

Comparative mitogenomics of marine angelfishes (F: Pomacanthidae)

Lauriane M. Baraf¹  | Julia Y. Hung¹ | Morgan S. Pratchett¹ | Peter F. Cowman^{1,2} 

¹College of Science and Engineering, James Cook University, Townsville, Queensland, Australia

²Biodiversity and Geosciences Program, Queensland Museum Tropics, Townsville, Queensland, Australia

Correspondence

Lauriane M. Baraf, College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia.

Email: lauriane.baraf@my.jcu.edu.au

Funding information

ARC DECRA Fellowships, Grant/Award Number: DE170100516; ARC Centre of Excellence Programme, Grant/Award Number: CE140100020

Abstract

The targeted capture of ultraconserved elements (UCEs) has substantially increased the amount of genetic data available for phylogenomic reconstructions. These capture datasets frequently contain mitochondrial DNA as a by-product, often in the form of complete mitogenomes. These can be efficiently harvested to expand existing datasets without additional costs. Here, we present new mitochondrial genomes for six marine angelfish species (F: Pomacanthidae), assembled and annotated from off-target UCE reads. We provide the first comparative analysis of all mitochondrial genomes available for the Pomacanthidae. Results showed that the average length of pomacanthid mitogenomes is 16.8 kbp. Total GC and AT content varied between 44.5% and 46.3%, and 53.7% and 55.5%, respectively. The architecture of angelfish mitogenomes was comparable to that seen in other fish species with 13 protein-coding genes (PCGs), 22 transfer RNA genes, two ribosomal RNA genes and the control region. All 13 PCGs evolved under purifying selection, highlighting a high level of selection pressure and gene expression to preserve genetic integrity. The ND6 and ATP8 genes had the highest ratio of non-synonymous (dN) to synonymous (dS) substitutions, indicating a relaxation of purifying selection constraints. Finally, these newly assembled mitogenomes will allow further investigations of the population genetics, systematics and evolutionary biology of one of the most prominent reef fish family in the aquarium trade.

KEYWORDS

bioinformatics, mitogenome, off-target reads, Pomacanthidae, reef fish, ultraconserved elements

TAXONOMY CLASSIFICATION

Genomics

1 | INTRODUCTION

Primarily designed to target specific regions of the genome, sequence capture (or enrichment) methods have the potential to

recover complete mitochondrial genomes as a by-product from off-target sequences (do Amaral et al., 2015; Miller et al., 2022; Samuels et al., 2013; Smith et al., 2014). Leveraging these off-target reads provides a valuable source of informative mitochondrial

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Ecology and Evolution* published by John Wiley & Sons Ltd.

genes that can have broad applications ranging from species identification and environmental biomonitoring through DNA barcoding, population genetics or phylogenetic reconstructions (DeSalle & Goldstein, 2019; Drummond et al., 2015; Miller et al., 2022; Pereira, 2000; Zarza et al., 2016). Without the need for additional sequencing or mtDNA-specific baits, DNA reference libraries can expand readily and rapidly at no further cost. This is particularly beneficial for non-model or taxonomically underrepresented organisms where genetic coverage is limited. Retrospectively, this approach can also help enrich existing multilocus genetic datasets to maximise their utility and taxonomic or genetic sampling. In general, multi-marker-based datasets were shown to yield more accurate species identification, biodiversity assessment and phylogenetic relationships than single-gene approaches (Baldauf et al., 2000; Hillis, 1996; Meyer & Paulay, 2005; Rubinoff & Holland, 2005). While the genomic era has provided the means for applying both reduced and whole genome methods in phylogenetic and evolutionary studies, there is still a place for comparing new datasets with published sequence data. In many cases, the extensive mitochondrial database available for species can be included in phylogenomic assessment through a holistic approach that involves extracting legacy markers from off-target reads (do Amaral et al., 2015; Quattrini et al., 2023).

Increased sequencing of mitochondrial material improves our understanding of the composition, structure and variation of mitogenomes, subsequently providing a framework to investigate genetic processes underlying the evolution of lineages (Mercer et al., 2011; Satoh et al., 2016; Taanman, 1999). The fish mitochondrial genome is a highly conserved, circular and double-stranded DNA molecule (Satoh et al., 2016) ranging from 15 to 24 kbp, as per NCBI GenBank Database in January 2024. It generally follows the same arrangement as other vertebrate mitogenomes, usually containing 37 genes of which 13 are protein-coding genes (PCGs), two are ribosomal RNA (rRNA) genes, 22 are transfer RNA (tRNA) genes and one is a non-coding control region (D-loop) (Satoh et al., 2016). As the most speciose group of vertebrates, fishes, have been of particular interest to scientists aiming to unravel evolutionary mechanisms driving diversification processes (Brawand et al., 2014; Dornburg & Near, 2021; Friedman et al., 2022; Near et al., 2013; Volff, 2005). This includes the highly diverse reef fish lineages that inhabit coral reef systems worldwide.

The family Pomacanthidae contains some of the most iconic and recognisable reef fish species. Their striking colour patterns and dramatic change in coloration while transitioning from juvenile to adult stage made them among the most valued specimens in the ornamental fish trade today (Wabnitz, 2003). There are 90 valid nominal species and seven genera of marine angelfishes (Fricke et al., 2023). Complete mitochondrial genomes are available on the NCBI Reference Sequence database for only 33 pomacanthid species across the seven genera, representing 37% taxonomic coverage for the family. Several studies have examined the systematics and ancestral biogeography of pomacanthids using mitochondrial data (Alva-Campbell et al., 2010; Baraf et al., 2019;

DiBattista et al., 2012; Gaither et al., 2014) but no comprehensive comparative analysis of their mitogenomes has been conducted. In addition to expanding the NCBI Reference sequence database for the family Pomacanthidae, additional mitogenomic resources will allow further research aiming to delve into the genetic diversity, differentiation and evolutionary history of reef fishes. Moreover, marine angelfishes exhibit of the highest occurrence of hybridisation recorded among reef fishes (Tea et al., 2020), a biological process that can be challenging to uncover using nuclear markers which undergo recombination (Wallis et al., 2017). Past introgression might be more readily identifiable using mitochondrial DNA (mtDNA) rather than nuclear DNA due to its maternal inheritance and non-recombining nature (Wallis et al., 2017). Although it will not be able to differentiate invasion-driven from selection-driven introgression (Seixas et al., 2018). Furthermore, asymmetric discordance between mitochondrial and nuclear markers can inform on degrees of sequence divergence and structural variation, helping to detect mitochondrial introgression (Toews & Brelsford, 2012; Wallis et al., 2017). Expanding the mitochondrial genome library for pomacanthids will serve as a valuable tool to help answer evolutionary questions about the family amidst the complexities introduced by hybridisation events.

In the present study, we assembled and describe new mitogenomes for six pomacanthid species (Figure 1) before conducting comparative and phylogenetic analyses across all mitogenomes available on the NCBI Reference Sequence Database.

2 | MATERIALS AND METHODS

2.1 | Taxon sampling, UCE sequencing and data handling

Target capture of ultraconserved elements (UCEs) was conducted across 68 samples representing 45 distinct nominal species of Pomacanthidae or 50% of the family. We extracted DNA from tissue samples of 45 pomacanthid species using DNeasy Blood and tissue kits (Qiagen) following the manufacturer's instructions. Each sample was then quantified with a Qubit 2.0 Fluorometer using Qubit dsDNA BR Assay kits and Nanodrop spectrophotometer (Thermo Fisher Scientific Inc.) to measure DNA quality. Genomic DNAs were shipped to Arbor Bioscience (US) for library preparation and targeted enrichment following MyBaits v.5.0 protocol using the *acanthomorph* probe set (Alfaro et al., 2018). Adapters were trimmed off of raw reads in the illumiprocessor v.0.39 wrapper program (Faircloth et al., 2012) using the *trimomatic* package (Bolger et al., 2014) with default parameters set for a minimum output sequence length of 40bp and Phred quality score of 33. The targeted UCE dataset has been analysed in a separate study (Baraf et al., in preparation). Here we concentrate on harvesting mitogenomic data from the off-target reads. To facilitate this, we built a reference mitogenomic database by downloading all 33 mitochondrial genomes for the family Pomacanthidae available on the NCBI Reference Sequence

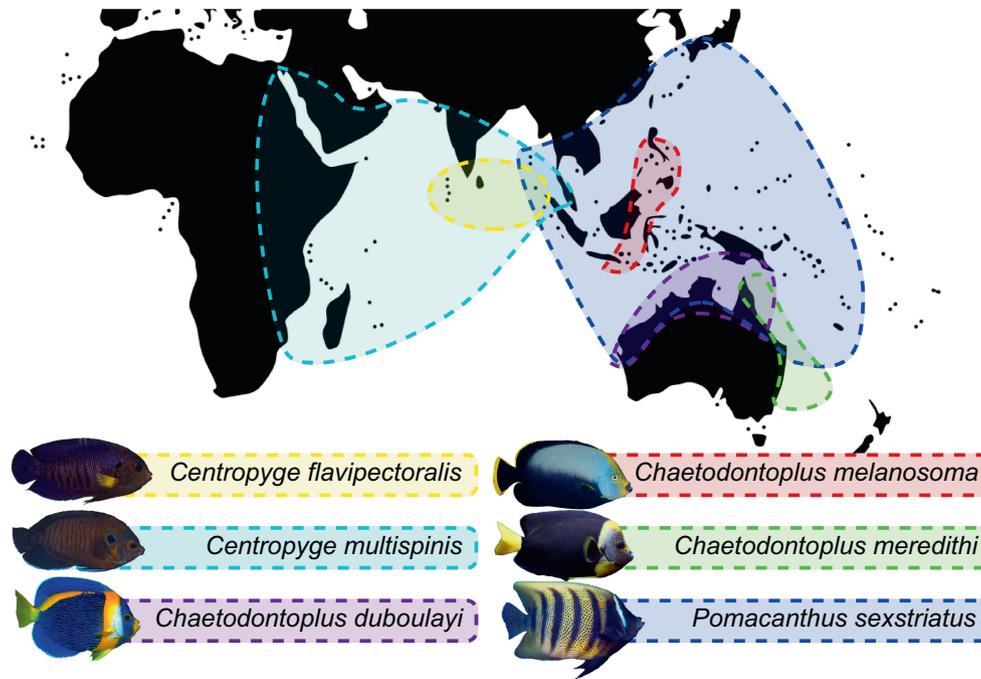


FIGURE 1 Geographical distribution of the six pomacanthid species sampled in this study. Pictures were provided by François Libert, Yi-Kai Tea, Jake Adams or obtained from the Wikimedia Images Commons. GenBank accession numbers for the six assembled and annotated mitogenomes are as follow: *Centropyge flavivectoralis* [PP316127], *Centropyge multispinis* [PP316128], *Chaetodontoplus duboulayi* [PP316126], *Chaetodontoplus meredithi* [PP316124], *Chaetodontoplus melanosoma* [PP316125] and *Pomacanthus sexstriatus* [PP316129].

database. We manually verified that gene nomenclature was consistent across all genomes and standardised sequence's names when needed (e.g. control region changed to D-loop).

2.2 | Mitogenome assembly and annotation

Cleaned raw reads were assembled and scanned for mitochondrial contigs, and annotated using MitoFinder v.1.4.1 (Allio et al., 2020). Due to the fragmented nature of mitochondrial sequences found in capture data, we estimated coverage for all contigs following the mapping workflow described in <https://phyluce.readthedocs.io/en/latest/daily-use/daily-use-4-workflows.html>, and set a minimum contig size of 500bp instead of the default 1000bp. Using the MitoFinder workflow, mitochondrial contigs were assembled with MetaSPAdes assembler (Nurk et al., 2017) and annotated using BLAST v.2.12.0 against the reference mitogenomic database with a percentage of overlap in BLAST best hit set at 20% and a e-value cut-off $\leq 1e-06$. Non-standard animal mitochondrial genes and tRNA genes were annotated with the `--new-genes` option and MitFi annotation pipeline (Juhling et al., 2012), respectively. A second round of annotation was conducted on the assembled mitochondrial contigs with the addition of the `--adjust-direction` option to establish the direction of contigs relative to those in the reference database. Final annotations were validated by assessing the consistency between annotations from MitoFinder and those from MitoAnnotator, a pipeline implemented in the Mitochondrial Genome Database of fish MitoFish v.3.90 (Iwasaki et al., 2013; Zhu et al., 2023). We

manually inspected the contigs and their alignments to the reference mitogenomes in Geneious v.2023.2.1. Finally, we extracted all COI sequences obtained from MetaSPAdes and blasted them against the NCBI database using BLASTN v. 2.14.1 (Zhang et al., 2000) and Megablast (Morgulis et al., 2008) to confirm species identification.

2.3 | Mitogenomes statistics

We calculated AT/GC-skew to assess the nucleotide bias and analyse strand asymmetry to describe overall patterns of nucleotide composition following the equation $AT-skew = (A - T)/(A + T)$ and $GC-skew = (G - C)/(G + C)$ (Perna & Kocher, 1995). We used AliGROOVE v.1.08 (Kück et al., 2014) to examine nucleotide composition and divergence patterns of 13 PCGs for all six newly assembled and the 33 existing RefSeq pomacanthid mitogenomes. To tag potentially unreliable species relationships, we used the phylogenetic tree inferred from IQTree (Nguyen et al., 2015) as a guiding tree and calculated reliabilities of single branches. To assess degree of sequence heterogeneity, we also calculated site score for amino acid to compare it to nucleotide-based results. To investigate evolutionary patterns in PCGs we calculated average non-synonymous (dN) and synonymous (dS) substitutions values of all pairwise comparisons, and dN/dS ratios between alignments of 13 PCGs for all pomacanthid species in DnaSP v.5 (Librado & Rozas, 2009). Additionally, we estimated pairwise genetic distances from mt PCGs using Kimura's two-parameter (K2P) model (Kimura, 1980) in MEGA v.11 (Tamura et al., 2021) and highest nucleotide diversity with R using the package *pegas*.

2.4 | Alignment and phylogenetic analysis

Protein-coding gene sequences for the 39 pomacanthid species and one outgroup species from the family Acanthuridae were aligned individually in Geneious using the MAFFT aligner (Kato et al., 2002; Kato & Standley, 2013) with the default parameters settings and manually curated when necessary. Multiple sequence alignments were then concatenated into a 75% complete matrix. After finding the best-fit partition model by running ModelFinder combined with the greedy algorithm from PartitionFinder2, we inferred a maximum likelihood (ML) tree from the concatenated alignments in IQTree (Kalyaanamoorthy et al., 2017; Lanfear et al., 2017; Nguyen et al., 2015). Branch support was estimated by Ultrafast Bootstrap approximation (Hoang et al., 2017) and SH-like approximate likelihood ratio test (Guindon et al., 2010) for 1000 bootstrap replicates.

3 | RESULTS

3.1 | Genome structure and nucleotide composition

From the 68 samples representing 45 pomacanthid species, UCE reads yielded complete and partial mitogenomes for 12 (or 18%

across samples) and 11 pomacanthid species, respectively. A total of 482 mitochondrial genes were found across all samples with the COI gene being the most frequently recovered marker (Figure 2). NCBI Megablast results on COI sequences had over 99% identity blast hits with conspecifics.

Lengths of the six assembled pomacanthid mitogenomes ranged from 16,496 for *Pomacanthus sexstriatus* to 17,544 bp, for *Centropyge multispinis* (Figures 1, 3 and 4), with an average length of 17 kbp. Size difference across mitogenomes could be attributed in majority to the length variation of the control region (Figure 4). All mitogenomes (except for *Chaetodontoplus meredithi* and *Pomacanthus sexstriatus*) were assembled from complementary strands of the reference DNA sequences and had a reversed direction. This likely resulted from the short read lengths of off-target sequences extracted from UCE captured data. Total GC and AT content varied from 44.6% to 46.3% and 53.7% to 55.5%, respectively, with a mean base composition of A: 27.7%, T: 26.6%, G: 16.5% and C: 29.1% (Table 1). The AT-skews were all positive (0.005 to 0.035) to the exception of *C. flavipectoralis* whereas GC-skews were all negative (−0.31 to −0.23, Table 1). All mitogenomes contained 37 genes – 13 PCGs (ND1, ND2, COXI, COXII, ATP8, ATP6, COXIII, ND3, ND4L, ND4, ND5, ND6 and CYTB), 22 tRNA genes, two rRNA (12S and 16S) genes and the control region (Figure 3). Mitochondrial ND6 and eight tRNA

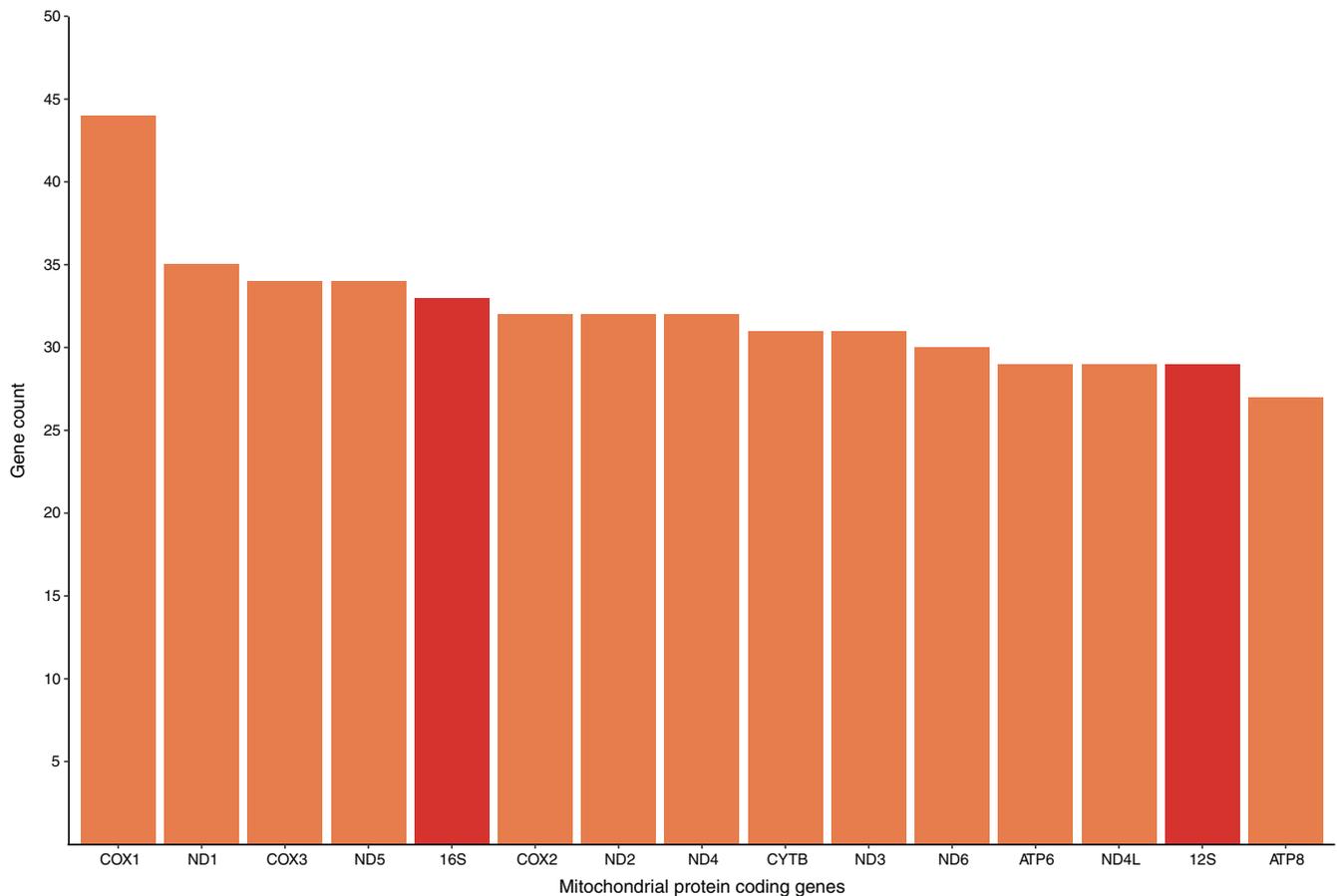


FIGURE 2 Total number of mitochondrial protein-coding genes (orange) and rRNAs (red) retrieved from UCE reads after two MitoFinder runs on 68 samples across 45 pomacanthid species.

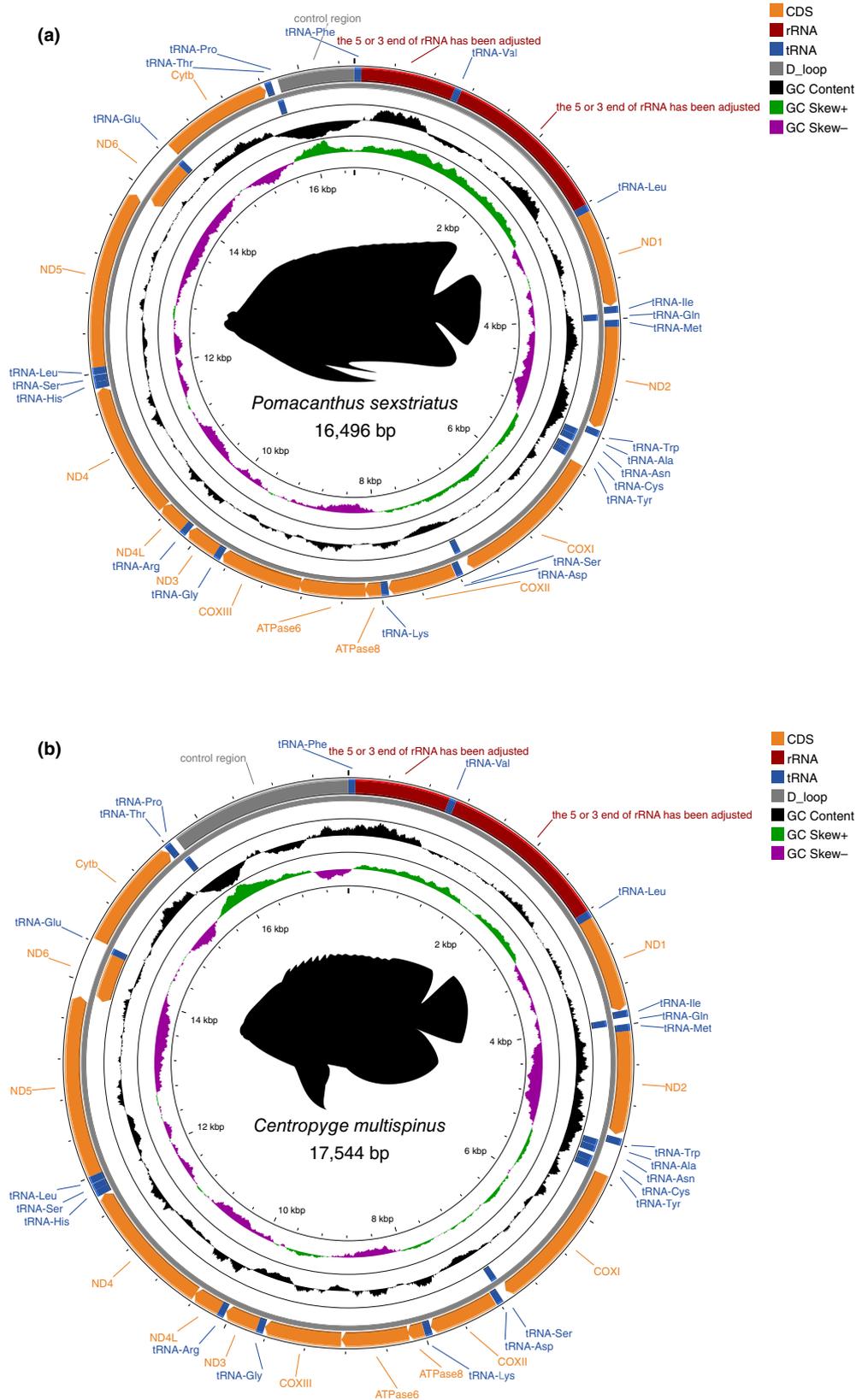


FIGURE 3 Mitochondrial genomes for two pomacanthid species, *Pomacanthus sexstriatus* and *Centropyge multispinus* of length 16,496 and 17,544 bp respectively. Circular gene maps were generated with cgview (Stothard & Wishart, 2005). Both mitogenomes contained 13 genes, two rRNAs and 22 tRNAs. Outer and inner circles represent the H-strand and L-strand of the mitochondrial DNA, respectively. Fish silhouette were drawn by L.M.B. Gene maps for the other four pomacanthid species are in Data S1.

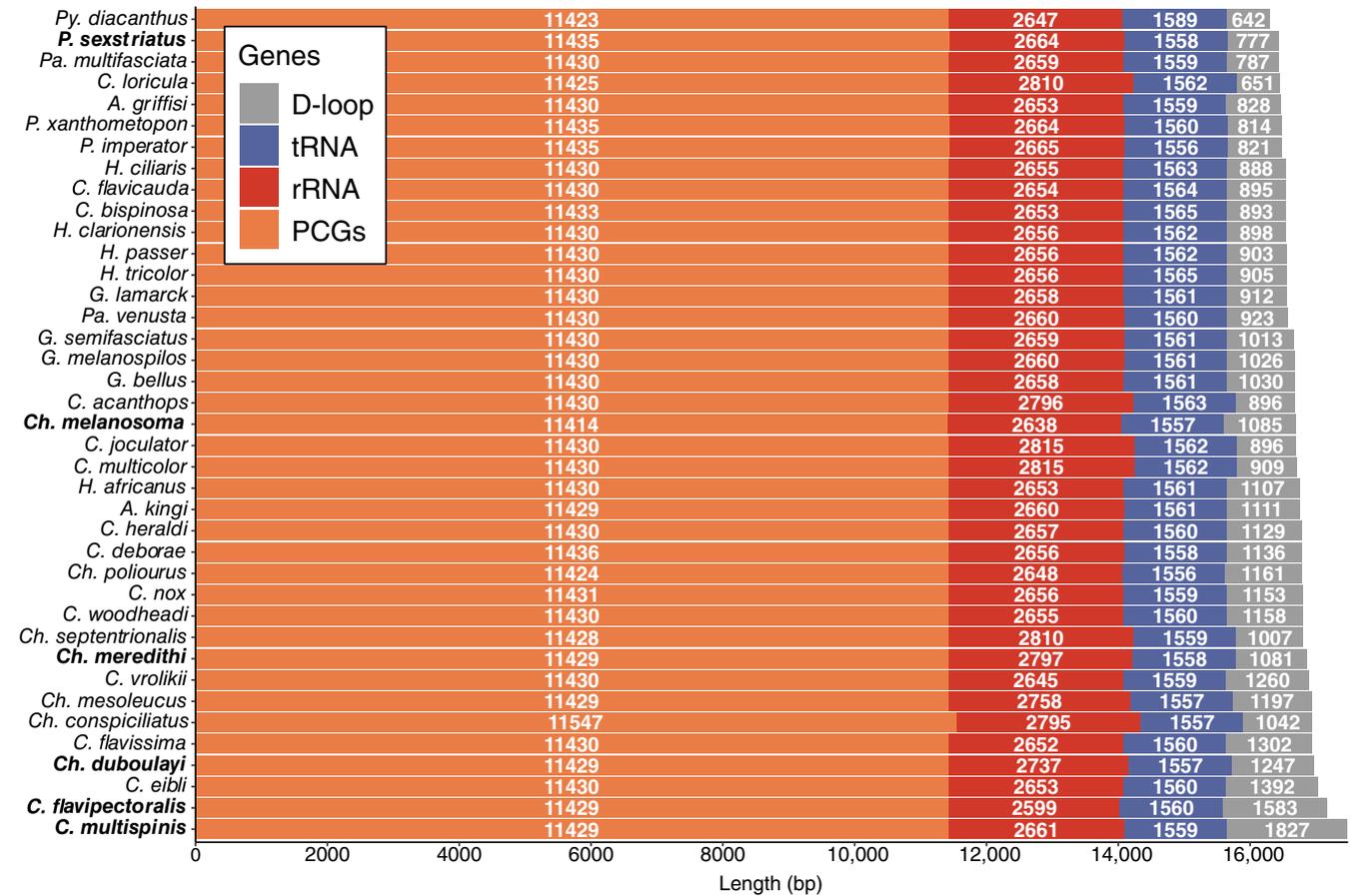


FIGURE 4 Summary of length of protein-coding genes, tRNAs, rRNAs and D-loop genes for six pomacanthid mitogenomes extracted from UCE data and 33 pomacanthid mitochondrial reference sequences obtained from NCBI GenBank. Abbreviations for genera are as follows: A, *Apolemichthys*; C, *Centropyge*; Ch, *Chaetodontoplus*; G, *Genicanthus*; H, *Holacanthus*; Pa, *Paracentropyge*; Po, *Pomacanthus* and Py, *Pygoplites*. Species names in bold are new assembled and annotated pomacanthid mitogenomes presented in this study. Accession numbers for mitogenomes obtained from NCBI GenBank are reported in Table S1.

TABLE 1 Summary of nucleotide composition of whole mitochondrial genomes for six pomacanthid species.

Species name	Size (bp)	GC (%)	AT (%)	Nucleotide composition				AT skew	GC skew
				A	T	G	C		
<i>Centropyge flavipectoralis</i>	17,251	45.4	54.3	27.1	27.5	17.4	28.0	-0.007326	-0.241685
<i>Centropyge multispinis</i>	17,544	46.1	53.9	27.1	26.8	17.3	28.7	0.00556586	-0.2478261
<i>Chaetodontoplus duboulayi</i>	17,021	46.3	53.7	27.6	26.1	15.9	30.4	0.02793296	-0.3131749
<i>Chaetodontoplus melanosoma</i>	16,771	45.1	55.1	28.2	26.8	16.1	28.9	0.02545454	-0.28444444
<i>Chaetodontoplus meredithi</i>	16,917	46.3	53.7	27.8	25.9	16.3	30.0	0.03538175	-0.2958963
<i>Pomacanthus sexstriatus</i>	16,496	44.6	55.4	28.7	26.8	15.9	28.7	0.03423423	-0.2869955

Note: Additional summary on gene features in Table S2. Strand asymmetry was calculated with the following formulas: AT skew = $[(A - T)/(A + T)]$ and GC skew = $[(G - C)/(G + C)]$ (Perna & Kocher, 1995).

genes (Gln, Ala, Asn, Cys, Tyr, Ser, Glu and Pro) were encoded on the light strand (L-strand) and the remaining 28 genes were encoded on the heavy strand (H-strand). The control region was positioned between tRNA-Pro and tRNA-Phe, and had the highest length variation of all mitochondrial genes, ranging from 777 to 1827 bp (Figure 4).

Intergenic spacers (IGS) were found between ND1 and tRNA-Ile (3–4 bp), ND2 and tRNA-Trp (8–44 bp), tRNA-Ala and tRNA-Asn

(1–2 bp), tRNA-Asn and tRNA-Cys (34–42 bp), tRNA-Tyr and COX1 (1 bp), COX1 and tRNA-Asp (6–8 bp), tRNA-Lys and ATP8 (1–2 bp), tRNA-Ser and tRNA-Leu (3–5 bp), tRNA-Glu and CytB (3–4 bp), D-loop and tRNA-Phe (1 bp). Lengths were generally consistent across species, with IGS between ND2 and tRNA-Trp genes located on the H-strand showing the most length variation (e.g. *Centropyge* species, Table S2). The largest IGS measured 42 bp and marked the separation between tRNA-Asn and tRNA-Cys (Table S2). We found

overlap (Table S2) between the adjacent PCGs ATP8-ATP6 (10 bp), ND4L-ND4 (7 bp), ND5-ND6 (4 bp) and ATP6-COX3 (1 bp, Table S2).

3.2 | Protein-coding genes

Combined length for the 13PCGs ranged from 11,414 to 11,430 bp across the six pomacanthid mitogenomes, representing around 68% of the mitogenome's total length (Figure 4). Twelve PCGs were encoded on the H-strand while only ND6 was encoded on the L-strand in reversed sequence (Figure 3). Two initiation codons for the amino acid methionine were identified (ATG/GTG) in the H-strand PCGs (Table S2), with ATG being the most common across all six mitogenomes. ND6 had different initiation codons that coded for the amino acid Leucine (TTA/CTA). Incomplete termination codons (T*/TA*) were detected for PCGs COX1, ND3, ND4, CytB, COX3 and ATP6 (Table S2). Pomacanthid mitogenomes included between 3738 and 3800 codons (excluding stop codons). Across all amino acids,

Leucine (669 ± 7.0) and Alanine (337 ± 3.5) had the highest frequency (Figure 5) with codons CTA/CTC/CTT and GCA/GCC found in highest proportions, respectively. Cysteine was the least predominant amino acid (26 ± 2.2). Overall, codon usage and amino acids were consistent across all pomacanthid species (Figure 5, Figure S2, Table S3).

There was no heterogeneity in sequence divergence across PCGs of the 39 pomacanthid species with all sites displaying positive mean similarity scores (Figure 6). The paraphyletic *Centropyge* genus had the highest degree of sequence heterogeneity across species with an averaged MSS of $0.81 (\pm 0.4)$. Heterogeneity was higher in the nucleotide dataset compared to amino-acid dataset (Figure S3), reflecting variation in codon usage but not in amino acid production. Ratios ω of non-synonymous (dN) to synonymous (dS) substitutions in the pomacanthid genomes revealed that all 13 PCGs evolved under purifying selection ($\omega < 1$, Table 2) and had a high degree of similarity ($\omega < 0.30$, Table 2). ND6 genes ($\omega = 0.27 \pm 0.004$) and ATP8 ($\omega = 0.18 \pm 0.005$) had the highest mutation rates while

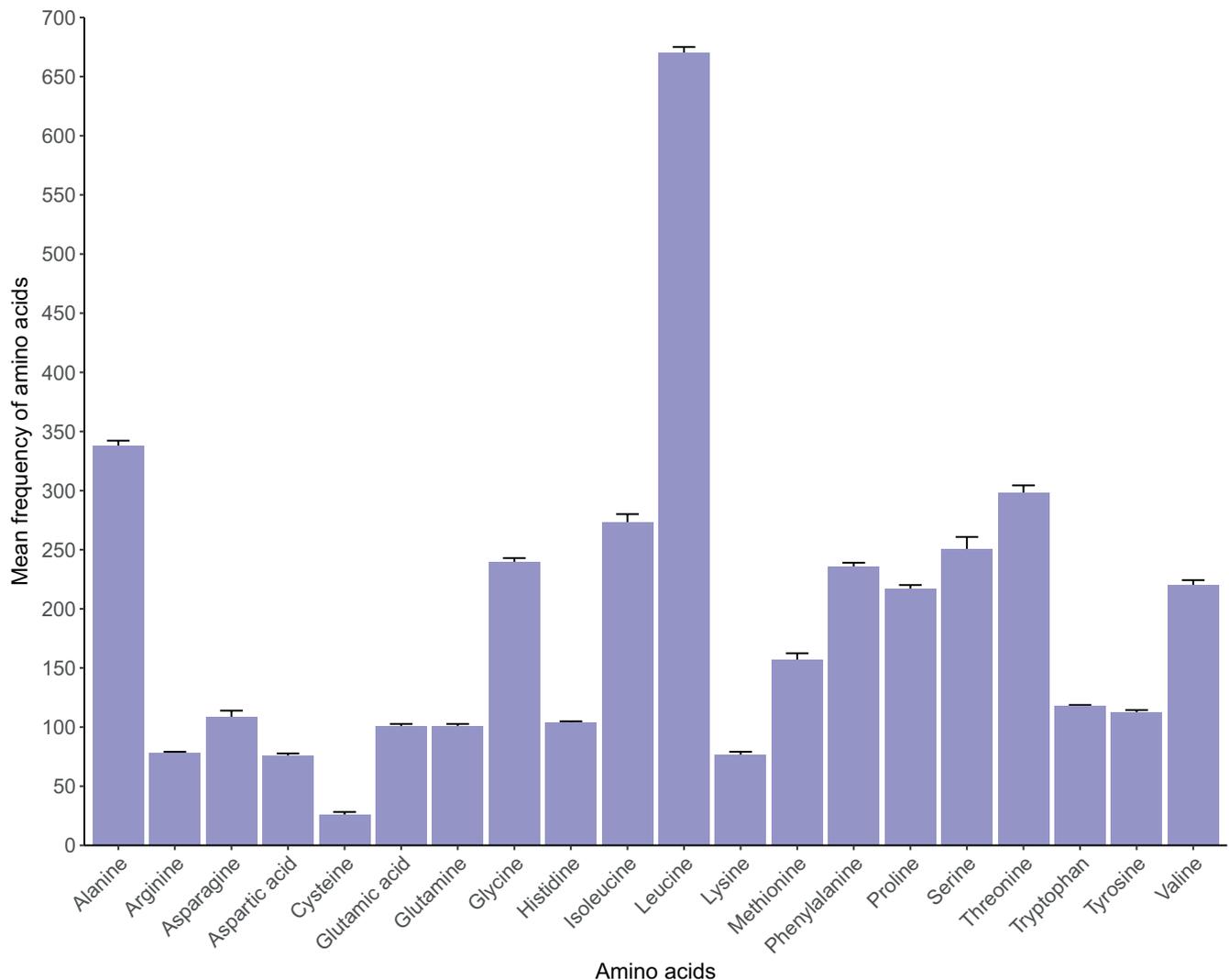


FIGURE 5 Mean frequency of amino acids in six pomacanthid mitogenomes. Error bars represent standard deviation. Details on codon usage can be found in the Table S3.

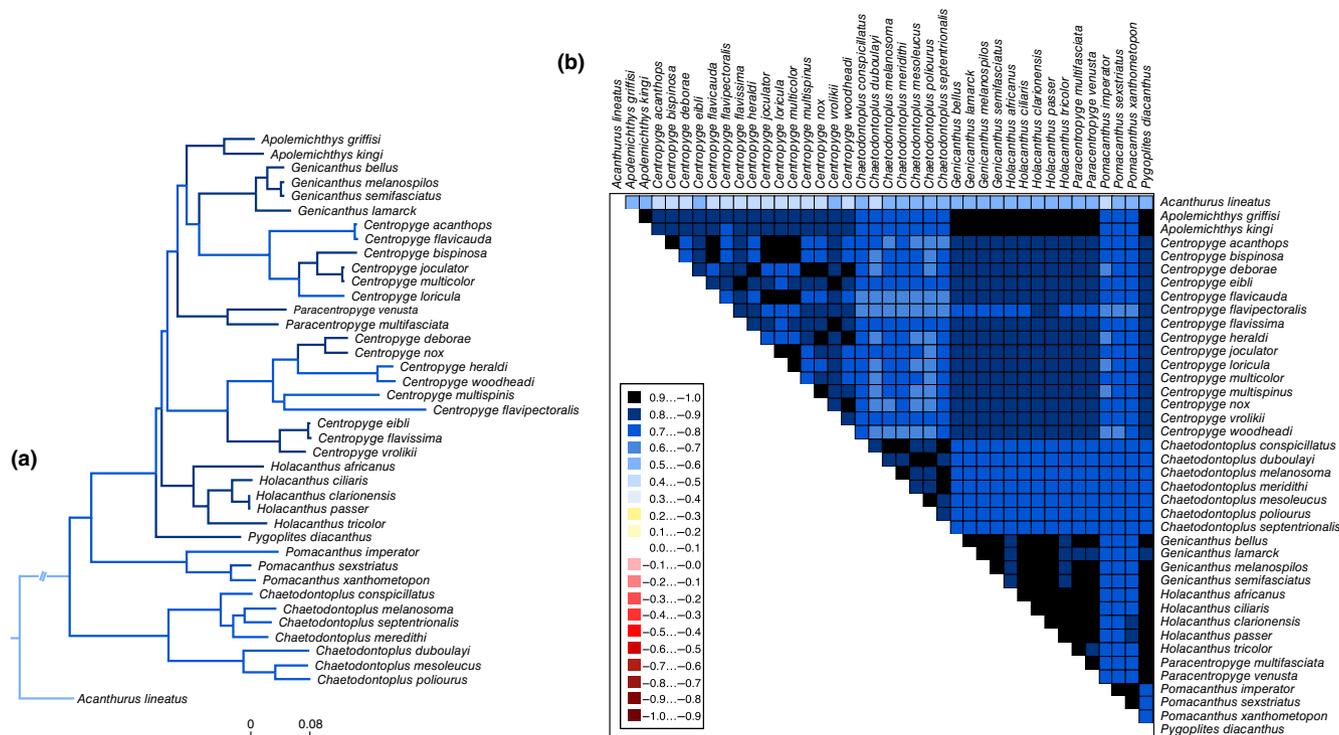


FIGURE 6 Sequence similarity score (a) guide tree and (b) matrix based on nucleotide composition of 13 PCGs for 39 complete mitochondrial genomes. The obtained mean similarity score (MSS) between sequences is represented by a coloured square. The site scores are ranging from -1 , indicating low mean similarity score in sequence composition (red colouring) to $+1$, indicating higher mean similarity score with other sequence composition (blue colouring). Coloured branches represent branch reliability and are based on the mean score between corresponding terminal taxon and all other taxa. Internal scores are calculated by the mean of the total pairwise similarity scores between taxa which are connected by a branch.

TABLE 2 Summary of average non-synonymous (dN) and synonymous (dS) substitutions values of all pairwise comparisons, and dN/dS ratios ω between alignments of 13 protein-coding genes for 39 pomacanthid species.

PCGs	Length (bp)	dN	dS	dN/dS (\pm SD)
ATP6	681	0.06394	1.20781	0.053 (\pm 0.029)
ATP8	168	0.17800	0.98076	0.181 (\pm 0.15)
CYTB	1134	0.04435	1.25660	0.035 (\pm 0.022)
COX1	1539	0.01178	1.06791	0.011 (\pm 0.007)
COX2	690	0.01788	0.93425	0.019 (\pm 0.013)
COX3	783	0.02171	0.90028	0.024 (\pm 0.0005)
ND1	975	0.04155	1.22241	0.034 (\pm 0.17)
ND2	1044	0.08364	1.11179	0.075 (\pm 0.038)
ND3	348	0.04419	1.18896	0.037 (\pm 0.021)
ND4	1380	0.04696	1.33760	0.041 (\pm 0.021)
ND4L	297	0.03586	1.16750	0.031 (\pm 0.026)
ND5	1836	0.06301	1.14503	0.055 (\pm 0.030)
ND6	521	0.17242	0.62733	0.275 (\pm 0.12)

the COX1 gene had the lowest ($\omega = 0.01 \pm 0.0002$) across all mitogenomes (Table 2). Comparably to observed patterns of dN/dS, ATP8 and ND6 genes had the largest K2P distance and highest nucleotide diversity among the 39 pomacanthid mitogenomes followed by

ND2 and ATP6 (Table 2, Figure 7). Interspecies K2P distances ranged from 0.153% (SD \pm 0.037) for the COX2 gene to 0.291% (SD \pm 0.141) for the ATP8 gene, which also showed the most variation in between species (Figure 7a). Nucleotide diversity ranged from 0.132 in the COX2 gene to 0.227 in the ATP8 gene but was consistent across species (Figure 7b).

3.3 | Transfer and ribosomal RNA genes, and control region

Across all mitogenomes, the 22tRNAs had lengths ranging from 65 to 77 bp (Table S2), adding up to total lengths varying between 1557 and 1560bp for *Ch. duboulayi* and *melanosoma*, and *C. flavipectoralis* respectively. All tRNAs were encoded on the H-strand. Small and large rRNA genes were also located on the H-strand, ranging from 949 to 958bp and 1650 to 1844bp respectively. They bordered the tRNA gene for Valine, as seen in other fish mitogenomes (Satoh et al., 2016). AT content for small and large rRNA genes varied between 50.7% and 55.2% [50.8/54.6 (*C. flavipectoralis*), 50.7/53.9 (*C. multispinis*), 51.9/53.8 (*Ch. duboulayi*), 51.5/55.0 (*Ch. melanosoma*), 51.3/54.8 (*Ch. meredithi*) and 51.2/55.2 (*P. sexstriatus*)]. The Control region or d-loop region had the most variation in length across all mitochondrial genes, ranging from 447 to 1827bp (Figure 2) and was located between tRNA-Thr and tRNA-Pro genes (Figure 3).

3.4 | Phylogenetic analyses

Final concatenated alignment comprised of 14,045bp. The topology obtained for the 39 pomacanthid delineated nine strongly supported clades (Figure 8) with the exception of *Centropyge* clade 1 (C1, SH-aLRT=93.5, UFBOOT=53), *Holacanthus* (SH-aLRT=53.3, UFBOOT=39) and *Pomacanthus*. *Chaetodontoplus* was the first clade to diverge and received strong support (SH-aLRT=100, UFBOOT=100). It was followed by the *Pomacanthus* lineage which had moderate support (SH-aLRT=80.4, UFBOOT=84) and monotypic *Pygoplites* (SH-aLRT=100, UFBOOT=100). The subsequent clade to diverge led to the *Holacanthus* genus, which had low support (SH-aLRT=53.3, UFBOOT=39). Comparably, the position of *Centropyge* C1 on the backbone of the pomacanthid tree was not well supported (SH-aLRT=93.5, UFBOOT=53). However, the separation between the two commonly identified sub-clades, delimiting species belonging to the *flavissima* 'complex' (C1.1) from the rest of the *Centropyge* species found within C1 (C1.2), was strongly supported (SH-aLRT=100, UFBOOT=100). Following the branching out of C1, *Paracentropyge*, *Apolemichthys*, *Genicanthus* and *Centropyge* C2 clades, were strongly supported (SH-aLRT=100, UFBOOT=100, Figure 8).

4 | DISCUSSION

Among the 13 complete pomacanthid mitogenomes obtained from off-target regions in UCE sequencing data, six were not on the NCBI Reference Sequence Database, bringing the taxonomic coverage for the family to 43%. All assembled mitogenomes had similar composition and arrangement as commonly found in other vertebrate and fish mitogenomes (Pereira, 2000; Satoh et al., 2016) and measured

around 16.8kbp on average. Mitochondrial genes followed a distribution between L and H-strands as previously observed in fish mitogenomes (Ruan et al., 2020) whereby ND6 and eight tRNA genes were encoded on the former while the remaining 28 genes were encoded on the H-strand. The control region was located between tRNA-Thr and tRNA-Pro genes as seen in other fishes and had higher AT content, from 57.5% to 69%, than the rest of the mitochondrial genes. GC and AT content were similar across mitogenomes, approximately 45.6% and 54.4% respectively (Table 1). GC skew was above -0.23 for all mitogenomes (Table 1), indicating an increase in cytosine (C) content over guanine (G) within the nucleotide composition. Lengths of IGS and overlap in adjacent gene sequences were consistent with those observed for other reef fish families (Gao et al., 2023; Tang et al., 2023; Wang et al., 2022). For all six pomacanthid mitogenomes, it contained a conserved sequence motif (5'-GCCGG-3') which was previously identified across some vertebrate species (Satoh et al., 2016), suggesting that this IGS can likely be linked to signals of initiation for L-strand replication (O_L ; Broughton et al., 2001; Clayton, 1991; Satoh et al., 2016). The origin of H-strand replication (O_H) is expected to be within the control region or D-loop gene (Clayton, 1991). As for many vertebrates, including fish, some incomplete termination codons might be completed by post-transcriptional polyadenylation occurring during mRNA maturation (Gissi et al., 2008; Li et al., 2021; Nagaike et al., 2008; Ruan et al., 2020).

4.1 | Comparative analysis

The 33 mitogenomes obtained from NCBI Reference Sequence database and the six newly assembled mitogenomes presented in this study were highly conserved in terms of their genome size,

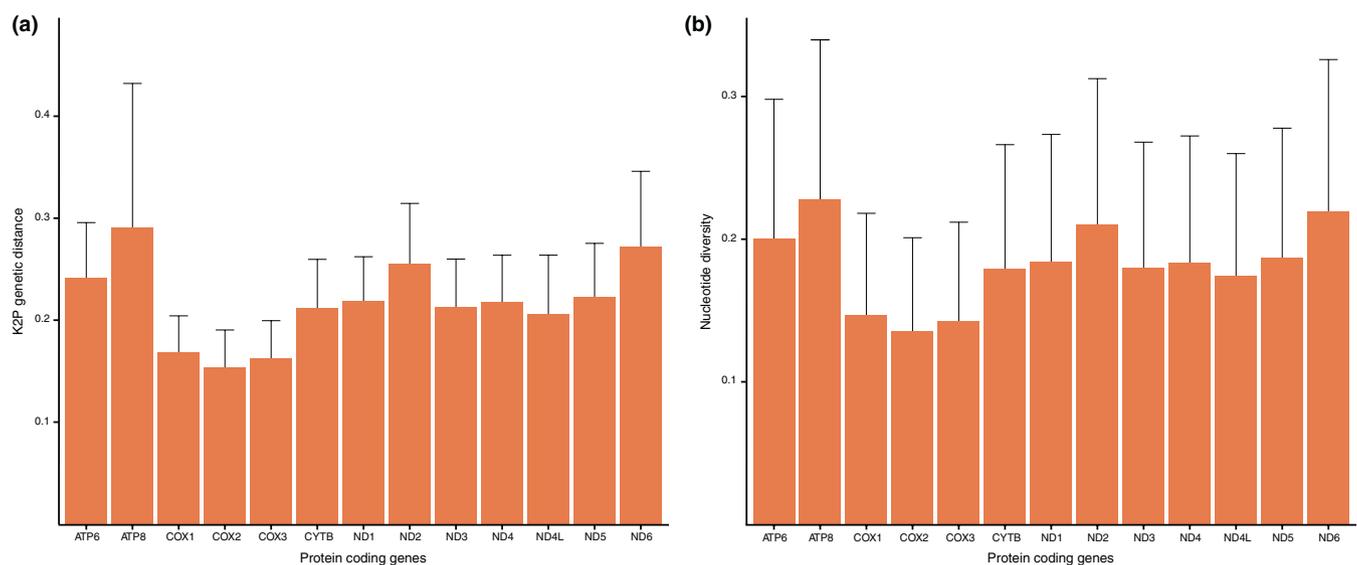


FIGURE 7 (a) Average nucleotide pairwise distances estimated in MEGA v.11 using Kimura 2-parameter model and (b) Average nucleotide diversity π calculated in R using the package *pegas*, for 13 protein-coding genes from 39 pomacanthid mitogenomes. Error bars represent standard deviations.

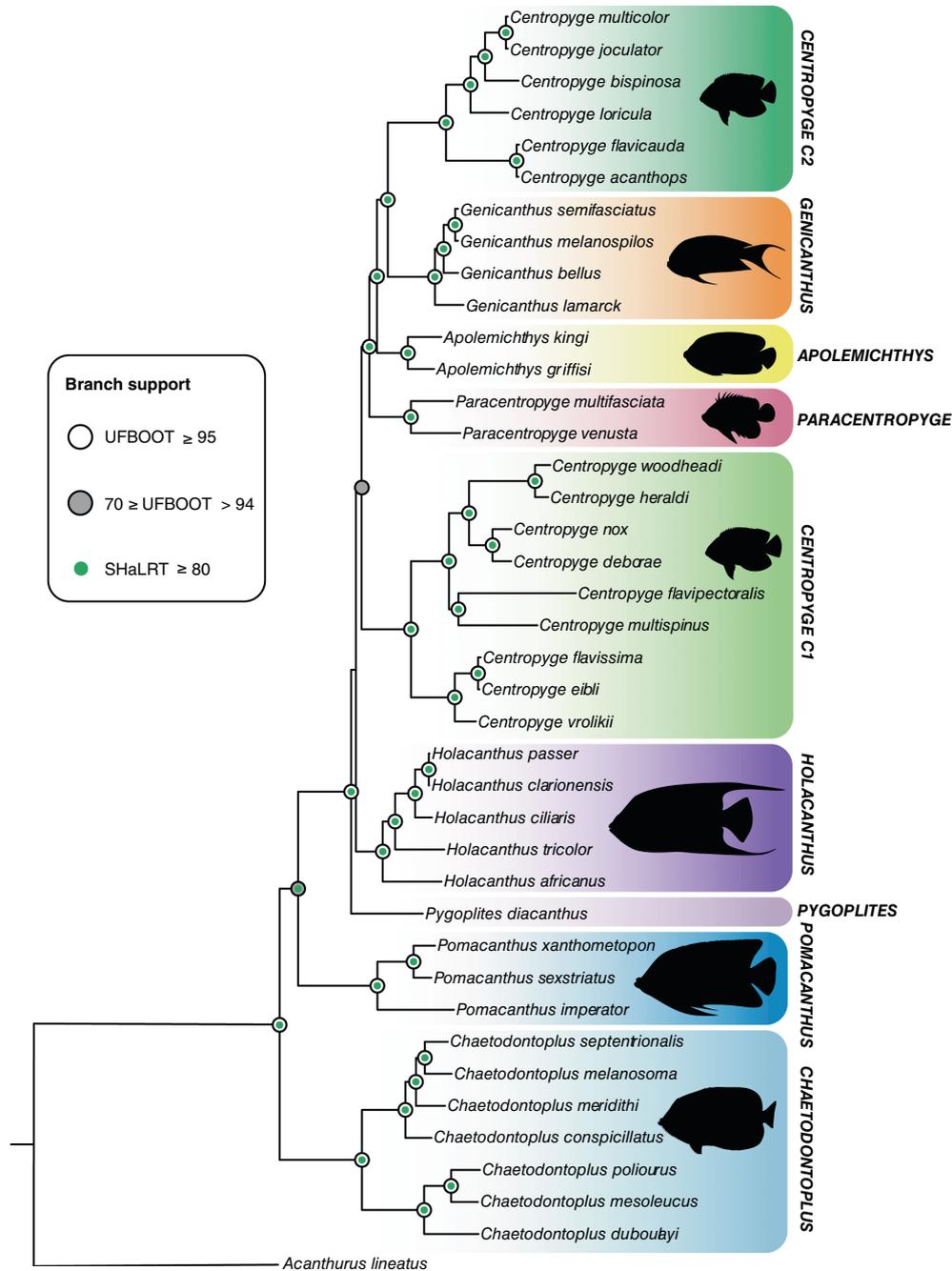


FIGURE 8 Inferred phylogeny of the Pomacanthidae generated from maximum likelihood (ML) analyses of 13 PCGs and two RNAs extracted from 39 mitochondrial genomes. Presented values correspond to SH-like approximate likelihood ratio test (SH-aLRT) and Ultrafast bootstrap (UFBoot2) analyses. Branch support values were selected following the IQ-TREE Manual (Nguyen et al., 2015).

gene content and sequence divergence (Figures 4–6). As observed in other fish mitogenomes (Sato et al., 2016), the control region showed the highest length variation of all mitochondrial genes, ranging from 777 bp for *P. sexstriatus* to 1827 bp for *C. multispinus* (Figure 4). All 13 PCGs evolved under negative or purifying selection (Table 2), which eliminates arising deleterious mutations, a pattern observed in other reef fish mitogenomes (Gao et al., 2023; Tang et al., 2023). This highlighted high level of selection pressure and gene expression to preserve the genetic integrity of

pomacanthid mitogenomes. Mitochondrial genes ND6 and ATP8 had the highest ratio ω of non-synonymous (dN) to synonymous (dS) substitutions (Table 2). This indicated that those genes were the most rapidly evolving PCGs in pomacanthid mitochondrial genomes. A relaxation of purifying selection constraints for those genes is a trend that has been observed in mitogenomes of sedentary fishes, which do not have as high energy demands as seen in more active fishes (Strohm et al., 2015). For instance, ND6 gene is involved in the production of the enzymatic protein NADH

dehydrogenase 6, which plays an essential role in the initiation phase of the electron transport chain process. ATP8 protein participate in the functioning of the mitochondrial proton pump. Both genes are involved in the process of oxidative phosphorylation, which is the primary source of ATP for metabolic processes providing an important energy source for biosynthesis.

4.2 | Phylogenetic analysis

The present analysis of 13 PCGs from 39 pomacanthid mitogenomes provides the most comprehensive assessment for the family Pomacanthidae using mitochondrial data to date. It was highly congruent with phylogenies previously inferred for the family based on limited and specific mtDNA and nuclear markers (Baraf et al., 2019; Frédérick et al., 2017) as well as the UCE-based phylogeny from which these off read mitochondrial resources originate from (Baraf et al., 2024), though these results might vary with future increases in taxon sampling. Our latest phylogeny strongly supported the position of *Chaetodontoplus* at the root of the tree (Figure 8), however this clade has sometimes been recovered as sister to *Pomacanthus* (Baraf et al., 2019; Frédérick et al., 2017), which only received moderate support from phylogenetic analyses. The present branching arrangement, whereby *Chaetodontoplus* is first to diverged followed by *Pomacanthus*, agrees with UCE-based phylogeny for the family (Baraf et al., 2024). Phylogenetic discordance for the divergence of these two pomacanthid lineages might result from ancient introgression involving extinct or outgroup lineages at the base of the pomacanthid tree (Baraf et al., in preparation). The most recent common ancestor of *Centropyge* C1 and *Holacanthus* received low support from phylogenetic analyses (Figure 8). The relationship between *Pygoplites* and *Holacanthus* genera relative to each other and the backbone of the pomacanthid tree has been challenging to resolve. Some phylogenetic reconstructions have one or the other lineage diverging first or forming a sister group relationship (Baraf et al., 2019; Frédérick et al., 2017), however none of these topological arrangements received strong support suggesting that underlying biological or evolutionary factors might be obscuring the recovery of a clear phylogenetic signal. Whole mitochondrial genomes yielded similar results whereby *Pygoplites* initially diverged and received strong support whereas the subsequent origins of the *Holacanthus* lineage was poorly supported. Strong incomplete lineage sorting for *Pygoplites* and *Centropyge* C1 has been identified as an important factor impacting the recovery of inter-generic relationships in phylogenetic reconstructions (Baraf et al., 2024). The closeness between these nodes and the *Holacanthus* branch (1 or 2 partitions) might in part explain for the low support of the clade in several phylogenies for the family. Note that the mitogenome available on the NCBI Reference Sequence Database for *Centropyge interrupta* (NC_026451.1) appeared to be misidentified as the species was recovered as part of the *Genicanthus* genus in phylogenetic analyses. A BLAST search of the COI sequence from

this mitogenome reference confirmed this misidentification and consequently, it was removed from our reference database.

5 | CONCLUSIONS

Pomacanthid mitogenome measured 16.8kbp on average and was highly conserved across species. All six mitogenomes assembled and presented in this study comprised of 37 genes, including 13 PCGs, 22 tRNA genes, two rRNA genes and one control region. It followed a genetic structure that has been observed in other vertebrates and fish mitogenomes. AT-skews were mostly all positives while GC-skews were all negatives. The most frequent start codon and amino acid were ATG and Leucine respectively. We found no heterogeneity in sequence divergence across pomacanthid species. While all PCGs were under purifying selection, ND6 and ATP8 genes showed some degree of relaxed purification and higher mutation rates, which might be reflecting the sedentary lifestyle of pomacanthid species. Moreover, the ATP8 gene showed the most genetic variation across the family. Downstream examination of mitogenomes will provide new insights into interspecific genetic variation, gene function and diversification processes of marine angelfish.

AUTHOR CONTRIBUTIONS

Lauriane M. Baraf: Conceptualization (lead); data curation (lead); formal analysis (lead); funding acquisition (supporting); investigation (lead); methodology (lead); visualization (lead); writing – original draft (lead). **Julia Y. Hung:** Resources (lead); writing – review and editing (supporting). **Morgan S. Pratchett:** Supervision (supporting); validation (equal); writing – review and editing (supporting). **Peter F. Cowman:** Conceptualization (supporting); funding acquisition (lead); supervision (lead); validation (lead); writing – review and editing (lead).

ACKNOWLEDGEMENTS

We thank Prof. David Bellwood for providing tissue samples for pomacanthid species and Dr. Rémi Allio for valuable conversations on the use of MitoFinder. We are grateful to François Libert, Yi-Kai Tea and Jake Adams for providing pictures of pomacanthid species. Finally, we thank the anonymous reviewer for insightful comments and suggestions on earlier versions of the manuscript. This work was funded by the ARC DECRA Fellowships [DE170100516] and ARC Centre of Excellence Programme [CE140100020] to P.F.C. and ARC Centre of Excellence for Coral Reef Studies – David Yellowlees Excellence in Research Award to L.M.B. Open access publishing facilitated by James Cook University, as part of the Wiley - James Cook University agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Complete mitochondrial genomes assembled and annotated in the present study are available on the NCBI GenBank Database. Accession

numbers are as follow: *Chaetodontoplus meredithi* [PP316124], *Chaetodontoplus melanosoma* [PP316125], *Chaetodontoplus duboulayi* [PP316126], *Centropyge flavivectoralis* [PP316127], *Centropyge multispinis* [PP316128] and *Pomacanthus sexstriatus* [PP316129]. Alignments and tree files are available on FigShare at <https://figshare.com/s/675722dd87d5be753ab5> and bioinformatics scripts on GitHub <https://github.com/Lavarchus/Mitogenomes.git>.

ORCID

Lauriane M. Baraf  <https://orcid.org/0009-0008-6063-6058>

Peter F. Cowman  <https://orcid.org/0000-0001-5977-5327>

REFERENCES

- Alfaro, M. E., Faircloth, B. C., Harrington, R. C., Sorenson, L., Friedman, M., Thacker, C. E., Oliveros, C. H., Černý, D., & Near, T. J. (2018). Explosive diversification of marine fishes at the Cretaceous–Palaeogene boundary. *Nature Ecology & Evolution*, 2, 688–696. <https://doi.org/10.1038/s41559-018-0494-6>
- Allio, R., Schomaker-Bastos, A., Romiguier, J., Prosdoci, F., Nabholz, B., & Delsuc, F. (2020). MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Molecular Ecology Resources*, 20, 892–905. <https://doi.org/10.1111/1755-0998.13160>
- Alva-Campbell, Y., Floeter, S. R., Robertson, D. R., Bellwood, D. R., & Bernardi, G. (2010). Molecular phylogenetics and evolution of *Holacanthus angelfishes* (Pomacanthidae). *Molecular Phylogenetics and Evolution*, 56, 456–461. <https://doi.org/10.1016/j.ympev.2010.02.014>
- Baldauf, S. L., Roger, A. J., Wenk-Siefert, I., & Doolittle, W. F. (2000). A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science*, 290, 972–977. <https://doi.org/10.1126/science.290.5493.972>
- Baraf, L. M., Pratchett, M. S., & Cowman, P. F. (2019). Ancestral biogeography and ecology of marine angelfishes (F: Pomacanthidae). *Molecular Phylogenetics and Evolution*, 140, 106596. <https://doi.org/10.1016/j.ympev.2019.106596>
- Baraf, L. M., Hung, J. Y., & Cowman, P. F. (2024). Phylogenomics of marine angelfishes: diagnosing sources of systematic discordance for an iconic reef fish family (F: Pomacanthidae). Manuscript submitted for publication.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.
- Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S., Simakov, O., Ng, A. Y., Lim, Z. W., Bezault, E., Turner-Maier, J., Johnson, J., Alcazar, R., Noh, H. J., Russell, P., Aken, B., Alföldi, J., Amemiya, C., Azzouzi, N., ... Di Palma, F. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature*, 513, 375–381. <https://doi.org/10.1038/nature13726>
- Broughton, R. E., Milam, J. E., & Roe, B. A. (2001). The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA. *Genome Research*, 11, 1958–1967. <https://doi.org/10.1101/gr.156801>
- Clayton, D. A. (1991). Replication and transcription of vertebrate mitochondrial DNA. *Annual Review of Cell Biology*, 7, 453–478. <https://doi.org/10.1146/annurev.cb.07.110191.002321>
- DeSalle, R., & Goldstein, P. (2019). Review and interpretation of trends in DNA barcoding. *Frontiers in Ecology and Evolution*, 7, 302. <https://doi.org/10.3389/FEVO.2019.00302>
- DiBattista, J. D., Waldrop, E., Bowen, B. W., & Rocha, L. A. (2012). Twisted sister species of pygmy angelfishes: Discordance between taxonomy, coloration, and phylogenetics. *Coral Reefs*, 31, 839–851. <https://doi.org/10.1007/s00338-012-0907-y>
- do Amaral, F. R., Neves, L. G., Resende, M. F. R., Jr., Mobili, F., Miyaki, C. Y., Pellegrino, K. C. M., & Biondo, C. (2015). Ultraconserved elements sequencing as a low-cost source of complete mitochondrial genomes and microsatellite markers in non-model amniotes. *PLoS One*, 10, e0138446. <https://doi.org/10.1371/JOURNAL.PONE.0138446>
- Dornburg, A., & Near, T. J. (2021). The emerging phylogenetic perspective on the evolution of actinopterygian fishes. *Annual Review of Ecology, Evolution, and Systematics*, 52, 427–452. <https://doi.org/10.1146/annurev-ecolsys-122120>
- Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C., Heled, J., Ross, H. A., Tooman, L., Grosser, S., Park, D., Demetras, N. J., Stevens, M. I., Russell, J. C., Anderson, S. H., Carter, A., & Nelson, N. (2015). Evaluating a multigenic environmental DNA approach for biodiversity assessment. *GigaScience*, 4, s13742-015. <https://doi.org/10.1186/s13742-015-0086-1>
- Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, 61, 717–726. <https://doi.org/10.1093/sysbio/sys004>
- Frédérich, B., Santini, F., Konow, N., Schnitzler, J., Lecchini, D., & Alfaro, M. E. (2017). Body shape convergence driven by small size optimum in marine angelfishes. *Biology Letters*, 13, 20170154. <https://doi.org/10.1098/rsbl.2017.0154>
- Fricke, R., Eschmeyer, W., & van der Laan, R. (2023). *Eschmeyer's catalog of fishes: Genera, species, references*. California Academy of Sciences.
- Friedman, S. T., Collyer, M. L., Price, S. A., & Wainwright, P. C. (2022). Divergent processes drive parallel evolution in marine and freshwater fishes. *Systematic Biology*, 71, 1319–1330. <https://doi.org/10.1093/sysbio/syab080>
- Gaither, M. R., Schultz, J. K., Bellwood, D. R., Pyle, R. L., DiBattista, J. D., Rocha, L. A., & Bowen, B. W. (2014). Evolution of pygmy angelfishes: Recent divergences, introgression, and the usefulness of color in taxonomy. *Molecular Phylogenetics and Evolution*, 74, 38–47. <https://doi.org/10.1016/j.ympev.2014.01.017>
- Gao, J., Li, C., Yu, D., Wang, T., Lin, L., Xiao, Y., Wu, P., & Liu, Y. (2023). Comparative mitogenome analyses uncover mitogenome features and phylogenetic implications of the parrotfishes (Perciformes: Scaridae). *Biology*, 12, 410. <https://doi.org/10.3390/biology12030410>
- Gissi, C., Iannelli, F., & Pesole, G. (2008). Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity*, 101, 301–320. <https://doi.org/10.1038/hdy.2008.62>
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59, 307–321.
- Hillis, D. M. (1996). Inferring complex phylogenies. *Nature*, 383, 130–131. <https://doi.org/10.1038/383130a0>
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Le, S. V. (2017). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35, 518–522.
- 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, 2531–2540. <https://doi.org/10.1093/molbev/mst141>
- Juhling, F., Putz, J., Bernt, M., Donath, A., Middendorf, M., Florentz, C., & Stadler, P. F. (2012). Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. *Nucleic Acids Research*, 40, 2833–2845.

- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*, 587–589.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780.
- Katoh, K., Misawa, K., Kuma, K.-i., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*, 3059–3066.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*, 111–120. <https://doi.org/10.1007/BF01731581>
- Kück, P., Meid, S. A., Groß, C., Wägele, J. W., & Misof, B. (2014). AliGROOVE – Visualization of heterogeneous sequence divergence within multiple sequence alignments and detection of inflated branch support. *BMC Bioinformatics*, *15*, 294. <https://doi.org/10.1186/1471-2105-15-294>
- Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, *34*, 772–773. <https://doi.org/10.1093/MOLBEV/MSW260>
- Li, B., Wang, H., Yang, L., Liu, S., & Zhuang, Z. (2021). Complete mitochondrial genome of *Pseudocaranx dentex* (Carangidae, Perciformes) provides insight into phylogenetic and evolutionary relationship among Carangidae Family. *Genes*, *12*, 1234. <https://doi.org/10.3390/genes12081234>
- Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*, 1451–1452. <https://doi.org/10.1093/bioinformatics/btp187>
- Mercer, T. R., Neph, S., Dinger, M. E., Crawford, J., Smith, M. A., Shearwood, A.-M. J., Haugen, E., Bracken, C. P., Rackham, O., Stamatoyannopoulos, J. A., Filipovska, A., & Mattick, J. S. (2011). The human mitochondrial transcriptome. *Cell*, *146*, 645–658. <https://doi.org/10.1016/j.cell.2011.06.051>
- Meyer, C. P., & Paulay, G. (2005). DNA barcoding: Error rates based on comprehensive sampling. *PLoS Biology*, *3*, e422. <https://doi.org/10.1371/JOURNAL.PBIO.0030422>
- Miller, C. D., Forthman, M., Miller, C. W., & Kimball, R. T. (2022). Extracting 'legacy loci' from an invertebrate sequence capture data set. *Zoologica Scripta*, *51*, 14–31. <https://doi.org/10.1111/zsc.12513>
- Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., & Schäffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics (Oxford, England)*, *24*, 1757–1764. <https://doi.org/10.1093/bioinformatics/btn322>
- Nagaike, T., Suzuki, T., & Ueda, T. (2008). Polyadenylation in mammalian mitochondria: Insights from recent studies. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1779*, 266–269. <https://doi.org/10.1016/j.bbagr.2008.02.001>
- Near, T. J., Dornburg, A., Eytan, R. I., Keck, B. P., Smith, W. L., Kuhn, K. L., Moore, J. A., Price, S. A., Burbrink, F. T., Friedman, M., & Wainwright, P. C. (2013). Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 12738–12743. <https://doi.org/10.5061/dryad.d3mb4>
- Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQTREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution*, *32*, 268–274.
- Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). metaSPAdes: A new versatile metagenomic assembler. *Genome Research*, *27*, 824–834. <https://doi.org/10.1101/gr.213959.116>
- Pereira, S. L. (2000). Mitochondrial genome organization and vertebrate phylogenetics. *Genetics and Molecular Biology*, *23*, 745–752. <https://doi.org/10.1590/S1415-47522000000400008>
- Perna, N. T., & Kocher, T. D. (1995). Patterns of Nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *Journal of Molecular Evolution*, *41*, 353–358.
- Quattrini, A. M., Snyder, K. E., Purov-Ruderman, R., Seiblit, I. G. L., Hoang, J., Floerke, N., Ramos, N. I., Wirshing, H. H., Rodriguez, E., & McFadden, C. S. (2023). Mito-nuclear discordance within Anthozoa, with notes on unique properties of their mitochondrial genomes. *Scientific Reports*, *13*, 7443. <https://doi.org/10.1038/s41598-023-34059-1>
- 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*, 1874.
- Rubinoff, D., & Holland, B. S. (2005). Between two extremes: Mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology*, *54*, 952–961. <https://doi.org/10.1080/10635150500234674>
- Samuels, D. C., Han, L., Li, J., Quanguo, S., Clark, T. A., Shyr, Y., & Guo, Y. (2013). Finding the lost treasures in exome sequencing data. *Trends in Genetics*, *29*, 593–599. <https://doi.org/10.1016/j.tig.2013.07.006>
- Satoh, T. P., Miya, M., Mabuchi, K., & Nishida, M. (2016). Structure and variation of the mitochondrial genome of fishes. *BMC Genomics*, *17*, 719. <https://doi.org/10.1186/s12864-016-3054-y>
- Seixas, F. A., Boursot, P., & Melo-Ferreira, J. (2018). The genomic impact of historical hybridization with massive mitochondrial DNA introgression. *Genome Biology*, *19*, 91. <https://doi.org/10.1186/s13059-018-1471-8>
- Smith, B. T., Harvey, M. G., Faircloth, B. C., Glenn, T. C., & Brumfield, R. T. (2014). Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. *Systematic Biology*, *63*, 83–95. <https://doi.org/10.1093/sysbio/syt061>
- Stothard, P., & Wishart, D. S. (2005). Circular genome visualization and exploration using CGView. *Bioinformatics (Oxford, England)*, *21*, 537–539. <https://doi.org/10.1093/bioinformatics/bti054>
- Strohman, J. H. T., Gwiazdowski, R. A., & Hanner, R. (2015). Fast fish face fewer mitochondrial mutations: Patterns of dN/dS across fish mitogenomes. *Gene*, *572*, 27–34. <https://doi.org/10.1016/j.gene.2015.06.074>
- Taanman, J.-W. (1999). The mitochondrial genome: Structure, transcription, translation and replication. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1410*, 103–123. [https://doi.org/10.1016/S0005-2728\(98\)00161-3](https://doi.org/10.1016/S0005-2728(98)00161-3)
- Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, *38*, 3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Tang, Q., Liu, Y., Li, C.-H., Zhao, J.-F., & Wang, T. (2023). Comparative mitogenome analyses uncover mitogenome features and phylogenetic implications of the reef fish family Holocentridae (Holocentridae). *Biology*, *12*, 1273. <https://doi.org/10.3390/biology12101273>
- Tea, Y.-K., Hobbs, J.-P. A., Vitelli, F., DiBattista, J. D., Ho, S. Y. M., & Lo, N. (2020). Angels in disguise: Sympatric hybridization in the marine angelfishes is widespread and occurs between deeply divergent lineages. *Proceedings of the Royal Society B: Biological Sciences*, *287*, 20201459. <https://doi.org/10.1098/rspb.2020.1459>
- Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, *21*, 3907–3930. <https://doi.org/10.1111/j.1365-294X.2012.05664.x>
- Volf, J. N. (2005). Genome evolution and biodiversity in teleost fish. *Heredity*, *94*, 280–294. <https://doi.org/10.1038/sj.hdy.6800635>
- Wabnitz, C. (2003). *From ocean to aquarium: The global trade in marine ornamental species*. UNEP/Earthprint.

- Wallis, G. P., Cameron-Christie, S. R., Kennedy, H. L., Palmer, G., Sanders, T. R., & Winter, D. J. (2017). Interspecific hybridization causes long-term phylogenetic discordance between nuclear and mitochondrial genomes in freshwater fishes. *Molecular Ecology*, *26*, 3116–3127. <https://doi.org/10.1111/mec.14096>
- Wang, C., Ye, P., Liu, M., Zhang, Y., Feng, H., Liu, J., Zhou, H., Wang, J., & Chen, X. (2022). Comparative analysis of four complete mitochondrial genomes of Epinephelidae (Perciformes). *Genes*, *13*, 660. <https://doi.org/10.3390/genes13040660>
- Zarza, E., Faircloth, B. C., Tsai, W. L. E., Bryson, R. W., Jr., Klicka, J., & McCormack, J. E. (2016). Hidden histories of gene flow in highland birds revealed with genomic markers. *Molecular Ecology*, *25*, 5144–5157. <https://doi.org/10.1111/mec.13813>
- Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, *7*, 203–214. <https://doi.org/10.1089/10665270050081478>
- Zhu, T., Sato, Y., Sado, T., Miya, M., & Iwasaki, W. (2023). MitoFish, MitoAnnotator, and MiFish pipeline: Updates in 10 years. *Molecular*

Biology and Evolution, *40*, msad035. <https://doi.org/10.1093/molbev/msad035>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Baraf, L. M., Hung, J. Y., Pratchett, M. S., & Cowman, P. F. (2024). Comparative mitogenomics of marine angelfishes (F: Pomacanthidae). *Ecology and Evolution*, *14*, e70127. <https://doi.org/10.1002/ece3.70127>