues-Filho LF, Souza RF, Rêgo PS, Schneider H, Sampaio I and Vallinoto M (2012) Inclusion of South American samples reveals new population structuring of the blacktip shark (*Carcharhinus limbatus*) in the western Atlantic. *Genet Mol Biol* 35:752-760.
- Song Y, Yu QF, Zhang D, Chen LG, Tan YH, Zhu W and Corlett RT (2024) New insights into the phylogenetic relationships within the Lauraceae from mitogenomes. *BMC Biol* 22:241.
- Sun CH, Liu HY, Xu N, Zhang XL, Zhang Q and Han BP (2021) Mitochondrial genome structures and phylogenetic analyses of two tropical characidae fishes. *Front Genet* 12:627402.
- Thapana W, Ariyaphong N, Wongtienchai P, Laopichienpong N, Singchat W, Panthum T and Srikulnath K (2022) Concerted and independent evolution of control regions 1 and 2 of water monitor lizards (*Varanus salvator macromaculatus*) and different phylogenetic informative markers. *Animals* 12:148.
- Triastuti J, Pursetyo KT, Monica A, Lutfiyah L and Budi DS (2018) Abnormalities of hybrid grouper (*Epinephelus fuscoguttatus* × *Epinephelus lanceolatus*) in Situbondo. *Earth Environ Sci* 137:012035.
- Utsunomia R, Ruiz-Ruano FJ, Silva DM, Serrano ÉA, Rosa IF, Scudeler PE and Foresti F (2017) A glimpse into the satellite DNA library in *Characidae* fish (Teleostei, Characiformes). *Front Genet* 8:103.
- Vaidya G, Lohman DJ and Meier R (2011) SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* 27:171-180.
- Wagner JT, Howe DK, Estes S and Denver DR (2020) Mitochondrial DNA variation and selfish propagation following experimental bottlenecks in two distantly related *Caenorhabditis briggsae* isolates. *Genes* 11:77.
- Winn JC, Maduna SN and Bester-van der Merwe AE (2024) A comprehensive phylogenomic study unveils evolutionary patterns and challenges in the mitochondrial genomes of Carcharhiniformes: A focus on Triakidae. *Genomics* 116:110771.
- Wynn EL and Christensen AC (2019) Repeats of unusual size in plant mitochondrial genomes: Identification, incidence and evolution. *G3 (Bethesda)* 9:549-559.
- Xiao S, Nguyen DT, Wu B and Hao W (2017) Genetic drift and indel mutation in the evolution of yeast mitochondrial genome size. *Genome Biol Evol* 9:3088-3099.
- Xu P, Yang T, Kundnani DL, Sun M, Marsili S, Gombolay AL and Storici F (2024) Light-strand bias and enriched zones of embedded ribonucleotides are associated with DNA replication and transcription in the human-mitochondrial genome. *Nucleic Acids Res* 52:1207-1225.
- Yano CF, Merlo MA, Portela-Bens S, Cioffi MDB, Bertollo LA, Santos-Júnior CD and Rebordinos L. (2020) Evolutionary dynamics of multigene families in *Triportheus* (Characiformes, Triportheidae): A transposon mediated mechanism? *Front Mar Sci* 7:6.
- Zardoya R, Garrido Pertierra A, Bautista JM, Villalta M, López-Pérez MJ and Montoya J (1995) Nucleotide sequence of the sheep mitochondrial DNA D-loop and its flanking tRNA genes. *Curr Genet* 28:94-96.
- Zhang M, Tang YW, Xu Y, Yonezawa T, Shao Y, Wang YG and Zhang WJ (2021) Concerted and birth-and-death evolution of 26S ribosomal DNA in *Camellia* L. *Ann Bot* 127:63-73.
- Zhang X, Li J, Chen S, Yang Q, Ji F, Luo H and Ding S (2024) The Hulong hybrid grouper (*Epinephelus fuscoguttatus* ♀ × *Epinephelus lanceolatus* ♂) has invaded the coastal waters of Hainan Island, China. *Front Mar Sci* 11:1438895.
- Zhao G, Li H, Zhao P and Cai W (2015) Comparative mitogenomics of the assassin bug genus *Peirates* (Hemiptera: Reduviidae: Peiratinae) reveal conserved mitochondrial genome organization of *P. atromaculatus*, *P. fulvescens* and *P. turpis*. *Plos One* 10:e0117862.
- Zhuang X, Qu M, Zhang X and Ding S (2013) A comprehensive description and evolutionary analysis of 22 grouper (Perciformes, Epinephelidae) mitochondrial genomes with emphasis on two novel genome organizations. *PloS One* 8:e73561.
- Wang X, Liu N, Zhang H, Yang X, Huang Y and Lei F (2015) Extreme variation in patterns of tandem repeats in mitochondrial control region of yellow-browed tits (*Sylviparus modestus*, Paridae). *Sci Rep* 5:13227.

Internet Resources

- Fricke R, Eschmeyer WN and Fong JD (2025) Eschmeyer's catalog of fishes: Genera, species, references, <http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp> (accessed 30 March 2025).

Supplementary Material

Figure S1 – Results of the analysis run at Tandem Repeats Finder.

Table S1 – Species of the Epinephelidae family included in the phylogenetic reconstruction.

Table S2 – Correlation analyses of the relationships between tandem repeat architecture and sequence length metrics in *Epinephelus* mitochondrial control regions.

Associate Editor: Carlos F.M. Menck

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License (type CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original article is properly cited.