Molecular Phylogenetics and Evolution 212 (2025) 108424

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Exon-capture data resolve relationships resulting from a rapid radiation within family Gobiidae

Kendall Johnson [a,*], Luke Tornabene [b], Chenhong Li [c], Lukas Rüber [d,e], Ulrich Schliewen [f], Derek Hogan [g], Frank Pezold [a,h]

[a] Department of Life Sciences, Texas A&M University – Corpus Christi, Corpus Christi, TX, USA
[b] School of Aquatic and Fishery Sciences, University of Washington, Seattle, USA
[c] Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources, Shanghai Ocean University, Ministry of Education, Shanghai, China
[d] Naturhistorisches Museum Bern, Bernastrasse 15, 3005 Bern, Switzerland
[e] Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland
[f] SNSB-Bavarian State Collection of Zoology, Münchhausenstr. 21, 81247 München, Germany
[g] .Fisheries and Oceans Canada, 114 Fish Hatchery Lane, French Village, New Brunswick, Canada
[h] Florida Museum of Natural History, University of Florida, Gainesville, FL, USA

ARTICLE INFO

Keywords:
Gobiidae
Exon-capture
Phylogenetics
Diversification
Gobioidei

ABSTRACT

The family Gobiidae is the most speciose family within the Gobioidei, and is also historically the most poorly resolved due to a short period of rapid diversification early in the evolutionary history of the group. Their taxonomic and ecological diversity makes them ideal subjects for the study of many topics of ecology and evolution, including the evolution of behavior and life history traits, ecological and morphological adaptation, and speciation. However, the ability to understand and study these topics in gobies will remain limited until the phylogenetic relationships within the family have been resolved. Previous studies have identified 14 lineages within the family, but the relationships among those lineages are ambiguous, and more data is necessary to confidently determine those relationships. This study used exon-capture to obtain sequence data for hundreds of protein-coding loci to create a high-resolution phylogeny of Gobiidae to confidently resolve the interrelationships of the major lineages within the family. The monophyly of all previously identified groups was well-supported, as were the interrelationships among all but three of the 14 lineages, and results were consistent among different subsets of our data as well as different phylogenetic inferences. The major area of uncertainty included two competing hypotheses regarding the placement of the Gobiosomatini-lineage, and alternative hypotheses for this group are discussed in light of previous phylogenetic studies. Overall, the results represent a promising example of how gene-capture can resolve the taxonomy of problematic relationships that result from periods of rapid radiation.

1. Introduction

Gobies (families Gobiidae and Oxudercidae) are small, predominantly benthic fishes that together form one of the largest clades of vertebrates (Gill et al., 2012) containing more than 2000 species in more than 160 genera. The two families previously formed a unified clade, until genetic evidence demonstrated they were deeply divided and split into sister clades (Tornabene et al., 2013a, Agorreta et al., 2013). Both families contain taxa that exhibit a wide variety of behaviors, life history traits, and ecological and morphological adaptations as a result of periods of rapid radiation (Rüber et al., 2003; Thacker 2014). Oxudercidae *sensu* Betancur-R et al. (2017; Gobionellidae *sensu* Thacker 2009) includes mostly amphidromous or euryhaline fishes that dominate coastal estuaries and freshwater streams, including stream habitats across isolated tropical islands. This group includes species that are capable of scaling waterfalls, fishes with reduced or absent eyes, with elongated bodies specialized for burrowing in the mud, and capable of living semi-terrestrially (Keith et al., 2010; Polgar et al., 2011; Prokofiev, 2015). Gobiidae *sensu* Thacker (2009) are notable as one of the most abundant components of tropical and subtropical marine reef fish communities (Winterbottom et al., 2011; Brandl et al., 2018), although some lineages appear in a variety of other habitats, incuding lakes,

* Corresponding author.
E-mail address: kjohnson47@islander.tamucc.edu (K. Johnson).

https://doi.org/10.1016/j.ympev.2025.108424
Received 1 January 2025; Received in revised form 20 June 2025; Accepted 24 July 2025
Available online 25 July 2025

rivers, seagrass beds, mangroves, oyster reefs, soft bottoms of the continental shelf, and tidepools. They also form commensal relationships with a variety of organisms. These include alpheid shrimps (*Cryptocentrus*), sponges (*Risor*), sea urchins (*Tigrigobius*), bryozoans (*Sueviota*), and corals (*Gobiodon*) (Patzner et al., 2011).

However, understanding evolutionary relationships among the gobiids is difficult because they underwent an early period of rapid diversification that has led to short internal branches and poor phylogenetic resolution early in the family's phylogeny (Thacker and Roje, 2011; Agorreta et al., 2013; Thacker, 2014; McCraney et al., 2025). In addition, a tendency toward miniaturization, reduction, or loss of several potentially informative characters has been postulated to have occurred several times independently, further confounding phylogenetic analysis of goby clades based on morphological characters (Winterbottom, 1990; Thacker, 2003). For example, many groups have independent losses or reductions in potentially informative characters such as the cephalic lateralis canals and associated pores, fused pelvic fins, scales, fusion of hypurals with the urostyle, sensory neuromasts, and many others (Birdsong et al., 1988**; Pezold, 1993; Van Tassell et al., 2011; Tornabene et al., 2018). It has also been shown that increased rates of speciation in Gobiidae and Oxudercidae are correlated with decreased morphological diversification rates (Thacker, 2014). Despite the lack of clear phylogenetic signal from morphological data, there are some clades that are well-supported by genetic evidence.

Molecular phylogenetic studies have attempted to elucidate relationships within Gobiidae. Thacker and Roje (2011) used data from three mitochondrial gene regions and two nuclear genes across 58 genera to identify 13 major lineages within the Gobiidae. The remaining unsampled genera were tentatively assigned to these lineages based on similarities in morphological characters, biogeographic distributions, or ecological characters. However, several of the clades in that study were not well supported, and the validity of the groups including the unsampled genera has not been tested. Agorreta et al. (2013) incorporated three nuclear and two mitochondrial genes regions and recovered 14 major lineages that partially corresponded to those identified in Thacker and Roje (2011). The 14th new lineage identified by Agorreta et al. (2013) was the *Aphia* lineage containing the genera *Aphia* and *Lesueurigobius*, neither of which were sampled in Thacker and Roje (2011). The comparable lineages are displayed in Table 1.

However, when comparing results to Thacker and Roje (2011), there were multiple cases of genera that moved across lineages identified in the two studies. These included *Discordipinna*, *Drombus*, *Gobiopsis*, and *Feia*. Notably, both studies found that extremely short internal branches near the base of the Gobiidae made it difficult to find additional conclusive support for relationships among the major gobiid lineages, despite the inclusion of 126 gobiid species across 94 genera by Agorreta et al. (2013). They concluded that faster-evolving nuclear markers may be needed to separate out the early cladogenetic events of Gobiidae (Agorreta et al., 2013).

Another approach examined the phylogeny of Gobiaria using 23 loci in a sparse supermatrix with expanded taxon sampling (50 outgroups and 777 ingroup taxa) obtained from data available on Genbank (McCraney et al., 2020). They used several methods of phylogenetic inference, but were unable to obtain robust support for many gobiid lineages across their different methods of tree construction. However, the few well-supported clades were consistent with previous findings from Thacker and Roje (2011) and Agorreta et al. (2013). These include clades containing the *Gobius* and *Gobiosoma* lineages, the *Asterropteryx* and *Lophogobius* lineages, the *Aphia* and *Valenciennea* lineages. and the *Glossogobius*, *Kraemeria*, *Cryptocentrus* and *Gobiopsis* lineages (McCraney et al., 2020). These clades consistently resolved within phylogenies, although not always with strong support or between phylogenetic inference methods.

Short internal branch lengths, like those observed near the base of the gobiids in phylogenetic studies, are considered a signature of rapid diversification that represents a major challenge in phylogeny

Table 1
Comparison of identified lineages in Thacker and Roje (2011) to Agorreta et al. (2013).

Clade name *sensu* Thacker and Roje (2011)	Genera sampled – Thacker and Roje (2011)	Clade name *sensu* Agorreta et al. (2013)	Genera sampled – Agorreta et al. (2013)
n/a	n/a	*Aphia*-lineage	*Aphia, Lesueurigobius*
Reef Shrimp	*Amblyeleotris, Asterropteryx, Ctenogobiops, Vanderhorstia*	*Asterropteryx*-lineage	*Amblyeleotris, Asterropteryx, Ctenogobiops, Vanderhorstia*
Flapheaded	*Callogobius*	*Callogobius*-lineage	*Callogobius*
Silt Shrimp	*Cryptocentrus, Mahidolia, Stonogobiops*	*Cryptocentrus*-lineage	*Cryptocentrus, Cryptocentroides, Discordipinna, Mahidolia*
Inshore	*Bathygobius, Psammogobius*	*Glossogobius*-lineage	*Bathygobius, Glossogobius, Psammogobius*
Coral	*Bryaninops, Eviota, Gobiodon, Paragobiodon*	*Gobiodon*-lineage	*Bryaninops, Eviota, Gobiodon*
Lagoon	*Acentrogobius, Afurcagobius, Arenigobius, Cabillus, Exyrias, Favonigobius, Istigobius, Oplopomus, Papillogobius*	*Gobiopsis*-lineage	*Acentrogobius, Amoya, Arenigobius, Drombus, Exyrias, Favonigobius, Gobiopsis, Istigobius, Porogobius, Silhouettea*
American 7 spined	*Barbulifer, Elacatinus, Gobiosoma, Microgobius, Nes, Ophiogobius, Risor, Tigrigobius*	Gobiosomatini-lineage	*Aboma, Akko, Aruma, Barbulifer, Bollmannia, Chriolepis, Elacatinus, Enypnias, Evermannichthys, Ginsburgellus, Gobiosoma, Gobulus, Gymneleotris, Microgobius, Nes, Ophiogobius, Parrella, Psilotris, Pycnomma, Risor, Tigrigobius*
Mediterranean & Ponto-Caspian & Eastern Atlantic Gobies	*Babka, Benthophilus, Mesogobius, Neogobius, Ponticola, Proterorhinus, Gobius, Zosterisessor, Caffrogobius, Coryogalops*	*Gobius*-lineage	*Babka, Benthophilus, Chromogobius, Corcyrogobius, Coryogalops, Didogobius, Gammogobius, Gobius, Gorogobius, Mauligobius, Millerigobius, Odondebuenia, Ponticola, Proterorhinus, Sufflogobius, Thorogobius, Vanneaugobius, Wheelerigobius, Zebrus, Zosterisessor*
Wormfishes & Dartfishes	*Cerdale, Gunnellichthys, Microdesmus, Nemateleotris, Ptereleotris*	*Gunnellichthys*-lineage	*Gunnellichthys, Microdesmus, Nemateleotris, Oxymetopon, Parioglossus, Ptereleotris, Schindleria*
Sanddivers	*Kraemeria*	*Kraemeria*-lineage	*Kraemeria*
Crested	*Coryphopterus, Fusigobius, Lophogobius, Rhinogobiops*	*Lophogobius*-lineage	*Coryphopterus, Fusigobius, Lophogobius*
Tiny Banded	*Lythrypnus, Priolepis, Trimma*	*Priolepis*-lineage	*Feia, Lythrypnus, Priolepis, Trimma*

(*continued on next page*)

Table 1 (*continued*)

Clade name *sensu* Thacker and Roje (2011)	Genera sampled – Thacker and Roje (2011)	Clade name *sensu* Agorreta et al. (2013)	Genera sampled – Agorreta et al. (2013)
Burrowing Paired	*Amblygobius, Signigobius, Valenciennea*	*Valenciennea*-lineage	*Amblygobius, Signigobius, Valenciennea*

reconstructions (Rokas et al., 2005). Sequencing a large number of loci can improve the accuracy of species trees with short branch lengths (Edwards et al., 2007). Exon capture is a way to efficiently target a large number of loci across the genomes of distantly related taxa for use in phylogenetics. The most prominent approaches utilize genomic regions that are strongly conserved across a highly diverged range of taxa and enrich DNA samples for these loci (Lemmon et al., 2012). The advantage of this method is that it captures orthologous loci for phylogenetically divergent taxa, allowing researchers to obtain numerous loci for use in addressing broad phylogenomic analyses (Lemmon and Lemmon, 2013). Additionally, single-copy exons evolve in ways that are well suited for existing evolutionary models and provide the opportunity to test for adaptive selection in ecologically distinct lineages (Daane et al., 2019). Thus, these methods produce a data set useful for the study's main goals of investigating the phylogeny of Gobiidae, while also providing data that can be used in future studies focusing on the diversification and adaptive radiations of gobiids.

Kuang et al. (2018) utilized exon-capture to assess the phylogeny of the Gobioidei, focusing on the broader relationships among the families in this group. While that study included only 36 gobioids, of which only 8 were within the Gobiidae, not enough to provide significant insight into family structure within the Gobiidae, nearly all nodes were perfectly supported, even when various degrees of missing data were included, thus demonstrating the utility of exon-capture data at this evolutionary scale. Exon-capture has been used with success in previous studies to clarify disputed relationships among other groups of fishes as well, including Gadiformes (cods and allies; Roa-Varon et al. 2021), Cichlidae (cichlids; Ilves et al. 2018; Alda et al. 2021), Characiformes (tetras and allies; Betancur-R et al., 2019), Pleuronectiformes (flatfishes; Atta et al. 2022), Clupeiformes (shads; Wang et al., 2022**), Labridae (wrasses and parrotfishes; Hughes et al. 2023), Lophiiformes (anglerfishes; Miller et al. 2024). Exon data were also used by Hughes et al. (2018) to reconstruct a phylogeny across all ray-finned fishes.

This study uses exon-capture to increase molecular information and taxon sampling to better assess intra-familial relationships within the Gobiidae. Representatives from the sister family Oxudercidae and other gobioids are included as outgroups. The inclusion of taxa from both Gobiidae and Oxudercidae will provide insight to both shallow relationships within each of these derived families and the broader relationships between them.

2. Materials and methods

2.1. Taxon sampling

The samples used for gene-capture were obtained from our own field collections or from museum loans. Samples were chosen with the aim of having multiple representatives for the lineages defined by Agorreta et al. (2013). A total of 107 samples were processed, which included 94 species and 72 genera that represented all of the 14 lineages identified by Agorreta et al. (2013), four species in the sister family Oxudercidae, as well as one species in the family Rhyacichthyidae (Table 2). Several genera had multiple species included in the taxon sampling. This was done to confirm monophyly of the group or decrease the chance that poor gene-capture success for a single sample would make it difficult to resolve the position of that genus. For some genera the only available tissues were from samples that had not been identified to the species level. In such cases, these samples were used rather than omit the genus from the study completely. Several non-gobiid outgroups were included. These were *Rhyacichthys aspro*, *Odontobutis potamophila*, *Milyeringa veritas*, *Eleotris acanthopoma*, and *Butis koilomatodon*; as well as four oxudercid outgroups *Paratrypauchen microcephalus*, *Redigobius leveri*, *Scartelaos histophorus*, and *Stiphodon caeruleus*. There is one non-Gobioid outgroup: *Oreochromis niloticus* (Cichlidae). All of these except *Rhyacichthys aspro* and the four oxudercid outgroups were included from a previous study on Gobioidei that used the same loci (Kuang et al. 2018). This final taxon sampling includes representatives of all the Gobioid families, except Thalasseleotrididae, and consisted of 108 samples.

2.2. Gene capture

RNA baits were designed using EvolMarkers to target approximately 12,000 exons present in eight model fish genomes (*Lepisosteus oculatus, Danio rerio, Oryzias latipes, Gasterosteus aculeatus, Tetraodon nigroviridis, Anguilla japonica, Gadus macrocephalus,* and *Oreochromis niloticus*) (Li et al., 2012). These sequences were screened so that the average length of each gene was 250 bp, and to ensure the loci were not too divergent from each other. Once these loci were identified, biotinylated RNA capture probes were designed from those sequences in the *Oreochromis niloticus* genome. The bait set generated was then compared to the mudskipper (*Boleophthalmus pectinirostris*; Oxudercidae) genome using BLAST. Sequences derived from that species were used to ensure a higher likelihood of capture success in the gobiid species being targeted.

To capture targeted loci, we followed modified protocol of protein-coding gene-capture across highly divergent species (Li et al., 2013). The method accomplished the capture of a broad array of targets by relaxing hybridization conditions and by including a second round of gene capture in which captured products from the first amplification are used as templates for the second round of gene-capture. This has been shown to increase the number of targets captured by 68 % on average as compared to just one round of gene capture.

Total genomic DNA was extracted from the preserved tissue samples using a Qiagen DNAeasy Blood and Tissue Kit (Qiagen, Valencia, California), and the quality of the DNA was assessed with gel electrophoresis. A total of 100 ng of DNA was put in a 270ul volume of PCR water and sheared to approximately 250 base-pairs on a modified Covaris machine using a custom protocol. This sheared DNA was used for library preparation. Library preparation and gene capture was carried out according to the protocol outlined in Li et al. (2013) using the custom bait set. An alternative protocol outlined in Li et al. (2013) for degraded DNA was used for samples with poor quality or fragmented DNA to maximize the number of loci captured. A unique combination of inline indexes was assigned to each sample to identify it, as well as reduce the probability of cross-contamination (Wang et al., 2022). After each step, quality checks were done with agarose gels and spectrophotometric quantification of the DNA. Any samples that did not have sufficient product upon completion of each step were redone if possible. Prior to sequencing, the library for each sample was size-selected by gel electrophoresis using BluePippin to include fragments within a range of 300–1000 bp. The majority of the sequences for the library were distributed in the 350–700 bp range, so selecting for sequences within this size range ensured that the majority of the gene-capture products were sequenced and that excluded sequencing products that were mainly adaptor, as the length of the adaptor was 124 bp.

2.3. Sequencing and read assembly

Samples were sequenced on an Illumina HiSeq 2500 or an Illumina HiSeq 4000. The total number of reads from the sequencing efforts was 884,954,276. The Assexon pipeline (Yuan et al., 2019) was used to assemble the raw gene-capture data. The Assexon pipeline requires a reference genome as one of its inputs used to verify the orthology of retrieved sequences (Yuan et al., 2019). Nile Tilapia (*Oreochromis*

Table 2
List of samples sequenced in this study. "Unknown" indicates a member of a genus not yet included in molecular studies that have placed it within a given lineage of the families Gobiidae or Oxudercidae.

Genus	Species	Lineage	Family	Catalog #	BioSample accessions #
Butis	*koilomatodon*	n/a	Butidae	Kuang et al., 2018	n/a
Eleotris	*acanthopoma*	n/a	Eleotridae	Kuang et al., 2018	n/a
Aphia	*minuta*	*Aphia-*	Gobiidae	NMBE 1066523	SAMN49091248
Lesueurigobius	*sanzi*	*Aphia-*	Gobiidae	NMBE 1066486	SAMN49091302
Amblyeleotris	*gymnocephala*	*Asterropteryx-*	Gobiidae	KU 29310	SAMN49091243
Amblyeleotris	*wheeleri*	*Asterropteryx-*	Gobiidae	Kuang et al., 2018	n/a
Amblyeleotris	*yanoi*	*Asterropteryx-*	Gobiidae	Kuang et al., 2018	n/a
Asterropteryx	*ensifera*	*Asterropteryx-*	Gobiidae	uncataloged	SAMN49091249
Asterropteryx	*semipunctatus*	*Asterropteryx-*	Gobiidae	UW 205843	SAMN49091250
Ctenogobiops	*feroculus*	*Asterropteryx-*	Gobiidae	UW 205849	SAMN49091268
Gladiogobius	*brevispinis*	*Asterropteryx-*	Gobiidae	UW 200470	SAMN49091284
Vanderhorstia	*ambanoro*	*Asterropteryx-*	Gobiidae	uncataloged	SAMN49091335
Callogobius	*hasseltii*	*Callogobius-*	Gobiidae	uncataloged	SAMN49091258
Callogobius	sp	*Callogobius-*	Gobiidae	UW 202123	SAMN49091259
Cryptocentrus	sp	*Cryptocentrus-*	Gobiidae	uncataloged	SAMN49091266
Cryptocentrus	*strigilliceps*	*Cryptocentrus-*	Gobiidae	UW 156932	SAMN49091267
Discordipinna	*griessingeri*	*Cryptocentrus-*	Gobiidae	BLIH 20103042	SAMN49091270
Mahidolia	*mystacina*	*Cryptocentrus-*	Gobiidae	NMBE 1066439	SAMN49091305
Stonogobiops	*nematodes*	*Cryptocentrus-*	Gobiidae	Kuang et al., 2018	n/a
Bathygobius	*geminatus*	*Glossogobius-*	Gobiidae	uncataloged	SAMN49091252
Glossogobius	*celebius*	*Glossogobius-*	Gobiidae	UW 200478	SAMN49091285
Glossogobius	*illimis*	*Glossogobius-*	Gobiidae	UW 202156	SAMN49091286
Psammogobius	*biocellatus* (1)	*Glossogobius-*	Gobiidae	UW 203283	SAMN49091321
Psammogobius	*biocellatus* (2)	*Glossogobius-*	Gobiidae	uncataloged	SAMN49091322
Kelloggella	*oligolepis*	*Glossogobius-*	Gobiidae	USNM 440506	SAMN49091300
Bryaninops	sp	*Gobiodon-*	Gobiidae	uncataloged	SAMN49091255
Eviota	*afelei*	*Gobiodon-*	Gobiidae	uncataloged	SAMN49091275
Eviota	*nigriventris*	*Gobiodon-*	Gobiidae	UW 205845	SAMN49091276
Eviota	*rubrisparsa*	*Gobiodon-*	Gobiidae	CAS-ICH 249459	SAMN49091277
Eviota	*sigillata*	*Gobiodon-*	Gobiidae	UW 205850	SAMN49091278
Eviota	sp	*Gobiodon-*	Gobiidae	uncataloged	SAMN49091279
Gobiodon	*oculolineatus*	*Gobiodon-*	Gobiidae	uncataloged	SAMN49091287
Gobiodon	*rivulatus*	*Gobiodon-*	Gobiidae	uncataloged	SAMN49091288
Gobiodon	*unicolor*	*Gobiodon-*	Gobiidae	UW 205844	SAMN49091289
Phyllogobius	*platycephalops*	*Gobiodon-*	Gobiidae	UW 205851	SAMN49091315
Pleurosicya	*muscarum*	*Gobiodon-*	Gobiidae	uncataloged	SAMN49091316
Sueviota	*lachneri*	*Gobiodon-*	Gobiidae	UW 205846	SAMN49091330
Acentrogobius	*janthinopterus*	*Gobiopsis-*	Gobiidae	UW 200352	SAMN49091241
Acentrogobius	*suluensis*	*Gobiopsis-*	Gobiidae	UW 200236	SAMN49091242
Amoya	*moloana*	*Gobiopsis-*	Gobiidae	LR12610 (NMBE accession in preparation)	SAMN49091247
Cabillus	*tongarevae*	*Gobiopsis-*	Gobiidae	uncataloged	SAMN49091256
Drombus	sp (1)	*Gobiopsis-*	Gobiidae	UW 200478	SAMN49091271
Drombus	sp (2)	*Gobiopsis-*	Gobiidae	UW 203299	SAMN49091272
Exyrias	*puntang*	*Gobiopsis-*	Gobiidae	UW 200504	SAMN49091280
Favonigobius	sp	*Gobiopsis-*	Gobiidae	UW 205852	SAMN49091281
Istigobius	*decoratus*	*Gobiopsis-*	Gobiidae	UW 205853	SAMN49091297
Istigobius	*ornatus*	*Gobiopsis-*	Gobiidae	UW 200526	SAMN49091298
Istigobius	*rigilius*	*Gobiopsis-*	Gobiidae	uncataloged	SAMN49091299
Macrodontogobius	*wilburi*	*Gobiopsis-*	Gobiidae	CAS 243557	SAMN49091304
Oplopomus	*oplopomus*	*Gobiopsis-*	Gobiidae	UW 202263	SAMN49091312
Porogobius	*schlegelii*	*Gobiopsis-*	Gobiidae	NMBE 1066461	SAMN49091318
Silhouettea	sp	*Gobiopsis-*	Gobiidae	NMBE 1066428	SAMN49091329
Yongeichthys	*nebulosus*	*Gobiopsis-*	Gobiidae	uncataloged	SAMN49091338
Barbulifer	*ceuthoecus*	Gobiosomatini-	Gobiidae	AMNH I-263710	SAMN49091251
Bollmannia	*communis*	Gobiosomatini-	Gobiidae	KU 30145	SAMN49091254
Elacatinus	*oceanops*	Gobiosomatini-	Gobiidae	KU IT 242	SAMN49091273
Evermannichthys	*spongicola*	Gobiosomatini-	Gobiidae	UF 180316	SAMN49091274
Gobiosoma	*bosc*	Gobiosomatini-	Gobiidae	uncataloged	SAMN49091290
Gobiosoma	*grosvenori*	Gobiosomatini-	Gobiidae	AMNH I-263713	SAMN49091291
Gobiosoma	*robustum*	Gobiosomatini-	Gobiidae	uncataloged	SAMN49091292
Risor	*ruber*	Gobiosomatini-	Gobiidae	USNM 349078	SAMN49091326
Tigrigobius	*macrodon*	Gobiosomatini-	Gobiidae	uncataloged	SAMN49091332
Benthophilus	sp	*Gobius-*	Gobiidae	NMBE 1066575	SAMN49091253
Caffrogobius	*caffer*	*Gobius-*	Gobiidae	KU IT 6494	SAMN49091257
Chromogobius	*britoi*	*Gobius-*	Gobiidae	NMBE 1066545	SAMN49091261
Didogobius	*kochi*	*Gobius-*	Gobiidae	NMBE 1066500	SAMN49091269
Gobius	*niger* (1)	*Gobius-*	Gobiidae	NMBE 1066473	SAMN49091293
Gobius	*niger* (2)	*Gobius-*	Gobiidae	Kuang et al., 2018	n/a
Gobius	*paganellus*	*Gobius-*	Gobiidae	LR00312 (NMBE accession in preparation)	SAMN49091294
Gorogobius	*nigricinctus*	*Gobius-*	Gobiidae	NMBE 1066476	SAMN49091295
Mauligobius	*maderensis* (1)	*Gobius-*	Gobiidae	NMBE 1066409	SAMN49091306
Mauligobius	*maderensis* (2)	*Gobius-*	Gobiidae	NMBE 1066440	SAMN49091307
Odondebuenia	*balearica*	*Gobius-*	Gobiidae	NMBE 1066446	SAMN49091311

(*continued on next page*)

Table 2 (*continued*)

Genus	Species	Lineage	Family	Catalog #	BioSample accessions #
Ponticola	*kessleri*	*Gobius-*	Gobiidae	LR00613 (NMBE accession in preparation)	SAMN49091317
Thorogobius	*ephippiatus*	*Gobius-*	Gobiidae	NMBE 1066435	SAMN49091331
Vanneaugobius	*canariensis*	*Gobius-*	Gobiidae	AMNH 233177	SAMN49091336
Wheelerigobius	*maltzani*	*Gobius-*	Gobiidae	NMBE 1066407	SAMN49091337
Zebrus	*zebrus*	*Gobius-*	Gobiidae	NMBE 1066408	SAMN49091339
Cerdale	*floridana*	*Gunnellichthys-*	Gobiidae	KUI 40264	SAMN49091260
Gunnellichthys	*monostigma*	*Gunnellichthys-*	Gobiidae	BLIH1999001	SAMN49091296
Microdesmus	*dorsipunctatus*	*Gunnellichthys-*	Gobiidae	Kuang et al., 2018	n/a
Microdesmus	*longipinnis*	*Gunnellichthys-*	Gobiidae	KUIT 2425	SAMN49091308
Nemateleotris	*magnifica* (1)	*Gunnellichthys-*	Gobiidae	uncataloged	SAMN49091309
Nemateleotris	*magnifica* (2)	*Gunnellichthys-*	Gobiidae	NMBE 1066445	SAMN49091310
Oxymetopon	*cyanoctenosum*	*Gunnellichthys-*	Gobiidae	NMBE 1066451	SAMN49091313
Parioglossus	*marginalis*	*Gunnellichthys-*	Gobiidae	AMS I.40838-013	SAMN49091314
Ptereleotris	*microlepis* (1)	*Gunnellichthys-*	Gobiidae	UW 200506	SAMN49091323
Ptereleotris	*microlepis* (2)	*Gunnellichthys-*	Gobiidae	uncataloged	SAMN49091324
Schindleria	*praematura*	*Gunnellichthys-*	Gobiidae	LR01351 (NMBE accession in preparation)	SAMN49091327
Kraemeria	sp	*Kraemeria-*	Gobiidae	KUIT 4898	SAMN49091301
Coryphopterus	*glaucofrenum*	*Lophogobius-*	Gobiidae	AMNH I-258507	SAMN49091262
Coryphopterus	*hyalinus*	*Lophogobius-*	Gobiidae	uncataloged	SAMN49091263
Cristatogobius	*aurimaculatus* (1)	Unknown	Gobiidae	uncataloged	SAMN49091264
Cristatogobius	*aurimaculatus* (2)	Unknown	Gobiidae	UW 203298	SAMN49091265
Fusigobius	*neophytus*	*Lophogobius-*	Gobiidae	UW 202261	SAMN49091283
Feia	*nympha*	*Priolepis-*	Gobiidae	BLIH 20070607	SAMN49091282
Lythrypnus	*nesiotes*	*Priolepis-*	Gobiidae	UF 183294	SAMN49091303
Priolepis	*cincta*	*Priolepis-*	Gobiidae	uncataloged	SAMN49091319
Priolepis	*nuchifasciata*	*Priolepis-*	Gobiidae	uncataloged	SAMN49091320
Amblygobius	*nocturnus*	*Valenciennea-*	Gobiidae	UW 200469	SAMN49091244
Amblygobius	*phalaena*	*Valenciennea-*	Gobiidae	UW 205842	SAMN49091245
Amblygobius	sp	*Valenciennea-*	Gobiidae	uncataloged	SAMN49091246
Signigobius	*biocellatus*	*Valenciennea-*	Gobiidae	NMBE 1066427	SAMN49091328
Valenciennea	*muralis*	*Valenciennea-*	Gobiidae	SLU550	SAMN49091333
Valenciennea	*puellaris*	*Valenciennea-*	Gobiidae	Kuang et al., 2018	n/a
Valenciennea	*sexguttata*	*Valenciennea-*	Gobiidae	AMNH166772	SAMN49091334
Valenciennea	*strigata*	*Valenciennea-*	Gobiidae	Kuang et al., 2018	n/a
Milyeringa	*veritas*	n/a	Milyeringidae	Kuang et al., 2018	n/a
Odontobutis	*potamophila*	n/a	Odontobutidae	Kuang et al., 2018	n/a
Redigobius	*leveri*	*Mugilogobius-*	Oxudercidae	xxx	n/a
Scartelaos	*histophorus*	*Periopthalmus-*	Oxudercidae	uncataloged	n/a
Stiphodon	*caeruleus*	*Stenogobius-*	Oxudercidae	CAS 235211	n/a
Paratrypauchen	*microcephalus*	Unknown	Oxudercidae	SAIAB 82351	n/a
Rhyacicthys	*aspro*	n/a	Rhyacichthyidae	xxx	SAMN49091325

niloticus) was used as the reference genome for the purposes of identifying paralogs and creating the coding DNA and amino acid reference sequences used during assembly. EvolMarkers was used to generate the OnehitCDSmarker file for Tilapia. Assembly consisted of six steps during which PCR duplicates were removed, reads were parsed to homologous targeted loci, de novo assembly was conducted, and exons were extracted. Potential paralogs were removed and sequences aligned with MAFFT. The Assexon pipeline includes a "reblast" (reciprocal BLAST paralogs filter) step that identifies and removes potential paralogs (Yuan et al., 2016). This, in combination with the baits designed to target single-copy loci, provided a high likelihood that the assembled loci were orthologous. Paralogs have been demonstrated to drive strongly supported conflicting hypotheses in phylogenies, so their removal is important (Siu-Ting et al., 2019). The program created three versions of the assembled data: coding regions with flanking regions, coding regions without flanking regions, and protein sequences for the coding sequences. Further phylogenetic analyses used the coding regions without flanking regions assembled dataset. This was done because it gave us a dataset of well-characterized protein-coding orthologs without as many issues of homology among taxa and less bioinformatic pre-processing needed to address possible uncertainties in alignment. Additionally, protein-coding sequences are better understood by models (Li et al., 2013). The trimmed, assembled sequences are available on the Dryad digital repository (https://doi.org/10.5061/dryad.s1rn8pkb8). The datasets that are further processed as detailed below are also available there. Raw reads (trimmed of adaptors) can be accessed at GenBank SRA PRJNA1276973 PRJNA1276973.

2.4. Quality control and filtering

Before the assembled data were used, they went through several quality-control steps. The first part of this quality control consisted of removing samples from the analyses. Samples with very few captured loci (less than 200) were removed from the overall dataset, which removed three samples. These samples were *Priolepis nuchifasciata*, *Oplopomus oplopomus*, and *Phyllogobius_platycephalops*. A custom script was used to detect potential cross-contamination between samples by examining percent similarity from taxa not specified a priori as closely related (such as those within the same genus). Samples that were determined to have significant cross-contamination were removed. This removed an additional three samples. These samples were *Gobiodon unicolor*, *Drombus* sp., and *Mauligobius maderensis*. Of the remaining 94 samples, including the *O. niloticus* outgroup, there were no genes shared among all the samples. Loci were subsampled from the remaining samples to create data matrices with the least amount of missing data while maintaining the largest combination of taxon sampling. We filtered out loci based upon whether they were present in at least a certain percentage of the total number of samples. This removed rare loci that were sequenced for only a few samples, and were thus the least phylogenetically informative. This strategy results in a more densely sampled matrix. On the other hand, removing many loci, or removing species that had small number of loci sampled, in order to get a more complete matrix could both result in the removal of potentially informative species as well as removing too much data, as it has been shown that filtering too much can result in biased phylogenetic trees (Chan et al., 2020). As a compromise, the datasets were filtered at two levels.

One was to include loci present in at least 50 % of the total number of samples. The other was to include loci present in at least 75 % of the total number of samples. (Table 3).

2.5. Phylogenetic inference

Several maximum-likelihood phylogenetic trees with bootstrap support were inferred with RAxML (Stamatakis, 2014) using the GTRGAMMA model and the concatenation of all loci included in the filtered data set produced from the previous steps. A phylogeny was produced for each filtering level. These were produced using the RAxML options -f a – (rapid Bootstrap analysis and search for best scoring ML tree in one program run) and with 500 bootstraps.

Partitioned analyses were conducted using IQTree2 (Nguyen et al., 2015; Minh et al., 2020a). Partition models for each codon were chosen using the included ModelFinder for model selection using an edge-proportional model in which the top 10 % of partition schemes were considered (-rcluster 10), and phylogenies were produced using 1000 ultrafast bootstrap replicates (Chernomor et al., 2016; Kalyaanamoorthy et al., 2017; Hoang et al., 2018).

Coalescent species trees were generated using ASTRAL version 5.7.3 (**Mirarab et al., 2014). Individual gene trees were generated with RAxML following the same procedure used to create the unpartitioned concatenated phylogenies. Additional trees were generated for ASTRAL using individual partitioned gene trees generated with IQTree2. Branches with low support (below 10 % bootstrap support) were collapsed using Newick utilities prior to use in ASTRAL (Junier and Zdobnov, 2010).

2.6. Topology tests and concordance factor analyses

When there were differences in topologies between the partitioned IQTREE analyses and unpartitioned analyses in RAxML, topology tests calculating the difference in site-wise log-likelihood score (ΔSLS) and the difference in gene-wise log-likelihood score (ΔGLS) between topologies were conducted as performed by Shen et al. (2017). To conduct them, a partitioned RAxML tree was created in an alternative topology that matched that of the IQTREE phylogenies. These were done to identify which genes supported which topology, and to see if either topology was being supported by only a few key genes. Gene concordance factor (gCF) and site concordance factor (sCF) analyses were conducted on the partitioned phylogenies produced with IQTree2 (Minh et al., 2020b).

3. Results

A total of 16,674 loci were captured and successfully assembled from the newly sequenced samples. The number of captured loci for each sample ranged from 19 to 11,889 (Fig. 1). The average number of loci captured per sample was 4508. When the dataset was filtered to 50 %, there were 1,775 loci included, and when it was filtered to 75 %, there were 245 loci included (Table 3).

3.1. Concatenated IQTree and RAxML analyses

The dataset was able to resolve the nodes connecting the 14 lineages within the Gobiidae recovered by Agorreta et al. (2013) with high bootstrap support; we use their nomenclature to describe the lineages in this study.

Table 3
Detail on the dataset created for the analyses.

# of Samples	Degree of Filtering	# of Included genes	Total length of concatenated sequence (bp)	% Missing data
102	50 %	1775	280,734	39.8 %
102	75 %	245	35,535	21.2 %

The partitioned IQTREE phylogenies produced with 50 %-filtered datasets had significantly higher bootstrap support values across the entire topology than the 75 %-filtered datasets, although both produced identical topologies. The 50 %-filtered data phylogenies had high (>70) ML bootstrap support values for all nodes connecting lineages, both within Gobiidae and Oxudercidae, whereas the 75 % filtered datasets had a few nodes where relationships between a few lineages were weakly supported. Given the 50 % filtered datasets produced the most well-supported phylogenies, we discuss those results below. Due to the congruency between the results we focus on the dataset filtered to include loci present in at least 50 % of the samples produced by IQTree as the main phylogeny (Fig. 2). The phylogeny for the 75 %-filtered analysis can be found in the Supplementary Fig. S1.

Within the Gobiidae, the phylogeny shows an initial split between a clade containing the *Glossogobius*, *Kraemeria*, *Gobiopsis*, and *Cryptocentrus* lineages, (henceforth clade A) and a clade containing the *Gobiodon*, *Lophogobius*, *Asterropteryx*, *Priolepis*, *Gobius*, Gobiosomatini, *Valenciennea*, *Aphia*, *Gunnellichthys*, and *Callogobius* lineages (henceforth clade B; Fig. 2). In Clade A, the *Glossogobius*-lineage was sister to a clade containing the other three lineages, within which *Kraemeria* was sister to a clade containing *Gobiopsis* and *Cryptocentrus*-lineages. Clade B was split into two other clades – one with the *Gobiodon*, *Asterropteryx*, and *Lophogobius*-lineages (clade B1), the other with the *Priolepis*, *Gobius*, Gobiosomatini, *Valenciennea*, *Aphia*, *Gunnellichthys*, and *Callogobius* lineages (clade B2). Clade B1 shows the *Gobiodon*-lineage sister to a clade consisting of the *Asterropteryx* and *Lophogobius*-lineages. Clade B2 had the *Callogobius* and then the *Gunnellichthys*-lineages as sisters to the remaining lineages. The remaining five lineages were split into two sister clades. One contained the Gobiosomatini-lineage as sister to the *Gobius* and *Priolepis*- lineages, and the other contained the *Aphia* and *Valenciennea*-lineages.

As with the partitioned analyses the topologies resolved by the unpartitioned RAxML 50 %-filtered and 75 %-filtered topologies were identical, so the results of the 50 %-filtered analyses are presented (Supplementary Fig. S2). The dataset resolved a slightly different topology in the unpartitioned RAxML analysis. The Gobiosomatini-lineage was sister to the *Aphia/Valenciennea*-clade, rather than the *Priolepis/Gobius* −clade (Fig. 3). The placement of the Gobiosomatini-lineage was the only difference between the topologies of the unpartitioned and partitioned analyses, with clades A and B1 remaining unchanged. There was only moderate support (76 bootstrap) for the new clade containing the Gobiosomatini-, *Aphia*-, and *Valenciennea*- lineages.

3.2. Topology tests and concordance factor analyses

Since the phylogeny displayed a different topology between the partitioned and unpartitioned analyses, it was investigated using ΔSLS and ΔGLS topology tests. The default concatenated topology had Gobiosomatini sister to the *Priolepis/Gobius*-clade, and the alternative topology had Gobiosomatini sister to the *Aphia/Valenciennea*-clade. The results of the ΔSLS and ΔGLS topology tests indicate that there is strong support for both topologies across multiple genes and multiple sites. The support was not localized to a small number of genes driving this conflicting topology (Supplementary Figs. 3 and 4). The highest support for the default topology was 4.187065 for ΔGLS and 4.190936 for ΔSLS. The highest support for the alternative topology was −6.39636 for ΔGLS and −3.441362 for ΔSLS. Overall, there was slightly more support for the topology showing Gobiosomatini sister to the *Priolepis/Gobius*-clade (sum of ΔGLS for all genes was 22.40274).

The gCF and sCF analyses revealed that despite high bootstrap support, there was discordance among the individual genes when compared to the resolved topology for the IQTree partitioned analyses. gCF values were low across lineage splits within Gobiidae with many of these nodes having gCF values of less than 10, in all phylogenies at both levels of

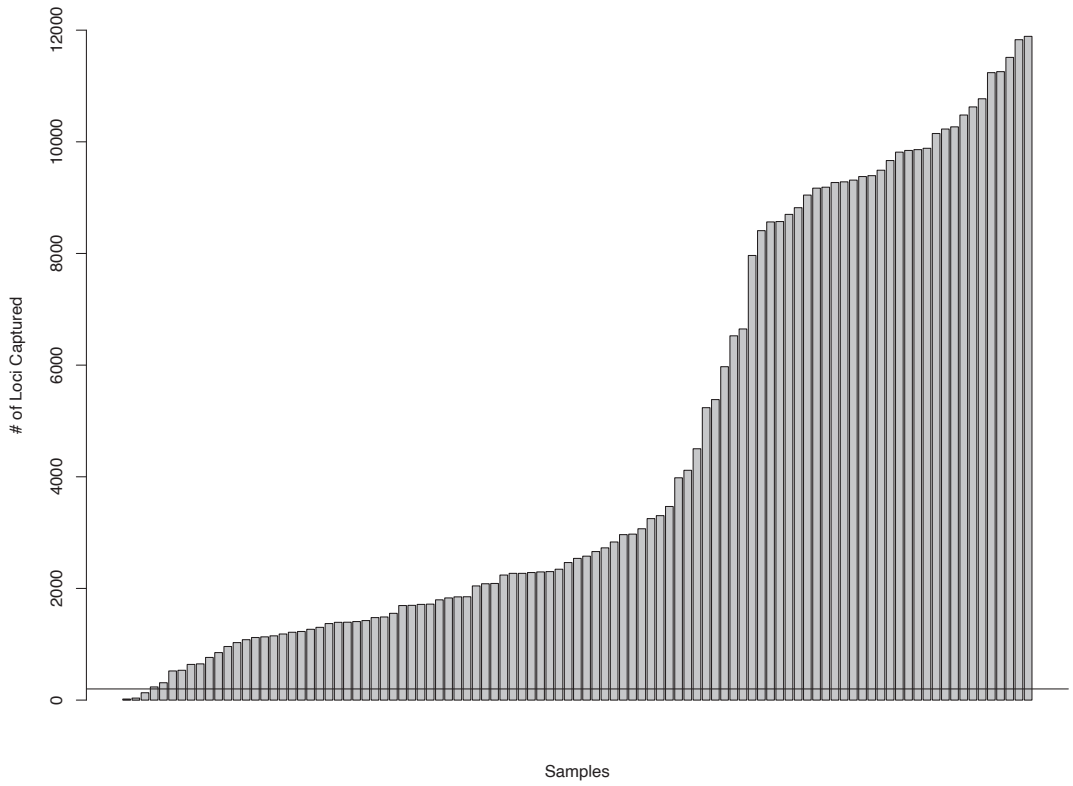

Fig. 1. Histogram of the distribution of the number of loci captured for each sample.

filtering. Overall, gCF values were low for all nodes excluding those between closely-related genera or species within the same genus. Low gCF values were always matched by high values of gDFP (gene discordance due to polyphyly), as opposed to high values of either gDF1 or gDF2 (gene discordance due to gene trees resolving either of two possible alternative topologies).

sCF values were similarly low across the phylogeny although few ever dropped below 30 %. This is expected given that the way sCF values are calculated results in their minimum value rarely being lower than 33 % (Minh et al., 2020b). They followed a similar pattern to the gCF values where their values were lowest at the nodes between lineages and increased for nodes between closely related samples.

3.3. ASTRAL analyses

The phylogenies produced by ASTRAL were less well-supported when compared to both the unpartitioned and partitioned RAxML analyses. Although most lineages were resolved consistently across phylogenies, several taxa including *Schindleria praematura* and the two *Callogobius* samples commonly resolved independently from their associated lineages from the RAxML tree. The phylogenies produced by the 75 %-filtered datasets were consistently less well-resolved than their 50 %-filtered counterparts.

The ASTRAL phylogeny produced with the 50 %-filtered dataset recovered a topology identical to that of the partitioned IQTree phylogeny, with the Gobiosomatini-lineage being sister to a clade of the *Gobius/Priolepis*-lineages (Supplementary Fig. S5). The Local Posterior Probability (LPP) branch support values also indicated relatively high support for many branches. The lowest support for a branch among lineages was 0.55 LPP, for the branch between the *Kraemeria*-lineage and the *Gobiopsis/Cryptocentrus*-lineages. The branch connecting the Gobiosomatini-lineage to the *Gobius/Priolepis*-lineages had 0.71 LPP, just above what can be considered the minimum threshold for confidence in a branch (Sayyari and Mirarab, 2016). This offered additional support for the Gobiosomatini and *Gobius/Priolepis* sister relationship shown by the partitioned RAxML topology.

The ASTRAL phylogenies produced from the partitioned gene trees in which low-support branches were contracted were similarly less well-resolved than the concatenated phylogenies. This was true for both the 50 % and 75 %-filtered datasets. The samples *Schindleria praematura*, *Kelloggella oligolepis,* and *Callogobius hasseltii* were frequently resolved outside of their associated lineages in the RAxML phylogeny. The dataset resolved the Gobiosomatini-lineage as sister to a clade of the *Gobius/Priolepis*-lineages and the *Aphia/Valenciennea*-lineages when filtered to include genes present in at least 50 % of the samples. However, when filtered to include genes present in at least 75 % of the samples, the Gobiosomatini-lineage was sister to the *Gobius/Priolepis*-lineages, although the LPP value was below the minimum threshold of confidence (0.58) (S6).

4. Discussion

Using gene-capture to obtain a large quantity of loci, this study was able to confidently resolve the relationships among lineages within the family Gobiidae, corroborating findings from previous studies and improving resolution across the family. Support values were very high across the majority of hypothesized relationships, and these relationships emerge among phylogenies created using different versions of the dataset and different methodologies. Specifically, the same general topologies were seen whether the phylogenies were made using genes present in at least 75 % or 50 % of the included taxa, although the 50 %-filtered datasets produced phylogenies with higher bootstrap support. It is possible that the increased filtering stringency in the 75 %-filtered dataset removed more localized loci shared only amongst closely related lineages that would have linked taxa across those lineages and provided more support for their relationships.

4.1. Gobiidae

Our results confirm the Gobiidae-Oxudercidae split among the

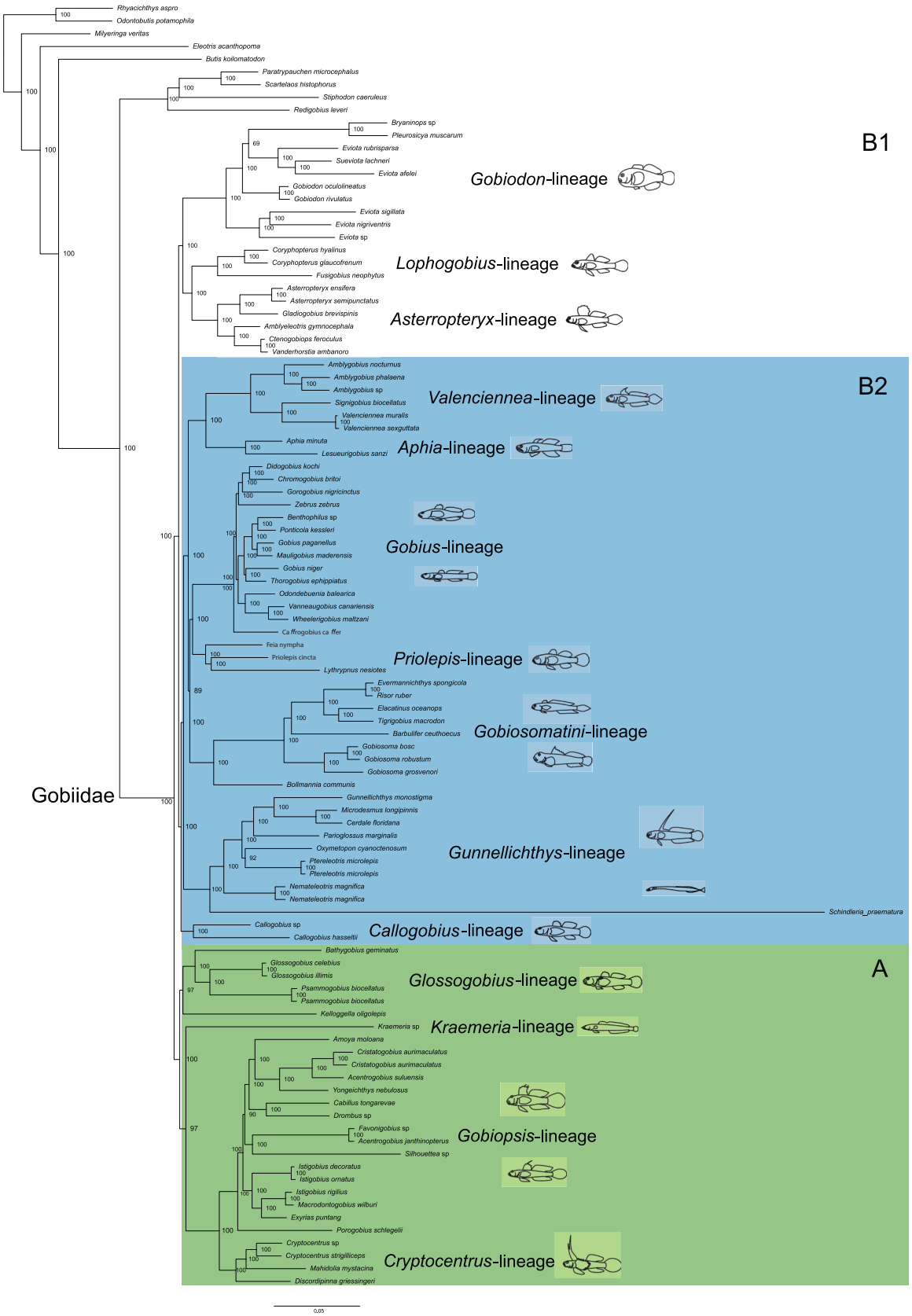

Fig. 2. Phylogeny produced with the IQTree partitioned by codon 50%-filtered dataset. Lineages within Gobiidae and Oxudercidae are highlighted in grey. Bootstrap support values are provided. The outgroup Oreochromis niloticus has been removed. Images from Agorreta et al. (2013).

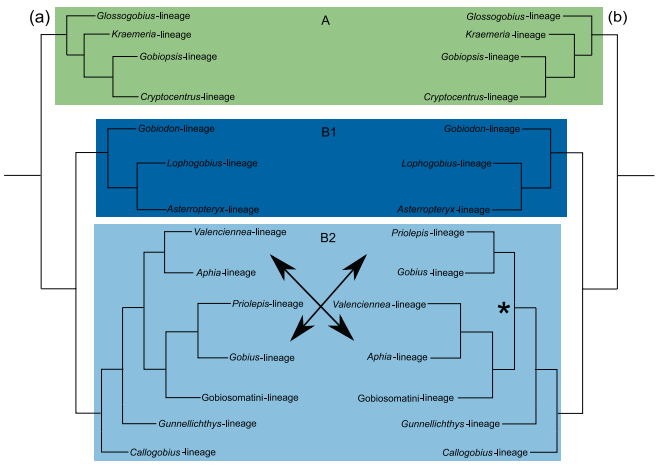

Fig. 3. Structure of clades resolved in the family Gobiidae for the (a) partitioned analyses (b) unpartitioned RAxML analysis. Asterisks indicate nodes that differed from those in partitioned analyses.

derived gobies and all taxa expected to resolve in Gobiidae or Oxudercidae based upon the findings of previous studies were found within their expected families. The lineages recovered in this study largely correspond to those identified in Agorreta et al. (2013), and their expansions in Thacker (2015). Agorreta et al. (2013) recovered sister relationships between the *Valenciennea-* and *Aphia*-lineages, and the *Asterropteryx-* and *Lophogobius*-lineages with high ML and BI support. Both of these were confirmed in this study. Agorreta et al. (2013) also recovered a sister relationship between the *Gobiodon*-lineage and the *Asterropteryx-* and *Lophogobius-* clade, although theirs was not well supported. McCraney et al. (2020) also resolved our clade "A" consisting of the *Glossogobius-*, *Kraemeria-*, *Cryoptocentrus-*, and *Gobiopsis-* lineages, as well as the *Asterropteryx-/Lophogobius-* clade and *Valenciennea-/Aphia-* clade.

The Gobiosomatini remains the only gobiid lineage for which this study was not able to resolve a single well-supported consistent placement. Our study indicates that the *Aphia-* and *Valenciennea-*, and *Gobius-* and *Priolepis-* lineages each form monophyletic clades, and that Gobiosomatini is part of a broader clade containing the other two clades. Within this broad clade, we identify two competing hypotheses for the position of the Gobiosomatini (Fig. 3). The first hypothesis is that the Gobiosomatini-lineage is sister to the *Aphia-* and *Valenciennea-* lineages (Fig. 3b), which is supported by the RAxML unpartitioned analyses. The second is that it is sister to the *Priolepis-* and *Gobius-* lineages (Fig. 3a), which is supported by the RAxML partitioned analyses, the IQTree partitioned analyses, as well as one of the ASTRAL analyses. A close relationship between the *Gobius*-lineage and the Gobiosomatini was previously hypothesized by both Agorreta et al. (2013) and McCraney et al. (2020). The conflicting phylogenetic signals for either topology reflect how both the unpartitioned and partitioned RAxML 50 %-filtered gobiid analyses resolved the node connecting Gobiosomatini to its sister clade with similar support – 76 bootstrap support for the unpartitioned phylogeny, and 79 bootstrap support in the partitioned phylogeny. The datasets had strong support for the placement of Gobiosomatini in the IQTree partitioned analyses when filtered to include genes in at least 50 % of the samples, however when filtered at the 75 % level the bootstrap support decreased to 66.

The dataset's increased number of genes likely contributed to the different topologies resolved for the Gobiosomatini- lineage between the unpartitioned and partitioned RAxML analyses. The different topologies between the datasets may be explained by the fact that additional genes included in the gobiid dataset provided more support for Gobiosomatini being sister to the *Gobius/Priolepis* clade. Thacker (2015) resolved Gobiosomatini sister to the *Gobius*-lineage, although her *Priolepis*-lineage was in a separate clade with the *Valenciennea-* and *Gunnellichthys*-lineages. However, she did infer that the Gobiosomatini-, *Priolepis-*, and *Gobius*-lineages all had ancestral distributions in the Western Indian Ocean. Those lineages also contain New World taxa present in the Atlantic or Mediterranean, with Gobiosomatini being exclusive to the Western Atlantic and Eastern Pacific. From a biogeographic perspective there is support for the Gobiosomatini-*Gobius/Priolepis* relationship, although the *Aphia*-lineage can also be found in those regions. There is also some morphological evidence for a relationship between Gobiosomatini and *Gobius*-lineages. Miller and Tortonese (1968) suggested a relation between Gobiosomatini and the genus *Odondebuenia*, a member of the *Gobius*-lineage, based upon similarities in squamation, sensory pore, and free neuromast patterns. However, patterns of general squamation, sensory pores, and free neuromasts are highly variable within the Gobiosomatini itself, as well as across the Gobiidae, and phylogenetic signal in these characters is suspect even at small phylogenetic scales (i.e. within genera or across closely related genera) (Tornabene et al., 2013b; Tornabene et al., 2016). The Gobiosomatini- and *Gobius*-lineages both contain taxa with modified basicaudal scales, such as *Aboma* in Gobiosomatini (Dawson, 1969) and *Didogobius*, *Odondebuenia*, and *Vanneaugobius* in *Gobius* (Van Tassell and Kramer, 2014). However, this same trait has been documented in *Cabillus caudimacula*, a member of the *Gobiopsis*-lineage (Greenfield and Randall, 2004).

The gCF values for the Gobiosomatini-lineage were low in both the 50 %-filtered datasets (gCF of 0.362) and the 75 %-filtered dataset (gCF value of 0.816), and both had gDFP values above 90 (98.57 and 98.78 respectively). The gene concordance factors do not provide support for an alternative topology for the Gobiosomatini- branches where they have low values, and do not provide any additional insight into the sister clade of that lineage. Although gCF values are similarly low for other gobiid lineages, it is only the Gobiosomatini-lineage that repeatedly resolves in different locations across the various phylogenies, with support for both placements divided evenly among genes as demonstrated by the ΔGLS topology tests. With this in mind we conclude that the placement of the Gobiosomatini-lineage remains to be concretely determined, although there is some preliminary evidence in favor of a sister relationship to the *Gobius/Priolepis*-clade. Including more members of the Microgobius group *sensu* Birdsong (1975) may help resolve the placement of Gobiosomatini. *Bollmannia* resolved as sister to the rest of the Gobiosomatini-lineage samples, which is expected given that it was the only representative of the *Microgobius* group *sensu* Birdsong (1975) included in this study, while the remaining samples were members of the *Gobiosoma* group (Rüber et al., 2003). This relationship among genera belonging to the *Microgobius* and *Gobiosoma* groups was also the case in Agorreta et al. (2013). The remaining relationships within the Gobiosomatini are largely congruent with those of Rüber et al. (2003) and Agorreta et al. (2013), except that *Barbulifer* is sister to the *Elacatinus/Tigrigobius/Risor/Evermannichthys* clade, versus being more closely related to *Gobiosoma*, as was the case in the former studies. Those studies contained far more species and genera in the Gobiosomatini which likely explains the differences observed in the topologies.

This study was able to resolve the location of the *Kraemeria*-lineage in Clade A with strong bootstrap support. *Kraemeria* was sister to a clade consisting of the *Gobiopsis-* and *Cryptocentrus*-lineages, all of which were sister to the *Glossogobius*-lineage. Previous studies have resolved *Kraemeria* in conflicting locations across the gobioids with poor support and it has been considered a "rogue taxon". Thacker (2003) recovered *Kraemeria* among oxudercid taxa, and Thacker (2009) as part of Gobiidae. Tornabene et al. (2018) resolved it as sister to *Kelloggella* with weak support, but when *Kraemeria* was removed from the analysis support for the relationship of *Kelloggella* and other taxa increased. This, along with lack of any unambiguous morphological synapomorphy between *Kraemeria* and *Kelloggella* led them to conclude that the genera were not each other's closest relative. Agorreta et al. (2013) resolved *Kraemeria* as sister to the *Gobius*-lineage with weak support, and excluding it from their analyses substantially increased support for multiple branches at

the base of the gobiids. One of these, however, was for a clade consisting of the *Gobius*-, *Priolepis*-, and Gobiosomatini-lineages. This may offer more support for the Gobiosomatini as sister to *Gobius/Priolepis* topology recovered in the partitioned analyses as the true topology, although the resolution among the lineages themselves was weak. The degree of resolution this study was able to obtain for *Kraemeria* is likely due to the greater amount of genetic information incorporated into the analyses. In the 50 % filtered data phylogeny there were 414 loci included for the *Kraemeria* sp. specimen. Agorreta et al. (2013)'s had five genes with around 5000–6000 bp combined, of which only about 2000 bp were parsimony-informative. When they removed *Kraemeria* from their analyses they found strong bootstrap support for a clade consisting of the *Glossogobius*-, *Cryptocentrus*-, and *Gobiopsis*-lineages as the sister to all other gobiid gobies. This, with the exception of the inclusion of *Kraemeria*, is identical to this study's recovered Clade A. While *Kraemeria* was resolved in a consistent location within Gobiidae the IQTree analyses had few genes that directly supported its placement in the topology The highest proportion of genes concordant with topology for *Kraemeria* sp. was 6 out of the 97 genes decisive for the branch containing it in the 75 %-filtered dataset (gCF = 6.19). Despite this, *Kraemeria* was also resolved in the same location as in Figs. 2 and 3 in our Clade A in the contracted ASTRAL phylogenies excluding the 75 %-filtered phylogeny, providing additional support for its relationship with the *Glossogobius*-, *Cryptocentrus*-, and *Gobiopsis*-lineages.

Kelloggella was found to resolve within the *Glossogobius*-lineage as its most basal member. Thacker (2015) expected it to be included in the *Gobiodon*-lineage, however it was not included in the sampling of Thacker and Roje (2011), Agorreta et al. (2013), nor McCraney et al. (2020). It was included in Tornabene et al. (2018), which found some molecular evidence for a relationship between it and the genera *Gobiodon*, *Eviota*, and *Bryaninops*. However, they were unable to find any significant morphological support for this group. The support for the node connecting *Kelloggella* to the rest of the *Glossogobius*-lineage in this study became weaker in the partitioned RAxML analyses. Both the unpartitioned and partitioned RAxML analyses on the dataset had lower support values for this node when it was filtered to include genes present in 75 % than 50 % of the samples, and this was also the case for the IQTree partitioned gobioid phylogenies. *Kelloggella oligolepis* resolved outside of the *Glossogobius*-lineage in all of the partitioned contracted ASTRAL analyses. The inclusion of other species from the genus into future phylogenetic studies may assist in providing additional support for the relationship of this genus to other gobiids.

This study found that the *Callogobius*-lineage was monogeneric as was the case in previous studies. The *Callogobius*-lineage was sister to a clade containing the remainder of the lineages in clade B2. Thacker and Roje (2011) found *Callogobius* as their "flaphead gobies" sister to a clade containing taxa in the *Asteropteryx*-, *Lophogobius*-, and *Valenciennea*-lineages, although there was no significant support for the nodes connecting any of these lineages. Agorreta et al. (2013) resolved it as sister to a clade containing the *Gobiodon*-, *Asteropteryx*-, and *Lophogobius*-lineages, although without significant support. McCraney et al. (2020) found the position of the *Callogobius*-lineage variable and were unable to confidently place it. This study's gene capture success with one of the *Callogobius* samples was relatively poor and was among the samples with the fewest genes included in the study. However, the datasets still included more genes for *Callogobius* than previous studies have, and with less sparseness than previous datasets. McCraney et al. (2020) included 23 genes in a dataset with 17 % completeness. The *Callogobius* sp. sample had at least 223 genes included in each phylogeny. The 50 %-filtered partitioned contracted ASTRAL phylogeny resolved both samples together, and there was strong support (LPP = 0.99) for the lineage within the topology. When both *Callogobius* samples were not resolved together, it was *Callogobius hasseltii* that resolved away from the expected placement of the lineage. This may be due to the lower number of genes with data available for that sample, as *Callogobius* sp. was always resolved where the *Callogobius*-lineage would be within Clade B2. The genus *Callogobius* has more than 40 species, so it is possible that including more representatives would have more confidently resolved this clade.

The topology of the *Gobiodon*-lineage differs from previous works. This study found that a monophyletic clade of *Eviota* was sister to another clade which included an *Eviota/Sueviota* clade sister to *Gobiodon, Bryaninops*, and *Pleurosicya*. The node separating the *Eviota* clade from the rest was very strongly supported (100 % bootstrap). This tree hypothesizes that *Eviota* is paraphyletic with *Sueviota, Gobiodon, Bryaninops*, and *Pleurosicya*. This is different from Agorreta et al. (2013) which showed *Gobiodon* as monophyletic and sister to a clade including one *Eviota* species and one *Bryaninops* species, although it did not include *Pleurosicya* or *Sueviota*. *Eviota* and the relationships of the *Gobiodon*-lineage were poorly-resolved in McCraney et al. (2020) and there are currently over 100 recognized species for the genus (Greenfield, 2017). Tornabene et al. (2013a) recovered a similar divide in their phylogeny of *Eviota* species, which they noted was divided into a clade possessing branched pectoral fin rays and a clade possessing unbranched pectoral fin rays. This morphological division of *Eviota* into two groups was first pointed out by Winterbottom and Hoese (1988). Additionally, they noted that not only was it unclear whether *Eviota* was monophyletic or not, it was not certain whether *Sueviota* was sister to all or only some *Eviota* species. They included *Sueviota* in the branched group, which is where it resolved in this study's phylogeny. Increased taxon sampling of *Eviota* in this study's dataset, which included five *Eviota* species, allows for the greater resolution of relationships within this lineage and implies that increased depth of taxon sampling within speciose genera is needed. Indeed, increased taxon sampling with the *Gobiodon*-lineage, specifically within the exceptionally diverse genus *Eviota* and several other genera that are un- or under-represented in large-scale phylogenies (e.g. *Paragobiodon*, *Pleurosicya*, *Luposicya*, etc.) is needed to clarify the monophyly of the genera in this lineage and address any taxonomic changes that may be needed.

Other genera that resolved as non-monophyletic in the phylogenies include *Gobius*, *Istigobius*, and *Acentrogobius*. *Thorogobius ephippiatus* was sister to *Gobius niger*, rather than *Gobius niger* being sister to the only other *Gobius* sample (*Gobius paganellus*). Since the phylogeny included only two species from the genus *Gobius*, which has at least 28 recognized species, it is difficult to formally assess if this is evidence for non-monophyly of *Gobius*. The *Istigobius ornatus* and *Istigobius decoratus* samples are sister to each other, however the third *Istigobius* sample, *Istigobius rigilius*, is sister to *Macrodontogobius wilburi*. The two *Acentrogobius* samples were broadly separated within the *Gobiopsis*-lineage, with *Acentrogobius suluensis* sister to the *Cristatogobius* clade and *Acentrogobius janthinopterus* being sister to *Favonigobius sp*. This also occurred in Agorreta et al. (2013). It is important to not use these results as the sole evidence for concluding that these genera are non-monophyletic. Taxon sampling of these genera was sparse in our phylogeny, and increased depth of taxon sampling for these genera will be necessary before making any firm conclusions about their status, however these results provide some initial evidence that many genera previously thought to be monophyletic may not be.

4.2. Branch lengths

Several samples had long branch lengths. These include *Kraemeria sp*., *Schindleria praematura*, *Fusigobius neophytus*, *Lythrypnus nesiotes*, and *Kelloggella oligolepis*. *Schindleria praematura* had by far the longest branch of all the gobies sampled. *Schindleria praematura* is paedomorphic and has a very short generation time with a lifespan of only two to three months. This could lead to accelerated rates of molecular evolution and may explain the exceedingly long branch length. The sample had a rather high number of successfully captured loci (208) so missing data is likely not a major factor. Alignments for *Schindleria praematura* were manually examined and there was no issue in alignment that would explain the long branch. Similarly, long branch lengths were present for

Schindleria praematura and *Schindleria pietschmanni* in Agorretta et al. (2013) and McCraney et al. (2025) which suggests that this long branch is not artificial.

Another cause for these long branch lengths could be a long period of evolution without any extant sister taxa or lineages emerging from it. This could be the case for *Kraemeria sp*. There are several taxa theorized to be closely related to *Kraemeria* (such as *Parkraemeria*, *Gobitrichonotus*, and *Schismatogobius*) but no samples for them were included in this study. Tornabene et al. (2018) observed similarly long branch lengths in their phylogeny for *Kelloggella*, *Kraemeria*, and *Schindleria*.

4.3. Concordance factors

Low concordance factors despite the high bootstrap support were common across all phylogenies produced with IQTree. While the lowest were among the gobiid lineages, even well-established and accepted branches such as that between Gobiidae and Oxudercidae had low gCF values (gCF = 11.8 in the 50 %-filtered phylogeny). It is expected that gCF values will be closer to zero the more gene trees are estimated from sequences with limited information or where branches are extremely short, as the number of gene trees that resolve non-monophyletic in regard to the overall topology increases (Minh et al. 2020b). Arcila et al. (2021) further demonstrated that gCF and sCF values are significantly correlated with branch length. Our individual genes are short and therefore do not provide much information on their own, and the low gCF values were paired with very high values of gDFP, indicating that they exhibited the behavior explained by Minh et al. (2020b) above. The lowest gCF values were found among the gobiid lineages, where there are extremely short branches; there is little phylogenetic signal that the individual genes would have contained to resolve them. Therefore, the low gCF values for the gobiid lineages are expected given the nature of this study's markers and dataset.

Some of the sCF values for nodes among lineages were close to or just below 33 %. One of these is the branch connecting the *Gunnellichthys*-lineage to the remaining lineages of clade B1 in the 75 %-filtered gobiid phylogeny (gCF = 0.448, sCF = 29.7). However, this same branch was resolved with high bootstrap support in all concatenated analyses, as well as with high LLP support in the contracted 50 %-filtered gobioid ASTRAL analysis and the gobiid 50 %-filtered uncontracted ASTRAL analysis. Minh et al. (2020b) proposed that branches with sCF values lower than 33 % may be under incomplete lineage sorting (ILS). However, much like the gCF values, these very low sCF values were also calculated for branches of well-known, established relationships, including the branch between *Milyeringa veritas* (Milyeringidae) and the remaining gobioid taxa in the 50 %-filtered phylogeny (gCF = 14.1, sCF = 31.2), With this in consideration, it is difficult to state that ILS is the sole reason for the low sCF values among the gobiid lineages or other branches in the phylogenies. Nevertheless, ILS is not unlikely in rapidly diverging clades, but still difficult to investigate for ancient divergences.

5. Conclusions

The relationships within the family Gobiidae *sensu* Thacker (2009) have long been investigated and long been unresolved, with studies concluding that additional data from genes across the genome would be necessary to resolve relationships among the major lineages. This study produced a dataset composed of single-copy protein-coding genes sampled from across the genome that was able to confidently resolve the interrelationships among almost all the previously identified gobiid lineages for the first time. The only ambiguity remains the exact placement of the unambiguously monophyletic Gobiosomatini-lineage, which was resolved as sister to either the *Aphia/Valenciennea*-clade or *Priolepis/Gobius*-clade in partitioned analyses depending on the dataset. Topology tests revealed that there was mixed support for either relationship across the genes and sites which resolved this ambiguity, and that these contentious relationships were not caused by a few select genes. This could be due to ILS or even hybridization between ancient lineages.

The family Gobiidae is not only biologically important as one of the most species-rich families of marine vertebrates and the dominant constituent of cryptobenthic reef fishes, but it is also ripe for macro-evolutionary studies. The well-supported phylogenetic tree can be used as a framework tracing the evolution of ecological and morphological characters across the tree and looking at how traits affect other evolutionary processes such as speciation and extinction rates. Opportunities for such studies have been largely limited due to the high level of phylogenetic uncertainty, which often restricts our ability to infer evolutionary trends across multiple gobiid lineages.

Future studies that utilize gene-capture could improve upon this by designing baits from a gobiid species. This might improve gene-capture efficiency, which negatively impacted this study resulting in low branch support values for a few taxa in which gene capture efficiency was very low. Collection efforts focusing on species that are not as well-studied could help by providing additional specimens for each species, as well as more species, that could assist in resolving the topologies with these clades. Additional specimens for species in exceptionally diverse genera are needed to properly investigate the possible non-monophyly of those genera. While this study resolved the relationships between goby lineages within the family Gobiidae, it did not include the depth of sampling needed to fully and confidently assess the nature of the relationships of some genera within those lineages. The results offer some evidence for non-monophyly of several of the genera, pointing at future studies to investigate the relationships within them directly. Additional research and increased taxon sampling will be needed to address those questions for genera such as *Gobius*, *Acentrogobius*, *Istigobus*, and *Eviota*.

This study's results not only provide a confident well-supported hypothesis for the phylogenetic structure of the family Gobiidae but demonstrates that exon-capture methodology can be used to unravel problematic relationships for clades that have long been considered unresolvable. By using enough data that connects at least half of the included taxa, it is possible to reveal patterns in the evolution of groups previously thought unresolvable due to a rapid period of explosive evolution. This holds promise for future studies to resolve the structure of other important clades that have long been eluded, including those where short branch lengths due to rapid radiations have hindered previous efforts.

CRediT authorship contribution statement

Kendall Johnson: Writing – review & editing, Writing – original draft, Visualization, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Luke Tornabene:** Writing – review & editing, Supervision, Data curation, Conceptualization. **Chenhong Li:** Writing – review & editing, Methodology, Investigation, Data curation. **Lukas Rüber:** Writing – review & editing, Resources, Methodology. **Ulrich Schliewen:** Writing – review & editing, Resources. **Derek Hogan:** Writing – review & editing, Methodology, Conceptualization. **Frank Pezold:** Writing – review & editing, Supervision, Resources, Methodology, Conceptualization.

Acknowledgements

This work was supported by the National Science Foundation's East Asia and Pacific Summer Institutes [award number 1714032]. Publication supported in part by an Institutional Grant [award number NA18OAR4170088, project number NA18-2019-E/GIA-2019-Johnson] to the Texas Sea Grant College Program from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. Constructive comments on the manuscript from Ainhoa Agoretta and Prosanta Chakrabarty.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2025.108424.

References

Agorreta, A., San Mauro, D., Schliewen, U., Van Tassell, J.L., Kovacic, M., Zardoya, R., Ruber, L., 2013. Molecular phylogenetics of Gobioidei and phylogenetic placement of European gobies. Mol. Phylogenet. Evol. 69, 619–633.

Alda, F., Ludt, W.B., Elias, D.J., McMahan, C.D., Chakrabarty, P., 2021. Comparing ultraconserved elements and exons for phylogenomic analyses of middle american cichlids: when data agree to disagree. Genome Biol. Evol. 13.

Arcila, D., Hughes, L.C., Melendez-Vazquez, B., Baldwin, C.C., White, W.T., Carpenter, K. E., Williams, J.T., Santos, M.D., Pogonoski, J.J., Miya, M., Orti, G., Betancur, R.R., 2021. Testing the utility of alternative metrics of branch support to address the ancient evolutionary radiation of tunas, stromateoids, and allies (Teleostei: Pelagiaria). Syst. Biol. 70, 1123–1144.

Atta, C.J., Yuan, H., Li, C., Arcila, D., Betancur, R.R., Hughes, L.C., Orti, G., Tornabene, L., 2022. Exon-capture data and locus screening provide new insights into the phylogeny of flatfishes (Pleuronectoidei). Mol. Phylogenet. Evol. 166, 107315.

Betancur, R.R., Wiley, E.O., Arratia, G., Acero, A., Bailly, N., Miya, M., Lecointre, G., Orti, G., 2017. Phylogenetic classification of bony fishes. BMC Evol. Biol. 17, 162.

Betancur, R.R., Arcila, D., Vari, R.P., Hughes, L.C., Oliveira, C., Sabaj, M.H., Orti, G., 2019. Phylogenomic incongruence, hypothesis testing, and taxonomic sampling: the monophyly of characiform fishes. Evolution 73, 329–345.

Brandl, S.J., Goatley, C.H.R., Bellwood, D.R., Tornabene, L., 2018. The hidden half: ecology and evolution of cryptobenthic fishes on coral reefs. Biol. Rev. Camb. Philos. Soc. 93, 1846–1873.

Chan, K.O., Hutter, C.R., Wood Jr., P.L., Grismer, L.L., Brown, R.M., 2020. Larger, unfiltered datasets are more effective at resolving phylogenetic conflict: Introns, exons, and UCEs resolve ambiguities in Golden-backed frogs (Anura: Ranidae; genus Hylarana). Mol. Phylogenet. Evol. 151, 106899.

Chernomor, O., von Haeseler, A., Minh, B.Q., 2016. Terrace aware data structure for phylogenomic inference from supermatrices. Syst. Biol. 65, 997–1008.

Daane, J.M., Dornburg, A., Smits, P., MacGuigan, D.J., Brent Hawkins, M., Near, T.J., William Detrich Iii, H., Harris, M.P., 2019. Historical contingency shapes adaptive radiation in Antarctic fishes. Nat. Ecol. Evol. 3, 1102–1109.

Dawson, C.E., 1969. A New Seven-Spined Goby, Gobiosoma (Austrogobius) polyporosum, from the Pacific Coast of Panamá. Copeia 1969, 510.

Edwards, S.V., Liu, L., Pearl, D.K., 2007. High-resolution species trees without concatenation. PNAS 104, 5936–5941.

Gill, A.C., and Randall D. Mooi, 2012. Thalasseleotrididae, new family of marine gobioid fishes from New Zealand and temperate Australia, with a revised definition of its sister taxon, the Gobiidae (Teleostei: Acanthomorpha). Zootaxa 3622, 41–52.

Greenfield, D.W., 2017. An overview of the dwarfgobies, the second most speciose coral-reef fish genus (Teleostei: Gobiidae: Eviota). J. Ocean Sci. Found. 29.

Greenfield, D.W., Randall, J.E., 2004. The Marine Gobies of the Hawaiian Islands. Proc. Calif. Acad. Sci. 55, 498–549.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.

Hughes, L.C., Orti, G., Huang, Y., Sun, Y., Baldwin, C.C., Thompson, A.W., Arcila, D., Betancur, R.R., Li, C., Becker, L., Bellora, N., Zhao, X., Li, X., Wang, M., Fang, C., Xie, B., Zhou, Z., Huang, H., Chen, S., Venkatesh, B., Shi, Q., 2018. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. PNAS 115, 6249–6254.

Hughes, L.C., Nash, C.M., White, W.T., Westneat, M.W., 2023. Concordance and discordance in the phylogenomics of the wrasses and parrotfishes (Teleostei: Labridae). Syst. Biol. 72, 530–543.

Ilves, K.L., Torti, D., Lopez-Fernandez, H., 2018. Exon-based phylogenomics strengthens the phylogeny of Neotropical cichlids and identifies remaining conflicting clades (Cichliformes: Cichlidae: Cichlinae). Mol. Phylogenet. Evol. 118, 232–243.

Junier, T., Zdobnov, E.M., 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26, 1669–1670.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Keith, P., Lord, C., Lorion, J., Watanabe, S., Tsukamoto, K., Couloux, A., Dettai, A., 2010. Phylogeny and biogeography of Sicydiinae (Teleostei: Gobiidae) inferred from mitochondrial and nuclear genes. Mar. Biol. 158, 311–326.

Kuang, T., Tornabene, L., Li, J., Jiang, J., Chakrabarty, P., Sparks, J.S., Naylor, G.J.P., Li, C., 2018. Phylogenomic analysis on the exceptionally diverse fish clade Gobioidei (Actinopterygii: Gobiiformes) and data-filtering based on molecular clocklikeness. Mol. Phylogenet. Evol. 128, 192–202.

Lemmon, E.M., Lemmon, A.R., 2013. High-Throughput Genomic Data in Systematics and Phylogenetics. Annu. Rev. Ecol. Evol. Syst. 44, 99–121.

Lemmon, A.R., Emme, S.A., Lemmon, E.M., 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst. Biol. 61, 727–744.

Li, C., Riethoven, J.J., Naylor, G.J., 2012. EvolMarkers: a database for mining exon and intron markers for evolution, ecology and conservation studies. Mol. Ecol. Resour. 12, 967–971.

Li, C., Hofreiter, M., Straube, N., Corrigan, S., Naylor, G.J., 2013. Capturing protein-coding genes across highly divergent species. Biotechniques 54, 321–326.

McCraney, W.T., Thacker, C.E., Alfaro, M.E., 2020. Supermatrix phylogeny resolves goby lineages and reveals unstable root of Gobiaria. Mol. Phylogenet. Evol. 151, 106862.

McCraney, W.T., Thacker, C.E., Faircloth, B.C., Harrington, R.C., Near, T.J., Alfaro, M.E., 2025. Explosion of goby fish diversity at the Eocene-Oligocene transition. Mol. Phylogenet. Evol. 207. https://doi.org/10.1016/j.ympev.2025.10834.

Miller, P.J., Tortonese, E., 1968. Distribution and systematics of the gobiid fish Odondebuenia balearica (Pellegrin Fage). Annali Del Museo Civico Di Storia Naturale Di Genova 77, 342–358.

Miller, E.C., Faucher, R., Hart, P.B., Rincon-Sandoval, M., Santaquiteria, A., White, W.T., Baldwin, C.C., Miya, M., Betancur-R, R., Tornabene, L., Evans, K., Arcila, D., 2024. Reduced evolutionary constraint accompanies ongoing radiation in deep-sea anglerfishes. Nat. Ecol. Evol.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020a. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534.

Minh, B.Q., Hahn, M.W., Lanfear, R., 2020b. New Methods to Calculate Concordance Factors for Phylogenomic Datasets. Mol Biol Evol 37, 2727-2733.Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30, i541–i548.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Patzner, R., Van Tassell, J.L., Kovacic, M., Kapoor, B.G., 2011. The Biology of Gobies. Taylor & Francis, St. Helier, Jersey, British Channel Islands.

Pezold, F., 1993. Evidence for a monophyletic Gobiinae. Copeia 3, 634–643.

Polgar, G., Malavasi, S., Cipolato, G., Georgalas, V., Clack, J.A., Torricelli, P., 2011. Acoustic communication at the water's edge: evolutionary insights from a mudskipper. PLoS One 6, e21434.

Prokofiev, A.M., 2015. An overview of gobies of the subfamily Amblyopinae (Gobiidae) from the Western South China Sea (Vietnam and Indonesia) and Pacific waters of the Philippines. J. Ichthyol. 55, 783–791.

Roa-Varon, A., Dikow, R.B., Carnevale, G., Tornabene, L., Baldwin, C.C., Li, C., Hilton, E. J., 2021. Confronting sources of systematic error to resolve historically contentious relationships: a case study using Gadiform fishes (Teleostei, Paracanthopterygii, Gadiformes). Syst. Biol. 70, 739–755.

Rokas, A., Kruger, D., Carroll, S.B., 2005. Animal evolution and the molecular signature of radiations compressed in time. Science 310, 1933–1938.

Rüber, L., Tassell, J.L., Zardoya, R., 2003. Rapid speciation and ecological divergence in the american seven-spined gobies (Gobiidae, Gobiosomatini) inferred from a molecular phylogeny. Evolution 57, 1584–1598.

Sayyari, E., Mirarab, S., 2016. Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. 33, 1654–1668.

Shen, X.X., Hittinger, C.T., Rokas, A., 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nat. Ecol. Evol. 1, 126.

Siu-Ting, K., Torres-Sanchez, M., San Mauro, D., Wilcockson, D., Wilkinson, M., Pisani, D., O'Connell, M.J., Creevey, C.J., 2019. Inadvertent paralog inclusion drives artifactual topologies and timetree estimates in phylogenomics. Mol. Biol. Evol. 36, 1344–1356.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Thacker, C.E., 2003. Molecular phylogeny of the gobioid fishes (Teleostei: Perciformes: Gobioidei). Mol. Phylogenet. Evol. 26, 354–368.

Thacker, C.E., 2009. Phylogeny of Gobioidei and placement within Acanthomorpha, with a new classification and investigation of diversification and character evolution. Copeia.

Thacker, C.E., 2014. Species and shape diversification are inversely correlated among gobies and cardinalfishes (Teleostei: Gobiiformes). Org. Divers. Evol. 14, 419–436.

Thacker, C.E., 2015. Biogeography of goby lineages (Gobiiformes: Gobioidei): origin, invasions and extinction throughout the Cenozoic. J. Biogeogr. 42, 1615–1625.

Thacker, C.E., Roje, D.M., 2011. Phylogeny of Gobiidae and identification of gobiid lineages. Syst. Biodivers. 9, 329–347.

Tornabene, L., Ahmadia, G.N., Berumen, M.L., Smith, D.J., Jompa, J., Pezold, F., 2013a. Evolution of microhabitat association and morphology in a diverse group of cryptobenthic coral reef fishes (Teleostei: Gobiidae: Eviota). Mol. Phylogenet. Evol. 66, 391–400.

Tornabene, L., Chen, Y., Pezold, F., 2013b. Gobies are deeply divided: phylogenetic evidence from nuclear DNA (Teleostei: Gobioidei: Gobiidae). Syst. Biodivers. 11, 345–361.

Tornabene, L., Van Tassell, J.L., Gilmore, R.G., Robertson, D.R., Young, F., Baldwin, C.C., 2016. Molecular phylogeny, analysis of character evolution, and submersible collections enable a new classification of a diverse group of gobies (Teleostei: Gobiidae:Nes subgroup), including nine new species and four new genera. Zool. J. Linn. Soc. 177, 764–812.

Tornabene, L., Deis, B., Erdmann, M.V., 2018. Evaluating the phylogenetic position of the goby genus Kelloggella (Teleostei: Gobiidae), with notes on osteology of the genus and description of a new species from Niue in the South Central Pacific Ocean. Zool. J. Linn. Soc. 183, 143–162.

Van Tassell, J.L., Kramer, A., 2014. A new species of Didogobius (Teleostei: Gobiidae) from the Canary Islands. Zootaxa 3793, 453–464.

Van Tassell, J.L., Tornabene, L. and Taylor, M, 2011. A History of Gobioid Morphological Systematics, in: R.A. Patzner, J.L.V.T., M. Kovacic and B.G. Kapoor (Ed.), The Biology of Gobies. CRC Press, Oxford & IBH Publishing Co., New Delhi.

Winterbottom, R., 1990. The Trimmatom nanus species complex (Actinopterygii, Gobiidae): phylogeny and progenetic heterochrony. Syst. Zool. 39, 253.

Winterbottom, R.H., Hoese, D.F., 1988. A new genus and four new species of fishes from the Indo-West Pacific (Pisces; Perciformes; Gobiidae), with comments on relationships. Royal Ontario Museum, Life Sciences Occasional Paper 37, 1–17.
Winterbottom, R., Alofs, K.M., Marseu, A., 2011. Life span, growth and mortality in the western Pacific goby Trimma benjamini, and comparisons with T. nasa. Environ. Biol. Fishes 91, 295–301.
Yuan, H., Jiang, J., Jimenez, F.A., Hoberg, E.P., Cook, J.A., Galbreath, K.E., Li, C., 2016. Target gene enrichment in the cyclophyllidean cestodes, the most diverse group of tapeworms. Mol. Ecol. Resour. 16, 1095–1106.
Yuan, H., Atta, C., Tornabene, L., Li, C., 2019. Assexon: assembling exon using gene capture data. Evol. Bioinformatics Online 15, 1176934319874792.