Molecular Phylogenetics and Evolution 189 (2023) 107935

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Interrelationships and biogeography of the New World pufferfish genus *Sphoeroides* (Tetraodontiformes: Tetraodontidae) inferred using ultra-conserved DNA elements

Elizabeth P. Hunt [a,*], Stuart C. Willis [a,b], Kevin W. Conway [c], David S. Portnoy [a]

[a] *Department of Life Sciences, Texas A&M University – Corpus Christi, 6300 Ocean Dr., Corpus Christi, TX 78412, USA*
[b] *Columbia River Inter-Tribal Fish Commission - Hagerman Genetics Lab, 3059-F National Fish Hatchery Road, Hagerman, ID 83332, USA*
[c] *Department of Ecology and Conservation Biology and Biodiversity Research and Teaching Collections, Texas A&M University, 534 John Kimbrough Blvd., College Station, TX 77843, USA*

ARTICLE INFO

Keywords:
Tetraodontidae
Molecular phylogenetics
Sphoeroides
Ultra-conserved DNA elements
Biogeography

ABSTRACT

Colonization of the New World by marine taxa has been hypothesized to have occurred through the Tethys Sea or by crossing the East Pacific Barrier. To better understand patterns and timing of diversification, geological events can be coupled with time calibrated phylogenetic hypotheses to infer major drivers of diversification. Phylogenetic relationships among members of *Sphoeroides*, a genus of four toothed pufferfishes (Tetraodontiformes: Tetraodontidae) which are found nearly exclusively in the New World (eastern Pacific and western Atlantic), were reconstructed using sequences from ultra-conserved DNA elements, nuclear markers with clear homology among many vertebrate taxa. Hypotheses derived from concatenated maximum-likelihood and species tree summary methods support a paraphyletic *Sphoeroides,* with *Colomesus* deeply nested within the genus. Analyses also revealed *S. pachygaster,* a pelagic species with a cosmopolitan distribution, as the sister taxon to the remainder of *Sphoeroides* and recovered distinct lineages within *S. pachygaster,* indicating that this cosmopolitan species may represent a species complex. Ancestral range reconstruction may suggest the genus colonized the New World through the eastern Pacific before diversifying in the western Atlantic, though date estimates for these events are uncertain due to the lack of reliable fossil record for the genus.

1. Introduction

There has been a considerable amount of research surrounding the colonialization of the New World by marine taxa, either through the Tethys Sea or the eastern Pacific before the closing of the Isthmus of Panama (e.g., Barber and Bellwood, 2005; Lin & Hastings, 2013; Floeter et al., 2008; Veron, 1995). The Tethys Sea began closing around 18 mya, with the final closuring occurring around 12 mya (Adams et al., 1999), and implications of the event can be seen in biogeographic patterns of both marine and freshwater taxa (Bellwood and Wainwright, 2002; George, 2005; Hrbek and Meyer, 2003; Nikulina et al., 2007). Floeter et al. (2008) analyzed data collected from 25 areas of the Atlantic to investigate how and when colonization of the Atlantic by reef fishes took place. They inferred that for some genera, geographic areas with heightened levels of species richness also appear in positions near the base of phylogenetic reconstructions, indicative of specific patterns of diversification. For example, the genera *Diplodus*, *Parabelnnius*, and *Scartella* are more speciose in the eastern Atlantic and available phylogenetic hypotheses indicate an eastern Atlantic origin with dispersal into the western Atlantic for each (Floeter et al., 2008; Summerer et al., 2001). A recent biogeographic study of the Syngnatharia (Teleostei) also concluded that the genus *Dactylopterus* colonized the western Atlantic from the eastern Atlantic, prior to the closure of the Tethys Sea, while goatfishes in the genera *Pseudupeneus* and *Mulloidichthys* likely colonized the western Atlantic from the eastern Pacific (Santaquiteria et al., 2021). Movement from the eastern Atlantic to the western Atlantic requires long-distance dispersal events which, while rare, can still be observed on contemporary timescales: for example, the surgeon fish *Acanthurus monroviae* is primarily distributed along the west coast of Africa but is occasionally found along the coast of Brazil (Luiz-Júnior et al., 2004).

The hypothesized alternative mode for colonization of the New World by reef fishes is via the eastern Pacific, with dispersal occurring

* Corresponding author.
E-mail address: Elizabeth.Hunt@tamucc.edu (E.P. Hunt).

https://doi.org/10.1016/j.ympev.2023.107935
Received 6 March 2023; Received in revised form 27 September 2023; Accepted 28 September 2023
Available online 29 September 2023

across the East Pacific Barrier and into the Caribbean prior to the closing of the Isthmus of Panama (Barber and Bellwood, 2005). In this scenario, the secession of geneflow caused by the closure (final closure ~ 3 mya; Bacon et al., 2015; O'Dea et al., 2016) would have then been followed by diversification on either side of the Isthmus and in the New World Tropics (Barber and Bellwood, 2005). This dynamic is supported by the observation of sister species distributed on either side of Isthmus across marine taxa, including invertebrates and fishes (Floeter et al., 2008; Gomon, 1997; Lessios et al., 2001; Santaquiteria et al., 2021). Because the closure was a gradual process, events leading to diversification likely occurred at different points in time, both before and after the closure (Thacker, 2017). Time-calibrated phylogenetic reconstructions lend insight into the most likely colonization and diversification scenarios in the New World, especially where they can then be compared to historic geologic processes. Lineages which are relics of dispersal from the Tethys Sea would be hypothesized to be older than 15 million years, while those from the eastern Pacific could be at least 3 million years old, often with geminate pairs found on either side of the Isthmus (Barber and Bellwood, 2005).

One marine group with high species diversity in the New World and that is appropriate for testing these hypotheses is pufferfishes in the genus *Sphoeroides* (Tetraodontidae). However, there is limited literature on the biogeography of tetraodontiform fishes, with previous work focusing mainly on Ostraciidae and Triacanthidae (Klassen, 1995; Santini and Tyler, 2002), and comparatively little study of the Tetraodontidae (Araujo et al. 2023). Pufferfishes (Actinopterygii: Tetraodontiformes: Tetraodontidae) are a charismatic group of fishes found circumglobally in tropical and subtropical waters, most notable for their ability to inflate their bodies when threatened, by pumping water into an expandable stomach (Brainerd, 1994, Wainwright and Turingan 1997). There are currently 192 described species of Tetraodontidae in 28 genera, with the majority of the species inhabiting marine environments and 12 species that are known to inhabit brackish and/or freshwater environments (Matsuura, 2015; Nelson et al., 2016; Frick et al. 2023a). Characteristics that unite the family include: fusion of the teeth into four tooth plates, lack of fusion between the premaxillae and dentaries at the midline, absence of ribs and intermuscular bones, scaleless bodies with the exception of small lappets or prickles, and an ability to inflate the body (Matsuura, 2015; Matsuura and Tyler, 1997; Fraser et al., 2012; Thiery et al., 2017). The family first appears in the fossil record in the Eocene (~50 mya) but seems to have begun to diversify during the Oligocene (Bannikov and Tyler, 2008; Benton, 1993; Santini and Tyler, 2002). Early systematic work on the Tetraodontidae in the 18th and 19th centuries resulted in a complex nomenclatural history driven by the lack of suitable specimens and failure to recognize important anatomical characters, as well as poor species descriptions (Shipp, 1974). Currently, extant species of tetraodontids are classified into two subfamilies, Tetraodontinae and Canthigastrinae, with the former comprising 154 species in 27 genera and the latter containing 39 species in one genus (Fricke et al., 2023a). The subfamily Tetraodontinae is united by a conspicuous lateral line (usually), a round body in cross section, one or two conspicuous nostrils on each side of the head, and a gill opening usually proceeding below the midpoint of the pectoral fin (Tyler, 1980). While some genera in the subfamily have been given significant attention due to their economic importance and suitability of their small, simple genomes to biomedical study (e.g., *Takifugu*; Aparicio et al., 2002; Gao et al., 2014; Liu et al., 2021; Song et al., 2001; Yamanoue et al., 2009), other genera, such as *Sphoeroides,* are not well-studied.

The distribution of *Sphoeroides* is interesting because most of the 22 described species are found in the western Atlantic (10) and eastern Pacific oceans (10), with the exception of *S. pachygaster*, which is distributed circumglobally in tropical and subtropical waters, and *S. marmoratus*, which is found in the eastern Atlantic (Robertson and Allen, 2015; Shipp, 1974; Walker and Bussing, 1996). Most species in the genus inhabit shallow waters in bays, estuaries, and reef areas, with the exception of *S. pachygaster*, which may be found at depths as great as ~ 400 m (Shipp, 1974, 2002). Although the current distributions highlight the richness of this genus in the New World, the means by which *Sphoeroides* colonized and diversified in this area remains unknown.

Interrelationships among species in the genus *Sphoeroides* have been assessed but mostly in a limited capacity by inclusion of a small number of species in molecular studies focused on family relationships within the order Tetraodontiformes (e.g., Amaral et al., 2013; Holcroft, 2005; Igarashi et al., 2013; Santini et al., 2013a, 2013b; Yamanoue et al., 2008). These studies included one (Santini et al., 2013a) to ten (Santini et al. 2013b) species of *Sphoeroides* and utilized one (Amaral et al, 2013) to 22 gene regions (two mitochondrial and twenty nuclear; Santini et al., 2013a) in phylogenetic reconstruction, and did not resolve comparable relationships among members of the genus due to incomplete and non-overlapping sampling. However, several notable similarities among these studies include the recovery of a paraphyletic *Sphoeroides* based on the placement of the Amazonian (exclusively freshwater) genus *Colomesus*, the placement of *S. pachygaster* at the base of the *Sphoeroides* clade, as well as a sister group relationship between *S. annulatus* (eastern Pacific) and *S. testudineus* (western Atlantic; Amaral et al., 2013; Igarashi et al., 2013; Santini et al. 2013b).

There are three studies in which relationships among the genus were looked at in greater depth. The first is a phylogenetic examination of *Sphoeroides* by Shipp (1974), where morphological and meristic characters, diet, and distributional records, but not molecular data, were assessed. This study included 13 of the 22 currently valid species of *Sphoeroides* and stressed variation in characters within species that occurred with geography. Notably, Shipp (1974) described population level phenotypic differences between eastern and western Atlantic populations of the cosmopolitan *S. pachygaster,* with differences in coloration and pattern, pectoral ray counts, eye size, and width of the interorbit, and inferred that there was uncertainty in species designation that required further inquiry. In addition, a study conducted by Troyer et al. (2022), examined paleoclimatic effects on body size of Tetraodontiformes. Their study used 1,103 genomic markers as well as 210 morphological characters to resolve relationships within the tetraodontids, including ten species of *Sphoeroides*. Five different analyses were performed (a likelihood approach, a Bayesian approach, both with and without the superfamily Plectocretacicoidea, and a time calibrated approach) and the relationships recovered among *Sphoeroides* changed with analysis type, as not all of the same taxa were included in each analysis. However, it is important to note that some analyses did share patterns with previous work, including the sister relationship between *S. annulatus* and *S. testudineus* and *S. pachygaster* as the sister taxon to all other *Sphoeroides*. Most recently, Araujo et al. (2023) assessed the phylogenetic relationships of *Sphoeroides* using a single mitochondrial gene (COI), proposing biogeographic hypotheses that explain how the genus diversified in the New World. Most notably, their study described a new species of *Sphoeroides* off the coast of Brazil (*S. camila*) and presented results that showed the genus *Colomesus* nested within *Sphoeroides*. Their biogeographic analysis suggested that *Sphoeroides* originated in the Caribbean and then colonized the eastern Pacific and eastern Atlantic, with an additional crossing from the eastern Atlantic to the western Atlantic off the coast of Brazil.

To better understand the evolutionary history of this genus, the phylogenetic relationships among species of *Sphoeroides* were investigated using ultra-conserved DNA elements (UCE). Compared to traditional phylogenetic approaches that use Sanger sequencing to characterize a few genes, UCEs provide greater information content for phylogenetic reconstruction due to the large number of loci that can be analyzed simultaneously (Faircloth et al., 2013; Gilbert et al., 2015). Another advantage of using UCEs is that these loci are conserved across evolutionarily distant taxa (Bejerano, 2004), allowing for resolution of phylogenetic relationships across different phylogenetic timescales (Faircloth et al., 2012; Gilbert et al., 2015). Both *Colomesus* and a broad

geographic sampling of *S. pachygaster* (western Atlantic, eastern Atlantic and Pacific) were included in this study to address potential problems with systematics and nomenclature noted by previous studies (Amaral et al., 2013; Igarashi et al., 2013; Santini et al. 2013b; Shipp, 1974; Araujo et al. 2023). Finally, ancestral range construction was used to understand patterns of colonization and diversification, adding to the growing body of literature on marine fish diversification in the New World.

2. Methods

2.1. Sampling

Tissues for this study were obtained from 45 individuals that represented 13 of the 22 species of *Sphoeroides* and are associated with museum-vouchered specimens. Additionally, tissues from species in two other genera, *Colomesus asellus*, *Lagocephalus laevigatus,* and *Balistes vetula,* were included. All specimens used in this study and their voucher information can be found in Table 1. To account for variation within *S. pachygaster* due to its circumglobal distribution, specimens from multiple ocean basins were included. Fin clips or muscle tissue from each specimen were stored in 95% non-denatured EtOH or DMSO-EDTA-NaCl buffer (Seutin et al., 1991) at room temperature until time of extraction.

Table 1
Available museum voucher numbers for samples used in the study. Sequence data can be found at NCBI, BioProject ID PRJNA821242. Museum collection abbreviations follow Sabaj (2020).

Specimen	Voucher	Latitude	Longitude
B. vetula	NA	NA	NA
C. asellus 1	AUM 3232	4.66765	−58.67883
C. asellus 2	AUM 3230	3.89000	−59.29370
L. laevigatus	TCWC 17524.01	28.43150	−93.44410
S. annulatus 1	SIO 004–85	25.80000	−111.33333
S. annulatus 2	SIO 007–91	23.38333	−110.20000
S. dorsalis 1	TCWC 17377.01	30.13110	−86.57610
S. dorsalis 2	TCWC 17378.02 A	29.50830	−87.11450
S. dorsalis 3	TCWC 17378.02C	29.50830	−87.11450
S. greeleyi 1	TCWC 17584.01	10.69950	−61.66476
S. greeleyi 2	TCWC 17587.01	10.69950	−61.66476
S. greeleyi 3	TCWC 17588.01	10.69950	−61.66476
S. greeleyi 4	TCWC 17580.01	10.69950	−61.66476
S. lispus 1	SIO 09–355	24.48333	−110.39500
S. lobatus 1	SIO 007–57	23.15000	−109.41667
S. lobatus 2	SIO 007–17	23.15000	−109.41667
S. lobatus 3	SIO 007–98	22.86667	−109.88333
S. maculatus 1	KU 1136	34.96670	−75.36670
S. maculatus 2	KU 1225	35.61670	−75.46670
S. maculatus 3	KU 1226	35.61670	−75.46670
S. marmoratus 1	USNM 093	15.75330	–23.09080
S. marmoratus 2	USNM 094	15.75330	–23.09080
S. nephelus 1	KU 5432	30.25886	−88.11328
S. pachygaster BE1	TCWC 19329.01	17.24376	−87.97325
S. pachygaster BE2	TCWC 19330.01	18.16466	−87.81406
S. pachygaster BE3	NA	NA	NA
S. pachygaster CV1	USNM 405,072	15.44330	–23.13250
S. pachygaster CV2	USNM 405,073	15.44330	–23.13250
S. pachygaster CV3	USNM 405,238	16.75580	−24.95830
S. pachygaster CV4	USNM 405,237	16.75580	−24.95830
S. pachygaster JP1	NA	39.68603	142.04998
S. pachygaster JP2	NA	39.68603	142.04998
S. pachygaster JP3	NA	39.68603	142.04998
S. pachygaster US	MCZ 171,869	35.78859	−74.92785
S. parvus 1	TCWC 17325.01 A	28.54340	−91.55910
S. parvus 2	TCWC 17307.02 A	28.38840	−94.41650
S. parvus 3	TCWC 17458.03 A	28.38330	−90.51670
S. parvus 4	TCWC 17458.03B	28.38330	−90.51670
S. parvus 5	TCWC 17458.03C	28.38330	−90.51670
S. spengleri 1	TCWC 17340.01B	29.12500	−84.08680
S. spengleri 2	TCWC 17340.01C	29.12500	−84.08680
S. testudineus 1	TCWC 17573.01	10.45536	−61.54018
S. testudineus 2	TCWC 17574.01	10.45536	−61.54018
S. testudineus 3	TCWC 17575.01	10.45536	−61.54018
S. trichocephalus 1	CIRUV 690	3.63330	−77.20000
S. trichocephalus 2	CIRUV 691	3.63330	−77.20000
S. trichocephalus 3	CIRUV 692	3.63330	−77.20000

2.2. Library preparation

Samples were processed using a modified in-solution hybrid enrichment protocol targeting 500 UCE loci (Faircloth et al., 2012). Briefly, total genomic DNA was extracted using Mag-Bind® Blood & Tissue DNA kits (Omega Bio-Tek) or phenol–chloroform extraction (Sambrook et al., 1989) and then quantitated using the AccuBlue High Sensitivity dsDNA Quantitation kit (Biotium) on a fluorometer. DNA was standardized to 8 ng/μl in 130 μl of 1X TE buffer before being randomly sheared to a length of ~ 800 bp (+/- 200 bp) using a M220 Focused-ultrasonicator (Covaris). After sonication, samples were purified using Ampure XP (Agencourt) in a 0.6X concentration to eliminate DNA fragments less than ~ 400 bp in length and electrophoresed on a 2% agarose gel to visually confirm the size distribution of DNA fragments. Samples were subsequently end repaired and blunt-end ligated to universal double-stranded oligonucleotide adapters using the NEB Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, NEB). Samples were amplified using polymerase chain reaction (PCR) and 'Adapterama' iTru5 and iTru7 indexed PCR primers (Glenn et al., 2016). The 50 μl PCR reactions contained water, 15 uL ligated DNA, NEBNext Ultra II Q5 Master Mix, and 50 pmol of both iTru5 and iTru7 indexed primers. PCR cycle conditions were as follows: initial denaturation at 98 °C for 45 s, followed by 18 cycles of 98 °C for 15 s, 60 °C for 30 s, and 72 °C for 75 s, followed by a final extension at 72 °C for 5 min. PCR reactions were subsequently cleaned, quantified and pooled for target enrichment of UCEs using a set of 2,001 baits designed to capture 500 UCE loci in actinopterygian fishes. Enrichment followed manufacturers protocols (MYbaits_Actinopts-UCE-0.5Kv1, Daicel Arbor Biosciences) and was followed by post-hybridization amplification using PCR. The 50 μl, post-enrichment PCR reactions contained 1X Phusion High Fidelity Buffer, 1.5 mM $MgCl_2$, 0.25 mM of each dNTP, 25 pmol of both P5 and P7 primers, 2.0 U/uL of NEB Phusion Taq, and 15 uL DNA. PCR cycle conditions were as follows: initial denaturation at 95 °C for 5 min, then 30 s at 98 °C, followed by 16 cycles of 98 °C for 15 s, 65 °C for 30 s, and 72 °C for 90 s, followed by a final extension at 72 °C for 5 min. To ensure that UCE enrichment was successful, pooled samples underwent quantitative PCR (qPCR) using GoTaq qPCR master mix (Promega) and a StepOnePlus Real-Time PCR System (Applied Biosystems, inc.). The 25 μl qPCR reactions contained water, GoTaq qPCR Master Mix, 25 pmol of qPCR primer, and 2 uL of DNA standardized to 1 ng/uL. qPCR conditions were as follows: initial denaturation at 95 °C for 5 min, followed by 40 cycles of 95 °C for 10 s, 60 °C for 20 s, and 72 °C for 30 s, followed by a final a cooling step at 40 °C for 10 s. Enrichment was considered successful when there was at least a 50X enrichment of target product as compared to pre-enriched pooled samples (Faircloth et al., 2012). When the 50X enrichment threshold was not met, samples were re-enriched following the aforementioned protocol. As a first step, 22 individuals across species were quantified, standardized, and pooled for paired-end (2 × 300 bp) sequencing on an Illumina MiSeq (Illumina, San Diego, CA). The resulting MiSeq reads were used to create a reference of UCE loci. Subsequent libraries were sequenced on an Illumina HiSeq 4000 (Illumina, San Diego, CA) and reads were mapped to the UCE reference.

2.3 Reference assembly and UCE identification

Demultiplexed reads from the MiSeq were trimmed for quality and adapter contamination using Trimmomatic (Bolger et al., 2014) and contigs assembled using Trinity (Grabherr et al., 2011). Assembled contigs were then matched to UCE loci using a modified pipeline that included utilities from PHYLUCE (Faircloth, 2016). Consensus sequences of contigs were clustered at an 80% similarity threshold using

CD-hit to reduce redundancy (Li and Godzik, 2006), then re-introduced to PHYLUCE for identification by probe similarity. A database was created that contained records of UCE loci present across individuals. FASTA sequences for each locus were extracted using the PHYLUCE *probe.matches.sqlite* database. For each locus, a multiple sequence alignment across individuals was created using MAFFT (Katoh and Standley, 2013), alignments were checked by eye, and a consensus sequence of the alignment created using EMBOSS with a strict consensus (Rice et al., 2000). The final consensus sequences for each locus were concatenated into a single reference for subsequent read mapping using Bowtie2 (Langmead and Salzberg, 2012).

Standard UCE bioinformatic processing involves *de novo* assembly of reads using the PHYLUCE pipeline (Faircloth, 2016), after assembling reads using a transcriptome assembler (e.g., Trinity). When employing this approach, the resulting assemblies were smaller than anticipated, with most of the fragments concentrated on the less variable UCE core region. This led to a reduction of the number of recovered UCE loci within individuals as well as a poor reconstruction of the more variable UCE flanking regions. An alternative strategy was, therefore, employed because it was found to recover target loci more consistently, across a greater number of species, with the recovery of longer sequences containing the flanking regions than the standard approach. Briefly, demultiplexed HiSeq reads were mapped directly to the MiSeq reference to assemble UCE sequences for each individual using Bowtie2 (Langmead and Salzberg, 2012). The resulting BAM file contained multiple reads at a given locus for an individual. Consensus sequences of mapped reads at each locus per individual were generated using SAMtools while applying filters for both minimum mapping quality and minimum base quality to ensure high quality reads and bases were retained (Danecek et al., 2021).

After loci were subsequently identified in each individual using similar steps as above, multiple sequence alignments for each locus were generated using the L-INS-i algorithm of MAFFT. Alignments for each locus were trimmed using the program trimAl with a 30% gap threshold to eliminate poor-quality and information-deficient bases at the ends of alignments (Capella-Gutierrez et al., 2009). Final datasets were filtered for different thresholds of missing data using the *phyluce_align_get_only_loci_with_min_taxa* script in PHYLUCE, resulting in five different datasets: "complete" (only including loci shared among all 45 individuals), ≤10% missing data allowed, ≤ 25% missing data allowed, ≤ 50% missing data allowed, and the "incomplete" dataset, which included any locus present across three or more individuals.

2.4. Phylogenetic analysis

Maximum-likelihood (ML) analysis was conducted on individual gene trees, using only those taxa present for each locus, as well as a single, concatenated sequence for each individual in each dataset (complete, ≤ 10% missing, ≤ 25% missing, ≤ 50% missing and incomplete), using RAxML v8.02 (Stamatakis, 2014), allowing for an assessment of the effects of missing data on the topology and node support. Trees were generated using the generalized time reversible (GTR) + Gamma substitution model in RAxML. To determine support for nodes, bootstopping in RAxML was employed. This method computes bootstrap thresholds at the time of analysis and determines that enough bootstrap replicates have been generated for sufficient support when the difference in branch support is smaller than 3% for at least 99% of the permutations (Pattengale et al., 2010; Stamatakis, 2014). Concatenated analyses were run independently five separate times using a different random starting seed, to avoid issues resulting from analysis getting stuck in a local optimum. The best trees (as determined in RAxML; Langmead and Salzberg, 2012) for each run were compared to one another to look for consensus in topologies and support values.

Estimating species trees solely based on concatenation of many loci can lead to over inflated support values for incorrect relationships within the tree (Kubatko and Degnan, 2007). To circumvent this, individual gene trees for each dataset were estimated using RAxML v.8.02 and each was analyzed in a "summary coalescent" framework using ASTRAL-III (Zhang et al., 2017). Simultaneous computation of gene trees and species tree is computationally demanding for high numbers of loci, so summary methods, such as that implemented in ASTRAL-III, are statistically consistent and appropriate for the type of phylogenomic data generated in this study (Mirarab et al., 2014; Nute et al., 2018; Sayyari and Mirarab, 2016). Support at each node was determined by the local posterior probablity (LPP) in ASTRAL-III. Local posterior probabitly is calculated for unrooted genetrees in a Bayesian posterior probabitly framework. The LPP is derived from gene tree quartet frequencies and quartet support values for the differening topologies around the quartet (Sayyari and Mirarab, 2016). This method has been shown to be accurate and consistent with the multispecies coalescent (Sayyari and Mirarab, 2016).

2.5. Ancestral range estimation

The resulting ASTRAL-III phylogeny from the complete dataset was selected for use in ancestral range estimation in BioGeoBEARS (Matzke, 2013a). This biogeographic analysis assumes the input tree is ultrametric, but the ability to time calibrate the inferred topology was constrained by the lack of reliable fossil data for the genus, whereas the use of geological calibrations would be circular with the objective of assessing alternate biogeographic hypotheses. As such, the R package APE was used to adjust the scaling of the tree so that all of the tips were the same height above the root node (Paradis et al., 2004) and the root age was arbitrarily set to one million years ago.

To perform the biogeographic analysis each taxon was assigned to one of five geographic regions: western Atlantic, eastern Atlantic, eastern Pacific, western Pacific, or the Amazon River. Species distribution area was determined using Fishbase (Froese and Pauly, 2022) and Shipp (2002). Several models were evaluated including DEC, DIVALIKE, and BayAreaLike analysis with and without the *J* parameter (Matzke, 2013a). In BioGeoBEARS each of these analyses are likelihood interpretations of previous biogeography reconstruction programs such as dispersal-extinction cladogenesis (DEC; LAGRANGE; Ree and Smith, 2008), parsimony dispersal-vicariance analysis (DIVA; Ronquist, 1997), and Bayesian estimates of ancestral ranges (BayArea; Landis et al., 2013). The DEC has two free parameters *d* (dispersal or range expansion) and *e* (extinction or range contraction), while also containing one fixed parameter *c* (cladogenesis; Matzke, 2013a). The DIVA model varies slightly from the DEC model by allowing for a variety of vicariance scenarios without allowing sympatric speciation (Matzke, 2013a). Finally, BayAreaLike uses a Bayesian approach to sample historic geographic ranges along phylogenetic branches without any specific range evolutionary process at cladogenesis (Matzke, 2013a). However, none of these models are able to account for founder-events leading to speciation. The inclusion of the *J* parameter in BioGeoBEARS allows for a daughter lineage to make a jump outside of the ancestral range, imitating a founder-speciation event (Matzke, 2013b). A more in-depth explanation of all the models and parameters included in BioGeoBEARS can be found in Matzke (2013a; 2013b).

3. Results

3.1. UCE datasets

A total of ~ 528 million demultiplexed HiSeq reads were generated with an average of ~ 5.5 million reads per sample. Demultiplexed sequence reads can be accessed through NCBI BioProject ID PRJNA821242. After trimming, an average of 52% of reads mapped to the UCE reference with a mean coverage across all loci of 862X (min = 16X, max = 2,717X). The frequency of variable sites across taxa followed the expected distribution for UCEs: with the core region of the UCE showing lower variability among taxa as compared to the flanking

regions, where more variable sites are expected to be found (Supplemental Fig.1; Faircloth et al. 2012). After pipeline filtering and UCE identification, 247, 429, 433, 436, and 437 loci were recovered for the complete, 10% missing data, 25% missing data, 50% missing data, and incomplete data sets, respectively. This resulted in concatenated datasets with lengths of 234,719 bp, 379,782 bp, 381,326 bp, 382,164 bp, and 382,425 bp for the complete, 10% missing data, 25% missing data, 50% missing data, and incomplete data sets, respectively. Final alignment information for the concatenated dataset for each missing data threshold including number of loci, alignment length, parsimony uninformative sites (PUI), and number of parsimony informative sites (PI) for each dataset are reported in Table 2.

3.2. Phylogenetic reconstruction

Resulting phylogenies for each dataset and analysis type contained 47 individuals representing 13 species of *Sphoeroides*, *Colomesus asellus* (n = 2), *Lagocephalus laevigatus* (n = 1), and *Balistes vetula* (n = 1). Topologies were congruent across all phylogenetic hypotheses generated using ML analysis of concatenated loci across all missing data thresholds, resolving six notable monophyletic groups. Nodal support varied across missing data thresholds, with most nodes well supported (ML bootstrap > 85%), although several terminal nodes had moderate to low support (Fig. 1A, Supplemental Figs. 2–5).

Within *Sphoeroides, S. pachygaster* formed a well-supported monophyletic group (A) in a sister relationship with all remaining species. Within this group, samples collected from the western Atlantic (off the coast of Virginia and Belize) grouped together and were sister to samples collected from the eastern Atlantic (off the coast of Cape Verde) and the western Pacific (Japan). *Colomesus asellus* and *S. greeleyi* also formed a well-supported monophyletic group (B) that was sister to the remaining species of *Sphoeroides*. Group C contained three species: *S. lispus*, *S. annulatus*, and *S. testudineus*. Group D contained three species: *S. dorsalis, S. spengleri, and S. marmoratus.* Group E solely contained *S. trichocephalus* and Group F contained *S. lobatus, S. maculatus*, *S. nephelus*, and *S. parvus*. Within this group, *S. nephelus* was resolved as sister to *S. maculatus.* Most of these groups received 100 % bootstrap support across analyses with varying levels of missing data, with the exception of Group F, which had at least 98 % support in each analysis.

The species tree analyses using ASTRAL yielded the same topologies as the concatenated ML analysis, with most of the nodes in the trees being fully supported (Fig. 1B, Supplemental Figures 6–9). Notably, the only instance of a node not being fully supported was in the analysis of the complete dataset (zero missing data; Fig. 1B), wherein the local posterior probability support dropped to 0.87 for a sister relationship between *S. maculatus* and *S. nephelus*, while all other species trees exhibited complete support for this relationship. Similarly, bootstrap support values for this node in the concatenated analyses were higher for data sets that allowed missing data (Supplemental Figs. 2–5). This suggests that despite increasing the overall amount of missing data, the additional gene trees (or bp of data for concatenated analyses) contained important coalescent information that facilitated greater resolution for this relationship.

3.3. Ancestral range estimation

DEC, DIVALIKE, and BAYAREALIKE models with and without the jump parameter (*J*) were tested using BioGeoBEARS to estimate ancestral ranges for the species of *Sphoeroides*. Comparison of the six models using AIC indicated the DIVALIKE + J to be the best fit model for the data (AIC = 50.45; Table 3). The DEC + J and BAYAREALIKE + J models had only marginally worse (higher) AIC values than the chosen model and cannot be completely ruled out, but their resulting ancestral range reconstructions were identical to those of the selected model. Results of the DIVELIKE + J model can be found in Fig. 2. Biogeographic reconstruction supported an eastern Pacific origin of the *S. annulatus-S. lobatus* group with diversification advancing eastward into the western and eastern Atlantic. However, three of the deepest nodes in the tree, including one at the base of Group A (*S. pachygaster*), did not show majority support for any one geographic range. Additionally, the reconstruction at the node shared by *S. greeleyi* and *S. asellus* was ambiguous but slightly favored an Amazon River origin.

4. Discussion

The current study focused on understanding the relationships within the genus *Sphoeroides* by generating data for 13 of the 22 species using UCEs and represents the most comprehensive molecular dataset produced for this taxon to date. Phylogenetic hypotheses generated herein recovered the same topologies in all analyses irrespective of analysis type or missing data allowance. Individuals of *Colomesus* were nested within *Sphoeroides* with high support values, deeming *Sphoeroides* paraphyletic as currently recognized. The cosmopolitan species *S. pachygaster* was resolved as a monophyletic group and sister to the remainder of *Sphoeroides* (including *Colomesus*). Within that group, western Atlantic *S. pachygaster* were resolved as divergent from eastern Atlantic and western Pacific *S. pachygaster*, which may be indicative of a species complex. While taxon sampling in this study was incomplete, which may change interpretation of relationships among species, divergence times of specific taxa, and biogeographic inferences, the current study encompasses taxa included in previous works, and many of the results obtained were consistent with previous work.

The current study, which involved a larger nuclear dataset, resolved *Colomesus* nested deeply within *Sphoeroides,* consistent with previous research that used few molecular markers (Amaral et al., 2013; Igarashi et al., 2013, Araujo et al. 2023) strengthening the argument that *Colomesus* should be subsumed into *Sphoeroides*. *Colomesus* currently contains three species, two of which (*C. asellus* and *C. tocantinensis*) inhabit freshwater and represent the only species of the Tetraodontidae found in the freshwaters of South America (Amaral et al., 2013; Reis et al., 2003). *Colomesus asellus,* the first species of *Colomesus* to be recognized, was originally described as a member of *Tetraodon* (Bloch and Schneider 1801) but was subsequently transferred to a separate genus (Gill 1884). Shipp (1974) noted that *Colomesus* was morphologically more similar to *Sphoeroides* and *Lagocephalus* than to *Tetraodon*, consistent with recent molecular findings. *Colomesus* likely arose from a marine *Sphoeroides* "like" ancestor as hypothesized by Fraser-Brunner (1943) and subsequently invaded the Amazon River. There have been multiple invasions into the Amazon River by marine lineages, including stingrays, flatfishes, and anchovies, resulting in diversification in freshwater (Bloom and Lovejoy, 2011, 2017; Lovejoy et al., 1998), occurring across geological time from the Eocene to the Pliocene (Bloom and Lovejoy, 2011, 2017; Lovejoy et al., 1998). Bloom and Lovejoy (2017) hypothesized that *Colomesus* invaded freshwater ~ 9.4–16.4 mya, but in that analysis assumed that *Colomesus* was sister to *Sphoeroides*. In the results of the analysis presented herein, there was a slightly higher probability that the ancestor of *C. asellus* and *S. greeleyi* inhabited the Amazon River than a marine coastal area, but this may be an artifact of analysis. BioGeoBEARS determines ancestral ranges based on user input species distribution and results are influenced by the defined distributions. For example, when species of *Sphoeroides* were assigned to either "marine" or "freshwater" the results indicate that *C. asellus* is derived from a marine ancestor (Supplemental figure 14). Given these trends there is no reason to expect that *Colomesus* is an exception to the currently known patterns of freshwater invasions by marine fishes into the Amazon basin. However, based on phylogenetic hypotheses presented here as well as others (i.e., Amaral et al., 2013; Igarashi et al., 2013, Araujo et al. 2023), it is clear that *Colomesus* should be considered a junior synonym of *Sphoeroides*.

Similar to previous studies, *S. pachygaster* was resolved as sister to the remaining species of *Sphoeroides* (Amaral et al., 2013; Igarashi et al., 2013; Santini et al., 2013b; Troyer et al., 2022; Yamanoue et al., 2008;

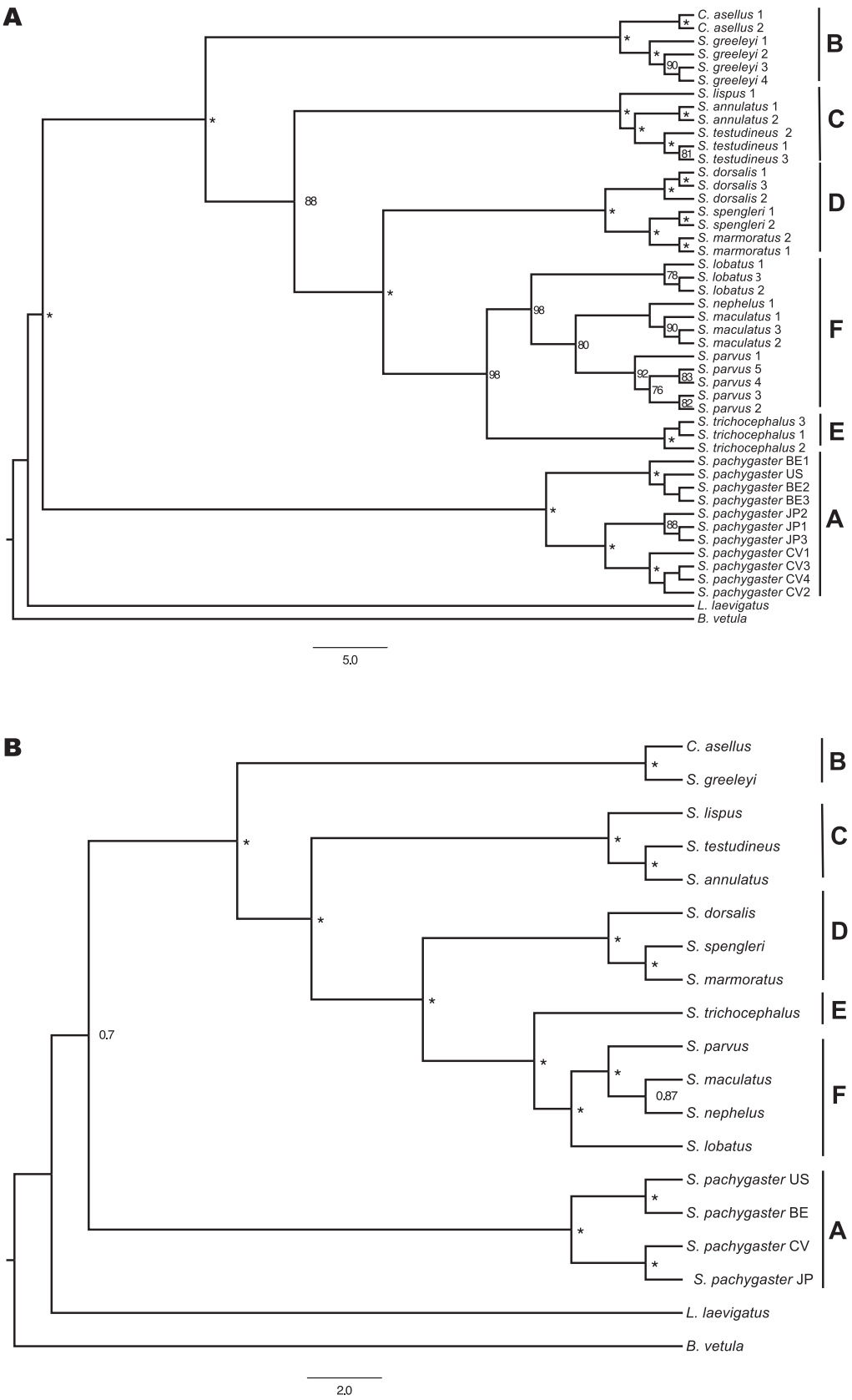

Fig. 1. A) RAxML v.8.02 maximum-likelihood analysis of the concatenated complete UCE dataset of 247 genes. Values at nodes represent bootstrap support. Support values of 100 are denoted with a (*) and those less than 70 were removed. **B)** ASTRAL-III species tree analysis of the complete UCE dataset of 247 genes. Values at nodes represent local posterior probability. Support values of 1.0 are denoted with a (*) and those less than 0.70 were removed. Location abbreviations for *S. pachygaster*: BE, Belize; CV, Cape Verde; JP, Japan; US, United States Atlantic Coast.

Table 2
Alignment information for each concatenated UCE dataset. Each dataset was filtered for varying percentages of missing data, resulting in five different datasets: complete (only including loci shared among all individuals), ≤10% missing data allowed, ≤25% missing data allowed, ≤50% missing data allowed, and incomplete. The incomplete dataset includes any locus present across three or more individuals. Abbreviations: n, number of loci; Length, total concatenated length in base-pairs; Min, minimum length of loci; Max, maximum length; Med, median length; PIS, parsimony informative sites; PUIS, parsimony uninformative sites; Mean PISITES/Locus, average number of parsimony informative sites per locus.

Dataset	n Loci	Length (bp)	Min (bp)	Max (bp)	Med(bp)	PISITES	PUISITES	Mean PISITES/Locus
Complete	247	234,719	290	1369	1026	22,431	212,288	91
10missing	429	379,782	189	1369	938	36,493	343,289	85
25missing	433	381,326	189	1369	934	36,651	344,675	85
50missing	436	382,164	189	1369	932	36,695	345,469	84
Incomplete	437	382,425	189	1369	932	36,701	345,724	101

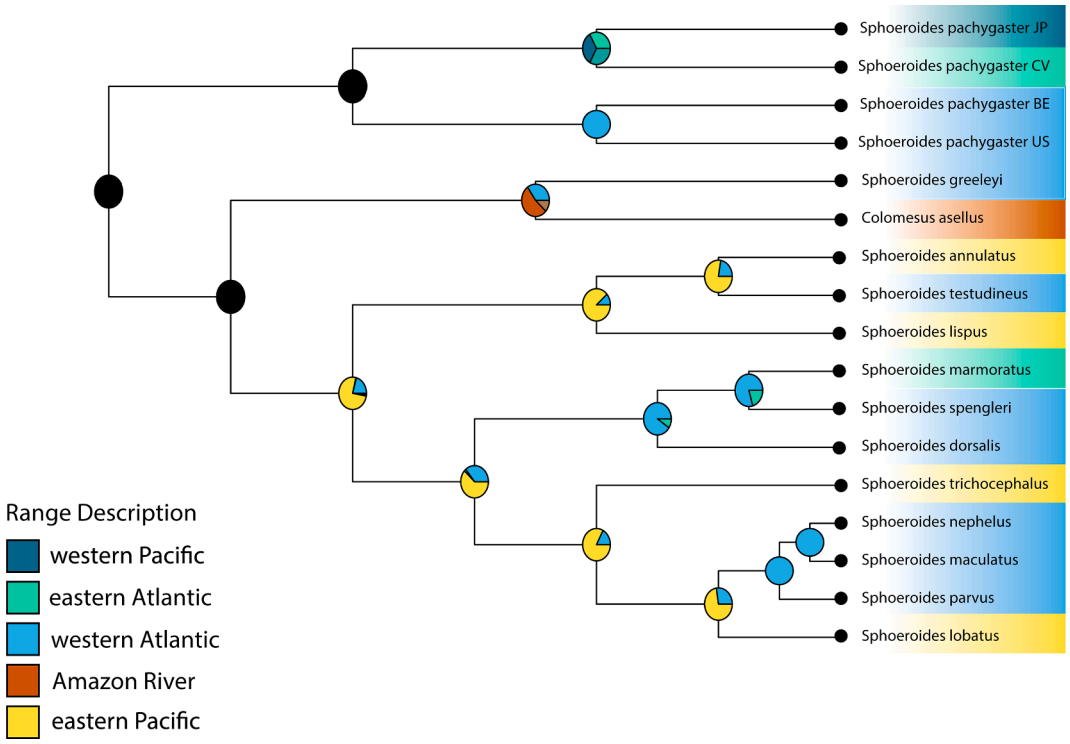

Fig. 2. Results of the DIVALIKE + J ancestral range reconstruction in BioGeoBEARS. Solid filled nodes indicate nodes where percentages were undiscernible. Abbreviations for *S. pachygaster*: BE, Belize; CV, Cape Verde; JP, Japan; US, United States Atlantic Coast.

Araujo et al. 2023). The species is distributed globally in temperate and tropical waters and has been described multiple times independently by different authors based on material from different ocean basins (see Fricke et al. 2023b). The current study included *S. pachygaster* from localities in both the eastern and western Atlantic, as well as the western Pacific and the results support that the species may harbor cryptic diversity (i.e., multiple similar-looking species currently recognized under a single name). Shipp (1974) noted morphological and meristic differences in *S. pachygaster* collected from populations on either side of the Atlantic Ocean but did not address whether they were different species. In this study, the eastern Atlantic samples were more closely related to those of the western Pacific than to the western Atlantic. This result is consistent with the findings of Amaral et al. (2013) who used COI to analyze samples of *S. pachygaster* across the species range including samples from the eastern and western Atlantic and Indian Ocean. Additionally, Araujo et al. (2023) resolved *S. pachygaster* samples from the southwest Atlantic as a distinct lineage from those sampled in the northwest Atlantic and Indo-Pacific, but they refrain from commenting on the evolutionary history of *S. pachygaster* due to sampling constraints. The biogeographic analysis presented here suggests that colonization of the western Atlantic by *S.pachygaster* may have proceeded from the eastern Pacific prior to the closing of the Isthmus of Panama, while the eastern Atlantic was colonized from the Indo-Pacific around the tip of Africa. This biogeographic pattern has been noted for other marine vertebrates, such as trumpetfishes (*Aulostomus;* Bowen et al., 2001) and ridley sea turtles (*Lepidochelys*; Bowen et al., 1998). To fully understand the potential for cryptic diversity in this species, multiple samples per region from the entirety of its global range would need to be collected and analyzed in both a morphological and a molecular context. If a more in-depth study confirmed species level differences between the genetic lineages of *S. pachygaster*, the name *S. pachygaster* would need to be applied to the lineage in the western Atlantic (the type locality of *S. pachygaster* is Barbados, West Indies). Several names are available from the synonymy of *S. pachygaster* (see Fricke et al. 2023b) and could be reinstated for lineages of *S. pachygaster* in other areas.

Across analyses the eastern Pacific species *S. lobatus*, *S. lispus*, and *S. trichocephalus* occupied basal positions with respect to western Atlantic species within the recovered clades, consistent with a pattern of west to east movement. Similarly, *S. dorsalis* occupied a more basal position in group D with represent to the eastern Atlantic species *S. marmoratus*. Furthermore, biogeographic ancestral range reconstruction, while lacking a proper time scale due to limited fossil information for the genus, statistically supported a westward directionality of divergence over multiple timescales. The most parsimonious

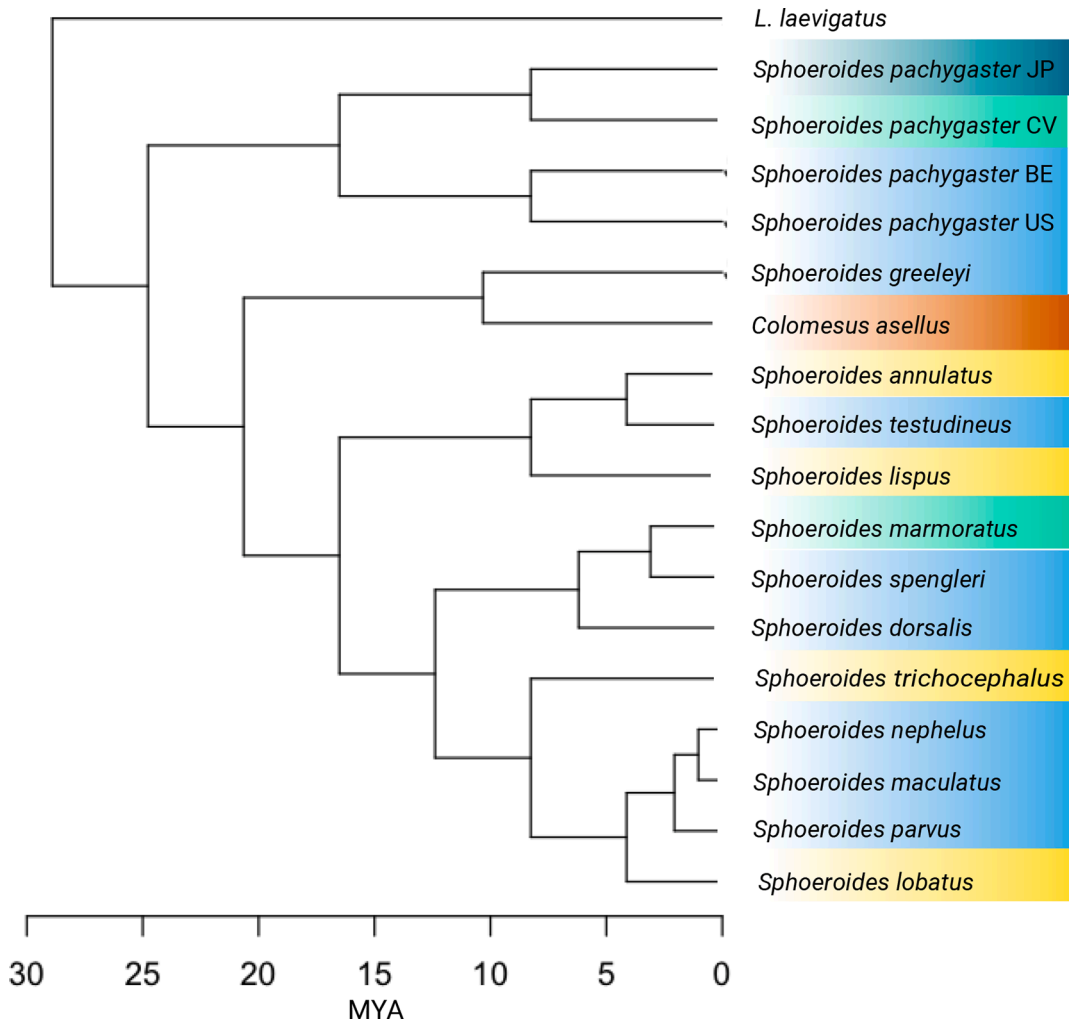

Fig. 3. Results of *post-hoc* time estimation of the *Sphoeroides* species tree using the R package APE. The time is show in millions of years ago (MYA). Abbreviations for *S. pachygaster*: BE, Belize; CV, Cape Verde; JP, Japan; US, United States Atlantic Coast.

Table 3
Results of Akaike information criterion (AIC) model testing using the R package BioGeoBEARS. Table header abbreviations are as follows: LnL, log likelihood; n, number of parameters; d, dispersal parameter; e, extinction parameter; j, jump parameter; AIC, AIC value, AIC w: AIC weight (calculated following Burnham and Anderson, 2002). Each taxon was coded as one of five geographical areas: a) western Atlantic, b) eastern Atlantic, c) eastern Pacific, d) western Pacific, or the e) Amazon River.

Model	LnL	n	d	e	j	AIC	AIC w
DEC	−33.9723	2	0.17	0.00	0.00	71.94	0.00
DEC + J	−23.0625	3	0.00	0.00	0.15	52.12	0.28
DIVALIKE	−30.6198	2	0.23	0.00	0.00	65.24	0.00
DIVALIKE + J	−22.2228	3	0.00	0.00	0.13	50.45	0.64
BAYAREALIKE	−48.5339	2	0.03	0.76	0.00	101.07	0.00
BAYAREALIKE + J	−24.3120	3	0.00	0.00	0.14	54.62	0.08

explanation given these results would be that a cosmopolitan ancestor, perhaps not unlike *S. pachygaster*, colonized the New World by crossing the Eastern Pacific Barrier prior to the closing of the Isthmus. This interpretation contrasts with Araujo et al. (2023), which suggested that *Sphoeroides* spread from the Caribbean. The differences in biogeographic inference between the studies are attributable to differences in the topology of the phylogenetic reconstructions, which are the result of taxon sampling and markers deployed. While increased taxon sampling is desirable for improving the ability to resolve phylogenetic relationships, marker number and type as well as model fit, can be more important (Reddy et al., 2017). Conserved regions of mtDNA, like COI, can have benefits for phylogenetic reconstruction (Rubinoff and Holland, 2005), but single gene trees do not always correctly represent species trees (Degnan and Rosenberg, 2006) and topologies based on mtDNA only may suffer from bias (Ballard and Rand, 2005; Shaw, 2002; Talavera and Vila, 2011). Ultimately, to address the conflicting topologies, a more complete taxon sampling coupled with increased numbers of markers and appropriate model testing would be required.

Once *Sphoeroides* entered the eastern Pacific and western Atlantic, incipient species may have been able to exploit new habitats that formed in association with the gradual rise of the Isthmus. Divergence of eastern Pacific and western Atlantic species of *Sphoeroides* likely began prior to the closing of the Isthmus of Panama, a common trend seen among marine organisms in the New World (e.g., Hodge and Bellwood, 2016; Lima et al., 2020; Thacker, 2017). There have been conflicting hypotheses on when the rising of the Isthmus would have spurred vicariant events, but the estimates range from ~ 20 mya and ~ 3.5 mya (Bacon et al., 2015; Lessios, 2008). To try to obtain a rough estimate for the timing of speciation in *Sphoeroides,* divergence dating was done *post-hoc* using the package *Ape* in R (Fig. 3), as generating a reliable time estimate for the genus is difficult due to a lack of reliable fossils assigned to the genus (Santini et al., 2013b). Time constraints on the node shared by *Lagocephalus* and *Sphoeroides* (38.5–26.4 mya) and the node shared by *C. asellus* and *S. greeleyi* (16.4–9.4 mya) were chosen based on

divergence dates provided by Santini et al. (2013b). With these rough estimates *S. pachygaster* may have diverged from the rest of *Sphoeroides* ~ 25–20 mya, prior to the closing of the Tethys Sea, and all eastern Pacific and Atlantic sister species may have diverged between ~ 12–4 mya, prior to the final closing of the Isthmus of Panama. By contrast, the split between sister species *S. nephelus* (found from northern Florida in the Atlantic, USA through the Gulf of Mexico and Caribbean) and *S. maculatus* (found from northern Florida, USA in the Atlantic to waters of Newfoundland, Canada; Supplemental Figure 10) was estimated < 2.5 mya, after the closing of the Isthmus, and divergence between these species may have resulted from changing currents and geomorphology of the Gulf of Mexico (Gold et al., 2011). One possibility is that the split coincides with the closing of Suwannee Straights, a powerful current that ran across peninsular Florida, connecting the northern Gulf of Mexico with the Atlantic as recently as 1.74 mya (Bert, 1986; Portnoy and Gold, 2012). Finally, the divergence between *S. spengleri* (western Atlantic) and *S. marmoratus* (eastern Atlantic) may have occurred < 4 mya and may be due to colonization of the eastern Atlantic via dispersal facilitated by changing currents of the Atlantic and the Gulf Stream driven by the rising of the Isthmus (Haug and Tiedemann, 1998), a phenomena noted in the seahorse species complex *Hippocampus* (Boehm et al., 2013). It is important to note that the divergence times presented here are rough approximations based on a previously published study (Santini et al., 2013b) and the timing of divergence events in *Sphoeroides* could be better estimated once more appropriate fossil evidence for the genus is available.

Ancestral range reconstruction favored models that included the *J* parameter as compared to the identical model without it, indicating that accounting for founder speciation events may be important. The inclusion of the *J* parameter has been shown to be important for other marine species such as butterflyfishes and syngnathids (DiBattista et al., 2018; Stiller et al., 2022). While most founder-speciation events usually elicit thoughts of island colonization, the scenario would also make sense for *Sphoeroides*. Current distributions of species of *Sphoeroides* in the New World may have likely begun with a founder event in the eastern Pacific by a wide-ranging species of *Sphoeroides*, allowing for diversification along the eastern Pacific. Following that, multiple founder events into the western Atlantic prior to the closure of the Isthmus of Panama may have occurred at varying times leading to speciation of western Atlantic sister taxa. The idea of a founder speciation dynamic is also consistent with the distribution of *S. marmoratus*, which is the only *Sphoeroides* found in the eastern Atlantic. Alternatively, it has been noted that inclusion of the *J* parameter is favored in situations where tip species occupy single ranges or in studies that cover large geographic ranges but have relatively few widespread species, potentially making the result artifactual (Matzke, 2013b). Considering these models do not take into account ranges for extinct species, and that there are also gaps in our taxon sampling, these results could change with the addition of new data.

In addition to biogeographic patterns observed in the data, there are interesting examples of ecomorphological differences between co-distributed species with recent common ancestry. For example, the eastern Pacific species *S. lobatus* and *S. trichocephalus* have overlapping distributions along Central America (Supplemental Figures 11 and 12) but occupy different niches. *Sphoeroides lobatus* has a larger body size than *S. trichocephalus* and occupies bays and estuaries, particularly sandy and weedy areas, while *S. trichocephalus* is found in turbid bays which contain soft and silty bottoms (Bussing, 1995; Robertson and Allen, 2015). Similarly, *S. dorsalis* and *S. spengleri* also have overlapping distributions (in the western Atlantic) that seem to be partitioned by habitat type (Supplemental Figure 11). *Sphoeroides dorsalis* is known to be found over soft, silty bottoms while *S. spengleri* can be found in reef and sea grass areas (Robins and Ray, 1986). This may be indicative of niche partitioning which is common in other organisms such as birds (Richman and Price, 1992), mammals (Di Bitetti et al., 2010), and other fishes (Brandl et al., 2020; Ross, 1986). Understanding the importance of niche partitioning in explaining patterns of diversity and the current distribution of *Sphoeroides* pufferfishes will require thorough taxon sampling paired with more in-depth data on the ecology of these species.

5. Conclusion

The current study provides the most comprehensive phylogenetic hypothesis, in term of molecular markers deployed, of the genus *Sphoeroides* to date, using ultra-conserved DNA elements. It also provides an alternative hypothesis on how this genus may have colonized the New World, adding to our understanding of colonization patterns in marine fishes and the potential biogeographic processes which may be responsible for their current distributions. For most marine taxa, colonization of the New World has been hypothesized to have proceeded from the Tethys Sea to the western Atlantic and then to the eastern Pacific (Barber and Bellwood, 2005; Bellwood, 1994). Contrary to the Tethys hypothesis, the results here support colonization of the New World proceeding from the eastern Pacific into the western Atlantic, with diversification occurring after or during the closure of the Isthmus of Panama consistent with hypotheses for some other marine organisms (Lessios et al., 1999; Santaquiteria et al., 2021).

While taxonomic revision of *Sphoeroides* is outside the scope of this study, this work also highlights issues that need further examination. For example, the cosmopolitan species currently known as *S. pachygaster* appears to harbor cryptic diversity. Due to advancements of molecular techniques, there has been an increase in the rate of species discovery in marine taxa (Leray and Knowlton, 2016; Randhawa et al., 2015) and the presence of undescribed diversity in *S. pachygaster* highlights the importance of these molecular techniques. However, further taxonomic research will be needed to fully understand the patterns of diversity in this group.

While thorough divergence dating was not the main focus of this study, due to the lack of usable fossils for the genus, rough dates paired with biogeographic processes are consistent with hypothesis about modes of diversifications. Further research is warranted to better understand the timing of diversification events within *Sphoeroides*, as better fossil data becomes available with more complete taxon sampling. In addition, the observation of niche differences between species may represent important areas of future research for this group.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank D. Werneke (AUM), A. Williston (MCZ), J. Tavera (CIRUV), A. Bently (KU), P. Chakrabarty and S. Parker (LSUMZ), P. Hastings and H.J. Walker (SIO), H. Prestridge (TCWC), C. Baldwin and D. Pitassy (USNM), K. Kingon (UTT), Ivy Baremore (MarAlliance), and S. Sakakura of the Japan Fisheries Research and Education Agency for providing tissues for this research. Additionally, we would like to thank the National Marine Fisheries Service Pascagoula vessel Oregon II staff, crew, and volunteers for collection of specimens used in this work. We thank E. Alcañiz and M. Jonuschat for allowing us to use their wonderful images of *Sphoeroides* in our graphical abstract. Finally, we would like to thank the anonymous reviewers for their helpful comments which led to the imporvement of this manuscript. KWC acknowledges support from TAMU Agrilife Research (HATCH TEX09452). This is publication number 35 of the TAMU-CC Marine Genomics Laboratory, number 128 of Genetic Studies in Fishes, and 1680 of the TAMU Biodiversity Research and Teaching Collections.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2023.107935.

References

Adams, C.G., Bayliss, D.D., Whittaker, J.E.P., 1999. The terminal tethyan event: a critical review of the conflicting age determinations for the disconnection of the Mediterranean from the Indian Ocean. In: Whybrow, P.J., Hill, A. (Eds.), Fossil Vertebrates of Arabia. Yale University Press, New Haven, pp. 477–484.

Amaral, C.R.L., Brito, P.M., Silva, D.A., Carvalho, E.F., 2013. A new cryptic species of South American freshwater pufferfish of the genus Colomesus (Tetraodontidae), based on both morphology and DNA data. PLoS One 8, e74397.

Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J., Dehal, P., Christoffels, A., Rash, S., Hoon, S., Smit, A., Gelpke, M.D.S., Roach, J., Oh, T., Ho, I.Y., Wong, M., Detter, C., Verhoef, F., Predki, P., Tay, A., Lucas, S., Richardson, P., Smith, S.F., Clark, M.S., Edwards, Y.J.K., Doggett, N., Zharkikh, A., Tavtigian, S.V., Pruss, D., Barnstead, M., Evans, C., Baden, H., Powell, J., Glusman, G., Rowen, L., Hood, L., Tan, Y.H., Elgar, G., Hawkins, T., Venkatesh, B., Rokhsar, D., Brenner, S., 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. Science 297, 1301–1310.

Araujo, G.S., Kurtz, Y.R., Sazima, I., Carvalho, P.H., Floeter, S.R., Vilasboa, A., Rotundo, M.M., Ferreira, C.E.L., Barreiros, J.P., Pitassy, D.E., Carvalho-Filho, A., 2023. Evolutionary history, biogeography, and a new species of Sphoeroides (Tetraodontiformes: Tetraodontidae): how the major biogeographic barriers of the Atlantic Ocean shaped the evolution of a pufferfish genus. Zool. J. Linn. Soc. 1–16.

Bacon, C.D., Silvestro, D., Jaramillo, C., Smith, B.T., Chakrabarty, P., Antonelli, A., 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. PNAS 112, 6110–6115. https://doi.org/10.1073/pnas.1423853112.

Ballard, J.W.O., Rand, D.M., 2005. The Population Biology of Mitochondrial DNA and Its Phylogenetic Implications. Annu. Rev. Ecol. Evol. Syst. 36, 621–642. https://doi.org/10.1146/annurev.ecolsys.36.091704.175513.

Bannikov, A.F., Tyler, J.C., 2008. A new species of the pufferfish Eotetraodon (Tetraodontiformes, Tetraodontidae) from the Eocene of the Northern Caucasus. Paleontol. J. 42, 526–530. https://doi.org/10.1134/S0031030108050080.

Barber, P.H., Bellwood, D.R., 2005. Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (Halichoeres: Labridae) in the Indo-Pacific and new world tropics. Mol. Phylogenet. Evol. 35, 235–253. https://doi.org/10.1016/j.ympev.2004.10.004.

Bejerano, G., 2004. Ultraconserved Elements in the Human Genome. Science 304, 1321–1325. https://doi.org/10.1126/science.1098119.

Bellwood, D.R., 1994. A phylogenetic study of the parrotfish family Scaridae (Pisces: Labroidea), with a revision of genera. Rec. Aust. Mus. Suppl. 20, 1–86. https://doi.org/10.3853/j.0812-7387.20.1994.51.

Bellwood, D.R., Wainwright, P.C., 2002. The history and biogeography of fishes on coral reefs. In: Sale, P.F. (Ed.), Coral Reef Fishes: Dynamics and Diversity on a Complex Ecosystem. Academic Press, New York, New York, pp. 5–32.

Benton, M.J. (Ed.), 1993. The Fossil Record 2, 1st ed. Chapman & Hall, London, UK.

Bert, T.M., 1986. Speciation in western Atlantic stone crabs (genus Menippe): the role of geological processes and climatic events in the formation and distribution of species. Mar. Biol. 93, 157–170. https://doi.org/10.1007/BF00508253.

Bloom, D., Lovejoy, N., 2011. The biogeography of marine incursions in South America. Historical Biogeography of Neotropical Freshwater Fishes. 137–144. https://doi.org/10.1525/california/9780520268685.003.0008.

Bloom, D.D., Lovejoy, N.R., 2017. On the origins of marine-derived freshwater fishes in South America. J. Biogeogr. 44, 1927–1938. https://doi.org/10.1111/jbi.12954.

Boehm, J.T., Woodall, L., Teske, P.R., Lourie, S.A., Baldwin, C., Waldman, J., Hickerson, M., 2013. Marine dispersal and barriers drive Atlantic seahorse diversification. J. Biogeogr. 40, 1839–1849. https://doi.org/10.1111/jbi.12127.

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.

Bowen, B.W., Clark, A.M., Abreu-Grobois, F.A., Chaves, A., Reichart, H.A., Ferl, R.J., 1998. Global phylogeography of the ridley sea turtles (Lepidochelys spp.) as inferred from mitochondrial DNA sequences. Genetica 101, 179–189.

Bowen, B.W., Bass, A.L., Rocha, L.A., Grant, W.S., Robertson, D.R., 2001. Phylogeography of the trumpetfishes (Aulostomus): ring species complex on a global scale. Evol. 55, 1029–1039. https://doi.org/10.1111/j.0014-3820.2001.tb00619.x.

Brainerd, E.L., 1994. Pufferfish inflation: functional morphology of postcranial structures in Diodon holocanthus (Tetraodontiformes). J. Morphol. 220, 243–261. https://doi.org/10.1002/jmor.1052200304.

Burnham, K.P., Anderson, D.R. (Eds.), 2002. Model Selection and Multimodel Inference, 2nd ed. Springer, New York, NY.

Bussing, W.A., 1995. In: Guia FAO Para Identification De Especies Para Lo Fines De La Pesca. Pacifico Centro-Oriental, FAO, Rome, pp. 1629–1637.

Capella-Gutierrez, S., Silla-Martinez, J.M., Gabaldon, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. GigaScience 10, giab008. https://doi.org/10.1093/gigascience/giab008.

Degnan, J.H., Rosenberg, N.A., 2006. Discordance of Species Trees with Their Most Likely Gene Trees. PLoS Genet. 2, e68.

DiBattista, J.D., Alfaro, M.E., Sorenson, L., Choat, J.H., Hobbs, J.-P.A., Sinclair-Taylor, T. H., Rocha, L.A., Chang, J., Luiz, O.J., Cowman, P.F., Friedman, M., Berumen, M.L., 2018. Ice ages and butterflyfishes: Phylogenomics elucidates the ecological and evolutionary history of reef fishes in an endemism hotspot. Ecol. Evol., 8, 10989–11008. https://doi.org/10.1002/ece3.4566.

Faircloth, B.C., 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32, 786–788.

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. Syst. Biol. 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.

Floeter, S.R., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W.F., Wirtz, P., Edwards, A.J., Barreiros, J.P., Ferreira, C.E.L., Gasparini, J.L., Brito, A., Falcón, J.M., Bowen, B.W., Bernardi, G., 2008. Atlantic reef fish biogeography and evolution. J. Biogeogr. 35, 22–47. https://doi.org/10.1111/j.1365-2699.2007.01790.x.

Fricke, R., Eschmeyer, W.N., Fong, J.D., 2023a. Eschmeyer's Catalog of Fishes: Genera/Species by Family/Subfamily. Eschmeyer's Catalog of Fishes. http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp (accessed 19 September 2023).

Froese, R., Pauly, D., 2022. FishBase. www.fishbase.org (accessed 01 February 2023).

Fricke, R., Eschmeyer, W.N., Van der Laan, R., 2023b. Eschmeyer's Catalog of Fishes: Genera. Eschmeyer's Catalog of Fishes, Species, References http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (accessed 19 September 2023.

Gao, Y., Gao, Q., Zhang, H., Wang, L., Zhang, F., Yang, C., Song, L., 2014. Draft sequencing and analysis of the genome of pufferfish *Takifugu flavidus*. DNA Res. 21, 627–637. https://doi.org/10.1093/dnares/dsu025.

George,, 2005. Tethys Sea fragmentation and speciation of *Panulirus* spiny lobsters. Crustacceana. 78, 1281–1309. https://doi.org/10.1163/156854005776759780.

Gilbert, P.S., Chang, J., Pan, C., Sobel, E.M., Sinsheimer, J.S., Faircloth, B.C., Alfaro, M. E., 2015. Genome-wide ultraconserved elements exhibit higher phylogenetic informativeness than traditional gene markers in percomorph fishes. Molecular Phylogenetics and Evolution 92, 140–146. https://doi.org/10.1016/j.ympev.2015.05.027.

Glenn, T.C., Nilsen, R.A., Kieran, T.J., Sanders, J.G., Bayona-Vásquez, N.J., Finger, J.W., Pierson, T.W., Bentley, K.E., Hoffberg, S.L., Louha, S., García-De León, F.J., Del Río-Portilla, M.A., Reed, K.D., Anderson, J.L., Meece, J.K., Aggrey, S.E., Rekaya, R., Alabady, M., Bélanger, M., Winker, K., Faircloth, B.C., 2016. Adapterama I: universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext) (preprint). Genomics. https://doi.org/10.1101/049114.

Gold, J.R., Voelker, G., Renshaw, M.A., 2011. Phylogenetic relationships of tropical western Atlantic snappers in subfamily Lutjaninae (Lutjanidae: Perciformes) inferred from mitochondrial DNA sequences: phylogeny of western Atlantic lutjanines. Biol. J. Linn. Soc. 102, 915–929. https://doi.org/10.1111/j.1095-8312.2011.01621.x.

Gomon, M.F., 1997. Relationships of fishes of the labrid tribe Hypsigenyini. Bull. Mar. Sci. 60, 789–871.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883.

Haug, G.H., Tiedemann, R., 1998. Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. Nature 393, 673–676. https://doi.org/10.1038/31447.

Hodge, J.R., Bellwood, D.R., 2016. The geography of speciation in coral reef fishes: the relative importance of biogeographical barriers in separating sister-species. J. Biogeogr. 43, 1324–1335. https://doi.org/10.1111/jbi.12729.

Holcroft, N.I., 2005. A molecular analysis of the interrelationships of tetraodontiform fishes (Acanthomorpha: Tetraodontiformes). Mol. Phylogenet. Evol. 34, 525–544. https://doi.org/10.1016/j.ympev.2004.11.003.

Hrbek, T., Meyer, A., 2003. Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). J. Evol. Biol. 16, 17–36. https://doi.org/10.1046/j.1420-9101.2003.00475.x.

Igarashi, Y., Doi, H., Yamanoue, Y., Kinoshita, S., Ishibashi, T., Ushio, H., Asakawa, S., Nishida, M., Watabe, S., 2013. Molecular phylogenetic relationship of *Tetraodon* pufferfish based on mitochondrial DNA analysis. Fish. Sci. 79, 243–250. https://doi.org/10.1007/s12562-013-0598-5.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.

Klassen, G.J., 1995. Phylogeny and biogeography of the Ostraciinae (Tetraodontiformes: Ostraciidae). Bull. Mar. Sci. 57, 393–441.

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56, 17–24. https://doi.org/10.1080/10635150601146041.

Landis, M.J., Matzke, N.J., Moore, B.R., Huelsenbeck, J.P., 2013. Bayesian analysis of biogeography when the number of areas is large. Syst. Biol. 62, 789–804. https://doi.org/10.1093/sysbio/syt040.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923.
Leray, M., Knowlton, N., 2016. Censusing marine eukaryotic diversity in the twenty-first century. Phil. Trans. r. Soc. B 371, 20150331. https://doi.org/10.1098/rstb.2015.0331.
Lessios, H.A., 2008. The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. Annu. Rev. Ecol. Evol. Syst. 39, 63–91.
Lessios, H.A., Kessing, B.D., Robertson, D.R., Paulay, G., 1999. Phylogeography of the Pantropical Sea Urchin Eucidaris in Relation to Land Barriers and Ocean Currents. Evolution 53, 806–817. https://doi.org/10.1111/j.1558-5646.1999.tb05374.x.
Lessios, H.A., Kessing, B.D., Pearse, J.S., 2001. Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. Evol. 55, 955–975. https://doi.org/10.1111/j.0014-3820.2001.tb00613.x.
Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658–1659. https://doi.org/10.1093/bioinformatics/btl158.
Lima, F.D., Strugnell, J.M., Leite, T.S., Lima, S.M.Q., 2020. A biogeographic framework of octopod species diversification: the role of the Isthmus of Panama. PeerJ 8, e8691.
Lin, H.-C., Hastings, P.A., 2013. Phylogeny and biogeography of a shallow water fish clade (Teleostei: Blenniiformes). BMC Evol. Biol. 13, 210.
Liu, B., Zhou, Z., Bai, Y., Yang, J., Shi, Y., Pu, F., Xu, P., 2021. Genome-scale phylogenetic and population genetic studies provide insight into introgression and adaptive evolution of takifugu species in east Asia. Front, Genet, p. 12.
Lovejoy, N.R., Bermingham, E., Martin, A.P., 1998. Marine incursion into South America. Nature 396, 421–422. https://doi.org/10.1038/24757.
Luiz-Júnior, O.J., Floeter, S.R., Gasparini, J.L., Ferreira, C.E.L., Wirtz, P., 2004. The occurrence of Acanthurus monroviae (Perciformes: Acanthuridae) in the south-western Atlantic, with comments on other eastern Atlantic reef fishes occurring in Brazil. J. Fish Biol. 65, 1173–1179. https://doi.org/10.1111/j.0022-1112.2004.00519.x.
Matsuura, K., 2015. Taxonomy and systematics of tetraodontiform fishes: a review focusing primarily on progress in the period from 1980 to 2014. Ichthyol. Res. 62, 72–113. https://doi.org/10.1007/s10228-014-0444-5.
Matsuura, K., Tyler, J.C., 1997. Tetraodontiform fishes, mostly from deep waters, of New Caledonia., in: Séret, B. (Ed.), Re'sultats Des Campagnes Musorstom. Mémoires Du Muséum National d'Histoire Naturelle. Zoologie, Paris, pp. 173–208.
Matzke, N.J., 2013a. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Front. Biogeogr. 5, 242–248.
Matzke, N.J., 2013b. Probabilistic Historical Biogeography: New Models for Founder-Event Speciation, Imperfect Detection, and Fossils Allow Improved Accuracy and Model-Testing (Doctoral). University of California, Berkeley, Berkeley, California.
Mirarab, S., Bayzid, M.S., Boussau, B., Warnow, T., 2014. Statistical binning enables an accurate coalescent-based estimation of the avian tree. Science 346, 1250463. https://doi.org/10.1126/science.1250463.
Nelson, J.S., Grande, T.C., Wilson, M.V.H., 2016. Fishes of the World, Fifth. ed. Wiley & Sons Inc, Hoboken, New Jersey.
Nikulina, E.A., Hanel, R., Schäfer, P., 2007. Cryptic speciation and paraphyly in the cosmopolitan bryozoan *Electra pilosa*—impact of the Tethys closing on species evolution. Mol. Phylogenet. Evol. 45, 765–776. https://doi.org/10.1016/j.ympev.2007.07.016.
Nute, M., Chou, J., Molloy, E.K., Warnow, T., 2018. The performance of coalescent-based species tree estimation methods under models of missing data. BMC Genom. 19 https://doi.org/10.1186/s12864-018-4619-8.
O'Dea, A., Lessios, H.A., Coates, A.G., Eytan, R.I., Restrepo-Moreno, S.A., Cione, A.L., Collins, L.S., de Queiroz, A., Farris, D.W., Norris, R.D., Stallard, R.F., Woodburne, M. O., Aguilera, O., Aubry, M.-P., Berggren, W.A., Budd, A.F., Cozzuol, M.A., Coppard, S.E., Duque-Caro, H., Finnegan, S., Gasparini, G.M., Grossman, E.L., Johnson, K.G., Keigwin, L.D., Knowlton, N., Leigh, E.G., Leonard-Pingel, J.S., Marko, P.B., Pyenson, N.D., Rachello-Dolmen, P.G., Soibelzon, E., Soibelzon, L., Todd, J.A., Vermeij, G.J., Jackson, J.B.C., 2016. Formation of the Isthmus of Panama. Sci. Adv. 2, e1600883.
Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412.
Pattengale, N.D., Alipour, M., O.R., Moret, B.M., Stamatakis, A.,, 2010. How many bootstrap replicates are necessary? J. Comput. Biol. 17, 337–354.
Portnoy, D.S., Gold, J.R., 2012. Evidence of multiple vicariance in a marine suture-zone in the Gulf of Mexico: vicariance in the Gulf of Mexico. J. Biogeogr. 39, 1499–1507. https://doi.org/10.1111/j.1365-2699.2012.02699.x.
Randhawa, H.S., Poulin, R., Krkošek, M., 2015. Increasing rate of species discovery in sharks coincides with sharp population declines: implications for biodiversity. Ecography 38, 96–107. https://doi.org/10.1111/ecog.00793.
Reddy, S., Kimball, R.T., Pandey, A., Hosner, P.A., Braun, M.J., Hackett, S.J., Han, K.-L., Harshman, J., Huddleston, C.J., Kingston, S., Marks, B.D., Miglia, K.J., Moore, W.S., Sheldon, F.H., Witt, C.C., Yuri, T., Braun, E.L., 2017. Why Do Phylogenomic Data Sets Yield Conflicting Trees? Data Type Influences the Avian Tree of Life more than Taxon Sampling. Syst. Biol. 66, 857–879. https://doi.org/10.1093/sysbio/syx041.
Ree, R.H., Smith, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57, 4–14. https://doi.org/10.1080/10635150701883881.
Reis, R.E., Kullander, S.O., Ferraris Jr, C.L., 2003. Check List of the Freshwater Fishes of South and Central America. EDIPUCRS, Porto Alegre, Brazil.
Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: The European molecular biology open software suite 276–277.
Richman, A.D., Price, T., 1992. Evolution of ecological differences in the old world leaf warblers. Nature 355, 817–821. https://doi.org/10.1038/355817a0.
Robertson, D.R., Allen, G.R., 2015. Shorefishes of the Tropical Eastern Pacific: online information system. https://biogeodb.stri.si.edu/sftep/en/pages (accessed 11 January 2023).
Robins, R.C., Ray, G.C., 1986. A field guide to Atlantic coast fishes: North America. Houghton Mifflin Harcourt, New York, New York.
Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Biol. 46, 195–203. https://doi.org/10.1093/sysbio/46.1.195.
Ross, S.T., 1986. Resource Partitioning in Fish Assemblages: A Review of Field Studies. Copeia 1986, 352–388. https://doi.org/10.2307/1444996.
Rubinoff, D., Holland, B.S., 2005. Between Two Extremes: Mitochondrial DNA is neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference. Syst. Biol. 54, 952–961. https://doi.org/10.1080/10635150500234674.
Sabaj, M.H., 2020. Codes for natural history collections in ichthyology and herpetology. Copeia 108. https://doi.org/10.1643/ASIHCODONS2020.
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
Santaquiteria, A., Siqueira, A.C., Duarte-Ribeiro, E., Carnevale, G., White, W.T., Pogonoski, J.J., Baldwin, C.C., Ortí, G., Arcila, D., Ricardo, B.-R., 2021. Phylogenomics and historical biogeography of seahorses, dragonets, goatfishes, and allies (Teleostei: Syngnatharia): assessing factors driving uncertainty in biogeographic inferences. Syst. Biol. 70, 1145–1162. https://doi.org/10.1093/sysbio/syab028.
Santini, F., Sorenson, L., Alfaro, M.E., 2013a. A new phylogeny of tetraodontiform fishes (Tetraodontiformes, Acanthomorpha) based on 22 loci. Mol. Phylogenet. Evol. 69, 177–187. https://doi.org/10.1016/j.ympev.2013.05.014.
Santini, F., Nguyen, M.T.T., Sorenson, L., Waltzek, T.B., Lynch Alfaro, J.W., Eastman, J. M., Alfaro, M.E., 2013b. Do habitat shifts drive diversification in teleost fishes? An example from the pufferfishes (Tetraodontidae). J. Evol. Biol. 26, 1003–1018. https://doi.org/10.1111/jeb.12112.
Santini, F., Tyler, J.C., 2002. Phylogeny and biogeography of the extant species of triplespine fishes (Triacanthidae, Tetraodontiformes). Zool. Scr. 31, 321–330. https://doi.org/10.1046/j.1463-6409.2002.00097.x.
Sayyari, E., Mirarab, S., 2016. Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. 33, 1654–1668. https://doi.org/10.1093/molbev/msw079.
Seutin, G., White, B.N., Boag, P.T., 1991. Preservation of avian blood and tissue samples for DNA analyses. Can. J. Zool. 69, 82–90. https://doi.org/10.1139/z91-013.
Shaw, K.L., 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proc. Natl. Acad. Sci. 99, 16122–16127. https://doi.org/10.1073/pnas.242585899.
Shipp, R.L., 1974. The Pufferfishes (Osteichthyes: Tetraodontidae) of the Atlantic Ocean. Publications of the Gulf Coast Research Laboratory Museum, Ocean Springs, Mississippi.
Shipp,, 2002. Tetraodontidae. In: Carpenter, K.E. (Ed.), The Living Marine Resources of the Western Central Atlantic. FAO, United Nations, Rome, pp. 1988–2006.
Song, L., Liu, B., Xiang, J., Qian, P.-Y., 2001. Molecular phylogeny and species identification of pufferfish of the genus Takifugu (Tetraodontiformes, Tetraodontidae). Mar. Biotechnol. 3, 398–406. https://doi.org/10.1007/s10126-001-0006-5.
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.
Stiller, J., Short, G., Hamilton, H., Saarman, N., Longo, S., Wainwright, P., Rouse, G.W., Simison, W.B., 2022. Phylogenomic analysis of Syngnathidae reveals novel relationships, origins of endemic diversity and variable diversification rates. BMC Biol. 20, 75. https://doi.org/10.1186/s12915 022-01271-w.
Summerer, M., Hanel, R., Sturmbauer, C., 2001. Mitochondrial phylogeny and biogeographic affinities of sea breams of the genus Diplodus (Sparidae). J. Fish Biol. 59, 1638–1652. https://doi.org/10.1111/j.1095-8649.2001.tb00227.x.
Talavera, G., Vila, R., 2011. What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta class phylogeny. BMC Evol. Biol. 11, 315. https://doi.org/10.1186/1471-2148-11-315.
Thacker, C.E., 2017. Patterns of divergence in fish species separated by the Isthmus of Panama. BMC Evol. Biol. 17, 111. https://doi.org/10.1186/s12862-017-0957-4.
Tyler, J.C., 1980. Osteology, phylogeny, and higher classification of the fishes of the order Plectognathi (Tetraodontiformes). NOAA Techn. Rep, NMFS, Circular, p. 434.
Tyler, J.C., Santini, F., 2002. Review and reconstructions of the tetraodontiform fishes from the Eocene of Monte Bolca, Italy, with comments on related tertiary taxa. Studi e Ricerche Sui Giacimenti Terziari Di Bolca 9, 47–119.
Veron, J.E.N., 1995. Corals in Space and Time: The Biogeography and Evolution of the Scleractinia. Cornell University Press.
Wainwright, P.C., Turingan, R.G., 1997. Evolution of Pufferfish Inflation Behavior. Evol. 51, 506–518. https://doi.org/10.1111/j.1558-5646.1997.tb02438.x.
Walker, H.J., Bussing, W.A., 1996. Two new pufferfishes of the genus *Sphoeroides* from the eastern Pacific. Copeia 1996, 677. https://doi.org/10.2307/1447532.
Yamanoue, Y., Miya, M., Matsuura, K., Katoh, M., Sakai, H., Nishida, M., 2008. A new perspective on phylogeny and evolution of tetraodontiform fishes (Pisces:

Acanthopterygii) based on whole mitochondrial genome sequences: Basal ecological diversification? BMC Evol. Biol. 8, 212. https://doi.org/10.1186/1471-2148-8-212.
Yamanoue, Y., Miya, M., Matsuura, K., Miyazawa, S., Tsukamoto, N., Doi, H., Takahashi, H., Mabuchi, K., Nishida, M., Sakai, H., 2009. Explosive speciation of *Takifugu*: another use of fugu as a model system for evolutionary biology. Mol. Biol. Evol. 26, 623–629. https://doi.org/10.1093/molbev/msn283.
Zhang, C., Sayyari, E., Mirarab, S., 2017. ASTRAL-III: increased scalability and impacts of contracting low support branches. In: Meidanis, J., Nakhleh, L. (Eds.), Comparative Genomics. Springer International Publishing, Cham, pp. 53–75. https://doi.org/10.1007/978-3-319-67979-2_4.