Received: 21 August 2021 | Accepted: 12 January 2022

DOI: 10.1111/jfb.14997

REGULAR PAPER

JOURNAL OF FISH BIOLOGY fsbi

Molecular phylogeny of the threadfin fishes (Polynemidae) using ultraconserved elements

Matthew G. Girard[1,2,3] | Matthew P. Davis[4] | Carole C. Baldwin[3] | Agnès Dettaï[5] | Rene P. Martin[1,2] | W. Leo Smith[1,2]

[1]Biodiversity Institute, University of Kansas, Lawrence, Kansas, USA

[2]Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas, USA

[3]Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, District of Columbia, USA

[4]Department of Biological Sciences, St. Cloud State University, St. Cloud, Minnesota, USA

[5]Département Systématique et Evolution, Muséum National d'Histoire Naturelle, Paris, FRA

Correspondence
Matthew G. Girard, Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, 10th & Constitution Ave. NW, Washington District of Columbia, 20560 USA.
Email: girardmg@si.edu

Funding information
American Museum of Natural History Lerner-Gray Grant for Marine Research; National Science Foundation (US), Grant/Award Numbers: DEB 1258141, DEB 1543654; University of Kansas, Grant/Award Number: 2105077; University of Kansas Biodiversity Institute Panorama Grant

Abstract

Threadfins (Teleostei: Polynemidae) are a group of fishes named for their elongated and threadlike pectoral-fin rays. These fishes are commonly found in the world's tropical and subtropical waters, and are an economically important group for people living in these regions, with more than 100,000 t harvested in recent years. However, we do not have a detailed understanding of polynemid evolutionary history such that these fishes can be monitored, managed and conserved as an important tropical food source. Recent studies hypothesize at least one genus of threadfins is polyphyletic, and no studies have focused on generating a hypothesis of relationship for the Polynemidae using DNA sequences. In this study, we analyse a genomic dataset of ultraconserved-element and mitochondrial loci to construct a phylogeny of the Polynemidae. We recover the threadfins as a clade sister to flatfishes, with the most taxonomically rich genus, *Polydactylus*, being resolved as polyphyletic. When comparing our dataset to data from previous studies, we find that a few recent broad-scale phylogenies of fishes have incorporated mislabelled, misidentified or chimeric terminals into their analyses, impacting the relationships of threadfins they recover. We highlight these problematic sequences, providing revised identifications based on the data sequenced in this study. We then discuss the intrarelationships of threadfins, highlighting morphological or ecological characters that support the clades we recover.

KEYWORDS
BOLD, COI, *Eleutheronema*, GenBank, paradise threadfins, *Polydactylus*

1 | INTRODUCTION

The family Polynemidae comprises fishes commonly called threadfins for their numerous, elongated and threadlike pectoral-fin rays (Figure 1). Threadfins are thought to use their filamentous fin rays to assist in the detection of food in areas with low visibility, as these fishes splay their fin rays away from the body while swimming near the bottom of their environment (Motomura *et al*., 2002). The recent study by Presti *et al*. (2020) identified tastebuds on the surface of these filaments, further supporting the hypothesis that threadfin fishes use their fin rays to detect food. Polynemids are found in many of the world's tropical and subtropical environments, with the majority of species inhabiting coastal areas within the Indo-Pacific (Motomura, 2004). Some species also exclusively inhabit freshwater environments (*e.g*., species of *Polynemus*), with other threadfins able to transition between environments of differing salinity levels (*e.g*., the fourfinger threadfin *Eleutheronema tetradactylum* (Shaw 1804), the paradise threadfin *Polynemus paradiseus* L.; David, 1954; Motomura, 2004). In the areas where

polynemids occur, these fishes are often highly sought after and form the basis for commercial, recreational and subsistence-based fishing. More than 100,000 t of threadfins were harvested across the world in 2019, in addition to roughly 12,000 t that were harvested through aquaculture (FAO, 2021). West African species of threadfins, including the lesser African threadfin *Galeoides decadactylus* (Bloch 1795), the royal threadfin *Pentanemus quinquarius* (L.) and the giant African threadfin *Polydactylus quadrifilis* (Cuvier 1829), make up the largest recorded proportions of polynemids harvested through fisheries in the world (~27,000, ~40,000 and ~41,000 t, respectively), with *E. tetradactylum* being the most heavily aquacultured species of threadfins (FAO, 2021). The management of polynemids is becoming increasingly important due to the commercial significance of these fishes, the world-wide concern of overfishing critical fish stocks and the difficulties associated with gathering fisheries data in many of the areas where threadfins are harvested (FAO, 2020). However, we lack a detailed understanding of their phylogeny and taxonomy, which is critical to integrating data on evolution and species richness with management and conservation efforts (Faith, 1992; Rolland *et al*., 2012).

The placement of the Polynemidae among teleosts has varied over the years, with morphology-based hypotheses allying the threadfins with percesocine fishes (Atheriniformes, Mugilidae, Sphyraenidae; Gosline, 1962, 1971; McAllister, 1968; Regan, 1912; Rosen, 1964) or with the drums and croakers (Sciaenidae; Freihofer, 1978; Johnson, 1993; Kang *et al*., 2017). In contrast, DNA-based hypotheses have recovered threadfins within a clade that includes jacks (Carangidae) and billfishes (Xiphioidea; *e.g*., Betancur-R *et al*., 2013a,b, 2017; Sanciangco *et al*., 2016; Smith *et al*., 2016), with some of the most recent studies recovering the flatfishes (Pleuronectoideo *sensu* Girard *et al*., 2020) as the sister group to threadfins (*e.g*., Alfaro *et al*., 2018; Harrington *et al*., 2016). The recent study by Girard *et al*. (2020) that combined morphological and genomic data similarly recovered a flatfish-threadfin sister group, thus analyses using genomic data, as well as analyses integrating genomic and morphological data, are converging on the placement of polynemids as sister to the flatfishes in the Carangiformes.

Within the Polynemidae, species of threadfins are classified among eight genera, with half of the family's species richness (21 of the 42 species) belonging to a single genus, *Polydactylus*. The remaining taxa are classified in four monotypic genera (*Galeoides*, *Leptomelanosoma*, *Parapolynemus* and *Pentanemus*) and three genera with fewer than 10 species each (*Eleutheronema*, *Filimanus* and *Polynemus*). The family has been the subject of many taxonomic studies (*e.g*., Feltes, 1991, 1993; Motomura & Iwatsuki, 2001a,b) that have provided additional information (*e.g*., description of the genus *Leptomelanosoma*; Motomura & Iwatsuki, 2001a) and impacted the current classification of threadfins (*e.g*., revision of *Filimanus*; Feltes, 1991). Four studies have focused on generating phylogenies for the threadfins (Feltes, 1986; Girard, 2021; Kang, 2017; Presti, 2019), but are currently unpublished masters theses or doctoral dissertations. The studies by Feltes (1986), Kang (2017) and Presti (2019) sampled hard- and soft-tissue characters, while the study by Girard (2021) sampled phenotypic and genotypic traits, some of which are included in the current study. While all four of these studies recover the most species-rich genus of threadfins, *Polydactylus*, as polyphyletic, the most recent three studies are only publicly available as abstracts with limited details on their hypotheses. In addition to these four works, studies focused on generating barcode data, mitogenomes or the overall relationships of fishes have, on occasion, sampled multiple threadfin taxa using DNA-sequence characters (*e.g*., Betancur-R *et al*., 2013a,b; Gopalakrishnan *et al*., 2021; Rabosky *et al*., 2018; Sanciangco *et al*., 2016) or used a combined-character approach (*e.g*., Mirande, 2016). Although not the focus of their analysis, the phylogeny of the Polynemidae received some attention by Rabosky *et al*. (2018), who sampled 13 species of threadfins (~30% of species) and recovered the genera *Eleutheronema* and *Polydactylus* as polyphyletic, but there are concerns about the data used in their analysis that we will discuss below.

In light of the possible nonmonophyly of at least one polynemid genus and the lack of a focused study testing the intrarelationships of threadfins using DNA-sequence characters, this study has three goals:

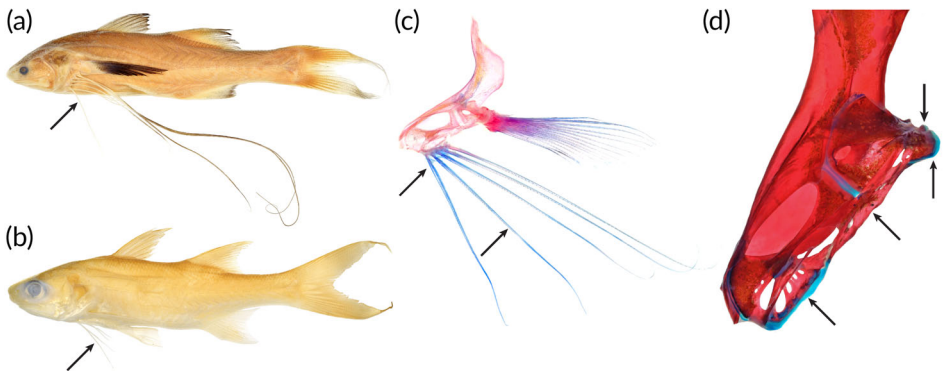

FIGURE 1 Images of whole-ethanol and cleared-and-stained specimens under white or daylight LED light. (a) Elongate, threadlike fin rays of the pectoral fin in *Polynemus melanochir* (UMMZ 245535), arrow, lateral view. (b) Elongate, threadlike fin rays of the pectoral fin in *Polydactylus quadrifilis* (OS 20526), arrow, lateral view. (c) Elongate, threadlike fin rays of the pectoral fin in *Polydactylus plebeius* (ANSP 172791), arrow, lateral view. (d) Pectoral radial osteology in *Polydactylus microstoma* (SU 38934), uppermost arrow indicates pectoral radials one and two, middle arrow indicates ventrally elongate third pectoral radial, lowermost arrow indicates fourth pectoral radial

(1) generate a genomic dataset of ultraconserved elements (UCEs) and cytochrome oxidase subunit 1 (COI) loci to construct a phylogeny of the threadfins; (2) test the monophyly of threadfin genera; and (3) resolve the relationships among the genera and clades when genera are resolved as para- or polyphyletic.

2 | MATERIALS AND METHODS

2.1 | Taxon sampling

We generated two datasets in this study. The first dataset, named the '32-terminal' dataset, consists of UCE data from 32 taxa, including 12 outgroup taxa and 20 of the 42 species (~47%) of the Polynemidae. To include as many polynemid species as possible, we constructed a second dataset, named the '42-terminal' dataset, that included all of the taxa and data from the 32-terminal dataset as well as an additional 10 species of the Polynemidae (30 of 42 species; ~71%) represented only by COI sequence data (see Tables 1 and 2, and Files S1, S2 and S3). Outgroup taxa included lineages that have been traditionally and more recently allied with the Polynemidae (*e.g.*, Alfaro *et al*., 2018; Harrington *et al*., 2016; Kang *et al*., 2017; Rosen, 1964). We included representatives from the Bedotiidae, Centrarchidae, Latidae, Mugilidae, Pleuronectidae, Psettodidae, Sciaenidae, Scombridae, Scophthalmidae and Sphyraenidae. The Atlantic mackerel *Scomber scombrus* L. served as the root for both datasets. Lists of taxa used in this study can be found in Table 1, with the localities of polynemid tissue samples shown in Figure 2. Symbolic codes for institutional resource collections follow Sabaj (2020) and common names follow Froese & Pauly (2021).

2.2 | DNA extraction, locus amplification and sequence alignment

Muscle or fin clips were preserved in 95% ethanol, RNA*later* Stabilization Solution (Invitrogen, Carlsbad, CA, USA) or frozen (fresh) prior to the extraction of DNA. We extracted DNA from tissue samples using either a DNeasy Tissue Extraction Kit (Qiagen, Germantown, MD, USA) or a Maxwell RSC Blood DNA Kit and Instrument (Promega, Madison, WI, USA) following manufacturer's extraction protocols, with the exception of replacing the Blood DNA Kit's lysis buffer with Promega's tissue lysis buffer. For extractions performed with a Qiagen DNeasy Kit only, the first or first and second elution from a Qiagen filter were combined and dried to a volume of 102 μL using a Savant DNA120 SpeedVac Concentrator (Thermo Scientific, Waltham, MA, USA). When eluted multiple times, the elutions were combined to increase the amount of DNA collected per extraction. For extractions performed with the Maxwell RSC only, extractions were eluted into a volume of 102 μL. For both types of preparations, 2 μL of the DNA extract was quantified using a Qubit Fluorometer 2.0 (Invitrogen) with the Qubit dsDNA BR Assay Kit (Invitrogen). Final quantified samples (100 μL in volume) were sent to Arbor Biosciences (Ann Arbor, MI, USA) for library preparation (*e.g.*, DNA shearing, size selection, clean-up), target capture, enrichment, sequencing on an Illumina HiSeq 2500 or NovaSeq 6000 (Illumina, San Diego, CA, USA) and demultiplexing. The 500 UCE actinopterygian-loci probe set (Faircloth *et al*., 2013) was used for target capture of UCE loci.

Demultiplexed sequence data from one or multiple runs were received in compressed FASTQ format from Arbor Biosciences. These data were uncompressed and combined into two read files per taxon. The data were then cleaned of adapter contamination and low-quality bases using the parallel wrapper illumiprocessor version 2.10 (Faircloth, 2013) around trimmomatic version 0.39 (Bolger *et al*., 2014). Cleaned sequencing reads were submitted to GenBank and have been assigned SRA Accession Numbers SRR17616018–SRR17616038 under BioProject PRJNA796495 (Table 1). The cleaned reads were combined with previously published UCE data obtained from Alfaro *et al*. (2018; BioProject Accession Number PRJNA348720; Table 1), Girard *et al*. (2020; BioProject Accession Number PRJNA604383; Table 1) and Harrington *et al*. (2016; BioProject Accession Number PRJNA341709; Table 1) for assembly. All clean reads were assembled using SPAdes version 3.14.1 (Prjibelski *et al*., 2020) under the default settings using the python script assemblo_spades.py within PHYLUCE version 1.7.0 (Faircloth *et al*., 2012; Faircloth, 2016). Taxon-specific contigs were identified within these loci by aligning and assembling the contigs into a relational file containing all probes using the python script match_contigs_to_loci.py within PHYLUCE and LASTZ version 1.0.4 (Harris, 2007). The minimum coverage and minimum identity for identifying UCE loci were set to 80%. The PHYLUCE script get_match_counts.py was used to search the relational database and generate a list of UCE loci shared among all taxa. This list was used as input for the PHYLUCE script get_fastas_from_match_counts.py to generate a single FASTA file containing all UCE sequence data for all taxa. Once constructed, the data in this file were divided by locus using explode_get_fastas_file.py within PHYLUCE and aligned using MAFFT version 7 (Katoh & Standley, 2013). Alignments that were minimally 65% complete at the level of individual locus were converted into PHYLIP-format files and prepared for downstream analyses.

For the 42-terminal dataset, COI gene fragments were sequenced anew or extracted from high-throughput cleaned sequencing reads. These sequences allowed for both identity verification of species sampled in our dataset and for the inclusion of polynemid taxa that have been sequenced using Sanger-based sequencing but were unavailable for high-throughput sequencing. For *Galeoides decadactylus*, the COI locus was obtained using the sequencing and editing protocol outlined in Weght *et al*. (2012). For the long-limb threadfin *Polydactylus longipes* Motomura Okamoto & Iwatsuki 2001, the COI locus was obtained using the sequencing and editing protocol outlined in Dettaï *et al*. (2011). Sequences from these two taxa were placed into a FASTA file to be combined with the other COI sequences. The remaining novel COI sequences analysed in this study were obtained by extracting gene fragments from high-throughput cleaned sequencing reads. To extract COI, cleaned reads were compared to existing sequence data from previously Sanger-sequenced COI loci of taxa within the same family or genus using the 'map to reference' function in Geneious version 11.1.5 (Kearse *et al*., 2012). Previously Sanger-

TABLE 1 Voucher information, GenBank accession numbers and statistics for UCE loci used in this study

Family	Species	Voucher for UCE dataset	Tissue for UCE dataset	SRA accession numbers	UCE loci collected	Total bps	Mean contig length (bps)	Median contig length (bps)	Minimum contig length (bps)	Maximum contig length (bps)	Contigs over 1000 bps	95% confidence interval	Voucher for COI dataset	Tissue for COI dataset	COI accession numbers	Locality number in Figure 2
Centrarchidae	*Lepomis cyanellus*	JFBM, uncat.	See voucher	SRR11016327	459	720,576	1569.88	1588	209	2934	442	12.19	Extracted from SRA	See voucher	–	–
Latidae	*Lates calcarifer*	AMNH 233713	See voucher	SRR11016328	457	668,734	1463.31	1432	606	3083	430	15.18	Extracted from SRA	See voucher	–	–
Leptobramidae	*Leptobrama muelleri*	H. Larson, pers. col.	See voucher	SRR11016326	449	633,923	1411.86	1429	369	3160	420	12.86	Extracted from SRA	See voucher	–	–
Melanotaeniidae	*Rheocles wrightae*	AMNH, uncat.	See voucher	SRR17616027	464	675,908	1456.70	1464	360	2535	447	10.61	AMNH, uncat.	See voucher	AY290803	–
Mugilidae	*Mugil curema*	AMNH, uncat.	See voucher	SRR11016345	463	621,267	1341.83	1343	530	2681	421	11.84	Extracted from SRA	See voucher	–	–
Pleuronectidae	*Lyopsetta exilis*	KUI 23754	KUIT 9346	SRR17616023	402	550,898	1370.39	1385	300	2641	349	16.67	Extracted from SRA	See voucher	–	–
Psettodidae	*Psettodes erumei*	LSUMZ 16779	LSUMZ F5295	SRR5237312	443	358,272	808.74	798	230	1668	52	8.65	Extracted from SRA	See voucher	–	–
Sciaenidae	*Micropogonias undulatus*	YPM ICH 023546	YFTC 021711	SRR4432445	433	335,302	774.37	754	239	2136	41	9.45	Extracted from SRA	See voucher	–	–
Sciaenidae	*Pareques acuminatus*	CAS PAC20	See voucher	SRR4432394	408	521,886	1279.13	1268	122	5672	343	19.42	Extracted from SRA	See voucher	–	–
Scombridae	*Scomber scombrus*	AMNH, uncat.	See voucher	SRR11016337	463	567,911	1226.59	1273	369	2013	391	10.83	Extracted from SRA	See voucher	–	–
Scophthalmidae	*Scophthalmus aquosus*	KUI 27107	KUIT 1253	SRR11016335	459	410,965	895.35	915	309	2115	138	9.34	Extracted from SRA	See voucher	–	–
Sphyraenidae	*Sphyraena barracuda*	UW 158204	UW ROA18221	SRR17616025	391	556,981	1424.50	1366	316	3212	348	18.91	Extracted from SRA	See voucher	–	–
Polynemidae	*Eleutheronema tetradactylum*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	Unavailable	Unavailable	MF281357	–
Polynemidae	*Eleutheronema rhadinum*	KUI 41529	KUIT 10585	SRR17616038	459	779,632	1698.54	1687	177	4236	443	19.63	Extracted from SRA	See voucher	–	16
Polynemidae	*Filimanus perplexa*	CSIRO, uncat.	CSIRO BW-A5778	SRR17616037	426	554,861	1302.49	1280	174	3289	354	16.85	Extracted from SRA	See voucher	–	14
Polynemidae	*Filimanus sealei*	USNM 408795	See voucher	SRR17616026	423	459,962	1087.38	1083	318	2208	304	11.82	Extracted from SRA	See voucher	–	18
Polynemidae	*Filimanus similis*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	Unavailable	Unavailable	MF281368	–
Polynemidae	*Filimanus xanthonema*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	CSIRO, uncat.	CSIRO BW-A9040	HQ564487	15
Polynemidae	*Galeoides decadactylus*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	USNM 405159	See voucher	–	7
Polynemidae	*Leptomelanosoma indicum*	USNM 444119	See voucher	SRR17616024	380	401,754	1057.25	1043	178	2404	229	14.68	Extracted from SRA	See voucher	–	13
Polynemidae	*Parapolynemus verekeri*	MAGNT S.17530–002	MAGNT A01566	SRR17616022	341	311,985	914.91	883	182	2268	93	14.51	Extracted from SRA	See voucher	–	20
Polynemidae	*Pentanemus quinquarius*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	Unavailable	Unavailable	LC484875	–
Polynemidae	*Polydactylus approximans*	SIO 98–176	See voucher	SRR17616021	460	599,217	1302.65	1341	222	2957	394	14.70	Extracted from SRA	See voucher	–	2
Polynemidae	*'Polydactylus' longipes*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	MNHN 2008–1644	See voucher	–	23
Polynemidae	*'Polydactylus' macrochir*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	CSIRO H4872-06	CSIRO BW-A1210	FOAC211-05	–
Polynemidae	*'Polydactylus' malagasyensis*	SIO 04–68	See voucher	SRR17616020	457	610,464	1335.81	1369	314	3596	386	16.35	Extracted from SRA	See voucher	–	9
Polynemidae	*'Polydactylus' microstoma*	USNM 408834	See voucher	SRR17616019	432	457,013	1057.90	1068	256	2326	288	12.11	Extracted from SRA	See voucher	–	19

TABLE 1 (Continued)

Family	Species	Voucher for UCE dataset	Tissue for UCE dataset	SRA accession numbers	UCE loci collected	Total bps	Mean contig length (bps)	Median contig length (bps)	Minimum contig length (bps)	Maximum contig length (bps)	Contigs over 1000 bps	95% confidence interval	Voucher for COI dataset	Tissue for COI dataset	COI accession numbers	Locality number in Figure 2
Polynemidae	*'Polydactylus' mullani*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	Unavailable	Unavailable	MF281374	–
Polynemidae	*'Polydactylus' multiradiatus*	MAGNT S.17508–007	MAGNT A01579	SRR17616018	189	62,633	331.39	307	80	833	0	8.29	Extracted from SRA	See voucher	–	21
Polynemidae	*'Polydactylus' nigripinnis*	MAGNT S.17552–001	MAGNT A02593	SRR17616036	375	268,318	715.51	703	142	1692	22	10.00	Extracted from SRA	See voucher	–	22
Polynemidae	*Polydactylus octonemus*	KUI 30108	KUIT 5105	SRR17616035	340	499,136	1468.05	1517	193	2879	309	19.37	Extracted from SRA	See voucher	–	3
Polynemidae	*Polydactylus oligodon*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	LBP 40523	See voucher	JQ365495	6
Polynemidae	*Polydactylus opercularis*	FMNH 143546	See voucher	SRR17616034	417	573,851	1376.14	1379	282	3397	367	16.81	Extracted from SRA	See voucher	–	5
Polynemidae	*'Polydactylus' plebeius*	USNM 439841	See voucher	SRR17616033	359	405,555	1129.68	1103	162	2749	279	15.63	Extracted from SRA	See voucher	–	1
Polynemidae	*'Polydactylus' quadrifilis*	OS 20961	OS GAB17-966	SRR17616032	393	432,874	1101.46	1080	227	2458	295	12.75	Extracted from SRA	See voucher	–	8
Polynemidae	*'Polydactylus' sexfilis*	SAIAB 77939	KUIT 6829	SRR5237313	446	543,916	1219.54	1204	333	4232	366	16.50	Extracted from SRA	See voucher	–	10
Polynemidae	*'Polydactylus' sextarius*	KUI 41535	KUIT 10667	SRR17616031	450	662,940	1473.20	1503	304	2916	403	18.16	Extracted from SRA	See voucher	–	17
Polynemidae	*Polydactylus virginicus*	USNM 349227	KUIT 9031	SRR11016339	464	738,121	1590.78	1624	424	3298	436	17.31	Extracted from SRA	See voucher	–	4
Polynemidae	*Polynemus aquilonaris*	ANSP 177984 (tag 1064)	ANSP t5607	SRR17616030	413	649,183	1571.87	1539	230	3589	394	19.42	Extracted from SRA	See voucher	–	12
Polynemidae	*Polynemus melanochir*	SIO 12–55	See voucher	SRR17616029	462	610,602	1321.65	1320	298	2751	415	14.04	Extracted from SRA	See voucher	–	–
Polynemidae	*Polynemus multifilis*	ANSP 177983 (tag 1099)	ANSP t5605	SRR17616028	438	683,248	1559.93	1590	199	2937	403	18.19	Extracted from SRA	See voucher	–	11
Polynemidae	*Polynemus paradiseus*	Unavailable	Unavailable	–	–	–	–	–	–	–	–	–	Unavailable	Unavailable	NC_026236	–

Note: Numbers in right-most column correspond to sampling locations in Figure 2. See Supporting Information File S4 for information about data partitioning.

TABLE 2 Previously generated mislabelled or misidentified DNA sequences for members of the Polynemidae

Original identification	GenBank number(s)	BOLD number(s)	Publication generating sequence	Publications(s) using sequence(s)	Revised species identification	Reasoning
Unverified *Eleutheronema rhadinum*	KX345093	ANGBF39349-19	–	Gopalakrishnan *et al*., 2021	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	EU595103	FSCS549-07	Zhang and Hanner, 2012	–	*Eleutheronema rhadinum*	§, Ω
Eleutheronema tetradactylum	EU595104	FSCS548-07	Zhang and Hanner, 2012	–	*Eleutheronema rhadinum*	§, Ω
Eleutheronema tetradactylum	EU595105	FSCS547-07	Zhang and Hanner, 2012	–	*Eleutheronema rhadinum*	§, Ω
Eleutheronema tetradactylum	EU595106	FSCS546-07	Zhang and Hanner, 2012	–	*Eleutheronema rhadinum*	§, Ω
Eleutheronema tetradactylum	EU595107	FSCS545-07	Zhang and Hanner, 2012	–	*Eleutheronema rhadinum*	§, Ω
Eleutheronema tetradactylum	EU595108	FSCS544-07	Zhang and Hanner, 2012	Rabosky *et al*., 2018	*Eleutheronema rhadinum*	§, Ω
Eleutheronema tetradactylum	FJ237985	FSCS767-08	Zhang and Hanner, 2012	–	*Eleutheronema rhadinum*	§, Ω
Eleutheronema tetradactylum	KC878730	GBMNA14573-19	–	–	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	KM401448	GBGCA9588-15	–	–	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	KM401449	GBGCA9587-15	–	–	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	KM401450	GBGCA9586-15	–	–	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	KT593869	GBMNA17106-19	Wang *et al*., 2016	–	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	MW388968	GBMND28778-21	–	–	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	MW388969	GBMND28779-21	–	–	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	MW388970	GBMND28780-21	–	–	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	MW388971	GBMND28781-21	–	–	*Eleutheronema rhadinum*	Ω
Eleutheronema tetradactylum	NC_021620	GBMTG3244-16	–	Mirande, 2016, Wang *et al*., 2016; Zhong *et al*., 2021	*Eleutheronema rhadinum*	Ω
Eleutheronema sp.	KX399435	ANGBF39360-19	–	–	*Eleutheronema tetradactylum*	§, Ω
Eleutheronema sp.	KX399436	GBMIN131352-17	–	–	*Eleutheronema tetradactylum*	§, Ω
Eleutheronema sp.	KX399438	GBMIN131353-17	–	–	*Eleutheronema tetradactylum*	§, Ω
Eleutheronema rhadinum	KU944081	ZOSKT774-16	Chang *et al*. 2017	–	*Eleutheronema tetradactylum*	Ω
Eleutheronema rhadinum	MW845829	–	Zhong *et al*. 2021	–	*Eleutheronema tetradactylum*	Ω
Eleutheronema rhadinum	–	FTW805-09	–	–	*Eleutheronema tetradactylum*	Ω
Eleutheronema rhadinum	–	FTW806-09	–	–	*Eleutheronema tetradactylum*	Ω

TABLE 2 (Continued)

Original identification	GenBank number(s)	BOLD number(s)	Publication generating sequence	Publications(s) using sequence(s)	Revised species identification	Reasoning
Polydactylus sextarius	JX983438	DBFN300-12	Khedkar *et al.* 2014	Gopalakrishnan *et al.*, 2021	*Eleutheronema tetradactylum*	§, Ω
Polydactylus sextarius	JX983439	DBFN290-12	Khedkar *et al.* 2014	–	*Eleutheronema tetradactylum*	§, Ω
Polydactylus sextarius	JX983440	DBFN291-12	Khedkar *et al.* 2014	Gopalakrishnan *et al.*, 2021	*Eleutheronema tetradactylum*	§, Ω
Filimanus heptadactyla	EF609523	WLIND395-07	Lakra *et al.*, 2011	–	*'Polydactylus' mullani**	†, §
Filimanus heptadactyla	EF609524	WLIND394-07	Lakra *et al.*, 2011	–	*'Polydactylus' mullani**	†, §
Filimanus heptadactyla	EF609525	WLIND393-07	Lakra *et al.*, 2011	–	*'Polydactylus' mullani**	†, §
Filimanus heptadactyla	EF609526	WLIND392-07	Lakra *et al.*, 2011	Rabosky *et al.*, 2018	*'Polydactylus' mullani**	†, §
Galeoides decadactylus	–	SAFW226-08	–	–	*'Polydactylus' quadrifilis*	†, §, Ω
Leptomelanosoma indicum	JQ937604, JQ937739, JQ937887, JQ938005, JQ938170, JQ938554, JQ938702, JQ939029, JQ939164, JQ939316, JQ939447, JQ939557, JQ939860, JQ940022	–	Betancur-R *et al.* 2013a	Mirande, 2016	*'Polydactylus' sexfilis*	†
Leptomelanosoma indicum	KC825655, KC826111, KC826836, KC827697, KC829172, KC830081, KC830378, KC830613, KC831367	–	Betancur-R *et al.* 2013b	Betancur-R *et al. 2017*; Mirande, 2016; Sanciangco *et al.*, 2016; Rabosky *et al.*, 2018	*'Polydactylus' sexfilis*	†
Polydactylus sexfilis	KY371977	SCS1319-16	–	–	*'Polydactylus' sextarius*	§, Ω
Polydactylus sexfilis	KY371978	SCS1318-16	–	–	*'Polydactylus' sextarius*	§, Ω
Polydactylus sexfilis	KY371979	SCS1320-16	–	–	*'Polydactylus' sextarius*	§, Ω
Polydactylus virginicus	–	BAHIA402-14	–	–	*Polydactylus oligodon*	§, Ω
Polydactylus virginicus	–	BAHIA403-14	–	–	*Polydactylus oligodon*	§, Ω
Polydactylus virginicus	JQ365495	MFSP365-10	Ribeiro *et al.*, 2012	Betancur-R *et al.*, 2013b, 2017; Gopalakrishnan *et al.*, 2021; Sanciangco *et al.*, 2016; Rabosky *et al.*, 2018	*Polydactylus oligodon*	§, Ω
Polydactylus virginicus	JQ365496	MFSP375-10	Ribeiro *et al.*, 2012	Gopalakrishnan *et al.*, 2021	*Polydactylus oligodon*	§, Ω
Polydactylus virginicus	KY402361	ANGBF39419-19	–	–	*Polydactylus oligodon*	Ω
Polydactylus virginicus	KY402362	ANGBF39420-19	–	–	*Polydactylus oligodon*	Ω
Polydactylus virginicus	KY402363	ANGBF39421-19	–	–	*Polydactylus oligodon*	Ω
Polydactylus virginicus	KY402364	ANGBF39422-19	–	–	*Polydactylus oligodon*	Ω
Polynemus dubius	MH721199	ANGBF54928-19	–	–	*Polynemus aquilonaris*	§, Ω

(Continues)

TABLE 2 (Continued)

Original identification	GenBank number(s)	BOLD number(s)	Publication generating sequence	Publications(s) using sequence(s)	Revised species identification	Reasoning
Polynemus dubius	NC_029710	ANGBF39423-19	–	Zhong *et al*., 2021	*Polynemus aquilonaris*	Ω
Polynemus dubius	KU199001	GBMNA14576-19	–	–	*Polynemus aquilonaris*	Ω

Note: Table limited to sequences without tentative identifications. Taxa with revised identifications followed by an asterisk (*) were also highlighted by Gopalakrishnan *et al*. (2021). Symbols in the right-most column indicate if the revised identity was based on examining the voucher or photo of the voucher (†), location from which the voucher was captured (§) and/or comparisons of sequence identity between those previously generated and those generated in this study (Ω). See text and Supporting Information Files S1, S2 and S3 for additional sequences and information.

sequenced COI loci were used as the 'reference sequence' and cleaned sequencing reads were mapped to the reference using the Geneious mapping algorithm with 'sensitivity' set to 'medium-low sensitivity' and 'fine tuning' set to 'iterate up to 10 times'. Homologous regions collected were inspected for and cleaned of ambiguities using Geneious. All COI sequences that were obtained from high-throughput reads and analysed in this study can be found in Supporting Information File S1. These COI sequences were collated into a single file with previously sequenced loci from published and unpublished studies, including Li *et al*. (2020), Shihab *et al*. (2017) and Sparks & Smith (2004; see Table 1 for GenBank Accession Numbers). Collated sequences aligned with MAFFT version 7 (Katoh & Standley, 2013) within Geneious and exported as a PHYLIP-format file for downstream analyses.

2.3 | Identification of mislabelled or misidentified threadfin sequences

As DNA-sequence datasets for phylogenetic inference become larger both in number of taxa and number of base pairs (bps) sampled, we have an opportunity to not only incorporate previously generated data into these datasets, but to also examine and identify sequences that are mislabelled, misidentified or are otherwise anomalous. While building the COI dataset and comparing the samples sequenced in this study to sequence data for threadfins databased and distributed through BOLD and/or GenBank, a number of previously generated sequences were found to be misidentified or mislabelled. These came from a variety of published (Betancur-R *et al*., 2013a,b; Chang *et al*., 2017; Horne *et al*., 2011; Khedkar *et al*., 2014; Lakra *et al*., 2011; Ribeiro *et al*., 2012; Shihab *et al*., 2017; Thu *et al*., 2019; Qu *et al*., 2020; Zhong *et al*., 2021) and unpublished works. The recent study by Gopalakrishnan *et al*. (2021) similarly highlighted mislabelled or misidentified samples of threadfins native to India. However, their study also incorporated anomalous sequences (see Supporting Information File S2), emphasizing the need to revise the identity of misleading sequences to study polynemids at various taxonomic levels. To identify and provide revised identifications for these sequences, COI data housed on BOLD or GenBank (released prior to 22 September 2021) were compared to the data generated from vouchered specimens in this study and the information provided in the works by Feltes and Motomura and colleagues (*e.g*., Feltes, 1986; Motomura, 2004; Motomura & Iwatsuki, 2001b). When the data housed on BOLD or Genbank could not be compared to sequences from vouchered specimens, voucher collection information was examined and GenBank's Basic Local Alignment Search Tool (BLAST) was used to assist in identification. A list of sequences that are mislabelled or misidentified along with revised identifications can be found in Table 2. A more complete table, including sequences that could only be tentatively identified, can be found in Supporting Information File S2. Identifiers are also included in Table 2 and Supporting Information File S2 that indicate if a revised identity was based on examining the voucher or photo of the voucher (†), location the voucher was captured from (§) and/or comparisons of sequence identity between those previously generated and those generated in this study (Ω). A more thorough discussion of identification revisions can be found in Supporting Information File S3.

2.4 | Partitioning schemes and nucleotide substitution models

A total of 450 aligned UCE loci were analysed in both analyses of this study. Mean sequence fragment length was approximately 1235 bps, with a range of 80–5672 bps (Table 1) across all UCE loci. The 450 UCE loci were concatenated into a single matrix that was minimally 65% complete at the level of individual loci. The final matrix was 514,958 bps in length, ~69% complete at the level of individual bps, with 90,856 parsimony-informative sites. This matrix was partitioned using the sliding-window site characteristics–entropy method (hereafter, SWSC-EN; Tagliacollo & Lanfear, 2018) to split each UCE locus into left and right flanking regions and the ultraconserved core (*i.e*., center segment). The resulting left, central and right segments from SWSC-EN were used as the input for PartitionFinder version 2.1.1 (Lanfear *et al*., 2014, 2017; Stamatakis, 2014) to find the best-fitting nucleotide substitution model for each data partition. The following parameters were used in PartitionFinder: branchlengths set to linked; models set to GTR, GTR + G and GTR + I + G; model_selection set to AICc; schemes search set to rclusterf; command line option raxml. PartitionFinder designated 555 subsets with associated models for these regions. For the 42-terminal dataset, the COI matrix was analysed in addition to the UCE matrix. The COI matrix, which was 656 bps in length (~98% complete) and contained 269 parsimony-informative sites,

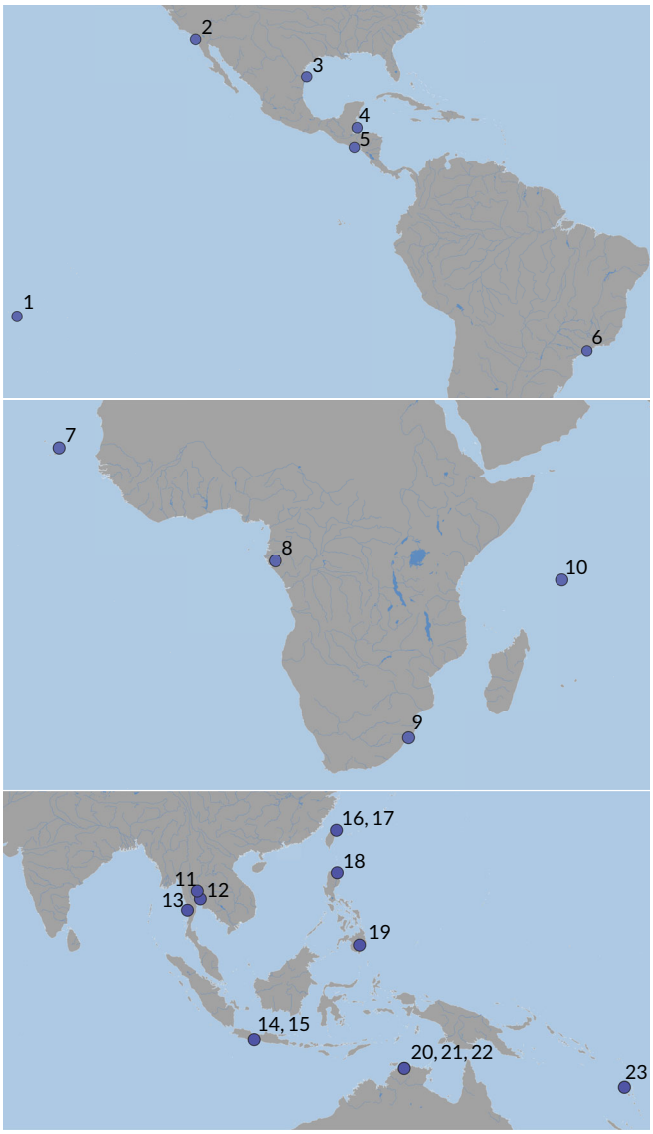

FIGURE 2 Collection localities of polynemid tissue samples. Map shown was generated in QGIS (QGIS Development Team, 2020) using the ocean, rivers + lake centerlines, lakes + reservoirs and land vector polygons from naturalearthdata.com. Numbers next to points correspond with numbers associated with the polynemid species sampled and are also listed in Table 1: 1, '*Polydactylus*' *plebeius*; 2, *Polydactylus approximans*; 3, *Polydactylus octonemus*; 4, *Polydactylus virginicus*; 5, *Polydactylus opercularis*; 6, *Polydactylus oligodon*; 7, *Galeoides decadactylus*; 8, '*Polydactylus*' *quadrifilis*; 9, '*Polydactylus*' *malagasyensis*; 10, '*Polydactylus*' *sexfilis*; 11, *Polynemus multifilis*; 12, *Polynemus aquilonaris*; 13, *Leptomelanosoma indicum*; 14, *Filimanus perplexa*; 15, *Filimanus xanthonema*; 16, *Eleutheronema rhadinum*; 17, '*Polydactylus*' *sextarius*; 18, *Filimanus sealei*; 19, '*Polydactylus*' *microstoma*; 20, *Parapolynemus verekeri*; 21, '*Polydactylus*' *multiradiatus*; 22,'*Polydactylus*' *nigripinnis*; 23, '*Polydactylus*' *longipes*

was broken into three partitions: one partition designated for each of the three codon positions in the protein-coding locus. These three partitions were also used as input for PartitionFinder version 2.1.1 using the same settings as above. PartitionFinder designated three subsets with associated models. A list of partitions and associated models for both UCE and COI loci can be found in Supporting Information File S4.

2.5 | Analysis of DNA data matrices

Following the assembly and partitioning of the matrices, the 32-terminal and 42-terminal datasets were analysed using IQ-Tree version 2.1.2 (Chernomor *et al*., 2016; Minh *et al*., 2020). Three analyses were conducted, a concatenated and a species-tree analysis for the 32-terminal dataset and a concatenated analysis for the 42-terminal dataset. For the 32-terminal concatenated analysis, the UCE matrix and an independent partition model file were used as inputs for IQ-Tree. For the 42-terminal concatenated analysis, the UCE matrix, COI matrix and an independent partition model file were used as inputs for IQ-Tree. The total number of partitions used in the 32-terminal concatenated analysis was 555 and the total number of partitions for the 42-terminal concatenated analysis was 558 (555 from the UCE matrix). Both concatenated analyses were performed by 20 independent executions of IQ-Tree with the perturbation strength (−pers) set to 0.2 and the number of unsuccessful iterations to stop (−nstop) set to 2000. Once these 20 analyses were completed, a second set of analyses were started using the resulting trees from the first 20 analyses as starting trees with the following settings: more thorough nearest-neighbour interchange search (−allnni), use starting tree (−t), number of trees in the candidate set to maintain during tree search (−nbest) to 25 and number of unsuccessful iterations to stop (−nstop) set to 5000. Support for the best-fitting topology of each dataset was generated using 200 standard bootstrap replicates (−bo) and reconciled with the most likely phylogeny using IQ-Tree (−con; Figures 3 and 4). For the 32-terminal species-tree analysis, individual UCE-locus trees were generated using IQ-Tree from individual UCE-locus NEXUS files obtained from the get_only_loci_with_min_taxa script of PHYLUCE. Individual locus files were analysed with the number of unsuccessful iterations to stop (−nstop) set to 1000 and without a designated model file, allowing IQ-Tree to find the best-fitting model for each locus using its ModelFinder function (Kalyaanamoorthy *et al*., 2017). A total of 450 best-fitting trees were consolidated and used as input data for ASTRAL-III version 5.7.7 (Zhang *et al*., 2018), resulting in a species tree and quadripartition supports for each branch (Figure 4). All resulting phylogenetic trees were visualized with FigTree version 1.4.4 (Rambaut, 2009).

3 | RESULTS

3.1 | Identification of mislabelled or misidentified DNA sequences

We provide revised identifications for 70 sequences of polynemids currently available on BOLD and/or GenBank (Table 2). These revisions were based on a combination of examining

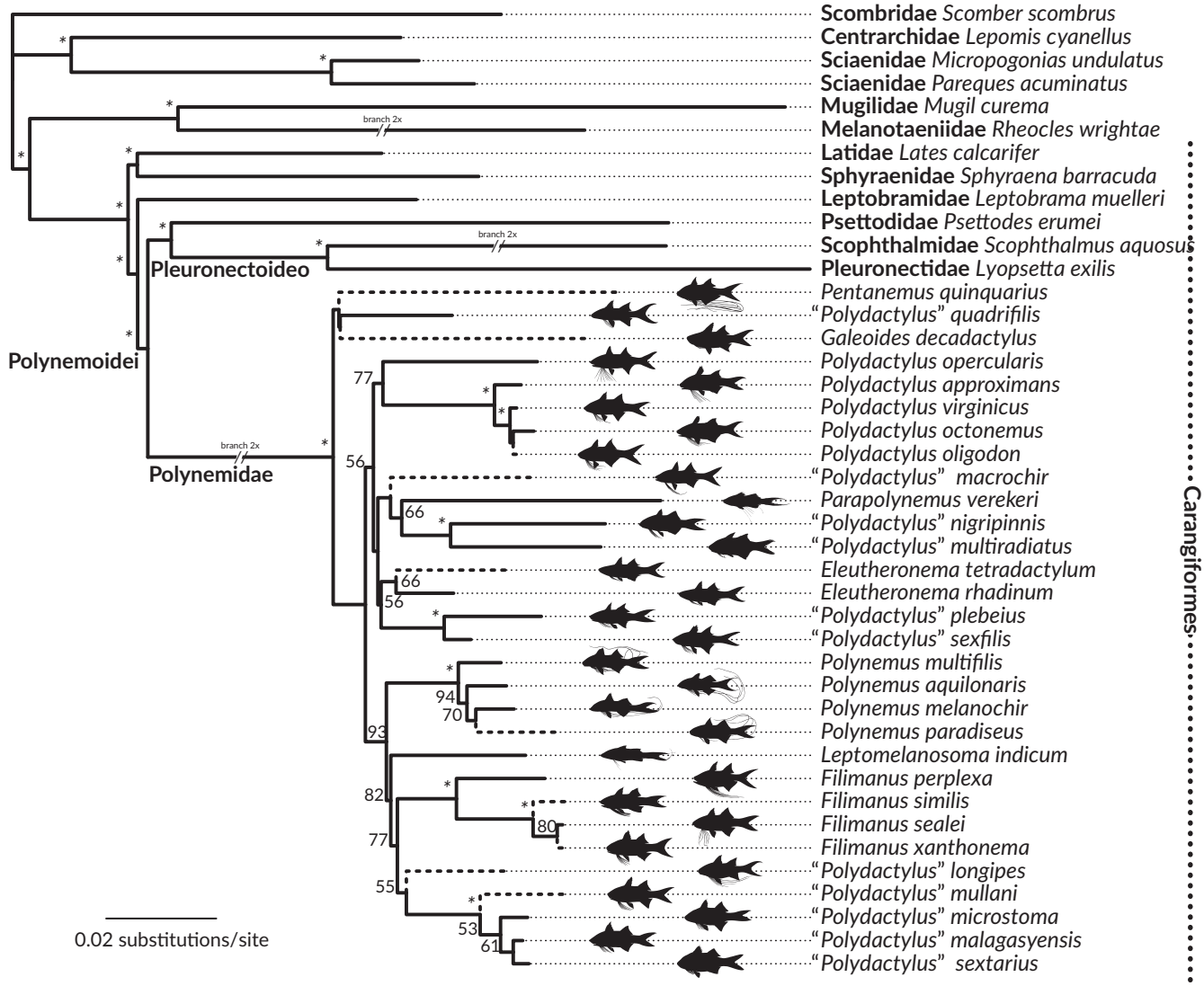

FIGURE 3 Hypotheses of relationships from partitioned likelihood analysis of the Polynemidae and outgroup taxa. Dataset included 450 ultraconserved element loci and one mitochondrial locus. Support was determined based on 200 bootstrap replicates. Clades with 50% or higher bootstrap support are noted on the nodes with their bootstrap percentage or an asterisk if support was 95% or more. Dashed branches indicate terminals that were represented only by mitochondrial data. Species of *Polydactylus* that fall outside of the monophyletic group that includes the type species of the genus are listed as '*Polydactylus*'

vouchers or photographs of the vouchers, locality information and/or comparisons of sequence identity. Additionally, we provide revised and tentative identifications for sequences currently identified as one species of *Filimanus* and one species of *Polydactylus*. Finally, we highlight more than 500 sequences belonging to lineages within the genus *Eleutheronema* that are genetically different at the COI locus to the sequences of the East Asian fourfinger threadfin *E. rhadinum* (Jordan & Everman 1902) and *E. tetradactylum* (see Supporting Information Files S2 and S3). As a majority of these sequences lack an associated museum-catalogued voucher, it is unclear if these sequences belong to species that are currently described (*i.e.*, the threefinger threadfin *E. tridactylum* (Bleeker 1849)), represent population-level structure rather than distinct species (Hillis *et al*., 2021; Sukumaran & Knowles, 2017) or if there are yet-to-be-described taxa within the family. However, as these sequences were most similar to sequences of *Eleutheronema rhadinum*, we labelled the sequences '*Eleutheronema cf. rhadinum*' followed by a geographic identifier (see Supporting Information Files S2 and S3).

3.2 | Phylogeny of the Polynemidae

The hypothesis of relationships recovered from all analyses are shown in Figures 3 and 4. The bootstrap values for the 32-terminal concatenated analysis yielded 29 nodes (of 29, 100%) with a bootstrap value of ≥80% and 27 nodes (~93%) with a bootstrap value of ≥95% (Figure 4). The species tree of the 32-terminal dataset yielded 28 nodes (of 29, ~96%) with a quadripartition support value of ≥80% and 26 nodes (~89%) with a support value of ≥95% (Figure 4). The

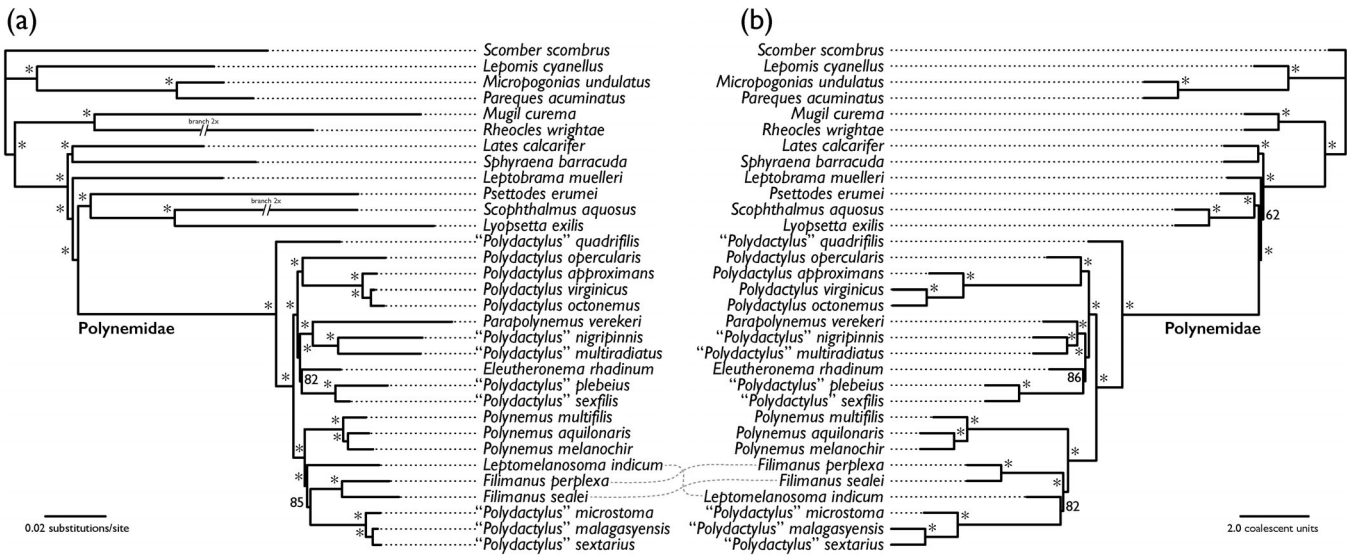

FIGURE 4 Hypotheses of relationships from concatenated and species-tree analyses of the 32-terminal dataset. Dashed lines drawn between adjacent phylogenies are to highlight taxa that are in different phylogenetic positions between the concatenated and species-tree analysis. Species of *Polydactylus* that fall outside of the monophyletic group that includes the type species of the genus are listed as '*Polydactylus*'. (a) Result of concatenated analysis of 450 ultraconserved-element loci. Support was determined based on 200 bootstrap replicates. Clades with 50% or higher bootstrap support are noted on the nodes with their bootstrap percentage or an asterisk if support was 95% or more. (b) Result of species-tree analysis of 450 ultraconserved-element loci. Branch lengths for terminal branches all identical (see ASTRAL). Support was determined based on quadripartition support. Clades with 50% or higher bootstrap support are noted on the nodes with their bootstrap percentage or an asterisk if support was 95% or more

bootstrap values for the 42-terminal concatenated analysis yielded 23 nodes (of 39, ~58%) with a bootstrap value of ≥80% and 19 nodes (~48%) with a bootstrap value of ≥95% (Figure 3). All nodes outside of the Polynemidae were supported with ≥98% bootstrap support in both concatenated analyses. In the species-tree analysis, the only node outside of the Polynemidae with lower than 99% quadripartition support was the clade including threadfins plus flatfishes (Polynemoidei *sensu* Girard *et al*., 2020) sister to the Leptobramidae (Figure 4) that had a support value of 62. Within the Polynemidae, fewer nodes were highly supported in the bootstrap analysis of the 42-terminal dataset. The lower support values in the 42-terminal dataset are likely due to the species only being represented by COI being easily perturbed and recovered in disparate locations (Figure 3, dashed terminals). While all 40 analyses of the 42-terminal dataset recovered the same topology with differing branch lengths, sampling a greater number of bps for the taxa represented by only COI in this study will undoubtedly produce a phylogeny with greater support values (*e.g*., Harrington *et al*., 2016; Martin *et al*., 2018; Roa-Varón *et al*., 2020).

The resulting topology from both the 32-terminal and 42-terminal analyses showed a monophyletic Polynemidae sister to the Pleuronectoideo and within a clade of carangiform fishes. The only differences between the concatenated and species tree 32-terminal analyses were the placement of *Leptomelanosoma* and *Filimanus*. *Leptomelanosoma* was recovered sister to a clade of *Filimanus* and three species of *Polydactylus* in the concatenated analysis while *Filimanus* was recovered sister to a clade of *Leptomelanosoma* and three species of *Polydactylus* in the species-tree analysis (Figure 4, dashed lines). As the relationships across the different concatenated analyses were the same for taxa that occurred in both datasets, as well as the greater number of taxa sampled in the 42-terminal dataset, we will be focusing on the relationships recovered from the 42-terminal analysis. The earliest diverging clade of polynemids comprises three species, including *Pentanemus quinquarius* as sister to a clade of *Galeoides decadactylus* and *Polydactylus quadrifilis* (Figure 3). Sister to the clade of *Galeoides*, *Pentanemus* and *Polydactylus quadrifilis* threadfins are recovered in two groups. Within the first clade, the yellow bobo *Polydactylus opercularis* (Gill 1863) is recovered sister to the blue bobo *P. approximans* (Lay & Bennett 1839), the Atlantic threadfin *P. octonemus* (Girard 1858), the littlescale threadfin *P. oligodon* (Günther 1860) and the barbu *P. virginicus* (L.). *Polydactylus approximans* is recovered sister to a clade of *P. octonemus*, *P. oligodon* and *P. virginicus*, where *P. octonemus* and *P. oligodon* are recovered as sister taxa. This clade of *Polydactylus approximans*, *P. octonemus*, *P. oligodon*, *P. opercularis* and *P. virginicus* is recovered sister to a clade of *Eleutheronema rhadinum*, *E. tetradactylum*, the dwarf paradise fish *Parapolynemus verekeri* (Saville-Kent 1889), the king threadfin *Polydactylus macrochir* (Günther 1867), the Australian threadfin *P. multiradiatus* (Günther 1860), the blackfin threadfin *P. nigripinnis* Munro 1964, the striped threadfin *P. plebeius* (Broussonet 1782) and the sixfinger threadfin *P. sexfilis* (Valenciennes 1831). Within this clade, *Polydactylus macrochir* is recovered sister to a grade of three taxa, including *Parapolynemus verekeri*, *Polydactylus multiradiatus* and *P. nigripinnis*. *Parapolynemus verekeri* is recovered sister to the clade of *Polydactylus multiradiatus* and *P. nigripinnis*. This clade of *Parapolynemus verekeri*, *Polydactylus macrochir*, *P. multiradiatus* and

P. nigripinnis is recovered sister to a clade of *Eleutheronema* and *Polydactylus*. Within this clade, *Eleutheronema tetradactylum* is recovered as the sister species to *Eleutheronema rhadinum*. Sister to the clade of *Eleutheronema*, we recovered a clade of *Polydactylus plebeius* sister to *P. sexfilis*. The second clade of threadfins includes all sampled species of *Filimanus* and *Polynemus*, as well as the Indian threadfin *Leptomelanosoma indicum* (Shaw 1804), *Polydactylus longipes*, the African blackspot threadfin *P. malagasyensis* Motomura & Iwatsuki 2001b, the smallmouth threadfin *P. microstoma* (Bleeker 1851), the Arabian blackspot threadfin *P. mullani* (Hora 1926) and the blackspot threadfin *P. sextarius* (Bloch & Schneider 1801). A monophyletic *Polynemus* is recovered as the earliest-diverging lineage of this clade. Within *Polynemus*, the elegant paradise fish *P. multifilis* Temminck & Schlegel 1843 diverges first and is sister to all other species sampled. The northern paradise fish *Polynemus aquilonaris* Motomura 2003 is recovered sister to the blackhand paradise fish *P. melanochir* Valenciennes 1831 and *P. paradiseus*. *Polynemus* is recovered sister to a clade comprising all species of *Filimanus* sampled, *Leptomelanosoma indicum*, *Polydactylus longipes*, *P. malagasyensis*, *P. microstoma*, *P. mullani* and *P. sextarius*. Within this clade, *Leptomelanosoma indicum* is recovered as the earliest-diverging lineage. The next clade to diverge includes all sampled species of *Filimanus*. The splendid threadfin *Filimanus perplexa* Feltes 1991 is recovered sister to a grade of the eightfinger threadfin *F. sealei* (Jordan & Richardson 1910), the Indian sevenfinger threadfin *F. similis* Feltes 1991 and the yellowthread threadfin *F. xanthonema* (Valenciennes 1831). *Filimanus sealei* and *F. xanthonema* are recovered as sister taxa. Sister to the clade of *Filimanus* is a clade composed of *Polydactylus longipes*, *P. malagasyensis*, *P. microstoma*, *P. mullani* and *P. sextarius*. *Polydactylus longipes* is recovered as the earliest diverging lineage of this clade. *Polydactylus mullani* is recovered sister to a clade of *P. malagasyensis*, *P. microstoma* and *P. sextarius*. *Polydactylus microstoma* is recovered sister to *P. malagasyensis* and *P. sextarius*. Considering the nonmonophyly of *Polydactylus*, the subsequent discussion will refer to species of *Polydactylus* that fall outside of the clade that includes the type species of the genus as '*Polydactylus*'.

4 | DISCUSSION

4.1 | Interrelationships of the Polynemidae

Through the analysis of both UCE and mitochondrial loci, we recovered the Polynemidae sister to members of the Pleuronectoideo in the carangiform suborder Polynemoidei. Although we sampled members of the Bedotiidae, Mugilidae, Sciaenidae and Sphyraenidae, we did not recover a sister-group relationship between the Polynemidae and these traditional percesocine or percoid groups as has been suggested in previous morphology-based studies. We recovered a similar placement of the Polynemidae to those in recent DNA-based (*e.g.*, Alfaro *et al.*, 2018; Harrington *et al.*, 2016) and morphology-and-DNA-based studies (Girard, 2021; Girard *et al.*, 2020) that have recovered flatfishes as the sister group of the threadfins. The studies by Girard (2021) and Girard *et al.* (2020) have provided several morphological characters that support a polynemid–pleuronectoid relationship and also have highlighted that some characters supporting alternative placements of the Polynemidae are homoplastic.

4.2 | Intrarelationships of the Polynemidae

Within the Polynemidae, we recovered all but one genus, *Polydactylus*, as monophyletic. While the recent study by Rabosky *et al.* (2018) recovered *Eleutheronema* as polyphyletic, our examination of previously generated sequence data for the Polynemidae uncovered a number of sequences mislabelled or misidentified, including those used in the study by Rabosky *et al.* (2018). The sequences used for *Eleutheronema tetradactylum*, *Filimanus heptadactyla*, *Leptomelanosoma indicum* and *Polydactylus virginicus* sampled in Rabosky *et al.* (2018) are from more than one threadfin species, making these analysed terminals chimeric. We have revised the identification of COI sequences for *Eleutheronema tetradactylum*, *Filimanus heptadactyla* and *Polydactylus virginicus* used in their study to *Eleutheronema rhadinum*, '*Polydactylus*' *mullani* and *Polydactylus oligodon*, respectively (see Table 2, and Supporting Information Files S2 and S3). However, the COI sequence used in the study by Rabosky *et al.* (2018) for *Leptomelanosoma indicum* is correctly identified, but the sequences from additional loci, which came from the studies by Betancur-R *et al.* (2013a,b), are misidentified. Herein, we recognize these misidentified sequences as '*Polydactylus*' *sexfilis* (see Table 2, and Supporting Information Files S2 and S3). Similar chimeric terminals for *Eleutheronema tetradactylum*, *Leptomelanosoma indicum* and '*Polydactylus*' *sextarius* were also used in the study by Mirande (2016; see Table 2 and Supporting Information File S2), with an additional sequence, GenBank Accession Number GU440471, used for both *Polydactylus approximans* and '*P.*' *plebeius* in his analyses (see appendix S2 of Mirande, 2016). Ultimately, it is unclear how these chimeric terminals have impacted the relationships recovered in previous DNA-based (*e.g.*, Betancur-R *et al.*, 2013a,b; Rabosky *et al.*, 2018; Sanciangco *et al.*, 2016) or combined studies (*e.g.*, Mirande, 2016), and our subsequent discussions will reference these topologies on a limited basis. Across both morphology- and DNA-based studies (*e.g.*, Feltes, 1986; Kang, 2017; Mirande, 2016; Rabosky *et al.*, 2018; Sanciangco *et al.*, 2016), all analyses have recovered a polyphyletic *Polydactylus*, as we do in this study. Some of the clades we recovered are congruent with groups of *Polydactylus* highlighted in previous works by Feltes and Motomura and colleagues (*e.g.*, Feltes, 1986; Motomura & Iwatsuki, 2001b).

4.2.1 | Black-spotted species of '*Polydactylus*'

One of the clades of '*Polydactylus*' we recovered comprises threadfin taxa that all possess a large, black, anterior lateral-line spot (Motomura & Iwatsuki, 2001b). This clade includes '*Polydactylus*' *malagasyensis*, '*P.*' *microstoma*, '*P.*' *mullani* and '*P.*' *sextarius* (see Figures 3

and 4). We expect one additional taxon, the Persian blackspot threadfin *P. persicus* Motomura & Iwatsuki 2001b to also be included in this clade, but this taxon was not available to be sampled in our study. Feltes (1986) suggested a close relationship between two species of black-spotted '*Polydactylus*', '*P.*' *microstoma* and '*P.*' *sextarius*, based on a lack of vomerine teeth, placement of the first haemal arch on the fifth vertebra and the shape of the posterior margin of the maxilla. Taking note of the lateral-line pigmentation in these members of '*Polydactylus*', Motomura & Iwatsuki (2001b) reviewed all members of the genus that exhibit this pigmentation pattern (see above). They noted several characters unique to these black-spotted '*Polydactylus*' in addition to those mentioned by Feltes (1986), including that all membranous pectoral-fin rays are branched except for the uppermost one or two fin rays and a trend toward a reduction in gas-bladder size across the group. Our study corroborates their finding that species of '*Polydactylus*' with an anterior lateral-line spot form a clade supported by the morphological characters listed by Feltes (1986) and Motomura & Iwatsuki (2001b). Based on the photographs of gas bladders provided by Motomura & Iwatsuki (2001b, figure 7), the topology of black-spotted threadfins we recovered also corroborates a trend toward a reduction in gas-bladder size within this clade. '*Polydactylus*' *mullani* was recovered as the earliest-diverging species in this clade, a taxon that possesses the most elongated gas bladder among black-spotted '*Polydactylus*' (see Motomura & Iwatsuki, 2001b, figure 7). Additionally, '*Polydactylus*' *sextarius*, which has the shortest and most-atrophied gas bladder among black-spotted '*Polydactylus*', is recovered in a derived position (see Figure 3). Future studies should look to include the remaining species of black-spotted '*Polydactylus*', '*P.*' *persicus*, and further investigate the trend in reduction of gas-bladder size within this clade of threadfins further.

4.2.2 | New World species of *Polydactylus*

The name-bearing species of *Polydactylus* is *P. plumierii* Lacepède 1803, which is a junior synonym of *P. virginicus* (Motomura, 2004). We recovered *Polydactylus virginicus* within a clade of threadfins that occur along the coasts of North and South America (Figures 3 and 4), including *Polydactylus approximans*, *P. octonemus*, *P. oligodon* and *P. opercularis*. Two of these taxa, *Polydactylus approximans* and *P. opercularis*, co-occur in largely overlapping ranges along the coasts of the eastern Pacific, both being found between California and Peru (Motomura, 2004). The remaining New World species of *Polydactylus* are found within the western Atlantic, with *P. octonemus* occurring from New York to the west coast of the Yucatán, *P. oligodon* from Florida to Brazil and *P. virginicus* occurring from New Jersey to Brazil (Motomura, 2004). Kang (2017) recovered a similar clade of New World *Polydactylus*, which included *P. approximans*, *P. octonemus*, *P. oligodon* and *P. virginicus*. Additionally, Sanciangco *et al*. (2016) and Rabosky *et al*. (2018) recovered similar clades of *Polydactylus* in their studies. In addition to their closely positioned distributions, *Polydactylus approximans*, *P. octonemus*, *P. oligodon* and *P. virginicus* possess a great deal of external morphological similarity. Among these four taxa, the lateral-line scales extend onto both the upper and lower lobes of the caudal fin, bifurcating slightly posterior to the proximal aspect of the caudal-fin rays (Motomura, 2004). Bifurcation of the lateral line on the caudal fin has been documented in only one other member of *Polydactylus*, the slender fivefinger threadfin *P. bifurcus* Motomura, Kimura & Iwatsuki 2001 (Motomura, 2004), which was not examined in this study. Other members of *Polydactylus*, such as black-spotted '*Polydactylus*', generally have pored lateral-line scales that extend exclusively onto the lower lobe of the caudal fin. While threadfins in other genera can have a single, straight extension (*Filimanus*, *Pentanemus* and *Polynemus*) or have an extension on just the upper lobe of the caudal fin (*Galeoides*), with some species of *Eleutheronema* exhibiting an additional bifurcation on the lower lobe of the caudal fin (Motomura, 2004). However, the remaining species of *Polydactylus* that occurs in the New World, *P. opercularis*, does not exhibit this bifurcated lateral-line morphology and is recovered sister to all other New World species of *Polydactylus* in our study. The placement of *Polydactylus opercularis* differs between our study and the topology presented by Kang (2017), where it was recovered as the sister taxon to species of *Eleutheronema*. While geographic overlap exists for species in the clade we recovered, it is unclear if morphological characters are present that support the relationship between all species of New World *Polydactylus*.

4.2.3 | Striped species of '*Polydactylus*'

A clade comprising '*Polydactylus*' *plebeius* and '*P.*' *sexfilis* was recovered in this study as well as the topology by Kang (2017). There are multiple similarities in the overall appearance and meristics of '*Polydactylus*' *plebeius* and '*P.*' *sexfilis*, including similar body depth at first dorsal-fin origin relative to standard length (25%–34% and 27%–34%), upper-jaw length relative to standard length (13%–16% for both), number of pored lateral-line scales (60–68 and 60–67) and the presence of dark longitudinal stripes along the flank scale rows (Motomura, 2004). While not included in our study, an additional species of '*Polydactylus*', the largemouth striped threadfin '*P.*' *siamensis* Motomura, Iwatsuki & Yoshino 2001, is similar in overall appearance and meristics to '*P.*' *plebeius* and '*P.*' *sexfilis*. We would expect '*Polydactylus*' *siamensis* to be closely related to '*P.*' *plebeius* and '*P.*' *sexfilis* when it is included in subsequent studies. Additionally, the recent study by Gopalakrishnan *et al*. (2021) highlighted a potentially new species of threadfin that is allied with striped species of '*Polydactylus*' (their *Polydactylus* sp. A). However, the authors did not provide additional descriptive information about the voucher of this sample in their study. Further investigation is needed to determine if the lineage highlighted by Gopalakrishnan *et al*. (2021) is a new species and if it is morphologically similar to striped species of '*Polydactylus*'.

4.2.4 | Remaining species of '*Polydactylus*'

While the three previously mentioned clades of species currently in the genus *Polydactylus* have been investigated or recovered in

previous studies, we recovered the remaining species within the genus sampled in this study in disparate locations throughout the phylogeny of polynemids (Figures 3 and 4). This is the first time '*Polydactylus*' *quadrifilis* has been sampled in a phylogeny, and we recovered the taxon sister to *Galeoides* and within the earliest diverging clade of threadfins along with *Pentanemus*. Little is known about the internal anatomy of '*Polydactylus*' *quadrifilis* or if any characters are present within this taxon that support a sister-group relationship with *Galeoides*. Three of the remaining species of '*Polydactylus*' included in this study, '*P.*' *macrochir*, '*P.*' *multiradiatus* and '*P.*' *nigripinnis*, have been recovered in different phylogenetic locations in previous studies. Feltes (1986) recovered '*Polydactylus*' *nigripinnis* sister to *Galeoides* and supported by one character state: shape of coracoid at ventral contact with cleithrum broad, with moderate ventral expansion and a reduction of the curve along the anterior aspect (his character state 39E). Kang (2017) recovered '*Polydactylus*' *macrochir*, '*P.*' *multiradiatus* and '*P.*' *nigripinnis* in a polytomy with all other species of *Eleutheronema*, *Galeoides*, *Leptomelanosoma* and '*Polydactylus*' sampled. In this study, we recovered '*Polydactylus*' *macrochir* as the earliest-diverging lineage of a clade that also includes *Parapolynemus*, '*Polydactylus*' *multiradiatus* and '*P.*' *nigripinnis*. This clade is somewhat surprising, as *Parapolynemus* has been recovered as the sister group to all species of *Polynemus* in previous studies (*e.g.*, Feltes, 1986; Kang, 2017). However, the relationship between '*Polydactylus*' *multiradiatus* and '*P.*' *nigripinnis* may be supported by the absence of a gas bladder, which is only seen in one other species of *Polydactylus* (*P. opercularis*), in species of *Eleutheronema* and in some species of *Polynemus*; Motomura, 2004). Finally, this is the first time '*Polydactylus*' *longipes* has been included in a phylogenetic study, and we recover it as sister to the clade of black-spotted species of '*Polydactylus*'. Only a few specimens of this species have been collected and catalogued into museum collections, making it one of the rarest and least-studied species of threadfin. Subsequent investigations regarding the morphology of '*Polydactylus*' *longipes*, along with additional bps, are needed to investigate the relationship between this taxon and black-spotted species of '*Polydactylus*'.

4.2.5 | Relationships among threadfin genera

The remaining nonmonotypic genera of the Polynemidae (*Eleutheronema*, *Filimanus* and *Polynemus*) were recovered as monophyletic in our study. However, the placements of these genera differ between our study and previous hypotheses. Feltes (1986) found a sister-group relationship between *Filimanus* and *Pentanemus* supported by seven character states, including posterior margin of ventral arm of nasal without a foramen and with mesial indentation (his character 32C) and a high number of gill rakers along the lateral surface of the first gill arch (his character 34B). Alternatively, Kang (2017) recovered *Filimanus* as sister to a clade of *Eleutheronema*, *Galeoides*, *Leptomelanosoma* and *Polydactylus*. We recovered *Filimanus* sister to the black-spotted species of '*Polydactylus*'. This clade may be supported by the lack of vomerine teeth (Motomura, 2004), but there are other threadfins that lack teeth on the vomer (*i.e.*, *Galeoides*, *Parapolynemus*, *Pentanemus* and '*Polydactylus*' *nigripinnis*). We also found *Polynemus* sister to a clade of *Filimanus*, *Leptomelanosoma* and the black-spotted species of '*Polydactylus*' in our analysis. Feltes (1986) recovered a relationship between *Parapolynemus* and *Polynemus* supported by 11 character states, including absent or reduced basisphenoid (his character state 8D) and posterior margin of preopercle with few serrations (his character state 16B). This relationship was later corroborated by Kang (2017). We recovered *Polynemus* in a clade of morphologically diverse threadfins that possess wide variation in the lengths of the pectoral fins (*e.g.*, species of *Polynemus* possess some of the longest free pectoral-fin rays among threadfins (Motomura, 2004)), differences in body size and shapes of the gas bladder (see Motomura & Iwatsuki, 2001a, figure 7, 2001b, figure 1). Additional research is needed to investigate the relationships among *Filimanus*, *Leptomelanosoma*, *Polynemus* and the black-spotted species of '*Polydactylus*' we recovered here. Finally, the placement of *Eleutheronema* has varied across previous studies. Feltes (1986) recovered *Eleutheronema* sister to a clade of *Filimanus*, *Galeoides*, *Pentanemus* and *Polydactylus* that was supported by one character state, the presence of gaps or foramina and struts in the fourth pectoral radial. *Eleutheronema* has also been recovered in a polytomy with *Galeoides*, *Leptomelanosoma* and *Polydactylus* (Kang, 2017), sister to *Pentanemus* and '*Polydactylus*' *macrochir* (Rabosky *et al.*, 2018), and sister to one black-spotted species of '*Polydactylus*', '*P.*' *sextarius* (Sanciangco *et al.*, 2016). We recovered *Eleutheronema* sister to the striped species of '*Polydactylus*'. Despite these groups being externally dissimilar, both *Eleutheronema tetradactylum* and '*Polydactylus*' *sexfilis* exhibit a similar sequence of sex change through life, where these species first mature as males then become hermaphroditic for a period of time before becoming fully functional females (May *et al.*, 1979; Santerre & May, 1977; Stanger, 1974). Information on life-history traits for threadfins is limited to a few species (see Feltes, 1986), but some polynemids that have been investigated are found to be protandrous with a brief period of hermaphroditism (Motomura, 2004). Although additional studies are needed, this prolonged period of hermaphroditism may be diagnostic of *Eleutheronema* and the striped species of '*Polydactylus*' clade we recovered.

4.3 | The need to revise mislabelled or misidentified sequences on public databases

Accurate taxonomic identification is paramount to not only phylogenetic inference but ecological studies and management-related efforts. Excluding sequences that we could only tentatively identify, the 70 publicly available sequences we provide revised identifications for are listed as a different threadfin taxon, a taxon from another family of fishes or a sequence from a lineage outside of Vertebrata altogether. At least 46 of these sequences have been used in at least 10 subsequent phylogenetic studies and impacted their findings. We also found more than 500 sequences that have a questionable

identification, particularly sequences that belong to lineages within the genus *Eleutheronema*. Despite uncovering these mislabelled, misidentified or otherwise anomalous sequences, we lack a tractable way for these revisions to be noted on the publicly available GenBank or BOLD accession numbers. Recent studies by Tang *et al*. (2021) and Phillips *et al*. (2021) highlight similar sets of mislabelled or misidentified sequences housed in public repositories for damselfishes (Pomacanthidae) and the marsupial genus *Myoictis*, respectively. Phillips *et al*. (2021) highlighted the current inability to flag or otherwise denote anomalous sequences in public repositories outside of contacting and having the sequence(s)-submitting author(s) agree to make revisions. We echo the call to GenBank by Phillips *et al*. (2021) and make a similar call to BOLD to allow for a flexible protocol to flag or denote anomalous sequences. This is equally critical for both databases, particularly as BOLD incorporates mitochondrial sequences from GenBank and direct links between the two databases are available for some sequences. While such a revisionary protocol presents many challenges, such as the required level of verification needed to obtain a flag and how these flags are shown on the website, some of which are discussed with possible solutions by Phillips *et al*. (2021), such a system will only improve the utility of these databases and allow for the dissemination of the most accurate data possible.

5 | CONCLUSIONS

Our analyses found that seven of the eight genera of threadfins are monophyletic. In contrast, the most species-rich genus, *Polydactylus*, was recovered as polyphyletic, with some of clades potentially being united by morphological characters that have been commented on by previous researchers, such as the lack of vomerine teeth or gas bladders. We also recovered new relationships among the genera of threadfins, with the traditionally allied genera *Parapolynemus* and *Polynemus* not being recovered as sister groups. In our examination of publicly available sequences, we encountered a number of mislabelled or misidentified sequences as well as others that may represent yet-to-be-described species or population-level structures within a single taxon. As the effectiveness of species management, monitoring and conservation is dependent on understanding the current biodiversity present, subsequent studies should examine the lineages of *Eleutheronema* we have highlighted to determine if these are different species. Additionally, the COI or 'barcode' reference repositories are a tremendous resource for the identification of fishes. Although future studies will use high-throughput sequencing to sample a diversity of loci, we encourage subsequent investigators to generate COI sequences from the vouchered specimens used in their datasets and make these generated data available. These barcode data are critical to the reliable identification of specimens and allow for an *a posteriori* detection of mislabelled, misidentified or otherwise anomalous sequences when compared to the larger database. Finally, we encourage subsequent authors to perform similar comparisons between their newly generated data and those that have been previously generated to identify sequences that may be problematic, as has been shown here and in Phillips *et al*. (2021) and Tang *et al*. (2021). Without revision, these mislabelled or misidentified sequences housed in data repositories will continue to impact our ability to resolve evolutionary relationships from available data and may affect subsequent studies focused on ecological and evolutionary questions more broadly.

ACKNOWLEDGEMENTS

We especially thank R. Smetana and K. Smith for helpful discussion and editing of this manuscript. We thank R. Arrindell, B. Brown and J. Sparks (AMNH), M. Arce H. and M. Sabaj (ANSP), D. Catania and L. Rocha (CAS), A. Graham (CSIRO), C. McMahan, S. Mochel and K. Swagel (FMNH), A. Bentley (KU), G. Dally and M. Hammer (MAGNT), B. Sidlauskas (OSU), B. Frable, P. Hastings and H. Walker (SIO), D. Nelson (UMMZ), J. Clayton, G. D. Johnson, K. Murphy, D. Pitassy, S. Raredon and J. Williams (USNM) and K. Maslenikov (UW) for providing access to tissue samples from their collections. We thank A. Hay and S. Reader (AMS) for providing specimen counts and photographs of specimens of *Eleutheronema* in their care. We thank M. Santos (Philippine Bureau of Fisheries and Aquatic Resources–National Fisheries Research and Development Institute) and K. Carpenter (ODU) for providing the sample of *Filimanus sealei*. We thank S. Planes (Centre de Recherche Insulaire et Observatoire de l'Environnement) for providing the sample of '*Polydactylus*' *plebeius*. We thank the members of the Santo 2006 survey on N.O. Alis led by P. Bouchet (D. Hinsinger and R. Fricke for identification) for providing the sample of '*Polydactylus*' *longipes* and sequencing of this specimen was part of agreement 2005/67 between the Genoscope and the Museum National d'Histoire Naturelle on the project 'Macrophylogeny of life' directed by Guillaume Lecointre. Research was funded by a Biodiversity Institute Panorama Grant (awarded to M.G.G.), American Museum of Natural History Lerner-Gray Grant for Marine Research (awarded to M.G.G.), National Science Foundation grants (DEB 1258141 and 1543654 awarded to M.P.D. and W.L.S.) and University of Kansas General Research Fund (#2105077 awarded to W.L.S.). Analyses were conducted using the KU Community Cluster in the Center for Research Computing at the University of Kansas.

AUTHOR CONTRIBUTIONS

M.G.G.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, visualization, writing—original draft and writing—review & editing. M.P.D.: conceptualization, data curation, investigation, funding acquisition, methodology and writing—review & editing. C.C.B.: data curation, resources and writing—review & editing. A.D.: data curation, resources and writing—review & editing. R.P.M.: methodology and writing—review & editing. W.L.S.: conceptualization, data curation, investigation, funding acquisition, methodology, resources and writing—review & editing.

ORCID

Matthew G. Girard https://orcid.org/0000-0003-3580-6808
Matthew P. Davis https://orcid.org/0000-0001-5349-417X
Carole C. Baldwin https://orcid.org/0000-0002-2875-0474
Agnès Dettaï https://orcid.org/0000-0002-8496-9737

Rene P. Martin https://orcid.org/0000-0003-0153-7160
W. Leo Smith https://orcid.org/0000-0001-8710-6673

REFERENCES

Alfaro, M. E., Faircloth, B. C., Harrington, R. C., Sorenson, L., Friedman, M., Thacker, C. E., … Near, T. J. (2018). Explosive diversification of marine fishes at the cretaceous-Palaeogene boundary. *Nature Ecology and Evolution*, *2*, 688–696. https://doi.org/10.1038/s41559-018-0494-6.

Betancur-R, R., Broughton, R. E., Wiley, E. O., Carpenter, K., López, J. A., Li, C., … Ortí, G. (2013b). The tree of life and a new classification of bony fishes. *PLoS Currents*, *5*. https://doi.org/10.1371/currents.tol.53ba26640df0ccaee75bb165c8c26288.

Betancur-R, R., Li, C., Munroe, T. A., Ballesteros, J. A., & Ortí, G. (2013a). Addressing gene tree discordance and non-stationarity to resolve a multi-locus phylogeny of the flatfishes (Teleostei: Pleuronectiformes). *Systematic Biology*, *62*, 763–785. https://doi.org/10.1093/sysbio/syt039.

Betancur-R, R., Wiley, E. O., Arratia, G., Acero, A., Bailly, N., Miya, M., … Ortí, G. (2017). Phylogenetic classification of bony fishes. *BMC Evolutionary Biology*, *17*, 162. https://doi.org/10.11886/s12862-017-0958-3.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Chang, C.-H., Shao, K.-T., Lin, H.-Y., Chiu, Y.-C., Lee, M.-Y., Liu, S.-H., & Lin, P.-L. (2017). DNA barcodes of the native ray-finned fishes in Taiwan. *Molecular Ecology Resources*, *17*, 796–805. https://doi.org/10.1111/1755-0998.12601.

Chernomor, O., Haeseler, A. von, & Minh, B. Q. (2016). Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, *65*, 997–1008. doi: https://doi.org/10.1093/sysbio/syw037

David, A. (1954). A preliminary survey of the fish and fisheries of a 5-mile stretch of the Hooghly River near Barrackpore. *Indian Journal of Fisheries*, *1*, 231–250.

Dettaï, A., Lautredou, A.-C., Bonillo, C., Goimbault, E., Busson, F., Causse, R., Couloux, … & Ozouf-Costaz, C. (2011). The actinopterygian diversity of the CEAMARC cruises: Barcoding and molecular taxonomy as a multi level tool for new findings. *Deep Sea Research II. 58*, 250–263 doi: https://doi.org/10.1016/j.dsr2.2010.05.021

Faircloth, B. C. (2013). *Illumiprocessor: A Trimmomatic wrapper for parallel adapter and quality trimming*. doi: https://doi.org/10.6079/J9ILL

Faircloth, B. C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, *32*, 786–788. https://doi.org/10.1093/bioinformatics/btv646.

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.

Faircloth, B. C., Sorenson, L., Santini, F., & Alfaro, M. E. (2013). A phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of ultraconserved elements (UCEs). *PLoS One*, *8*, e65923. https://doi.org/10.1371/journal.pone.0065923.

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*, 1–10. https://doi.org/10.1016/0006-3207(92)91201-3.

FAO. (2020). *The state of world fisheries and aquaculture 2020*. Sustainability in action. Rome. //doi.org/https://doi.org/10.4060/ca9229en

FAO. (2021). *Fishery and aquaculture statistics 2019*. Rome: FAO Yearbook.

Feltes, R. M. (1986). *A systematic revision of the Polynemidae (Pisces) (unpublished doctoral dissertation)*. Columbus, Ohio: The Ohio State University.

Feltes, R. M. (1991). Revision of the polynemid fish genus *Filimanus*, with the description of two new species. *Copeia*, *1991*, 302–322. https://doi.org/10.2307/1446580.

Feltes, R. M. (1993). *Parapolynemus*, a new genus for the polynemid fish previously known as *Polynemus verekeri*. *Copeia*, *1993*, 207–215. https://doi.org/10.2307/1446312.

Freihofer, W. C. (1978). Cranial nerves of a percoid fish, *Polycentrus schomburgkii* (family Nandidae), a contribution to the morphology and classification of the order Perciformes. *Occasional Papers of the California Academy of Sciences*, *128*, 1–178.

Froese, R., & Pauly, D. (Eds.) (2021). FishBase. World Wide Web electronic publication. Retrieved from https://www.fishbase.org

Girard, M. G. (2021). *Systematics and macroevolutionary dynamics of carangiform fishes: Integrating genomic, morphological, and spatial data (unpublished doctoral dissertation)*. Lawrence, Kansas: University of Kansas.

Girard, M. G., Davis, M. P., & Smith, W. L. (2020). The phylogeny of carangiform fishes: Morphological and genomic investigations of a new fish clade. *Copeia*, *108*, 265–298. https://doi.org/10.1643/ci-19-320.

Gopalakrishnan, A., Vineesh, N., Shihab, I., Akhilesh, K. V., Bineesh, K. K., Menon, M., & Vijayagopal, P. (2021). An insight into the threadfin (Perciformes: Polynemidae) diversity of Indian waters using mitochondrial COI signatures. *Thalassas: An International Journal of Marine Sciences*, *37*, 689–700. https://doi.org/10.1007/s41208-021-00332-1.

Gosline, W. A. (1962). Systematic position and relationships of the percesocine fishes. *Pacific Science*, *16*, 207–217.

Gosline, W. A. (1971). *Functional morphology and classification of Teleostean fishes*. Honolulu, Hawaii: University of Hawaii Press.

Harrington, R. C., Faircloth, B. C., Eytan, R. I., Smith, W. L., Near, T. J., Alfaro, M. E., & Friedman, M. (2016). Phylogenomic analysis of carangimorph fishes reveals flatfish asymmetry arose in a blink of the evolutionary eye. *BMC Evolutionary Biology*, *16*, 224. https://doi.org/10.1186/S22862-016-0786-x.

Harris, R. S. (2007). *Improved pairwise alignment of genomic DNA (unpublished doctoral dissertation)*. State College, Pennsylvania: The Pennsylvania State University.

Hillis, D. M., Chambers, E. A., & Devitt, T. J. (2021). Contemporary methods and evidence for species delimitation. *Ichthyology & Herpetology*, *109*, 895–903. https://doi.org/10.1643/h2021082.

Horne, J. B., Momigliano, P., Welch, D. J., Newman, S. J., & Herwerden, L. V. (2011). Limited ecological population connectivity suggests low demands on self-recruitment in a tropical inshore marine fish (*Eleutheronema tetradactylum*: Polynemidae). *Molecular Ecology*, *20*, 2291–2306. https://doi.org/10.1111/j.1365-294x.2011.05097.x.

Johnson, G. D. (1993). Percomorph phylogeny—Progress and problems. *Bulletin of Marine Science*, *52*, 3–28.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Haeseler, A. von, & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*, 587–589. doi: https://doi.org/10.1038/nmeth.4285

Kang, S. (2017). *Comparative morphology and phylogenetic relationships of the family Polynemidae (Pisces: Perciformes) (unpublished doctoral dissertation)*. Hokkaido: Hokkaido University.

Kang, S., Imamura, H., & Kawai, T. (2017). Morphological evidence supporting the monophyly of the family Polynemidae (Teleostei: Perciformes) and its sister relationship with Sciaenidae. *Ichthyological Research*, *65*, 29–41. https://doi.org/10.1007/S20228-017-0591-6.

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. https://doi.org/10.1093/molbev/mst010.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of

sequence data. *Bioinformatics*, *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/btS299.

Khedkar, G. D., Jamdade, R., Naik, S., David, L., & Haymer, D. (2014). DNA barcodes for the fishes of the Narmada, one of India's longest rivers. *PLoS One*, *9*, e101460. https://doi.org/10.1371/journal.pone.0101460.

Lakra, W. S., Verma, M. S., Goswami, M., Lal, K. K., Mohindra, V., Punia, P., … Hebert, P. (2011). DNA barcoding Indian marine fishes. *Molecular Ecology Resources*, *11*, 60–71. https://doi.org/10.1111/j.1755-0998.2010.02894.x.

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., & Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, *14*, 82. https://doi.org/10.1186/1471-2148-14-82.

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, Y.-C., Tamemasa, S., Zhang, J.-Y., & Sato, H. (2020). Phylogenetic characterisation of seven Unicapsula spp. (Myxozoa: Myxosporea: Multivalvulida) from commercial fish in southern China and Japan. *Parasitology*, *147*, 448–464. https://doi.org/10.1017/s0031182019001793.

Martin, R. P., Olson, E. E., Girard, M. G., Smith, W. L., & Davis, M. P. (2018). Light in the darkness: New perspective on lanternfish relationships and classification using genomic and morphological data. *Molecular Phylogenetics and Evolution*, *121*, 71–85. https://doi.org/10.1016/j.ympev.2017.12.029.

May, R. C., Akiyama, G. S., & Santerre, M. T. (1979). Lunar spawning of the threadfin, *Polydactylus sexfilis*, in Hawaii. *Fishery Bulletin*, *76*, 900–904.

McAllister, D. E. (1968). Evolution of branchiostegals and classification of teleostome fishes. *Bulletin of the National Museum of Canada*, *221*, 1–239.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Haeseler, A. von, & 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: https://doi.org/10.1093/molbev/msaa015

Mirande, J. M. (2016). Combined phylogeny of ray-finned fishes (Actinopterygii) and the use of morphological characters in large-scale analyses. *Cladistics*, *33*, 333–350. https://doi.org/10.1111/cla.12171.

Motomura, H. (2004). Threadfins of the world (family Polynemidae) an annotated and illustrated catalogue of polynemid species known to date. In *Species catalogue for fishery purposes*. 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. https://doi.org/10.1007/S20228-001-8112-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. https://doi.org/10.1007/S20228-001-8157-y.

Motomura, H., Sado, T., & Kimura, S. (2002). Feeding behavior of *Polydactylus plebeius* (Perciformes: Polynemidae) in an aquarium (in Japanese). *Japanese Journal of Ichthyology*, *49*, 156–157.

Phillips, M. J., Westerman, M., & Cascini, M. (2021). The value of updating GenBank accessions for supermatrix phylogeny: The case of the new Guinean marsupial carnivore genus *Myoictis*. *Molecular Phylogenetics and Evolution*, *166*, 107328. https://doi.org/10.1016/j.ympev.2021.107328.

Presti, P. (2019). In *Phylogenetic relationships and evolution of the musculoskeletal system of Polynemidae (Teleostei: Percomorphacea: Perciformes) (unpublished masters thesis)*. São Paulo: Museu de Zoologia da Universidade de São Paulo.

Presti, P., Johnson, G. D., & Datovo, A. (2020). Anatomy and evolution of the pectoral filaments of threadfins (Polynemidae). *Scientific Reports*, *10*, 17751. https://doi.org/10.1038/s41598-020-74896-y.

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., & Korobeynikov, A. (2020). Using SPAdes *de novo* assembler. *Current Protocols in Bioinformatics*, *70*, e102. https://doi.org/10.1002/cpbi.102.

QGIS Development team. (2020). QGIS 3.10 geographic information system. *Open Source Geospatial Foundation Project*. Retrieved from. http://qgis.org.

Qu, Z., Nong, W., Yu, Y., Baril, T., Yip, H. Y., Hayward, A., & Hui, J. H. L. (2020). Genome of the four-finger threadfin *Eleutheronema tetradactylum* (Perciforms: Polynemidae). *BMC Genomics*, *21*, 726. https://doi.org/10.1186/S22864-020-07145-1.

Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., & Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, *559*(7714), 392–395. https://doi.org/10.1038/s41586-018-0273-1

Rambaut, A. (2009). FigTree (Version 1.4). Retrieved from http://tree.bio.ed.ac.uk/software/figtree/

Regan, C. T. (1912). XXVIII.—The classification of the teleostean fishes of the order Pediculati. *Journal of Natural History*, *9*, 277–289.

Ribeiro, A. d. O., Caires, R. A., Mariguela, T. C., Pereira, L. H. G., Hanner, R., & Oliveira, C. (2012). DNA barcodes identify marine fishes of São Paulo state, Brazil. *Molecular Ecology Resources*, *12*, 1012–1020. https://doi.org/10.1111/1755-0998.12007.

Roa-Varón, A., Dikow, R. B., Carnevale, G., Tornabene, L., Baldwin, C. C., Li, C., & Hilton, E. J. (2020). Confronting sources of systematic error to resolve historically contentious relationships: A case study using gadiform fishes (Teleostei, Paracanthopterygii, Gadiformes). *Systematic Biology*, *70*, 739–755. https://doi.org/10.1093/sysbio/syaa095.

Rolland, J., Cadotte, M. W., Davies, J., Devictor, V., Lavergne, S., Mouquet, N., … Winter, M. (2012). Using phylogenies in conservation: New perspectives. *Biology Letters*, *8*, 692–694. https://doi.org/10.1098/rsbl.2011.1024.

Rosen, D. E. (1964). The relationships and taxonomic position of the halfbeaks, killifishes, silversides, and their relatives. *Bulletin of the American Museum of Natural History*, *127*, 217–268.

Sabaj, M. H. (2020). Codes for natural history collections in ichthyology and herpetology. *Copeia*, *108*, 593–669. https://doi.org/10.1643/asihcodonS3020.

Sanciangco, M. D., Carpenter, K. E., & Betancur-R, R. (2016). Phylogenetic placement of enigmatic percomorph families (Teleostei: Percomorphaceae). *Molecular Phylogenetics and Evolution*, *94*, 565–576. https://doi.org/10.1016/j.ympev.2015.10.006.

Santerre, M. T., & May, R. C. (1977). Some effects of temperature on laboratory-reared eggs and larvae of *Polydactylus sexfilis* (Pisces: Polynemidae). *Aquaculture*, *10*, 341–351.

Shihab, I., Gopalakrishnan, A., Vineesh, N., Muktha, M., Akhilesh, K. V., & Vijayagopal, P. (2017). Histological profiling of gonads depicting protandrous hermaphroditism in *Eleutheronema tetradactylum*. *Journal of Fish Biology*, *90*, 2402–2411. https://doi.org/10.1111/jfb.13324.

Smith, W. L., Stern, J. H., Girard, M. G., & Davis, M. P. (2016). Evolution of venomous cartilaginous and ray-finned fishes. *Integrative and Comparative Biology*, *56*, 950–961. https://doi.org/10.1093/icb/icw070.

Sparks, J. S., & Smith, W. L. (2004). Phylogeny and biogeography of the Malagasy and Australasian rainbowfishes (Teleostei: Melanotaenioidei): Gondwanan vicariance and evolution in freshwater. *Molecular Phylogenetics and Evolution*, *33*, 719–734. https://doi.org/10.1016/j.ympev.2004.07.002.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

Stanger, J. D. (1974). *A study of the growth, feeding and reproduction of the threadfin, Eleutheronema tetradactylum (Shaw) (unpublished honors thesis)*. Townsville: James Cook University.

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, *114*, 1607–1612. https://doi.org/10.1073/pnas.1607921114.

Tagliacollo, V. A., & Lanfear, R. (2018). Estimating improved partitioning schemes for Ultraconserved elements. *Molecular Biology and Evolution*, *35*, 1798–1811. https://doi.org/10.1093/molbev/msy069.

Tang, K. L., Stiassny, M. L. J., Mayden, R. L., & DeSalle, R. (2021). Systematics of damselfishes. *Ichthyology & Herpetology*, *109*, 258–318. https://doi.org/10.1643/i2020105.

Thu, P. T., Huang, W.-C., Chou, T.-K., Quan, N. V., Chien, P. V., Li, F., … Liao, T.-Y. (2019). DNA barcoding of coastal ray-finned fishes in Vietnam. *PLoS One*, *14*, e0222631. https://doi.org/10.1371/journal.pone.0222631.

Wang, G.-H., Hao, R.-C., Yang, G.-Z., & Zan, L.-S. (2016). The complete mitochondrial genome sequence of Eleutheronema tetradactylum (Mugiliformes: Polynemidae) and phylogenetic studies of Mugiliformes. *Mitochondrial DNA Part A*, *27*(6), 4457–4458. https://doi.org/10.3109/19401736.2015.1089569

Weigt, L. A., Driskell, A. C., Baldwin, C. C., & Ormos, A. (2012). DNA barcoding fishes. In W. J. Kress & D. L. Erickson (Eds.), *DNA barcodes: Methods and protocols* (pp. 109–126). Totowa, New Jersey: Humana Press. https://doi.org/10.1007/978-1-61779-591-6_6.

Zhang, J., & Hanner, R. (2012). Molecular approach to the identification of fish in the south China sea. *PLoS ONE*, *7*(2), e30621. https://doi.org/10.1371/journal.pone.0030621

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, *19*, 153. https://doi.org/10.1186/S22859-018-2129-y.

Zhong, L., Wang, M., Li, D., Tang, S., & Chen, X. (2021). Mitochondrial genome of *Eleutheronema rhadinum* with an additional non-coding region and novel insights into the phylogenetics. *Frontiers in Marine Science*, *8*, 746598. https://doi.org/10.3389/fmars.2021.746598.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Girard, M. G., Davis, M. P., Baldwin, C. C., Dettaï, A., Martin, R. P., & Smith, W. L. (2022). Molecular phylogeny of the threadfin fishes (Polynemidae) using ultraconserved elements. *Journal of Fish Biology*, *100*(3), 793–810. https://doi.org/10.1111/jfb.14997