Unlocking the genomic potential of historical and formalin-fixed specimens: phylogenetic insights from museum-preserved threadfin fishes (Teleostei: Polynemidae)

Matthew G. Girard[1,2] and Kevin R. Chovanec[2]

[1] Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, United States of America

[2] Biodiversity Institute, University of Kansas, Lawrence, KS, United States of America

ABSTRACT

DNA sequencing continues to revolutionize our understanding of biodiversity, ecology, and evolution. While analyzing sequence data allows us to address countless questions, most of the world's vertebrate museum specimens have been historically inaccessible for genetic sampling. This is partially due to the absence of modern genetic samples and/or the impact formalin has on DNA during the preservation of specimen vouchers. Recent studies have shown successful extraction of DNA from historic museum specimens using additional chemicals and/or exposing the sample(s) to heat, with these advances enhancing the possibility of capturing genomic information from type specimens, characterizing genetic diversity within species complexes, and incorporating rare samples into phylogenetic analyses. However, questions remain about the reliability of these data and utility of historic DNA in modern phylogenetic analyses. In this study, we use a commercial extraction kit that targets formalin-fixed, paraffin-embedded samples to successfully extract DNA from historic museum specimens of threadfin fishes (Teleostei: Polynemidae). These specimens represent rare, genetically uncharacterized taxa that have yet to be included in a phylogenetic analysis. Low-depth shotgun sequencing is then used to sequence mitochondrial loci from the historic samples. The resulting sequence data are assembled, validated, and incorporated into a newly generated mitochondrial dataset that is simultaneously analyzed with a previously published ultraconserved-element dataset to construct a phylogenetic framework. We then explore new and previously described morphological variation within this new evolutionary framework for threadfins, identifying several shared characters that warrant revision of the generic-level classification. These findings add to the growing body of literature that demonstrates sequencing historical DNA from museum specimens and analyzing these data with complementary datasets of molecular markers from modern genetic samples can provide reliable and comprehensive assessments of biodiversity, ecology, and evolution across fishes and other vertebrates.

Submitted 6 May 2025
Accepted 13 August 2025
Published 30 September 2025

Corresponding author
Matthew G. Girard, mgirard@ku.edu

Academic editor
Barbara Nowak

Additional Information and Declarations can be found on page 18

DOI **10.7717/peerj.20029**

Copyright
2025 Girard and Chovanec

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Aquaculture, Fisheries and Fish Science, Biodiversity, Evolutionary Studies, Marine Biology, Taxonomy

Keywords Combined analyses, Mitochondrial genome, *Polydactylus*, Ultraconserved elements, UCEs

How to cite this article Girard MG, Chovanec KR. 2025. Unlocking the genomic potential of historical and formalin-fixed specimens: phylogenetic insights from museum-preserved threadfin fishes (Teleostei: Polynemidae). *PeerJ* **13**:e20029 http://doi.org/10.7717/peerj.20029

INTRODUCTION

Molecular characters have provided a wealth of information about the planet's biodiversity. Since the development of standardized primers to amplify mitochondrial loci (*e.g.*, *Kocher et al., 1989*), researchers have employed DNA barcoding to assess biodiversity (*e.g.*, *Hebert et al., 2003*), metagenomics to track ecosystem change (*e.g.*, *Stein et al., 1996*), and DNA-based phylogenetics to gain insights into evolutionary history. Within vertebrates, DNA-based phylogenetics has enabled researchers to identify novel hypotheses of relationships in amphibians and reptiles (*e.g.*, *Feller & Hedges, 1998*; *Hedges & Poling, 1999*), birds (*e.g.*, *Edwards & Wilson, 1990*; *Edwards, Arctander & Wilson, 1991*), fishes (*e.g.*, *Chen, Bonillo & Lecointre, 2003*; *Miya et al., 2003*), and mammals (*e.g.*, *Madsen et al., 2001*; *Murphy et al., 2001*). As new molecular techniques emerged, the questions posed simultaneously expanded in scope and scale, leading to the development of datasets that included more loci and phylogenies that encompassed a broader range of taxa (*e.g.*, *Smith & Craig, 2007*; *Gardner et al., 2010*; *Pyron & Wiens, 2011*; *Near et al., 2012*; *Near et al., 2013*; *Betancur et al., 2017*; *Foley et al., 2023*; *Stiller et al., 2024*). These datasets and analyses continue to contribute to our understanding of the intra- and interrelationships of vertebrate groups and address questions about large-scale evolutionary patterns (*e.g.*, *Alfaro et al., 2018*; *Rabosky et al., 2018*; *Ghezelayagh et al., 2022*; *Christmas et al., 2023*). While genomic data continues to have a profound impact, these data have generally been sequenced from recently collected specimens that have associated genetic samples. Many species lack modern genomic resources, leaving numerous questions about biodiversity, ecology, and evolution unanswered. Hereafter, the usage of "modern" for genetic samples and other genomic resources applies to samples that were purposefully collected for genetic analyses (*i.e.,* sampled prior to specimen fixation and preserved in ethanol, liquid nitrogen, or similar).

Natural history museums house millions of preserved specimens representing the diversity of taxa from across the globe, including abundant, rare, described, and undescribed species. These libraries of life document our pursuit to understand the biodiversity of our planet and hold the answers to countless questions within biodiversity sciences. Early attempts to extract DNA from fluid-preserved museum specimens yielded mixed success (*e.g.*, *Chakraborty, Sakai & Iwatsuki, 2006*; *Zhang, 2010*), often due to the heavy crosslinking that occurs between strands of DNA during formalin fixation. Degradation and fragmentation of historic DNA (hDNA) can also impact the accuracy of base calls in the resulting reads, yielding unreliable results (*e.g.*, *Cooper, Smith & Westneat, 2009*, but see *Bernardi, 2011*). Recent studies have shown promising results for hDNA extraction using various methods and technological improvements (*e.g.*, *Ruane & Austin, 2017*; *Silva et al., 2019*; *Appleyard et al., 2022*; *Hawkins et al., 2022*; *Sullivan et al., 2022*; *Speer et al., 2022*), including sampling DNA-rich liver samples (*e.g.*, *Hykin, Bi & McGuire, 2015*; *Zacho et al., 2021*; *Hahn et al., 2022*), exposing samples to elevated temperatures for various amounts of time to remove crosslinking (*e.g.*, *Straube et al., 2021*), or lysing samples in alkaline environments to enhance enzymatic digestion (*e.g.*, *Hahn et al., 2024*). Some of these new techniques can be found within commercially available extraction kits targeting

formalin-fixed, paraffin-embedded (FFPE) samples (*e.g.*, exposing samples to elevated temperatures), and commercially produced single-stranded library preparation kits are increasingly available for the conversion of short and damaged DNA. These advances in extraction and library preparation are allowing researchers to sequence genomic information from type specimens, characterize genetic diversity within species complexes, and incorporate rare samples into phylogenetic analyses (*e.g.*, *Sullivan et al., 2022*; *Bernstein et al., 2023*; *Quattrini et al., 2024*; *Muschick, Rüber & Matschiner, 2025*).

Commonly called threadfins for their elongate and thread-like pectoral-fin rays, the fish family Polynemidae includes eight genera, with 21 of the 42 species placed in a single genus, *Polydactylus*. Four previous studies have targeted the intrarelationships of this family, with three analyzing morphological characters (*i.e., Feltes, 1986*; *Kang, 2017*; *Presti, Johnson & Datovo, 2023*) and one (*Girard et al., 2022a*) analyzing a combination of mitochondrial and ultraconserved element (UCE) loci. While these studies have sampled between 19 (*Presti, Johnson & Datovo, 2023*) and 30 species (*Girard et al., 2022a*), recovering different relationships among them, all have consistently recovered the genus *Polydactylus* as polyphyletic. Hereafter, species of *Polydactylus* that are recovered outside of the monophyletic group that includes the type species of the genus (*i.e., P. virginicus*) are listed as "*Polydactylus.*" *Girard et al. (2022a)* proposed names for several clades of "*Polydactylus*" (*i.e.,* black-spotted species, striped species) but did not revise the classification or the taxonomy of the genus, as multiple key taxa were missing from their analyses due to a lack of modern genetic samples. Some of these unsampled species have been allied with groups of "*Polydactylus*" in previous taxonomic works (*e.g.*, "*P.*" *persicus* with black-spotted species, *Motomura & Iwatsuki, 2001b*). Other key taxa have morphological characters that could ally them with several species groups within the family, such as the highly elongate threadlike pectoral-fin rays in "*P.*" *macrophthalmus*. Highly elongate pectoral-fin rays are found in several species of threadfins in at least three genera, including *Filimanus,* "*Polydactylus*," and all species of *Polynemus* (*Motomura, 2004*) making the placement of "*P.*" *macrophthalmus* difficult based on morphology alone. This and other unsampled species of "*Polydactylus*" are ideal targets for new hDNA techniques because they lack modern genetic resources, they have yet to be included in any phylogenetic analyses, and their confusing suites of anatomical characters obfuscate their phylogenetic placement within the family when analyzing morphology alone.

To build on the advances in hDNA extraction and library preparation, leverage the ease of commercially available FFPE extraction and single-stranded library preparation kits, and test the utility of hDNA data in modern phylogenetic analyses, we targeted two enigmatic species of "*Polydactylus*" for hDNA sampling. We successfully extracted and sequenced mitochondrial loci from the livers of two historical museum specimens representing "*P.*" *bifurcus* and "*P.*" *macrophthalmus*. As previous studies have shown, analyzing datasets of mitochondrial loci alongside complementary multigene nuclear datasets of representative taxa can produce robust phylogenetic hypotheses both within polynemids (*Girard et al., 2022a*) and other groups of animals (*e.g.*, *Talavera et al., 2021*). We apply a similar integrative approach, analyzing the hDNA mitochondrial loci alongside a novel mitochondrial dataset and the ultraconserved elements (UCE) dataset published by *Girard*

Specimen selection:
- Collaborate with collection staff and curators to understand collection concerns and select specimen(s) for sampling.
- Larger specimen(s) and/or those with previous damage or incisions are often preferred.
- Ensure any/all planned sample extraction minimizes impact on specimen long-term integrity and conservation.

Sample extraction:
- In sterile environment, using sterilized tools, extract sample from museum specimen (e.g., liver).
  - Ensure sample does not exceed maximum weight listed in manufacturer extraction protocol (e.g., 25 mg).
- Section extracted sample into 2–4 equal parts to promote digestion.
- Store sample in 95% ethanol for at least 24 hours prior to extraction.

DNA extraction and quantification:
- In ultra-clean facility with technician using gloves, face mask, and hooded coverall suit, extract DNA using formalin-fixed paraffin-embedded (FFPE) DNA extraction kit of choice and 24–hour digestion.
  - Macerate sample using a sterilized glass rod after 12 hours of sample digestion.
  - Add an additional 10 µL of proteinase K after 12 hours of sample digestion.
- Following completion of extraction protocol, quantify samples.
  - Yield may be low (e.g., as low as 0.1 ng/µL).

Library preparation and sequencing:
- In ultra-clean facility, prepare libraries using SRSLY PicoPlus ssDNA library kit.
  - Additional shearing of libraries may not be needed due to extraction protocol and/or DNA degradation.
  - Pooling libraries equally may be difficult given uneven degradation patterns and sequence lengths across libraries.
- Sequence libraries (e.g., low-depth shotgun sequencing, also known as genome skimming).
  - Adjust sequencing depth and framework based on target loci and library quantifications.

Library quality assessment and assembly:
- Clean data of adapter contamination and assess sequence quality of cleaned reads (e.g., fastp).
- Extract and assemble mitogenomes or other target loci using reference-based method.
- Align merged and overlapping reads to final assemblies (e.g., bwa aln; Li & Durbin, 2009).
- Assess sequence damage and degradation (e.g., mapdamage2; Jónsson et al., 2013).
- Annotate final contig(s) based on target loci, if necessary.

Figure 1 Generalized workflow for sequencing, assembling, and validating DNA from historic museum specimens used in this study. Additonal information and specifics relating to this study can be found in the Materials & Methods section.

Full-size DOI: 10.7717/peerj.20029/fig-1

et al. (2022a) to generate a hypothesis of relationships. Given the recovered placement of the hDNA-sampled taxa, we explore new and previously described morphological variation within threadfins, highlighting several shared characters that support the placement of these taxa with their recovered sister groups. With these newly identified synapomorphies, the taxonomy of the family is revised, including the description of a new genus.

MATERIALS & METHODS

All taxa used in this study, along with associated museum catalog information, can be found in File S1, with codes for museum collections following *Sabaj (2020)*. A visualization of sampling, extraction, library preparation, and sequencing protocols used in this study is shown in Fig. 1.

Sampling, extraction, and sequencing of DNA from historic samples

The specimen of "*P.*" *bifurcus* (USNM 76627; 125 mm standard length [SL]) was collected by F. Baker in Kao-Hsiung, Taiwan, on 3–4 December 1914. The specimen has been stored in 70–75% ethanol and lacks records for methods of fixation. The specimen of "*P.*" *macrophthalmus* (ZRC 39003; 208 mm SL) was collected by H. H. Ng and colleagues at a market in Jambi, southern Sumatra, Indonesia, in June 1995. The specimen was fixed in 4% formalin for ∼3 weeks prior to transfer to ∼70% ethanol for long-term storage.

Samples of liver were dissected from each of these specimens through an abdominal incision. Dissections occurred in sterile environments, using sterilized scalpels and forceps. Liver was chosen for extraction due to its higher DNA yield among formalin-fixed samples (*e.g.*, *Hahn et al., 2022*). Liver samples weighing between 13 ("*P.*" *bifurcus*) and 24 mg ("*P.*" *macrophthalmus*) were divided into 2–4 equal parts after dissection and stored in two mL tubes of 95% ethanol for at least 24 h prior to hDNA extraction. Extractions were performed in the Historic DNA Laboratory at the Laboratories of Analytical Biology (Smithsonian Institution) with the individual performing the extraction mitigating contamination by wearing gloves, a face mask, a hair net, and a hooded coverall suit. Samples were extracted using a Quick-DNA/RNA FFPE Miniprep Kit (Zymo Research) following manufacturer's extraction protocols for "Total Nucleic Acid Purification," with the following modifications: omitting the deparaffinization preparation step; macerating the sample using a sterilized glass rod after 12 h of the 24-hour sample-digestion step; adding 10 μL of proteinase K after 12 h of the 24-hour sample-digestion step. Extracted samples were quantified using 2 μL of the hDNA extract and the dsDNA HR Assay Kit (Thermo Fisher Scientific) with Qubit Fluorometer (Thermo Fisher Scientific). Final 98 μL samples were sent to Daicel Arbor Biosciences for library preparation using the SRSLY PicoPlus ssDNA library kit (Claret Bioscience), sequencing, and demultiplexing. Libraries were sequenced using a low-depth whole-genome-sequencing approach commonly applied in genome skimming (*e.g.*, *Hoban et al., 2022*; *Quattrini et al., 2024*). This sequencing approach was chosen as it yields reads for many loci while bypassing the complications of amplifying target regions of the genome from degraded samples. Further, mitochondrial loci have been shown to be abundant in modern DNA samples that were genome skimmed. Mitochondrial locus lengths, arrangements, and start and stop codons have been well studied in vertebrates (*e.g.*, *Miya et al., 2003*), making orthologous loci from congener taxa readily available for comparison to hDNA samples. Samples were sequenced in a 150-base-pair (bp) paired-end framework on a NovaSeq 6000 (Illumina) targeting 20 million reads per sample. Sample extraction, library quantification, and sequencing results can be found in Table 1.

Sampling, extraction, and sequencing of DNA from modern genetic samples

To test the utility of hDNA in phylogenetic analyses, we analyzed mitochondrial and nuclear loci in combination to generate a phylogeny (*e.g.*, *Martin et al., 2018*; *Talavera et al., 2021*; *Smith et al., 2022*; *Smith et al., 2024*). Taxon sampling was primarily based on the dataset published by *Girard et al. (2022a)*, as their study includes the greatest number of threadfin species to date. We used the '32 terminal' UCE dataset in *Girard et al. (2022a)* that included 20 species of polynemids (~47% family diversity) and 12 outgroup taxa. A complementary mitochondrial dataset was generated in this study that sampled all taxa represented in the '32 terminal' dataset as well as an additional 10 species of threadfins. In total, the combined datasets sample 32 of 42 species (≈76%) of threadfins and 12 outgroup taxa (File S1). Outgroup taxa include representatives from the Bedotiidae, Centrarchidae, Latidae, Mugilidae, Pleuronectidae, Psettodidae, Sciaenidae, Scombridae, Scophthalmidae,

Table 1 Sequencing statistics and GenBank information for genome-skimmed samples.

Family	Species	Type of genetic sample	Concentration of DNA after extraction (ng/μ L)	Concentration of DNA after library prep (ng/μ L)	Total clean reads sequenced	GC% clean reads sequenced	Reference mitogenome used for mapping	Reads mapped to reference mitogenome	Percentage of clean reads mapped	Annotated mitogenome length	Mean coverage depth (×)
Mugilidae	*Mugil curema*	Modern DNA	0.159	2.1	12,797,874	41	KP018403	51,464	0.40%	16,828	292.4
Polynemidae	*Filimanus perplexa*	Modern DNA	2.688	34.9	16,226,057	40	PV590078	94,009	0.58%	17,117	457.4
Polynemidae	*Filimanus xanthonema*	Modern DNA	0.305	1.6	13,948,392	41	PV590078	159,487	1.14%	17,428	855.9
Polynemidae	*Filistriatus bifurcus* gen. nov.	Historic DNA	0.355	9.5	37,034,349	46	NC_026235	22,057*	0.06%*	16,564*	111.2*
Polynemidae	*Galeoides decadactylus*	Modern DNA	20.466	40.9	24,947,149	40	NC_027088	49,144	0.20%	16,804	283.1
Polynemidae	*Leptomelanosoma macrophthalmus*	Historic DNA	0.118	28.9	86,551,031	43	PV590078	13,622*	0.01%*	16,346*	62.9*
Polynemidae	*Parapolynemus verekeri*	Modern DNA	5.881	52.9	16,256,888	40	NC_026236	20,426	0.13%	16,698	101.3
Polynemidae	*Pentanemus quinquarius*	Modern DNA	13.845	55.4	22,943,533	41	NC_057649	61,636	0.27%	16,681	315.1
Polynemidae	*Polydactylus approximans*	Modern DNA	8.181	49.1	21,701,244	41	OP056931	11,583	0.05%	16,713	51.0
Polynemidae	*"Polydactylus" macrochir*	Modern DNA	3.467	45.1	13,887,832	41	MW630081	15,028	0.11%	16,670	56.3
Polynemidae	*"Polydactylus" multiradiatus*	Modern DNA	0.83	10.8	12,758,097	43	NC_026235	19,231	0.15%	16,711	102.6
Polynemidae	*"Polydactylus" nigripinnis*	Modern DNA	6.31	50.5	22,426,340	41	PV590087	11,259	0.05%	16,683	53.7
Polynemidae	*Polydactylus octonemus*	Modern DNA	5.985	47.9	17,041,943	41	OP056931	49,929	0.29%	16,665	250.0
Polynemidae	*Polydactylus opercularis*	Modern DNA	8.353	50.1	21,888,387	40	MW630081	48,432	0.22%	16,606	257.9
Polynemidae	*"Polydactylus" quadrifilis*	Modern DNA	9.83	49.2	20,401,301	39	PV590078	55,678	0.27%	17,108	244.2
Polynemidae	*Polynemus melanochir*	Modern DNA	2.195	28.5	17,134,260	40	NC_026236	31,514	0.18%	16,834	143.7
Polynemidae	*Polynemus multifilis*	Modern DNA	2.835	36.8	20,976,546	40	NC_026236	43,712	0.21%	16,848	217.5

Notes.

Information for mitochondrial genomes extracted from SRAs not included in table (see File S1).

*indicate values after bwa aln and deduplication.

and Sphyraenidae. DNA was extracted from modern genetic samples using an AutoGenPrep 965 automated DNA extraction robot, following the manufacturer's protocols. Libraries were prepared using the NEB Ultra II FS DNA library prep kit (New England Biolabs), following the manufacturer's protocol, and iTru y-yoke adapter and dual indices (*Glenn et al., 2019*). Libraries were quantified using the fluorometer and assay kit listed above and pooled for sequencing. Extractions and library preparation were performed in the Laboratories of Analytical Biology (Smithsonian Institution). Pooled libraries were sent to Oklahoma Medical Research Foundation NGS Core for sequencing using a NovaSeq 6000 (Illumina) and a paired-end 150 bp framework targeting 13 million reads per sample. Sample extraction, library quantification, and sequencing results can be found in Table 1.

Mitogenome assembly, annotation, and quality assessment

Demultiplexed sequence data from multiple runs were received in compressed FASTQ format. These data were uncompressed into two read files per taxon and uploaded to GenBank (SRA Accession Numbers [SRR33326234–SRR33326248]; BioProject PRJNA720393; File S1). The data were cleaned of adapter contamination using fastp (*Chen et al., 2018*) and FastQC version 0.12.1 (Babraham Bioinformatics) was used to assess sequence quality (Fig. 2). Cleaned reads were collated with previously published SRA data obtained from GenBank for assembly (BioProjects PRJNA341709, PRJNA604383, PRJNA720393, PRJNA796495; File S1). A reference-based method within Geneious version 11.1.5 (*Kearse et al., 2012*) was used to assemble reads into mitochondrial contigs. Multiple reference sequences were tested for assembly, with the putative closest threadfin

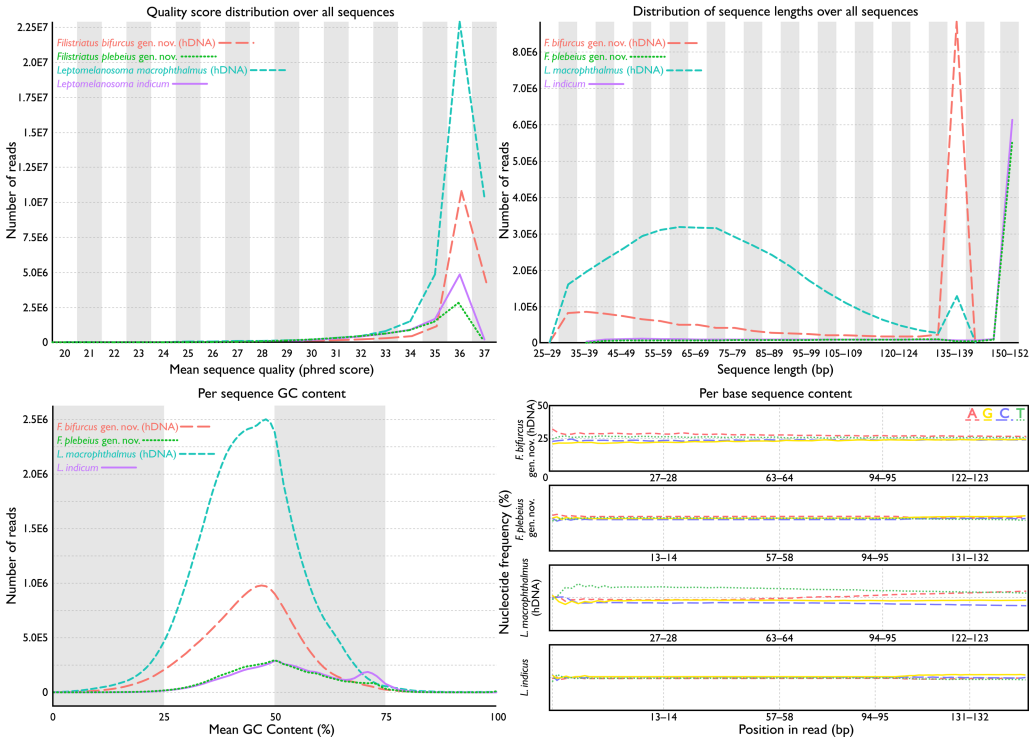

Figure 2 **FastQC assessments of sequence quality for historic *versus* modern DNA.** Taxa selected for comparison based on phylogenetic relatedness (see phylogeny). Historic DNA libraries were prepared for genome skimming. Modern DNA libraries were prepared using target-capture approach. Modern DNA libraries prepared for genome skimming are similar in quality score distributions, per-sequence GC content, and per-base sequence content to modern DNA libraries prepared for target capture. Distribution of sequence lengths were more variable for genome-skimmed libraries than target-capture libraries.

Full-size DOI: 10.7717/peerj.20029/fig-2

taxon based on the phylogeny presented in *Girard et al.* (*2022a*; if available) generally used as the reference sequence for assembly (see Table 1). The 'map to reference' function in Geneious was set to 'medium-low sensitivity' and 'iterate up to five times.' To assess DNA damage in historic samples, we aligned merged, overlapping read pairs to final assemblies using bwa aln (*Li & Durbin, 2009*), removed duplicates and reads with MAPQ scores <20, and analyzed resulting bam files with mapdamage2 (*Jónsson et al., 2013*). Historical and modern mitogenomes were then annotated using MitoAnnotator (*Iwasaki et al., 2013*; *Sato et al., 2018*; *Zhu et al., 2023*). Annotated mitogenomes were submitted to GenBank and assigned accession numbers (PV590073–PV590096; PV590343–PV590344; PV593525; see File S1).

Partitioning schemes and phylogenetic analyses

For the UCE dataset, we used the same partitioning scheme of 555 subsets and associated models as listed in *Girard et al. (2022a)*. For the mitochondrial dataset, orthologous loci for two rRNA and 13 protein-coding regions of sequenced mitogenomes were collated into individual FASTA files and aligned with MAFFT version 7 (*Katoh & Standley, 2013*). Additionally, COI 'barcode' sequences from *Filimanus similis* (MF281368), "*P.*" *longipes*

(PV590860), "*P.*" *mullani* (MF281374), and *P. oligodon* (JQ365495) were included in the COI alignment (as in *Girard et al., 2022a*). Lengths of alignments were as follows, with completeness at the level of individual bp in parentheses: 12S 1,050 bp (89%); 16S 1,900 bp (88%); ATPase6 702 bp (96%); ATPase8 189 bps (88%); COI 1,566 bp (93%); COII 695 bp (99%); COIII 786 bp (96%); CytB 1,165 bp (96%); ND1 978 bp (99%); ND2 1,102 bp (93%); ND3 349 bp (100%); ND4 1,385 bp (98%); ND4L 297 bp (100%); ND5 1,872 bp (97%); ND6 536 bp (98%). Individual matrices were concatenated for partitioning and phylogenetic inference. The final alignment was 14,572 bp in length and 89% complete at the level of individual bp. IQ-TREE version 2.2.2.6 (TESTMERGEONLY with rclusterf 50; *Chernomor, Von Haeseler & Minh, 2016*; *Kalyaanamoorthy et al., 2017*; *Minh et al., 2020*) recovered an optimal partitioning scheme of 12 partitions and associated models. Model assessment and resulting partition information can be found in Files S2 and S3. The mitochondrial dataset, UCE dataset, and associated partitioning schemes were analyzed simultaneously using IQ-TREE. Twenty tree searches were performed with the perturbation strength (-pers) set to 0.2 and the number of unsuccessful iterations to stop (-nstop) set to 2,000. Support for the best-fitting topology was generated using 500 standard bootstrap replicates (-bo) and reconciled with the most likely phylogeny using IQ-TREE (-con; Files S4 and S5).

Morphological investigation

We explored new and previously described (*e.g.*, *Marathe & Bal, 1958*; *Kang, 2017*; *Presti, Johnson & Datovo, 2023*) morphological variation as they related to the placement of "*P.*" *bifurcus* and "*P.*" *macrophthalmus* in the phylogeny. Microcomputed tomography (μCT) was used to examine internal osteology of museum specimens. Specimens were scanned using a GE Phoenix v|tome| x M 240/180 kV Dual Tube μCT at the National Museum of Natural History, Smithsonian Institution. Scan settings were 90–130 kV, 130–190 μA, 250–500 ms exposure time, and 25–57 μm voxel size. Resulting scans are available through MorphoSource project ID [000735880] and media identifiers (*L. indicum* USNM 357761 [000735950], *Pentanemus quinquarius* UF 221610 [000167767], "*P.*" *bifurcus* USNM 76627 [000735955], "*Polydactylus*" *macrophthalmus* UMMZ 171714 [000735937], "*P.*" *plebeius* USNM 403223 [000735960], "*P.*" *sexfilis* CSIRO C261 [000735945], and *P. virginicus* FMNH 104648 [000735940]). All scan data were visualized and segmented using the SlicerMorph module (*Rolfe et al., 2021*) in 3D Slicer (*Fedorov et al., 2012*) and the protocol described in *Girard et al. (2022b)*. All other specimen imaging was performed using equipment and protocols listed in *Girard, Davis & Smith (2020)*.

Nomenclature

The electronic version of this article in Portable Document Format (PDF) will represent a published work according to the International Commission on Zoological Nomenclature (ICZN), and hence the new names contained in the electronic version are effectively published under that Code from the electronic edition alone. This published work and the nomenclatural acts it contains have been registered in ZooBank, the online registration system for the ICZN. The ZooBank LSIDs (Life Science Identifiers) can

be resolved and the associated information viewed through any standard web browser by appending the LSID to the prefix http://zoobank.org/. The LSID for this publication is: urn:lsid:zoobank.org:pub:E0E70F44-3417-4BD9-87EF-F4F2CD157F68. The online version of this work is archived and available from the following digital repositories: *PeerJ*, PubMed Central SCIE and CLOCKSS.

RESULTS

Assessment and annotation of sequencing reads

Total cleaned reads generated from hDNA samples of "*P.*" *bifurcus* and "*P.*" *macrophthalmus* equaled 37,034,349 and 86,551,031 bp, respectively. Mean number of reads from modern genetic samples equaled 18,834,790 bp (range: 12,758,097–24,947,149 bp; see Table 1). Analyses of reads *via* FastQC showed similarly acceptable per-base sequence quality assessments, per-sequence quality scores, per-base N content, and sequence duplication levels across all samples (Fig. 2). FastQC flagged GC content for 18 of the 27 samples. These flags were generally associated with samples where the percent GC content was under 42. Mean percent GC content across all samples was 41, with a range of 39–46 (Table 1). Distribution of sequence lengths was more variable from libraries prepared for genome skimming than those for target capture, with a greater abundance of shorter reads in the genome-skimmed libraries. The greatest variability in sequencing-read length came from the hDNA sample of "*P.*" *macrophthalmus* (Fig. 1).

Assembled and adjusted contigs from hDNA samples were 16,346 ("*P.*" *macrophthalmus*) and 16,564 bp ("*P.*" *bifurcus*). After quality filtering and duplicate removal, mean contig coverage range was 63.9–111.2× and mean fragment length range was 77.3–83.5 bp (Fig. 3, Table 1). As expected (*e.g.*, *Dabney, Meyer & Pääbo, 2013*), terminal deamination was found on the 5′ and 3′ ends of hDNA reads, with deamination frequencies ranging between 5% and 8% in "*P.*" *macrophthalmus* and "*P.*" *bifurcus*, respectively (see Files S6).

Mitogenomes for 23 species of threadfins were assembled and annotated from newly sequenced modern samples and previously sequenced SRAs. Contig lengths ranged from 16,606–17,428 bps, with mean contig coverage ranging from 51–855.9× (Table 1). Across modern and historic samples, all polynemid mitogenomes encoded 13 protein-coding loci, two rRNAs, and one D-loop. All but one sample encoded 22 tRNAs (tRNA-Ser and tRNA-Leu removed from "*P.*" *macrophthalmus*). Locus orientation and order match those found in previously sequenced species of the Polynemidae (*Miya et al., 2003*). For two of the outgroup taxa in the dataset that did not have mitogenomes publicly available (Files S1), only a subset of mitochondrial loci were able to be extracted from previously published SRAs sequenced from modern genetic samples. These include: *Leptobrama muelleri* (2 rRNA, 11 protein-coding loci) and *Rheocles wrightae* (1 rRNA, 8 protein-coding loci; Files S1).

Phylogenetic analyses

The hypothesis of relationships recovered from the analysis is shown in Fig. 4. The bootstrap analysis yielded 32 nodes (of 40; ≈80%) with a value of ≥80% and 28 nodes (≈70%) with a value of ≥95% (Fig. 4). All nodes outside of the Polynemidae were supported with ≥93%

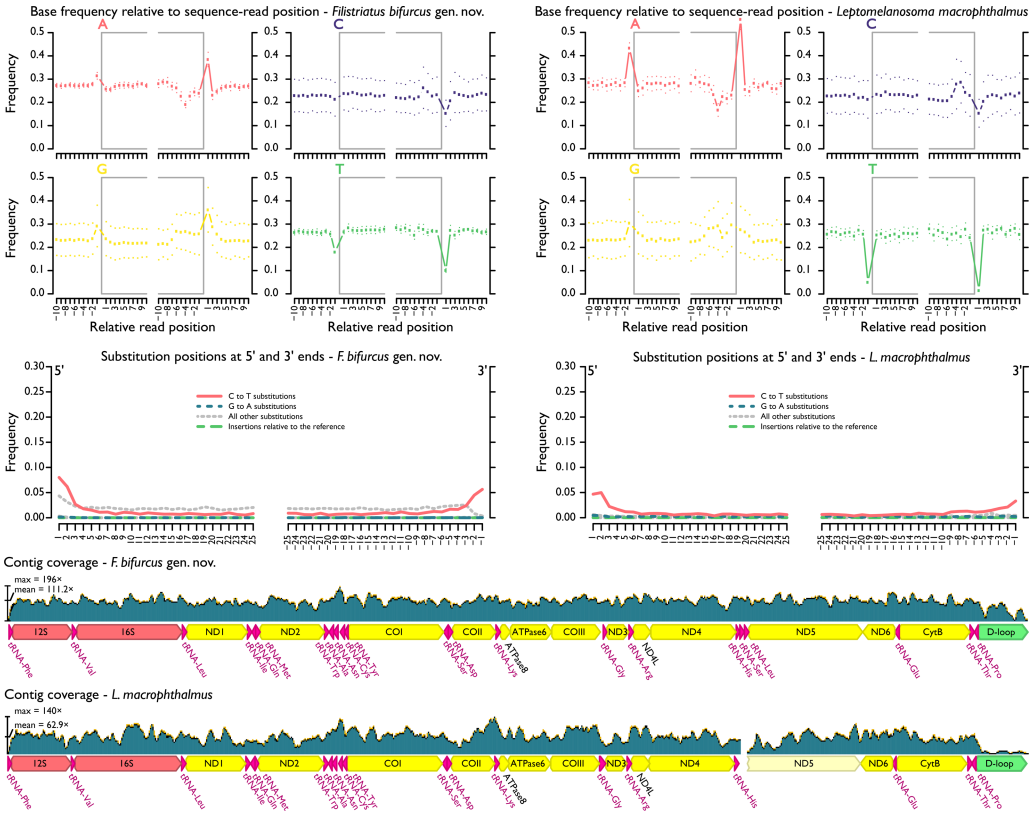

Figure 3 **MapDamage2 assessments of base pairs and contig coverage for historic DNA.** Upper plots highlight position and frequency of nucleotides both within and outside the sequence read, with grey boxes corresponding to the read. Strong deviation in base-call frequency immediately outside of the read indicates read damage. Middle plots highlight position of specific substitutions from the 5′ and the 3′ ends. Overall, misincorporations followed general trends found in other studies (*e.g.*, *Dabney, Meyer & Pääbo, 2013*) and were concentrated at the ends of sequencing reads. Lower plots show distribution of coverage across hDNA contigs. Additional files output from mapDamage2 not shown here are included in the supplement.

Full-size DOI: 10.7717/peerj.20029/fig-3

bootstrap support. Within the Polynemidae, bootstrap supports were higher than those found by *Girard et al. (2022a)* in their combined analysis. This is likely due to this analysis sampling multiple mitochondrial loci *versus* only one mitochondrial locus as in *Girard et al. (2022a)*.

The resulting topology showed a similar topology to that of *Girard et al. (2022a)* except for "*Polydactylus*" *quadrifilis*, *Pentanemus quinquarius*, and *Galeoides decadactylus* being recovered as an early diverging grade sister to the rest of the Polynemidae rather than a clade. "*Polydactylus*" *macrophthalmus* is recovered sister to *Leptomelanosoma indicum*. "*Polydactylus*" *bifurcus* is recovered sister to a clade of "*P.*" *plebeius* and "*P.*" *sexfilis*. Given the non-monophyly of *Polydactylus* in multiple studies (*e.g.*, *Kang, 2017*; *Girard et al., 2022a*; *Presti, Johnson & Datovo, 2023*) and the type species of the genus (*P. virginicus*) recovered in a distant clade, we describe a new genus for "*P.*" *bifurcus*, "*P.*" *plebeius*, and "*P.*" *sexfilis* and reclassify "*P.*" *macrophthalmus* in the genus *Leptomelanosoma*.

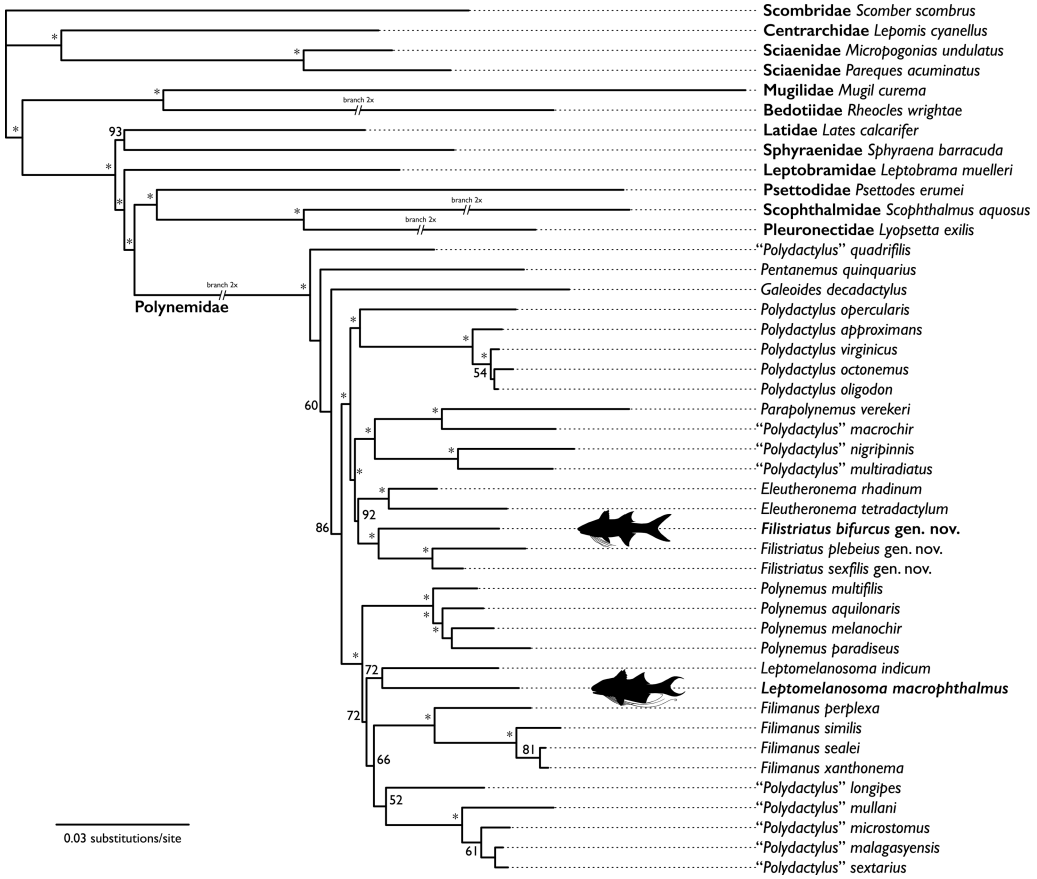

Figure 4 Hypothesis of relationships from likelihood analysis of polynemid and outgroup taxa based on mitochondrial and UCE datasets. Bootstrap support values <50 not shown. Bootstrap support values indicated by "*" indicate support value ≥95. Taxa with representative silhouettes highlight samples sequenced from historic museum specimens.

Full-size DOI: 10.7717/peerj.20029/fig-4

Taxonomic modifications to the Polynemidae

***Filistriatus* gen. nov. Girard** urn:lsid:zoobank.org:act:983362F8-2F4E-4730-A9BE-BE24F2784474

Diagnosis: a genus of small-to-moderately-sized threadfins differentiated from all other genera of the Polynemidae by the following combination of characters: 7–9 dark stripes along the longitudinal scale rows above the lateral line, 7–9 faint stripes along the longitudinal scale rows below the lateral line, reduced lateral pore of the supraorbital canal, reduced sensory canal commissure on the dorsal surface of the frontal, 5–6 free pectoral-fin rays, and 54–72 lateral-line scales.

Type species: *Polynemus sexfilis* Valenciennes 1831

Included species: *Polydactylus bifurcus* Motomura, Kimura & Iwatsuki 2001, *Polydactylus plebeius* (Broussonet 1782), *Polydactylus sexfilis* (Valenciennes 1831), *Polydactylus siamensis* Motomura, Iwatsuki & Yoshino 2001.

Etymology: the generic name refers to the thread-like fin rays of the pectoral fin and the diagnostic stripes along the lateral flanks [*fili* (Latin) = thread and *striatus* (Latin) = striped]. Gender masculine.

Remarks: "*Polydactylus*" *siamensis* was characterized by several longitudinal dark stripes along the flank, 5 pectoral filaments, and 54–58 lateral-line scales (*Motomura, Iwatsuki & Yoshino, 2001*). Although we could not sample this taxon in this study, the species has the diagnostic characters of *Filistriatus*. The species is also reassigned to the new genus.

Reassessment of *Leptomelanosoma* Motomura & Iwatsuki 2001

Diagnosis: a genus of the Polynemidae with the following combination of characters: ethmoid not covered dorsally by frontals, basisphenoid posteriorly displaced, prootic excluded from rear margin of orbit by pronounced expansion of pterosphenoid.

Type species: *Polydactylus indicus* (Shaw 1804)

Included species: *Polydactylus indicus* (Shaw 1804), *Polydactylus macrophthalmus* (Bleeker 1858)

Remarks: although not included in the description by *Motomura & Iwatsuki (2001a)*, the lateral-line scales in specimens of *L. indicum* are poorly pigmented and nearly white along the length of the flank in both fresh and preserved specimens. Further, the scales immediately above and below the lateral line are densely pigmented, accentuating the lateral-line scales in this taxon. In specimens of *L. macrophthalmus*, a similar condition exists in the lateral line, particularly in the scales anterior to the anal fin of preserved specimens (see *Motomura et al, 2001*, fig. 2). We did not include this lateral-line pigmentation as diagnostic for the genus as we could not locate images of freshly caught specimens of *L. macrophthalmus*. Subsequent works should explore this character as a diagnostic feature of the genus.

DISCUSSION

Species and relationships of *Filistriatus*

Filistriatus bifurcus is a medium-sized threadfin (119–272 mm SL; *Motomura, 2004*) currently only known from shallow waters along the southern coast of Indonesia and the southwestern coast of Taiwan. The species was described from a single specimen captured near Lombok Island, Indonesia, based on differences in fin-ray and lateral-line scale counts, robustness of the second dorsal-fin spine, and bifurcation of the lateral line on the caudal fin. The bifurcation of the lateral line on the caudal fin has been documented in a few species of threadfins, including two species of *Eleutheronema* and species of *Polydactylus* that occur in the New World (*Motomura, 2004*). Some species of *Eleutheronema* have two bifurcations of the lateral line on the caudal fin, with a secondary bifurcation occurring on the lower lobe (*Motomura, 2004*). *Filistriatus bifurcus* was further characterized by 8–9 dark stripes above the lateral line and 8–9 faint stripes below the lateral line, along longitudinal flank scale rows. Only three species of threadfin are known to have similar striped patterns above and below the lateral line: *F. plebeius*, *F. siamensis*, and *F. sexfilis*. These three species are largely sympatric over their respective ranges (*Feltes, 1986*; *Motomura, Iwatsuki & Yoshino, 2001*; *Motomura, Iwatsuki & Kimura, 2001*; *Motomura, 2004*), with *F. plebeius* occurring as

far west as the eastern coast of Africa, *F. sexfilis* occurring as far east as the Hawaiian Islands, and *F. siamensis* found only in the waters around Thailand. Feltes (*1986*: 516) notes that *F. plebeius* and *F. sexfilis* are "so closely related" that differentiating individual specimens can be difficult. *Filistriatus plebeius* and *F. sexfilis* have been included in three analyses on polynemid intrarelationships, being recovered either as a clade (*Kang, 2017*; *Girard et al., 2022a*) or as close allies in a clade with *Eleutheronema* and "*P.*" *opercularis* (*Presti, Johnson & Datovo, 2023*). However, *F. bifurcus* and *F. siamensis* have yet to be included in an analysis.

The mitogenome of *F. bifurcus* was obtained through the extraction of hDNA from the liver of a 100+-year-old museum specimen. The combined analysis of these data with mitochondrial and UCE datasets recovered *F. bifurcus* in a clade with *F. plebeius* and *F. sexfilis*. Several counts and measurements among species of *Filistriatus* support the recovered relationship, including similar body depth at first dorsal-fin origin relative to SL (*F. bifurcus*: 26–28%; *F. plebeius*: 25–34%; *F. sexfilis*: 27–34%), upper-jaw length relative to SL (*F. bifurcus*: 14%; *F. plebeius* and *F. sexfilis*: 13–16% for both), and number of pored lateral-line scales (*F. bifurcus*: 69–72; *F. plebeius*: 60–68; *F. sexfilis*: 60–67). Further, the presence of 7–9 dark longitudinal stripes above and below the lateral line (see above; *Motomura, Kimura & Iwatsuki, 2001*; *Motomura, 2004*) is unique among members of the family. Characters in the neurocranium also support these taxa being closely related, including a reduced lateral pore of the supraorbital canal and reduced canal commissure on the dorsal surface of the frontal. In their assessment of morphological variation across species of threadfins that occur near Bombay, *Marathe & Bal (1958)* noted morphological variation in the supraorbital sensory canal in all species of threadfins they examined. Typically, this canal has a pronounced lateral pore that is visible when viewing the lateral aspect of the neurocranium and a medial commissure that is elevated and open mesially. In species of *Filistriatus*, the lateral pore is reduced, not visible within the lateral triangular aperture of the frontal (Fig. 5). Additionally, the canal commissure is reduced, with a distinct lowering of the dorsal surface compared to other species of *Polydactylus* (Fig. 5). Based on these morphological features, overlapping meristics, and unique pigmentation pattern among threadfins, we find strong support for a clade of *Filistriatus*, as recovered in the analysis of mitochondrial and UCE loci.

Species and relationships of *Leptomelanosoma*

Leptomelanosoma indicum is a widely distributed species of threadfin, found from Pakistan to Papua New Guinea (*Motomura, 2004*). The species was described as *Polynemus indicus* Shaw 1804 based on an illustrated specimen from India (see *Russell, 1803*: 68, fig. 184) and later attributed to the genus *Polydactylus* by *Myers (1936)*. *Marathe & Bal (1958)* noted several characters that are unique to the species among threadfins, including a large ethmoid extending anteriorly and exposed dorsally by the frontals, an oval-shaped vomerine tooth plate, a posteriorly displaced basisphenoid, a dorsally exposed sphenotic, and an enlarged palatine. *Motomura & Iwatsuki (2001a)* described a new genus, *Leptomelanosoma*, for the taxon based on the poorly developed lip adjacent to the lower jaw, the gas bladder with numerous lateral appendages, the ethmoid exposed dorsally by frontals, the anterior



Figure 5 **Morphology relating to the monophyly of *Filistriatus* gen. nov.** (A) Fresh specimen of *F. plebeius* gen. nov. (USNM 471329) in lateral view. (B) Neurocranium of *F. bifurcus* gen. nov. (USNM 76627) in dorsolateral view. (C) Neurocranium of *F. sexfilis* gen. nov. (CSIRO C261) in dorsolateral view. (D) Neurocranium of *P. virginicus* (FMNH 104648) in dorsolateral view. (E) Neurocranium of *L. indicum* (USNM 357716) in dorsolateral view. Left frontal and medial supraorbital canal commissure colored in (B–E), with isolated colored frontals below each whole neurocranium. Arrow I indicates dark stripes along the longitudinal scale rows above the lateral line. Faint stripes also present along the longitudinal scale rows below the lateral line. Arrow II indicates reduced lateral pore of the supraorbital canal. Arrow III indicates reduced medial commissure of sensory canal on the dorsal surface of the frontal. Arrow IV indicates pronounced lateral pore of the supraorbital canal that is visible when viewing the side of the neurocranium. Arrow V indicates elevated medial commissure of sensory canal. Scale bars = five mm.

Full-size DOI: 10.7717/peerj.20029/fig-5

one-third of lower jaw with small teeth extending onto the lateral surface, the sphenotics exposed on the margin of neurocranium, filamentous caudal-fin rays, the vomer with oval-shaped tooth plate, and the wide tooth plates on palatine and ectopterygoid. While diagnostic to *Leptomelanosoma*, the poorly developed lower-jaw lip and filamentous caudal-fin rays were listed as homoplastic with the monospecific genus *Parapolynemus* by *Motomura & Iwatsuki (2001a)*, and several other characters, such as the gas bladder with lateral appendages and sphenotics visible on the dorsal margin of the neurocranium, were found to be autapomorphic among threadfins (*Motomura & Iwatsuki, 2001a*). *Presti, Johnson & Datovo (2023)* recovered *L. indicum* sister to a clade of *Parapolynemus* and *Polynemus* based on nine morphological characters. They also identified 14 autapomorphic characters for *Leptomelanosoma*, including articulation between the metapterygoid and endopterygoid (their character 38) and fusion of the endopterygoid and ectopterygoid (their character 48). The study by *Girard et al. (2022a)* recovered *L. indicum* sister to a clade of *Filimanus* and the black-spotted species of "*Polydactylus*" but did not report any morphological characters that supported this hypothesis.

Leptomelanosoma macrophthalmus (Bleeker 1858) was originally described in the genus *Polynemus* based on two specimens from Indonesia. Only known to occur in rivers on two Indonesian islands, the Kapuas River in Kalimantan and the Batanghari and Musi Rivers in Sumatra, this species is thought to have the most-restricted distribution of any species within the family (*Motomura, Iwatsuki & Kimura, 2001*). The taxon was considered a species of *Polynemus* until *Feltes (1993)* included it in *Polydactylus*. *Motomura, Iwatsuki & Kimura (2001)* redescribed the species, diagnosing it by the presence of villiform teeth in broad bands on the vomer, palatine, and ectopterygoid, elongate pectoral-fin rays, and a well-developed gas bladder, among other characters. No previous studies have included this taxon in a phylogenetic analysis.

The mitochondrial genome of *L. macrophthalmus* was acquired through the extraction of hDNA from the liver of a 20+-year-old museum specimen. The combined analyses of these data with mitochondrial and UCE loci recovered *L. macrophthalmus* sister to *L. indicum*. When examining the morphology of *L. macrophthalmus*, we find four characters in the neurocranium that support a sister-group relationship between these taxa, including characters once considered to be unique to *L. indicum* (Fig. 6). The typical condition in polynemid neurocrania is for the anterior margin of the frontal to extend rostrally above the ethmoid and to meet at a discrete point (see *Marathe & Bal, 1958*, fig. 2; *Motomura & Iwatsuki, 2001a*, fig. 4). One of the diagnostic characters of *Leptomelanosoma* is the truncation of the frontal and separation from the opposing frontal rostrally, such that the dorsal margin of the ethmoid is exposed (*Marathe & Bal, 1958*; *Motomura & Iwatsuki, 2001a*; Fig. 6). When examining the neurocranium of *L. macrophthalmus*, the anterior margin of the frontal is also truncated and distinctly separated from the opposing frontal, with the dorsal margin of the ethmoid visible (Fig. 6). This character has not been previously mentioned for this or any other species of polynemid outside of *L. indicum* (*Motomura, Iwatsuki & Kimura, 2001*; *Motomura, 2004*). Across polynemid genera, several character states are found within the posterior margin of the orbit, the position of the basisphenoid, the shape of the prootic, and the interaction between the parasphenoid,

prootic, and pterosphenoid. The typical condition for species of *Polydactylus* is for the basisphenoid to be largely vertical and positioned anterior to the posterior margin of the orbit. Further, the prootic typically contributes to the posterior margin of the orbit, separating the parasphenoid from the pterosphenoid. *Marathe & Bal (1958)* noted that the basisphenoid is not visible in lateral view in *L. indicum* and we find a similar condition in *L. macrophthalmus* (Fig. 6). The basisphenoid in both taxa is posteriorly displaced and reclined within the cranial vault, but remains in contact with the pterosphenoid and prootic, as is typical in polynemids broadly (*Feltes, 1986*; *Feltes, 1991*; *Feltes, 1993*). *Presti, Johnson & Datovo (2023)* found that the prootic was excluded from the posterior margin of the orbit in *L. indicum* as well as *Eleutheronema*, *Parapolynemus*, *Pentanemus*, and *Polynemus*. *Feltes (1993)* noted that the condition in *Eleutheronema* was different than that of *Parapolynemus* and *Polynemus*, as exclusion of the prootic in *Eleutheronema* was by a slight connection between the anteroventral corner of the pterosphenoid and anterodorsal corner of the parasphenoid *versus* the extensive expansion of the pterosphenoid, dorsal extension of the parasphenoid, and broad point of contact between the parasphenoid and pterosphenoid in *Parapolynemus* and *Polynemus*. In the specimens we examined, we did not find the prootic excluded from the posterior margin of the orbit in *Pentanemus*. The contact between the pterosphenoid and parasphenoid is broader in *Leptomelanosoma* than in *Eleutheronema*, with the pterosphenoid having a substantial anteroventral arm that extends to and interacts with the parasphenoid. The lateral arm of the prootic that reaches towards the sphenotic is also shortened, not extending to the posterior margin of the orbit. The contact between the parasphenoid and pterosphenoid in *Leptomelanosoma* is not as broad as what is seen in *Parapolynemus* and *Polynemus*, where the pterosphenoid is extensively expanded (*Feltes, 1993*). Based on these morphological features, we find strong support for the inclusion of *L. macrophthalmus* in the genus *Leptomelanosoma*, as recovered with molecular data.

CONCLUSIONS

Historical DNA can reliably inform modern phylogenetic analyses

This study builds on the recent success of sequencing DNA from historic and/or formalin-fixed specimens, successfully obtaining sequence data from fluid-preserved museum specimens using a commercial-kit-based approach. The resulting mitochondrial data from these specimens allowed for two enigmatic taxa, *F. bifurcus* and *L. macrophthalmus*, to be placed in an evolutionary context for the first time. This new phylogenetic framework allowed for the further study of internal morphology and subsequent modifications to the classification of "*Polydactylus*" based on both genomic and anatomical characters, including the description of a new genus. This result further supports analyzing historic samples alongside a robust dataset of molecular markers generated from modern samples can inform modern phylogenetic analyses. The methods described here add to the growing body of literature of methods supporting the utility of hDNA to genetically characterize type specimens, to aid in accounting for genetic diversity within species complexes, and to improve our understanding of taxonomy and classification across groups of animals that have millions of historically inaccessible specimens for genomic approaches.

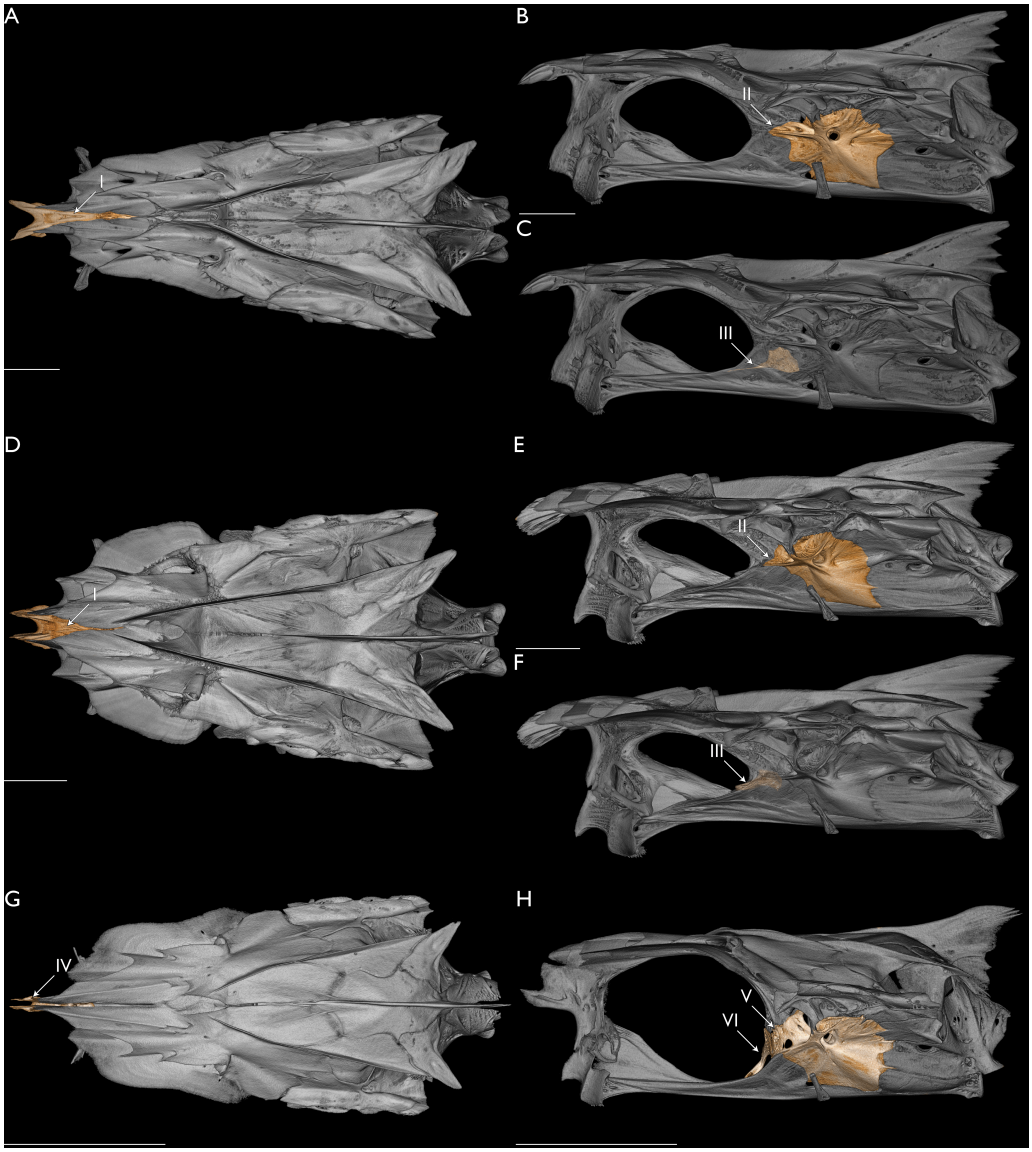

Figure 6 **Morphology relating to the monophyly of *Leptomelanosoma*.** Neurocranium of *L. macrophthalmus* (UMMZ 171714) in dorsal (A) and lateral (B and C) views. Neurocranium of *L. indicum* (USNM 357716) in dorsal (D) and lateral (E and F) views. Neurocranium of *Polydactylus virginicus* (FMNH 104648) in dorsal (G) and lateral (H) views. (A, D, G) Ethmoid colored. (B, E, H) Prootic colored. (C, F, H) Basisphenoid colored. Opacity of neurocranium reduced in (C) and (F) so basisphenoid position could be visualized. Arrow I indicates ethmoid not covered by frontals and exposed dorsally. Arrow II indicates prootic excluded from rear margin of orbit by pronounced expansion of pterosphenoid. Arrow III indicates basisphenoid posteriorly displaced. Arrow IV indicates ethmoid covered by frontals. Arrow V indicates prootic included in rear margin of orbit, separating parasphenoid from pterosphenoid. Arrow VI indicates basisphenoid largely vertical and anterior to the posterior margin of the orbit. Scale bars = five mm.

Full-size DOI: 10.7717/peerj.20029/fig-6

Collections concerns surrounding the genomic potential of museum specimens

Targets for destructive hDNA sampling typically represent rare, extinct, taxonomically, and historically significant specimens (*e.g.*, *Palandačić et al., 2024*; *Muschick, Rüber & Matschiner, 2025*). As museum specimens are irreplaceable records of biodiversity, researchers and collection staff should work collaboratively to balance the growing demands for destructive genomic research with the long-term conservation of specimens for other research needs. We performed minimally invasive dissections to extract liver samples through small abdominal incisions. Prioritizing both specimen integrity and genetic sampling allowed for genomic and morphological approaches to be applied, classification and taxonomy to be revised, and our understanding of threadfin evolution to be improved. While the techniques presented in this study can be applied to specimens greater than 50 mm SL, successfully sampling hDNA from smaller specimens and maintaining an intact morphological vouchers presents additional challenges (*e.g.*, *Muschick, Rüber & Matschiner, 2025*). As technology and methods continue to improve, the likelihood of acquiring viable hDNA from muscle punches or external features will increase, decreasing the need for invasive sampling from internal organs. Until those methods are developed, it is critical that both answering new questions and specimen preservation are balanced, ensuring these unique records are available to future researchers.

ACKNOWLEDGEMENTS

We would like to thank C Baldwin, K Bemis, MTR Hawkins, GD Johnson [deceased], T Munroe, D Pitassy, and WL Smith for discussions and feedback on this manuscript; C Craig and J Steier for performing modern DNA library preparation; JJ Hill for facilitating μCT scanning; N Redmond for assistance with project management; A Bentley, G Dally, A Dettaï, B Frable, A Graham, M Hammer, C McMahan, S Mochel, K Murphy, D Nelson, D Pitassy, A Reft, R Robbins, M Sabaj, B Sidlauskas, R Singer, K Swagel, Tan HH, J Williams, and L Willis, for providing specimen support and/or access to specimens in their care.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This study was supported by a Smithsonian Institution Barcode Network (SIBN) grant at NMNH (FY23 award cycle) and the Herbert R. and Evelyn Axelrod Endowment for Systematic Ichthyology at NMNH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Smithsonian Institution Barcode Network: FY23.
Herbert R. and Evelyn Axelrod Endowment for Systematic Ichthyology NMNH.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Matthew G. Girard conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Kevin R. Chovanec conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The annotated mitochondrial genomes are available at GenBank: PV590073–PV590096; PV590343–PV590344; PV593525; SRR33326234–SRR33326248.

Data Availability

The following information was supplied regarding data availability:

The annotated mitochondrial genomes are available at GenBank: PV590073–PV590096; PV590343–PV590344; PV593525; SRR33326234–SRR33326248.

The models are available at MorphoSource: Available at https://www.morphosource.org/projects/000735880.

Users are required to request a loan of the scans either by contacting the museums or using the MorphoSource "Request Media" or "Requests Download" functionality:

– https://doi.org/10.17602/M2/M735960: Diane Pitassy, pitassyd@si.edu;

– https://doi.org/10.17602/M2/M735955: Diane Pitassy, pitassyd@si.edu;

– https://doi.org/10.17602/M2/M735950: Diane Pitassy, pitassyd@si.edu;

– https://doi.org/10.17602/M2/M735945: Helen O'Niell, helen.oneill@csiro.au; Nicole Fisher, nicole.fisher@csiro.au;

– https://doi.org/10.17602/M2/M735940: Caleb McMahan, cmcmahan@fieldmuseum.org;

– https://doi.org/10.17602/M2/M735937: Hernan Lopez-Fernandez, hlopezf@umich.edu.

Phylogenetic analyses were performed on the High Performance Cluster (SI/HPC), Smithsonian Institution https://doi.org/10.25572/SIHPC.

New Species Registration

The following information was supplied regarding the registration of a newly described species:

Publication LSID: urn:lsid:zoobank.org:pub:E0E70F44-3417-4BD9-87EF-F4F2CD157F68.

Filistriatus Genus LSID: urn:lsid:zoobank.org:act:983362F8-2F4E-4730-A9BE-BE24F2784474.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.20029#supplemental-information.

REFERENCES

Alfaro ME, Faircloth BC, Harrington RC, Sorenson L, Friedman M, Thacker CE, Oliveros CH, Cerný D, Near TJ. 2018. Explosive diversification of marine fishes at the Cretaceous Palaeogene boundary. *Nature Ecology and Evolution* **2**:688–696 DOI 10.1038/41559-018-0494-6.

Appleyard SA, Maher S, Miskiewicz AG, Lara-Lopez A, Matis P, Fielder DS, Suthers IM. 2022. Genetic and morphological identification of formalin fixed, preserved larval fishes; can we have the best of both worlds? *Journal of Experimental Marine Biology and Ecology* **553**:151763 DOI 10.1016/j.jembe.2022.151763.

Bernardi G. 2011. Monophyletic origin of brood care in damselfishes. *Molecular Phylogenetics and Evolution* **59**:245–248 DOI 10.1016/j.ympev.2010.12.021.

Bernstein JM, De Souza H, Murphy J, Voris H, Brown R, Myers E, Harrington S, Shanker K, Ruane S. 2023. Phylogenomics of fresh and formalin specimens resolves the systematics of Old World mud snakes (Serpentes: Homalopsidae) and expands biogeographic inference. *Bulletin of the Society of Systematic Biologists* **2**:1–24 DOI 10.18061/bssb.v2i1.9393.

Betancur RR, Wiley EO, Arratia G, Acero A, Bailly N, Miya M, Lecointre G, Ortí G. 2017. Phylogenetic classification of bony fishes. *BMC Evolutionary Biology* **17**:162 DOI 10.1186/s12862-017-0958-3.

Chakraborty A, Sakai M, Iwatsuki Y. 2006. Museum fish specimens and molecular taxonomy: a comparative study on DNA extraction protocols and preservation techniques. *Journal of Applied Ichthyology* **22**:160–166 DOI 10.1111/j.1439-0426.2006.00718.x.

Chen WJ, Bonillo C, Lecointre G. 2003. Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. *Molecular Phylogenetics and Evolution* **26**:262–288 DOI 10.1016/s1055-7903(02)00371-8.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**:i884–i890 DOI 10.1093/bioinformatics/bty560.

Chernomor O, Von Haeseler A, Minh BQ. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology* **65**:997–1008 DOI 10.1093/sysbio/syw037.

Christmas MJ, Kaplow IM, Genereux DP, Dong MX, Hughes GM, Li X, Sullivan PF, Hindle AG, Andrews G, Armstrong JC, Bianchi M, Breit AM, Diekhans M, Fanter C, Foley NM, Goodman DB, Goodman L, Keough KC, Kirilenko B, Kowalczyk A, Lawless C, Lind AL, Meadows JRS, Moreira LR, Redlich RW, Ryan L, Swofford R, Valenzuela A, Wagner F, Wallerman O, Brown AR, Damas J, Fan K, Gatesy J, Grimshaw J, Johnson J, Kozyrev SV, Lawler AJ, Marinescu VD, Morrill KM, Osmanski A, Paulat NS, Phan BN, Reilly SK, Schäffer DE, Steiner C, Supple MA, Wilder AP, Wirthlin ME, Xue JR, Consortiumß Z, Birren BW, Gazal S, Hubley RM, Koepfli K-P, Marques-Bonet T, Meyer WK, Nweeia M, Sabeti PC, Shapiro B, Smit AFA, Springer MS, Teeling EC, Weng Z, Hiller M, Levesque DL, Lewin HA,

Murphy WJ, Navarro A, Paten B, Pollard KS, Ray DA, Ruf I, Ryder OA, Pfenning AR, Lindblad-Toh K, Karlsson EK, Andrews G, Armstrong JC, Bianchi M, Birren BW, Bredemeyer KR, Breit AM, Christmas MJ, Clawson H, Damas J, Palma FD, Diekhans M, Dong MX, Eizirik E, Fan K, Fanter C, Foley NM, Forsberg-Nilsson K, Garcia CJ, Gatesy J, Gazal S, Genereux DP, Goodman L, Grimshaw J, Halsey MK, Harris AJ, Hickey G, Hiller M, Hindle AG, Hubley RM, Hughes GM, Johnson J, Juan D, Kaplow IM, Karlsson EK, Keough KC, Kirilenko B, Koepfli K-P, Korstian JM, Kowalczyk A, Kozyrev SV, Lawler AJ, Lawless C, Lehmann T, Levesque DL, Lewin HA, Li X, Lind A, Lindblad-Toh K, Mackay-Smith A, Marinescu VD, Marques-Bonet T, Mason VC, Meadows JRS, Meyer WK, Moore JE, Moreira LR, Moreno-Santillan DD, Morrill KM, Muntané G, Murphy WJ, Navarro A, Nweeia M, Ortmann S, Osmanski A, Paten B, Paulat NS, Pfenning AR, Phan BN, Pollard KS, Pratt HE, Ray DA, Reilly SK, Rosen JR, Ruf I, Ryan L, Ryder OA, Sabeti PC, Schäffer DE, Serres A, Shapiro B, Smit AFA, Springer M, Srinivasan C, Steiner C, Storer JM, Sullivan KAM, Sullivan PF, Sundström E, Supple MA, Swofford R, Talbot J-E, Teeling E, Turner-Maier J, Valenzuela A, Wagner F, Wallerman O, Wang C, Wang J, Weng Z, Wilder AP, Wirthlin ME, Xue JR, Zhang X. 2023. Evolutionary constraint and innovation across hundreds of placental mammals. *Science* **380**:eabn3943 DOI 10.1126/science.abn3943.

Cooper WJ, Smith LL, Westneat MW. 2009. Exploring the radiation of a diverse reef fish family: phylogenetics of the damselfishes (Pomacentridae), with new classifications based on molecular analyses of all genera. *Molecular Phylogenetics and Evolution* **52**:1–16 DOI 10.1016/j.ympev.2008.12.010.

Dabney J, Meyer M, Pääbo S. 2013. Ancient DNA damage. *Cold Spring Harbor Perspectives in Biology* **5**:a012567 DOI 10.1101/cshperspect.a012567.

Edwards SV, Arctander P, Wilson AC. 1991. Mitochondrial resolution of a deep branch in the genealogical tree for perching birds. *Proceedings of the Royal Society B* **243**:99–107 DOI 10.1098/rspb.1991.0017.

Edwards SV, Wilson AC. 1990. Phylogenetically informative length polymorphism and sequence variability in mitochondrial DNA of Australian songbirds (*Pomatostomus*). *Genetics* **126**:695–711 DOI 10.1093/genetics/126.3.695.

Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller JV, Pieper S, Kikinis R. 2012. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging* **30**:1323–1341 DOI 10.1016/j.mri.2012.05.001.

Feller AE, Hedges SB. 1998. Molecular evidence for the early history of living amphibians. *Molecular Phylogenetics and Evolution* **9**:509–516 DOI 10.1006/mpev.1998.0500.

Feltes RM. 1986. A systematic revision of the Polynemidae (Pisces). D. Phil. Thesis, The Ohio State University.

Feltes RM. 1991. Revision of the polynemid fish genus *Filimanus*, with the description of two new species. *Copeia* **1991**:302–322 DOI 10.2307/1446580.

Feltes RM. 1993. Parapolynemus, a new genus for the polynemid fish previously known as *Polynemus verekeri*. *Copeia* **1993**:207–215 DOI 10.2307/1446312.

Foley NM, Mason VC, Harris AJ, Bredemeyer KR, Damas J, Lewin HA, Eizirik E, Gatesy J, Karlsson EK, Lindblad-Toh K, Consortium ‡ Z, Springer MS, Murphy WJ, Andrews G, Armstrong JC, Bianchi M, Birren BW, Bredemeyer KR, Breit AM, Christmas MJ, Clawson H, Damas J, Palma FD, Diekhans M, Dong MX, Eizirik E, Fan K, Fanter C, Foley NM, Forsberg-Nilsson K, Garcia CJ, Gatesy J, Gazal S, Genereux DP, Goodman L, Grimshaw J, Halsey MK, Harris AJ, Hickey G, Hiller M, Hindle AG, Hubley RM, Hughes GM, Johnson J, Juan D, Kaplow IM, Karlsson EK, Keough KC, Kirilenko B, Koepfli K-P, Korstian JM, Kowalczyk A, Kozyrev SV, Lawler AJ, Lawless C, Lehmann T, Levesque DL, Lewin HA, Li X, Lind A, Lindblad-Toh K, Mackay-Smith A, Marinescu VD, Marques-Bonet T, Mason VC, Meadows JRS, Meyer WK, Moore JE, Moreira LR, Moreno-Santillan DD, Morrill KM, Muntané G, Murphy WJ, Navarro A, Nweeia M, Ortmann S, Osmanski A, Paten B, Paulat NS, Pfenning AR, Phan BN, Pollard KS, Pratt HE, Ray DA, Reilly SK, Rosen JR, Ruf I, Ryan L, Ryder OA, Sabeti PC, Schäffer DE, Serres A, Shapiro B, Smit AFA, Springer M, Srinivasan C, Steiner C, Storer JM, Sullivan KAM, Sullivan PF, Sundström E, Supple MA, Swofford R, Talbot J-E, Teeling E, Turner-Maier J, Valenzuela A, Wagner F, Wallerman O, Wang C, Wang J, Weng Z, Wilder AP, Wirthlin ME, Xue JR, Zhang X. 2023. A genomic timescale for placental mammal evolution. *Science* **380**:eabl8189 DOI 10.1126/science.abl8189.

Gardner JL, Trueman JWH, Ebert D, Joseph L, Magrath RD. 2010. Phylogeny and evolution of the Meliphagoidea, the largest radiation of Australasian songbirds. *Molecular Phylogenetics and Evolution* **55**:1087–1102 DOI 10.1016/j.ympev.2010.02.005.

Ghezelayagh A, Harrington RC, Burress ED, Campbell MA, Buckner JC, Chakrabarty P, Glass JR, McCraney WT, Unmack PJ, Thacker CE, Alfaro ME, Friedman ST, Ludt WB, Cowman PF, Friedman M, Price SA, Dornburg A, Faircloth BC, Wainwright PC, Near TJ. 2022. Prolonged morphological expansion of spiny-rayed fishes following the end-Cretaceous. *Nature Ecology & Evolution* **6**:1211–1220 DOI 10.1038/s41559-022-01801-3.

Girard MG, Davis MP, Baldwin CC, Dettaï A, Martin RP, Smith WL. 2022a. Molecular phylogeny of the threadfin fishes (Polynemidae) using ultraconserved elements. *Journal of Fish Biology* **100**:793–810 DOI 10.1111/jfb.14997.

Girard MG, Davis MP, Smith WL. 2020. The phylogeny of carangiform fishes: Morphological and genomic investigations of a new fish clade. *Copeia* **108**:265–298 DOI 10.1643/ci-19-320.

Girard MG, Davis MP, Tan HH, Wedd DJ, Chakrabarty P, Ludt WB, Summers AP, Smith WL. 2022b. Phylogenetics of archerfishes (Toxotidae) and evolution of the toxotid shooting apparatus. *Integrative Organismal Biology* **4**:obac013 DOI 10.1093/iob/obac013.

Glenn TC, Nilsen RA, Kieran TJ, Sanders JG, Bayona-Vásquez NJ, Finger JW, Pierson TW, Bentley KE, Hoffberg SL, Louha S, Leon FJG-D, Portilla MA del R, Reed KD, Anderson JL, Meece JK, Aggrey SE, Rekaya R, Alabady M, Belanger M, Winker K,

Faircloth BC. 2019. Adapterama I: universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext). *PeerJ* **7**:e7755 DOI 10.7717/peerj.7755.

Hahn EE, Alexander MR, Grealy A, Stiller J, Gardiner DM, Holleley CE. 2022. Unlocking inaccessible historical genomes preserved in formalin. *Molecular Ecology Resources* **22**:2130–2147 DOI 10.1111/1755-0998.13505.

Hahn EE, Alexander MR, Stiller J, Grewe PM, Holleley CE. 2024. Hot alkaline lysis gDNA extraction from formalin-fixed archival tissues. *PLOS ONE* **19**:e0296491 DOI 10.1371/journal.pone.0296491.

Hawkins MTR, Flores MFC, McGowen M, Hinckley A. 2022. A comparative analysis of extraction protocol performance on degraded mammalian museum specimens. *Frontiers in Ecology and Evolution* **10**:984056 DOI 10.3389/fevo.2022.984056.

Hebert PDN, Cywinska A, Ball SL, De Waard JR. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**:313–321 DOI 10.1098/rspb.2002.2218.

Hedges SB, Poling LL. 1999. A molecular phylogeny of reptiles. *Science* **283**:998–1001 DOI 10.1126/science.283.5404.998.

Hoban ML, Whitney J, Collins AG, Meyer C, Murphy KR, Reft AJ, Bemis KE. 2022. Skimming for barcodes: rapid production of mitochondrial genome and nuclear ribosomal repeat reference markers through shallow shotgun sequencing. *PeerJ* **10**:e13790 DOI 10.7717/peerj.13790.

Hykin SM, Bi K, McGuire JA. 2015. Fixing formalin: a method to recover genomic-scale DNA sequence data from formalin-fixed museum specimens using high-throughput sequencing. *PLOS ONE* **10**:e0141579-16 DOI 10.1371/journal.pone.0141579.

Iwasaki W, Fukunaga T, Isagozawa R, Yamada K, Maeda Y, Satoh TP, 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 DOI 10.1093/molbev/mst141.

Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**:1682–1684 DOI 10.1093/bioinformatics/btt193.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Haeseler von A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**:587–589 DOI 10.1038/nmeth.4285.

Kang S. 2017. Comparative morphology and phylogenetic relationships of the family Polynemidae (Pisces: Perciformes). D. Phil. Thesis, Hokkaido University.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**:772–780 DOI 10.1093/molbev/mst010.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious basic: an integrated and extendable desktop software platform

for the organization and analysis of sequence data. *Bioinformatics* **28**:1647–1649 DOI 10.1093/bioinformatics/bts199.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America* **86**:6196–6200 DOI 10.1073/pnas.86.16.6196.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**:1754–1760 DOI 10.1093/bioinformatics/btp324.

Madsen O, Scally M, Douady CJ, Kao DJ, De Bry RW, Adkins R, Amrine HM, Stanhope MJ, de JongWW, Springer MS. 2001. Parallel adaptive radiations in two major clades of placental mammals. *Nature* **409**:610–614 DOI 10.1038/35054544.

Marathe VB, Bal DV. 1958. A brief comparative account of the axial skeleton of six polynemids from Bombay waters. *Journal of the University of Bombay* **26**:139–151.

Martin RP, Olson EE, Girard MG, Smith WL, Davis MP. 2018. Light in the darkness: new perspective on lanternfish relationships and classification using genomic and morphological data. *Molecular Phylogenetics and Evolution* **121**:71–85 DOI 10.1016/j.ympev.2017.12.029.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Haeseler von A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**:1530–1534 DOI 10.1093/molbev/msaa015.

Miya M, Takeshima H, Endo H, Ishiguro NB, Inoue JG, Mukai T, Satoh TP, Yamaguchi M, Kawaguchi A, Mabuchi K, Shirai SM, Nishida M. 2003. Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **26**:121–138 DOI 10.1016/s1055-7903(02)00332-9.

Motomura H. 2004. *Threadfins of the world (Family Polynemidae) an annotated and illustrated catalogue of polynemid species known to date*. Rome: Food and Agriculture Organization of the United Nations.

Motomura H, Iwatsuki Y. 2001a. A new genus, Leptomelanosoma, for the polynemid fish previously known as *Polydactylus indicus* (Shaw, 1804) and a redescription of the species. *Ichthyological Research* **48**:13–21 DOI 10.1007/s10228-001-8112-y.

Motomura H, Iwatsuki Y. 2001b. Review of *Polydactylus* species (Perciformes: Polynemidae) characterized by a large black anterior lateral line spot, with descriptions of two new species. *Ichthyological Research* **48**:337–354 DOI 10.1007/s10228-001-8157-y.

Motomura H, Iwatsuki Y, Kimura S. 2001. Redescription of *Polydactylus sexfilis* (Valenciennes in Cuvier and Valenciennes, 1831), a senior synonym of *P. kuru* (Bleeker, 1853) with designation of a lectotype (Perciformes: Polynemidae). *Ichthyological Research* **48**:83–89 DOI 10.1007/s10228-001-8120-y.

Motomura H, Iwatsuki Y, Yoshino T. 2001. A new species, Polydactylus siamensis, from Thailand and redescription of *P. plebeius* (Broussonet, 1782) with designation of a neotype (Perciformes: Polynemidae). *Ichthyological Research* **48**:117–126 DOI 10.1007/s10228-001-8126-5.

Motomura H, Kimura S, Iwatsuki Y. 2001. Polydactylus bifurcus, a new species of threadfin from Lombok Island, Indonesia (Perciformes: Polynemidae). *Ichthyological Research* **48**:299–305 DOI 10.1007/s10228-001-8150-5.

Motomura H, Oijen van MJP, Isbrücker IJH, Iwatsuki Y. 2001. Redescription of a rare threadfin (Perciformes: Polynemidae). Polydactylus macrophthalmus (Bleeker, 1858), with designation of a lectotype and notes on distributional implications. *Ichthyological Research* **48**:289–294 DOI 10.1007/s10228-001-8148-z.

Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* **409**:614–618 DOI 10.1038/35054550.

Muschick M, Rüber L, Matschiner M. 2025. Museomics reveals the phylogenetic position of an enigmatic vertebrate family (Lophiiformes: Lophichthyidae). *Molecular Phylogenetics and Evolution* **211**:108406 DOI 10.1016/j.ympev.2025.108406.

Myers GS. 1936. A new polynemid fish collected in the Sadong River, Sarawak, by Dr. William T. Hornaday, with notes on the genera of Polynemidae. *Journal of the Washington Academy of Sciences* **26**:376–382.

Near TJ, Dornburg A, Eytan RI, Keck BP, Smith WL, Kuhn KL, Moore JA, Price SA, Burbrink FT, Friedman M, Wainwright PC. 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 DOI 10.1073/pnas.1304661110.

Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, Wainwright PC, Friedman M, Smith WL. 2012. Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences of the United States of America* **109**:13698–13703 DOI 10.1073/pnas.1206625109.

Palandačić A, Chai MJ, Shandikov GA, Akkari N, Frade PR, Randolf S, Berg H-M, Mikschi E, Bogutskaya NG. 2024. An annotated catalogue of selected historical type specimens, including genetic data, housed in the Natural History Museum Vienna. *ZooKeys* **1203**:253–323 DOI 10.3897/zookeys.1203.117699.

Presti P, Johnson GD, Datovo A. 2023. Untangling the threads: phylogenetic relationships of threadfins (Percomorphacea: Perciformes: Polynemidae). *Zoological Journal of the Linnean Society* **199**:156–205 DOI 10.1093/zoolinnean/zlad020.

Pyron RA, Wiens JJ. 2011. A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Molecular Phylogenetics and Evolution* **61**:543–583 DOI 10.1016/j.ympev.2011.06.012.

Quattrini AM, McCartin LJ, Easton EE, Horowitz J, Wirshing HH, Bowers H, Mitchell K, González-García MdelP, Sei M, McFadden CS, Herrera S. 2024. Skimming genomes for systematics and DNA barcodes of corals. *Ecology and Evolution* **14**:e11254 DOI 10.1002/ece3.11254.

Rabosky DL, Chang J, Title PO, Cowman PF, Sallan L, Friedman M, Kaschner K, Garilao C, Near TJ, Coll M, Alfaro ME. 2018. An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**:392–395 DOI 10.1038/s41586-018-0273-1.

Rolfe S, Pieper S, Porto A, Diamond K, Winchester J, Shan S, Kirveslahti H, Boyer D, Summers A, Maga AM. 2021. SlicerMorph: an open and extensible platform to retrieve, visualize and analyse 3D morphology. *Methods in Ecology and Evolution* **12**:1816–1825 DOI 10.1111/2041-210x.13669.

Ruane S, Austin CC. 2017. Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. *Molecular Ecology Resources* **17**:1003–1008 DOI 10.1111/1755-0998.12655.

Russell P. 1803. *Descriptions and figures of two hundred fishes: collected at Vizagapatam on the coast of Coromandel.* London: W. Bulmer and Co., Shakspeare Press.

Sabaj MH. 2020. Codes for natural history collections in ichthyology and herpetology. *Copeia* **108**:593–669 DOI 10.1643/asihcodons2020.

Sato Y, Miya M, Fukunaga T, Sado T, Iwasaki W. 2018. MitoFish and MiFish pipeline: a mitochondrial genome database of fish with an analysis pipeline for environmental DNA metabarcoding. *Molecular Biology and Evolution* **35**:1553–1555 DOI 10.1093/molbev/msy074.

Silva PC, Malabarba MC, Vari R, Malabarba LR. 2019. Comparison and optimization for DNA extraction of archived fish specimens. *MethodsX* **6**:1433–1442 DOI 10.1016/j.mex.2019.06.001.

Smith WL, Craig MT. 2007. Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid Fishes. *Copeia* **2007**:35–55 DOI 10.1643/0045-8511(2007)7[35:ctpnwt]2.0.co;2.

Smith WL, Ghedotti MJ, Domínguez-Domínguez O, McMahan CD, Espinoza E, Martin RP, Girard MG, Davis MP. 2022. Investigations into the ancestry of the Grape-eye Seabass (*Hemilutjanus macrophthalmos*) reveal novel limits and relationships for the Acropomatiformes (Teleostei: Percomorpha). *Neotropical Ichthyology* **20**:e210160 DOI 10.1590/1982-0224-2021-0160.

Smith WL, Girard MG, Walker Jr HJ, Davis MP. 2024. The phylogeny of bristlemouths, lightfishes, and portholefishes with a revised family-level classification of the dragonfishes (Teleostei: Stomiiformes). *NOAA Professional Paper NMFS* **24**:167–184 DOI 10.7755/pp.24.13.

Speer KA, Hawkins MTR, Flores MFC, McGowen MR, Fleischer RC, Maldonado JE, Campana MG, Muletz-Wolz CR. 2022. A comparative study of RNA yields from museum specimens, including an optimized protocol for extracting RNA from formalin-fixed specimens. *Frontiers in Ecology and Evolution* **10**:953131 DOI 10.3389/fevo.2022.953131.

Stein JL, Marsh TL, Wu KY, Shizuya H, De Long EF. 1996. Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *Journal of Bacteriology* **178**:591–599 DOI 10.1128/jb.178.3.591-599.1996.

Stiller J, Feng S, Chowdhury A-A, Rivas-González I, Duchêne DA, Fang Q, Deng Y, Kozlov A, Stamatakis A, Claramunt S, Nguyen JMT, Ho SYW, Faircloth BC, Haag J, Houde P, Cracraft J, Balaban M, Mai U, Chen G, Gao R, Zhou C, Xie Y, Huang Z, Cao Z, Yan Z, Ogilvie HA, Nakhleh L, Lindow B, Morel B, Fjeldså J, Hosner PA,

da FonsecaRR, Petersen B, Tobias JA, Székely T, Kennedy JD, Reeve AH, Liker A, Stervander M, Antunes A, Tietze DT, Bertelsen MF, Lei F, Rahbek C, Graves GR, Schierup MH, Warnow T, Braun EL, Gilbert MTP, Jarvis ED, Mirarab S, Zhang G. 2024. Complexity of avian evolution revealed by family-level genomes. *Nature* **629**:851–860 DOI 10.1038/s41586-024-07323-1.

Straube N, Lyra ML, Paijmans JLA, Preick M, Basler N, Penner J, Rödel M, Westbury MV, Haddad CFB, Barlow A, Hofreiter M. 2021. Successful application of ancient DNA extraction and library construction protocols to museum wet collection specimens. *Molecular Ecology Resources* **21**:2299–2315 DOI 10.1111/1755-0998.13433.

Sullivan JP, Hopkins CD, Pirro S, Peterson R, Chakona A, Mutizwa TI, Mulelenu CM, Alqahtani FH, Vreven E, Dillman CB. 2022. Mitogenome recovered from a 19th Century holotype by shotgun sequencing supplies a generic name for an orphaned clade of African weakly electric fishes (Osteoglossomorpha, Mormyridae). *ZooKeys* **1129**:163–196 DOI 10.3897/zookeys.1129.90287.

Talavera G, Lukhtanov V, Pierce NE, Vila R. 2021. DNA barcodes combined with multilocus data of representative taxa can generate reliable higher-level phylogenies. *Systematic Biology* **71**:392–395 DOI 10.1093/sysbio/syab038.

Zacho CM, Bager MA, Margaryan A, Gravlund P, Galatius A, Rasmussen AR, Allentoft ME. 2021. Uncovering the genomic and metagenomic research potential in old ethanol-preserved snakes. *PLOS ONE* **16**:e0256353 DOI 10.1371/journal.pone.0256353.

Zhang J. 2010. Exploiting formalin-preserved fish specimens for resources of DNA barcoding. *Molecular Ecology Resources* **10**:935–941 DOI 10.1111/j.1755-0998.2010.2838.x.

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 DOI 10.1093/molbev/msad035.