Syst. Biol. 70(6):1145–1162, 2021

DOI:10.1093/sysbio/syab028
Advance Access publication April 23, 2021

Phylogenomics and Historical Biogeography of Seahorses, Dragonets, Goatfishes, and Allies (Teleostei: Syngnatharia): Assessing Factors Driving Uncertainty in Biogeographic Inferences

AINTZANE SANTAQUITERIA[1], ALEXANDRE C. SIQUEIRA[2,3], EMANUELL DUARTE-RIBEIRO[1], GIORGIO CARNEVALE[4], WILLIAM T. WHITE[5], JOHN J. POGONOSKI[5], CAROLE C. BALDWIN[6], GUILLERMO ORTÍ[6,7], DAHIANA ARCILA[1,8] AND RICARDO BETANCUR-R.[1,*]

[1]*Department of Biology, The University of Oklahoma, 730 Van Vleet Oval, Norman, OK 73019, USA;* [2]*Research Hub for Coral Reef Ecosystem Functions, College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia;* [3]*ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD 4811, Australia;* [4]*Dipartimento di Scienze della Terra, Università degli Studi di Torino, via Valperga Caluso 35, 10125, Torino, Italy;* [5]*CSIRO Australian National Fish Collection, National Research Collections Australia, Hobart, TAS, Australia;* [6]*Department of Vertebrate Zoology, Smithsonian National Museum of Natural History, 10th St. & Constitution Ave. NW, Washington, DC 20560, USA;* [7]*Department of Biological Sciences, George Washington University, 2029 G St. NW, Washington, DC 20052, USA;* [8]*Department of Ichthyology, Sam Noble Oklahoma Museum of Natural History, 2401 Chautauqua Ave, Norman, OK 73072, USA*

**Correspondence to be sent to: Department of Biology, The University of Oklahoma, 730 Van Vleet Oval, Norman, OK 73019, USA; E-mail: ricardo.betancur@ou.edu.*

Received 14 December 2020; reviews returned 15 April 2021; accepted 19 April 2021
Associate Editor: Matt Friedman

Abstract.—The charismatic trumpetfishes, goatfishes, dragonets, flying gurnards, seahorses, and pipefishes encompass a recently defined yet extraordinarily diverse clade of percomorph fishes—the series Syngnatharia. This group is widely distributed in tropical and warm-temperate regions, with a great proportion of its extant diversity occurring in the Indo-Pacific. Because most syngnatharians feature long-range dispersal capabilities, tracing their biogeographic origins is challenging. Here, we applied an integrative phylogenomic approach to elucidate the evolutionary biogeography of syngnatharians. We built upon a recently published phylogenomic study that examined ultraconserved elements by adding 62 species (total 169 species) and one family (Draconettidae), to cover *ca.* 25% of the species diversity and all 10 families in the group. We inferred a set of time-calibrated trees and conducted ancestral range estimations. We also examined the sensitivity of these analyses to phylogenetic uncertainty (estimated from multiple genomic subsets), area delimitation, and biogeographic models that include or exclude the jump-dispersal parameter (*j*). Of the three factors examined, we found that the *j* parameter has the strongest effect in ancestral range estimates, followed by number of areas defined, and tree topology and divergence times. After accounting for these uncertainties, our results reveal that syngnatharians originated in the ancient Tethys Sea *ca.* 87 Ma (84–94 Ma; Late Cretaceous) and subsequently occupied the Indo-Pacific. Throughout syngnatharian history, multiple independent lineages colonized the eastern Pacific (6–8 times) and the Atlantic (6–14 times) from their center of origin, with most events taking place following an east-to-west route prior to the closure of the Tethys Seaway *ca.* 12–18 Ma. Ultimately, our study highlights the importance of accounting for different factors generating uncertainty in macroevolutionary and biogeographic inferences.[Historical biogeography; jump-dispersal parameter; macroevolutionary uncertainty; marine fishes; syngnathiformes; ultraconserved elements].

Molecular phylogenetic studies are steadily resolving long-lasting uncertainties in the Fish Tree of Life, most notably within percomorphs, a clade of spiny-finned fishes often referred to as the "bush at the top" (Nelson 1989) that is characterized by spectacular morphological and taxonomic diversity (*ca*. 18,000 species). Interrelationships and composition of major percomorph lineages remained controversial after decades of anatomical studies (Johnson 1993; Johnson and Patterson 1993), but recent phylogenetic analyses based on molecular evidence have unambiguously supported the resolution of the percomorph bush into nine supraordinal clades (Near et al. 2012; Betancur-R et al. 2013; Sanciangco et al. 2016; Betancur-R et al. 2017; Alfaro et al. 2018; Hughes et al. 2018). Most of these clades have never been inferred on the basis of morphological evidence, and therefore the origin and mechanisms shaping the evolutionary radiation of percomorphs remain poorly explored (Alfaro et al. 2018). Recent studies have suggested an association between the Cretaceous-Palaeogene (K-Pg) mass extinction and the origin of five of these nine percomorph crown groups, implying an important effect of extinctions on the evolutionary dynamics that resulted in the clades' astonishing diversity (Alfaro et al. 2018; Ribeiro et al. 2018).

One such percomorph clade has been classified as the series Syngnatharia, which includes *ca*. 670 described species arranged in 10 families, four suborders (Callionymoidei, Dactylopteroidei, Mulloidei, and Syngnathoidei) and a single order (Syngnathiformes; Betancur-R et al. 2017). Best-known among syngnatharians are seahorses and pipefishes (family Syngnathidae), forming a clade with over 320 species (Fricke et al. 2020). The fossil record of Syngnatharia dates back to the Late Cretaceous and includes representatives from all suborders (Carnevale et al. 2006; Bannikov 2014; Cantalice and Alvarado-Ortega 2016; Carnevale and Bannikov 2019). Most of these fossils have been found in the Eocene strata of Monte Bolca (Italy), at the ancient western Tethys Sea—one of the most important localities bearing exquisitely preserved fossil teleosts known from the paleontological record (Bannikov 2014;

Carnevale et al. 2014; Friedman and Carnevale 2018). Many extant syngnatharian families are circumglobally distributed, occurring in tropical and temperate marine waters, although some species inhabit brackish and freshwater environments (Whitfield 1999; York et al. 2006; IUCN 2019; OBIS 2019). Syngnatharians live in diverse habitats from soft substrates (e.g., goatfishes, seamoths, flying gurnards) to more complex environments such as coral and rocky reefs (seahorses, trumpetfishes) and mangrove forests (seahorses); other groups also inhabit sea-grass beds (pipefishes) and drifting macroalgae (*Sargassum*-associated pipefishes) (Froese and Pauly 2019). Reflecting this diversity in habitats, and the associated challenges in locomotion, reproduction and feeding ecology, fishes in this group display a variety of body plans. Some of the most conspicuous morphological traits and behaviors observed in syngnatharians include snout and body elongation (e.g., pipefishes, pipehorses, trumpetfishes, and cornetfishes; Bergert and Wainwright 1997), vertical swimming (e.g., shrimpfishes and seahorses; Atz 1962; Lin et al. 2016; Fish and Holzman 2019), prehensile tails (seahorses and pipehorses; Neutens et al. 2014; Hamilton et al. 2017), hyoid barbels (goatfishes; Gosline 1984), and, most remarkably, male pregnancy (pipefishes, pipehorses and seahorses; Berglund et al. 1986). Although there is no morphological evidence that unifies the disparate array of clades included in Syngnatharia (e.g., goatfishes and seahorses), the monophyly of the group is consistently resolved by different molecular studies, whether based on mitochondrial markers alone (Kawahara et al. 2008; Song et al. 2014), a handful of nuclear and mitochondrial markers (Chen et al. 2003; Near et al. 2012; Betancur-R et al. 2013) or genome-scale data sets (Longo et al. 2017; Alfaro et al. 2018; Hughes et al. 2018).

Despite robust support for the monophyly of Syngnatharia, phylogenetic relationships among early branching (suborder-level) clades have been controversial due to a seemingly explosive radiation during the early stages of their evolution (Near et al. 2013; Betancur-R et al. 2017; Longo et al. 2017; Alfaro et al. 2018). Short internodes along the backbone of the syngnatharian phylogeny make it challenging to resolve these relationships due to high levels of incomplete lineage shorting (Maddison 1997) and/or low signal-to-noise ratios (Rokas and Carroll 2006; Townsend et al. 2012). Whereas previous studies using sparse taxonomic sampling or few nuclear markers failed to provide robust resolution at the suborder level (Near et al. 2012; Betancur-R et al. 2013; Song et al. 2014; Hughes et al. 2018), more recent phylogenetic analyses based on genome-scale data coupled with dense-taxonomic sampling (Longo et al. 2017) delineated the four major, morphologically supported suborders (Kim 2002; Wiley and Johnson 2010; Nelson et al. 2016; see also Betancur-R et al. 2017), making progress towards the resolution of this spectacular radiation.

Resolving syngnatharian relationships will open new avenues to address a variety of macroevolutionary inferences, such as historical biogeography. Previous attempts to investigate the biogeographic history of this group have been restricted to particular families or genera (e.g., Syngnathidae, *Hippocampus, Mulloidichthys*; Bowen et al. 2001; Teske et al. 2004, 2007; Boehm et al. 2013; Lessios and Robertson 2013; Hamilton et al. 2017; Li et al. 2021). These studies have consistently identified the Indo-Pacific (IP) as the center of origin for these geographically widespread clades, but their results differed regarding the inferred routes and timing of the colonization of the Atlantic and the eastern Pacific (EP).

These previous studies, and others that have examined the historical biogeography of other clades, do not typically account for factors generating uncertainty in comparative analyses, such as variations in tree topology and divergence times (e.g., Batista et al. 2020) or alternative delineation of biogeographic areas and models (e.g., Richardson et al. 2004). Recent advances in statistical approaches for ancestral range estimations now allow the implementation of alternative biogeographic models—e.g., dispersal-extinction-cladogenesis (DEC; Ree and Smith 2008), dispersal-vicariance–analyses (DIVA; Ronquist 1997), and Bayesian inference of historical biogeography for discrete areas (BayAREA; Landis et al. 2013)—and parameters (e.g., the jump-dispersal or founder-event speciation [*j*]) with different assumptions regarding the dispersal abilities of groups between areas at different time slices (e.g., LaGrange: Ree and Smith 2008; BioGeoBEARS: Matzke 2013). Remarkably, among the alternative biogeographic models available for analyses, variants that include the *j* parameter have been recently criticized because their likelihood can be artificially inflated leading to an overestimation of jump-dispersal events (Ree and Sanmartín 2018; but see Klaus and Matzke 2020).

Here, we assess the evolutionary and biogeographic history of Syngnatharia using a variety of sensitivity analyses in a phylogenomic comparative framework. We first expanded a recently constructed molecular phylogeny for Syngnatharia that examined ultraconserved element (UCE) loci (Longo et al. 2017) to include 62 additional species (169 species total), and inferred robust time-calibrated trees by integrating paleontological and geological information.

To account for topological and divergence-time uncertainty in downstream comparative analyses, we also partitioned the complete matrix into genomic subsets. With the resulting, comprehensive phylogenetic framework, we explored the global biogeography of the group by incorporating geographic distribution data from extant and fossil species in three different time slices. Additionally, we examined uncertainties in ancestral range estimations using different sets of predefined area schemes, biogeographic models (including and excluding the *j* parameter), and topologies. Finally, in light of these results, we inferred possible routes through which different lineages colonized the EP and the Atlantic.

Materials and Methods

Taxonomic Sampling and Genomic Data.—We describe the main methods here, but also refer readers to the extended Materials and Methods in the Supplementary material available on Dryad at https://doi.org/10.5061/dryad.4xgxd2580. Our genomic data set was built upon a recently published phylogenomic analysis of Syngnatharia based on UCE data for 113 species (Longo et al. 2017), including a scombroid species as the outgroup (*Taractichthys longipinnis*). We added 78 newly sequenced specimens (62 species), including one additional family (Draconettidae) thereby covering all 10 representative families in the group. We initially assembled a data set of 190 individuals. We updated the total number of species after implementing steps for sample quality-control (see Supplementary material available on Dryad), based on CO1 sequence comparisons to the Barcode of Life Database (BOLD) and National Center for Biotechnology Information (NCBI) databases. These steps resulted in the reidentification of six taxa, including four sequenced by (Longo et al. 2017) that turned out to be duplicates (see Supplementary material available on Dryad). Our final data set consisted of 184 individuals comprising 169 syngnatharian species (107 previously published and 62 newly sequenced) or 25.3% of the clade's diversity (Supplementary Appendix 1 available on Dryad).

DNA Extractions, UCE Library Preparation and Sequencing.—We extracted DNA from tissue samples in a 96-well plate format on a GenePrep, following manufacturer's instructions at the Laboratory of Analytical Biology at the Smithsonian National Museum of Natural History. We checked the quality of DNA extractions by visually inspecting high molecular weight DNA stained with GelRed (Biotium) on a 1% agarose gel. High quality DNA extractions were sent to Arbor Biosciences for target enrichment using the MYbaits UCE Capture Kits, a custom bait set of approximately 1340 UCE loci identified in acanthomorph fishes (McGee et al. 2016); available from http://ultraconserved.org). Enriched libraries were quantified with qPCR (Kapa Biosystems) and sequenced using a single lane of Illumina HiSeq 2500 at the U. Chicago Genomics Facility.

Data assembly and Alignment.—We used the PHYLUCE computational pipeline (Faircloth 2016) to process the raw sequence reads (http://phyluce.readthedocs.io/en/latest/tutorial-one.html). This pipeline generated a "monolithic" FASTA file with all UCEs for all newly sequenced individuals. After this step, we added to the "monolithic" file UCE data for the 113 species (112 syngnatharians and the scombroid outgroup) previously sequenced (Longo et al. 2017). We "exploded" this file to obtain individual UCE loci and aligned them using MAFFT (Katoh and Standley 2013) based on a maximum divergence of 0.2. We trimmed the resulting alignments using Gblocks v0.91b (Castresana 2000) to remove ambiguously aligned flanking regions. Edited alignments consisted of 1186 UCE loci that we used to generate two subsets, each including all taxa examined: a 75% completeness matrix with 932 UCEs (142 taxa contain data in all gene alignments) and a 90% completeness matrix with 346 UCEs (171 taxa in all gene alignment). The resulting alignments included nine of ten syngnatharian families, with only Draconettidae missing. To cover all representative families, we probed the raw data from a specimen of *Draconetta xenica* that was sequenced using exon capture approaches (Hughes et al. 2020) to identify shared UCE loci using the map to reference function implemented in Geneious v.11.1.2 program (Kearse et al. 2012). We recovered a total of 17 and 50 loci present in the 90% and 75% completeness matrices, respectively. See Supplementary material available on Dryad for additional details.

Phylogenomic Analyses.—For both matrices, we determined the best-fit partitioning scheme as well as the best-fit model for each partition using the Bayesian Information Criterion (BIC) in Partition Finder v2.1.1 (Lanfear et al. 2017). We used the sliding-window approach and entropy site characteristic (SWSC-EN), a partition method proposed for UCE data (Tagliacollo and Lanfear 2018). The SWSC-EN produces partitions for each locus based on a core and two flanking regions. We estimated concatenation-based maximum likelihood (ML) trees in RAxML v8.2.4 (Stamatakis 2014) using the best-fit partitioning schemes and the GTRGAMMA substitution model. We ran 30 independent ML searches and used nonparametric bootstrapping via autoMRE (number of bootstrap replicates automatically determined) to assess edge support. We also conducted coalescent-based species-tree analyses in ASTRAL-II (Mirarab and Warnow 2015) using RAxML gene trees as input. Gene trees were inferred using the UCE core-flank partitions and the same parameters applied for concatenation analyses. Before phylogenetic dating analyses, we pruned duplicate individuals per species from the corresponding trees.

To account for variation in topology and divergence times in biogeographic analyses (see below), we also assembled 12 largely independent subsets (subsampled from the 75% matrix), each with a sufficient number of genes to overcome sampling error. Preliminary tests including a higher number of subsets, each with fewer genes (18 subsets), resulted in high levels of topological discordance, particularly for trees estimated with ASTRAL-II. We thus reduced the number of subsets to 12 (two with 89 loci, six with 90 loci, and four with 91 loci), all of which produced trees with lower levels of topological incongruence compared to those obtained using fewer genes. To maintain *Draconetta xenica* across all trees, all subsets overlapped in 17 anchor UCE markers that include this taxon. We conducted phylogenetic analyses using the 12 subsets in RAxML and ASTRAL-II, as explained above, producing a total of 24 trees (two per subset). Finally, we used the 28 trees

TABLE 1. Alternative biogeographic schemes used in BioGeoBEARS

Scheme	Number of areas	Areas	Maximum range size parameter
1. Six-area	6 + Tethys Sea	WIO, CIP, CP, EP, WA, EA	6
2. Seven-area	7 + Tethys Sea	WIO, CIP, CP, TA, EP, WA, EA	7
3. Eight-area	8 + Tethys Sea	WIO, CIP, CP, TA, TNP, TEP, WA, EA	7

WIO = Western Indian Ocean; CIP = Central Indo-Pacific; CP = Central Pacific; TA = Temperate Australasia; EP = Eastern Pacific; TNP = Temperate Northeast Pacific; TEP = Tropical Eastern Pacific; WA = Western Atlantic; EA Eastern Atlantic.

inferred (two each with the 75% and 90% completeness matrices, and 24 with the subsets) as input for divergence time estimations in MCMCTree.

Phylogenomic Dating.—We estimated divergence times using the MCMCTree package implemented in PAML v4.9 (Yang 2007), which can analyze genome-scale data sets in a Bayesian framework (dos Reis and Yang 2019). A drawback of MCMCTree, however, is that it cannot jointly estimate topology and divergence times, requiring instead predefined topologies as input, for which we used the 28 topologies inferred in the previous step. Because running time in MCMCTree depends more on the number of defined partitions rather than the number of genes, we divided each subset into two partitions (flanks and core UCEs). We used a total of 11 calibration points, including a secondary root calibration (Pelagiaria + Syngnatharia), six fossil calibrations, and four geological calibrations (see Supplementary material available on Dryad). We ran the 75% and 90% matrices for 10–50 million generations and the genomic subsets for 3–20 million generations until convergence was reached based on effective sampling size values (>200). We used using the approximate likelihood method and the HKY85 model. Prior parameters for the MCMCTree runs were as follow: independent rate relaxed-clock model, BDparas: 1, 1, 0.27; kappa_gamma: 6, 2; alpha_gamma: 1, 1; rgene_gamma: 2, 200, 1; sigma2_gamma: 2, 5, 1. We conducted two independent runs for each data set. To check for convergence, we used Tracer v1.7.1 (Rambaut et al. 2018) to examine trace plots and ESS values for each parameter, after a 10% burn-in.

Biogeographic Analyses.—We ran biogeographic analyses using the R package BioGeoBEARS (Matzke 2013), which compares models of range evolution in a phylogenetic framework. We used the tree inferred in RAxML with the 75% completeness matrix ("master tree" hereafter) as the summary phylogeny. We also implemented three different biogeographic schemes to account for different levels of resolution for delimitation of biogeographic regions (Table 1). The first scheme is based on six recognized marine biogeographic regions for tropical fishes proposed by Kulbicki et al. (2013): Western Indian Ocean (WIO), Central Indo-Pacific (CIP), Central Pacific (CP), Tropical Eastern Pacific (TEP), Western Atlantic (WA), and Eastern Atlantic (EA). For the second scheme, based on Spalding et al. (2007), we added the Temperate Australasia (TA) area to the previous six-area scheme to account for species restricted to the temperate waters of Australia and New Zealand (these species were lumped into the CIP in the six-area scheme). Finally, for the third scheme, we further added an area to the seven-area scheme to discriminate species that are largely endemic to the Temperate Northeast Pacific (TNP; lumped into the EP in the 6- and 7-area schemes). We built a presence/absence matrix for each scheme by coding each extant species according to their geographic ranges primarily based on a combination of the IUCN Red List (IUCN 2019) and Ocean Biogeographic Information System (OBIS 2019) databases. We also used paleogeographic domain information as biogeographic constraints using the six fossils used to calibrate our tree. These constraints were placed in the nodes into which fossil calibration points were assigned (Table 2). Five of the six fossil species, placed in Syngnatharia, Syngnathidae/Solenostomidae, *Hippocampus*, and †Ramphosidae (the sister family of Pegasidae), were present in the western Tethys or Paratethys and thus coded as an additional area, the ancient Tethys Sea (calibrations 2–4, 6, and 7; Table 2 and Fig. 1). The sixth fossil used for Aulostomoidea (calibration 5; Table 2 and Fig. 1) was described from the Gulf of Mexico/Proto Caribbean Sea and thus coded as WA.

We tested 12 biogeographic models in a maximum likelihood framework, previously used for marine fishes (Siqueira et al. 2019; Rincon-Sandoval et al. 2020). These models include DEC (Ree and Smith 2008), DIVA (Ronquist 1997), and BayAREA (Landis et al. 2013). Each model was run with and without the founder-speciation event (*j*) (Matzke 2014) and the dispersal matrix power exponential (*w*) parameters (Dupin et al. 2017). The *j* parameter allows the colonization of a new area by a daughter lineage while the splitting-sister lineage stays at the ancestral area (Matzke 2014). The *w* parameter is used to infer the optimal dispersal multiplier matrix, acting as an exponent on the matrix (Dupin et al. 2017). We set the *w* parameter to be free in order to allow the model to adjust the matrices according to the data. We analyzed each model using three time slices (100–12 Ma, 12–2.8 Ma, and 2.8–0 Ma) according to different geological events that span the evolutionary history of the clade. The Tethys Sea region was only added to the analyses for the first time slice (100–12 Ma) to reflect the existence of this ancient basin. A dispersal-multiplier matrix for each scheme was assembled to account for the dynamics of biogeographic barriers over time. The connectivity between areas was determined by three dispersal probability categories: 1.0 for well-connected areas, 0.05 for relatively separated areas, and 0.0001 for widely separated or disconnected areas. From 100 to 12 Ma, we allowed high dispersal probability

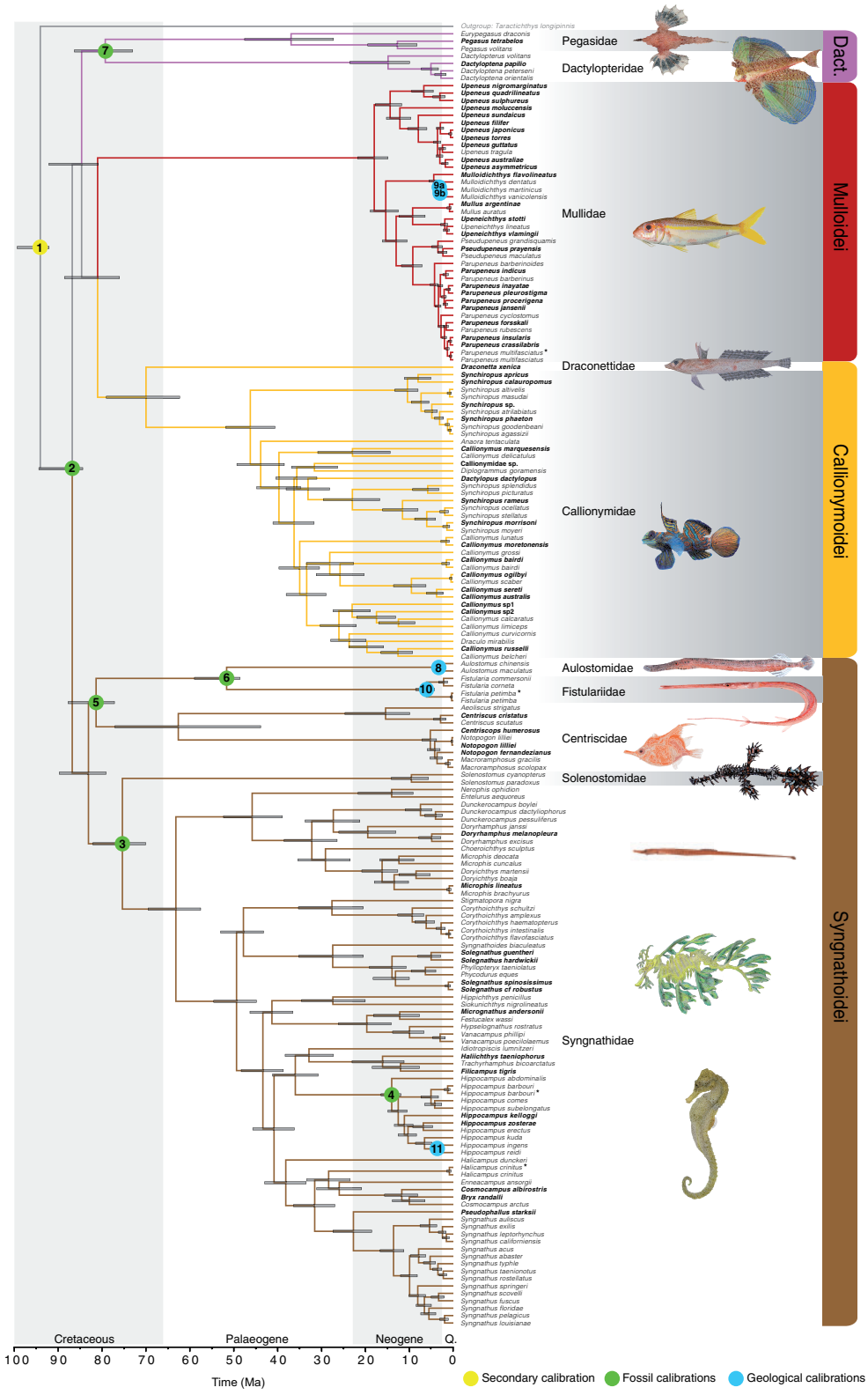

FIGURE 1. Time-calibrated phylogeny ("master tree") for 169 species of Syngnatharia estimated in MCMCTree. The topology reflects the maximum likelihood RAxML tree based on 932 UCEs (75% completeness matrix). Gray bars at nodes represent the 95% HPD intervals. Circles at the nodes indicate the 11 calibration points used in MCMCTree, which are colored according to the type of calibration used (see Supplementary material available on Dryad). For support values see Supplementary Fig. S3 available on Dryad. Dact. = Dactylopteroidei; Q. = Quaternary; Ma = millions of years. New taxa added for this study marked are shown in bold; asterisks (*) denote taxa examined by (Longo et al. 2017) that were reidentified.

TABLE 2. Fossil geographic distributions used as biogeographic constraints in BioGeoBEARS

Fossil calibrations	MRCA	Paleogeographic domain	Area code
(2) Syngnatharia	*Syngnathus louisianae*, *Eurypegasus draconis*	Western-central Tethys	Tet
(3) Syngnathidae/Solenostomidae	*Solenostomus cyanopterus*, *Syngnathus louisianae*	Western Tethys	Tet
(4) *Hippocampus*	*Hippocampus abdominalis*, *Hippocampus kuda*	Pannonian Basin, Central Paratethys	Tet
(5) Aulostomoidea	*Aulostomus maculatus*, *Aeoliscus strigatus*	Gulf of Mexico/Proto Caribbean Sea	WA
(6) Fistulariidae	*Fistularia corneta*, *Aulostomus maculatus*	Western Tethys	Tet
(7) Pegasidae	*Pegasus volitans*, *Dactylopterus volitans*	Western Tethys	Tet

Tet = Tethys Sea; WA = Western Atlantic.

between WIO and EA through the Tethys Seaway. The Terminal Tethyan Event (TTE), which opened and closed intermittently between 12 and 18 Ma, divided the western and the eastern Tethys Sea (Steininger and Rögl 1979; Adams et al. 1983; Rögl 1998). Thus, from 12 Ma (final closure) onwards we used a low probability value between WIO and EA to allow dispersal through southern Africa (Rocha et al. 2005). To account for the final closure of the Isthmus of Panama, which may have occurred as early as 2.8 Ma as stated above (O'Dea et al. 2016), we assigned a very low dispersal probability between WA and EP. Finally, for all time slices, we set a high dispersal probability between adjoining areas of the Indo-Pacific and a low dispersal probability between CP and EP to reflect dispersal limitations associated with the crossing of the EP Barrier (Bellwood and Wainwright 2002; Lessios and Robertson 2006). Using the "master tree" as the input phylogeny, we calculated Akaike Information Criterion scores corrected for small sample size (AICc) for each biogeographic model and for each biogeographic scheme independently. The best-fitting model (DEC, DIVA, and BayAREA, each with a combination of $\pm j$and $\pm w$ parameters) was then selected for each scheme (Supplementary Table S4 available on Dryad).

Accounting for Uncertainty in Biogeographic Analyses.—We assessed sensitivity of ancestral range estimations to three major sources of variation: topology and divergence times, area schemes, and biogeographic models. First, based on a recently proposed approach (Rincon-Sandoval et al. 2020), we used the set of 28 trees inferred using the 75% and 90% completeness matrices as well as the 12 genomic subsets. Resulting trees reflect uncertainty in divergence times and phylogenetic relationships based on different underlying data. This approach fundamentally differs from the common practice of conducting comparative analyses using "pseudo-replicated" trees obtained from a Bayesian posterior distribution estimated with a single data set, typically consisting of a handful of genes (Huelsenbeck et al. 2000). To assess topological disparity, we estimated tree space plots for the 28 trees using a multidimensional scaling (MDS) visualization implemented in the R package *phytools* (Revell 2012). To assess the effects of phylogenetic variation in biogeographic inferences, we used a code produced by (Matzke 2019) to summarize ancestral range estimates from multiple trees by selecting the "master tree" as the topology upon which the results from all 28 trees were overlain. This approach allowed us to obtain averaged probabilities across the different trees for compatible nodes present on the "master tree." For comparison, we also estimated ancestral ranges based on an alternative tree estimated with ASTRAL-II using the 75% completeness matrix ("alternative tree" hereafter) and the 28-tree averaging approach described above. The final set of analyses aimed at accounting for phylogenetic uncertainty involved running BioGeoBEARS without the averaging approach, using only the "master tree" (see also previous section) and the "alternative tree."

Second, we compared the biogeographic results obtained with the three different area schemes defined (6 areas, 7 areas, and 8 areas), using both the master and alternative trees with and without the 28-tree averaging approach. Finally, given recent criticisms on the implementation of the jump-dispersal parameter (*j*) parameter (Ree and Sanmartín 2018), we interpreted different biogeographic histories based on analyses of the best-fitting models selected for different trees and area schemes, with (+*j*) and without (-*j*) this parameter (Supplementary Table S4 available on Dryad).

In cases where colonization events of oceanic basins inferred from these different types of analyses produced incongruent results, we assessed the relative probabilities of these histories by conducting biogeographic stochastic mapping (BSM), as implemented in BioGeoBEARS (Dupin et al. 2017). For BSM analyses, a total of four possible routes were assessed (Floeter et al. 2008): (i) Tethyan relicts, (ii) lineages with Indo-Pacific origin that crossed the Tethys Seaway before its closure, (iii) lineages with Indo-Pacific origin that colonized via the Cape of Good Hope, southern Africa, and (iv) lineages with Indo-Pacific origin that crossed the EP Barrier. We simulated 1000 stochastic histories on the "master tree" based on the best-fit biogeographic model (with and without the *j* parameter) and calculated the probability for alternative routes. These alternatives are only considered (depicted in maps) when the probability for a major colonization event is less than 70%.

RESULTS

Phylogenomic Inference, Tree Uncertainty and Divergence Times.—We conducted phylogenomic analyses using maximum likelihood (ML; RAxML) and coalescent-based (ASTRAL-II) approaches applied to the two

assembled matrices—the 75% completeness matrix (932 UCEs, 268,279 sites, 11.6% missing data) and the 90% completeness matrix (346 UCEs, 119,467 sites, 6% missing data). Overall, phylogenetic relationships among the four suborders previously defined are congruent and highly supported (>75%) based on the concatenation-based ML analyses (Fig. 1 and Supplementary Figs. S2, S3, and S5 available on Dryad). However, trees inferred with ASTRAL-II using both matrices (75% and 90%; Supplementary Figs. S4 and S6 available on Dryad) did not resolve the monophyly of Dactylopteroidei (Dactylopteridae + Pegasidae). Additionally, the suborder Callionymoidei, a clade comprising the families Draconettidae and Callionymidae, which have a strong morphological affinity (Gosline 1984; Wiley and Johnson 2010; Nelson et al. 2016), was not resolved as monophyletic with ASTRAL-II using the 90% matrix. These results suggest that lower gene coverage for Draconettidae (only 17 and 50 UCE loci present in the 90% and 75% matrices, respectively) may have affected ASTRAL-II analyses more than concatenation-based inferences. Trees estimated with RAxML had higher average bootstrap support values than those estimated with ASTRAL-II (96.8-98.3% vs. 90.5-92.4%, respectively). Likewise, trees estimated with the 75% matrix resulted in clades with higher support values relative to the 90% matrix (mean support 92.4-98.3% vs. 90.5-96.8%, respectively). All families were resolved as monophyletic in all inferred trees. Similar topologies were obtained using the additional 12 subsets (24 trees), except for the suborders Dactylopteroidei (15 trees), Callionymoidei (4 trees), and Syngnathoidei (2 trees), which were not resolved as monophyletic in some trees, mostly those estimated using ASTRAL-II (13 ASTRAL-II trees vs. 5 RAxML trees; Supplementary Fig. S7 available on Dryad). The relative placement of the family Centriscidae, most often resolved as a sister group to the clade composed of Aulostomidae + Fistulariidae, also varies in eight subset-based trees.

The MDS plots of assessment of topological disparity between the 28 trees inferred by different methods show that, regardless of the number of genes, RAxML and ASTRAL-II trees fall in opposite areas of tree space, never overlapping (Supplementary Fig. S8 available on Dryad). The ASTRAL-II trees reveal, however, a greater degree of topological disparity than RAxML trees, a pattern that is probably the result of gene-tree error affecting ASTRAL-II reconstructions. While trees inferred with more than 300 loci (75% and 90% completeness matrices) tend to be more tightly clustered in the tree space relative to subset-based trees, RAxML and ASTRAL-II topologies obtained with the same genomic data set or subset differ substantially (see Supplementary Fig. S8 available on Dryad).

The time-calibrated phylogeny of syngnatharians based on 11 calibration points in MCMCTree is shown in Fig. 1 (RAxML "master tree"); results obtained with the 28 trees are summarized in Supplementary Table S3 and Fig. S9 available on Dryad. Our inferred evolutionary timescale places the origin of total group Syngnatharia at 94.1 Ma (95% highest posterior density [HPD] 92.0–99.3 Ma), whereas the crown group age is dated at 86.8 Ma (HPD 84.4–94.4 Ma) in the Late Cretaceous. With the exception of the long-stemmed Mulloidei, which originated at 18.0 Ma (HPD 14.9–21.8 Ma), the divergence of all other major suborder-level clades also took place in the Late Cretaceous (~70–83 Ma), long before the Cretaceous-Palaeogene (K-Pg) mass extinction event (~66 Ma).

Sensitivity of Biogeographic Analyses to Tree Uncertainty, Area Schemes and the j Parameter.—Sensitivity analyses to the three sources of variation examined reveal that the use of the *j* parameter has the strongest effect on ancestral range estimations, followed by the number of areas defined (see Fig. 3 and Supplementary Table S5 available on Dryad for details). Tree variance, by contrast, has a relatively smaller effect on the inferences (Supplementary Tables S5 and S6 available on Dryad). The biogeographic reconstructions conducted to account for estimation error show that approximately one-third of the colonization routes vary depending on the analysis (Fig. 3). These biogeographic patterns tend to be more similar between different area schemes using the same model rather than within each area scheme using different models (i.e., including or excluding the *j* parameter; Fig. 3 and Supplementary Table S5 available on Dryad).

As expected, ancestral range estimates that use the BayAREA model along with the *j* parameter tend to identify more long-distance and recent dispersal events than those using the BayAREA model alone, which are otherwise more consistent with a Tethys Sea origin for many clades implying fewer colonization events due to widespread ancestors despite a lower model fit overall (AICc 1082–1313 for BayAREA vs. AICc 1015–1255 for BayAREA +*j*; Figs. 2 and 3). At least nine major differences in colonization routes are observed between analyses that include or exclude the *j* parameter, four of which are observed in the family Mullidae alone (Fig. 3). According to the -*j* inferences, this family colonized the WA from the Tethys Sea/IP at 17.6 Ma (HPD 16.1–19.2 Ma) followed by dispersal of *Mulloidichthys* from the IP to the EP via the EP Barrier at 3.9 Ma (HPD 3.2–4.6 Ma), and a subsequent dispersal at ~2.9 Ma into the WA through the Central American Seaway prior to the emergence of the Isthmus of Panama. This inference also suggests that *Pseudupeneus* colonized the EP from the WA at 2.9 Ma (HPD 2.2–3.5 Ma). In contrast, +*j* range estimates show that both *Mulloidichthys* and *Pseudupeneus* colonized the WA through the EP Barrier (3.9 Ma [HPD 3.2–4.6 Ma] and 6.6 Ma [HPD 3.5–9.6 Ma], respectively), whereas *Mullus* dispersed at 5.3 Ma (HPD 1–9.6 Ma) into that basin via the Cape of Good Hope (southern Africa; Fig. 3).

Ancestral range estimations also differ based on the number of areas used, but these are also largely confounded by the inclusion or exclusion of the *j*

parameter. For example, biogeographic analyses based on the six- and eight-area schemes (±*j*) or the seven-area scheme (-*j*) indicate that the genera *Synchiropus* and *Hippocampus* are Tethyan relicts that colonized the WA at 27.9 Ma (HPD 10.1–45.0 Ma) and at 13.2 Ma (HPD 12.3–14.0 Ma), respectively. Conversely, according to the seven-area +*j* inferences, *Synchiropus* took a different route via southern Africa to colonize the WA at 5.9 Ma (HPD 4.–7.1 Ma), whereas *Hippocampus* colonized that basin in two independent dispersal events. The most recent common ancestor (MRCA) of *Hippocampus zostera* and *Hippocampus erectus* (8.8 Ma, HPD 7.1–10.5 Ma) dispersed via southern Africa, followed by the WA colonization of the MRCA of *Hippocampus reidi* and *Hippocampus ingens* (5.1 Ma, HPD 6.5–3.8 Ma), most likely through southern Africa (65.4% probability) rather than via the Central American Seaway after crossing the EP Barrier (26.9% probability). Noteworthy, while the six- and eight-area (±*j*) or the seven-area (-*j*) inferences suggest that *H. ingens* crossed the Central American Seaway and colonized the EP at ∼3.8 Ma, the seven-area +*j* inference supports this colonization event but there is also a smaller probability (26.9%) that *H. reidi* could have crossed the seaway from the EP to the WA (vs. 65.4% through southern Africa). Lastly, the six- and eight-area schemes -*j* identify an additional colonization of the EP through the Central American Seaway in *Cosmocampus*. In summary, we find that for most clades the differences observed among area schemes are most striking when the *j* parameter is used (Fig. 3), particularly with the seven-area scheme.

Finally, analyses using the "master tree" (Fig. 2 and Supplementary Figs. S10–S14 available on Dryad) and the "alternative tree," with and without the 28-tree averaging approach, resulted in rather similar biogeographic histories (Supplementary Tables S5 and S6 available on Dryad) and colonization routes, with a few exceptions. For instance, in one estimation (7 areas, +*j*) the "alternative tree" supports the colonization of crown Mullidae into the WA ∼18 Ma, whereas the "master tree" suggests that colonization of this area took place in three different mullid lineages (*Mulloidichthys, Mullus,* and *Pseudupeneus*) much later (∼3.9–9.6 Ma). Likewise, only a few major differences are found between ancestral range estimates based on 28-tree averaging vs. single tree approaches. For example, inferences based on all 28 trees (summarized on either the master or the alternative trees), indicate that the Tethys Sea is the ancestral area state for the MRCA of *Dactylopterus* + *Dactyloptena*, crown *Aulostomus*, and several other lineages. In contrast, those nodes appear to be more widespread based on estimates that used either the "master tree" or the "alternative tree" alone (Supplementary Tables S5 and S6 available on Dryad).

Ancestral Range Estimation and Colonization of the Atlantic and Eastern Pacific.—Because of the uncertainties noted above, in this section we focus on identifying emergent patterns that are congruent across the different analyses to explain the biogeographic history of syngnatharians (Figs. 2 and 3). We chose to illustrate ancestral range estimates obtained on the basis of the seven-area scheme (Fig. 2) given that an important fraction (9.7%) of the species diversity in syngnatharians are endemic to TA (e.g., *Upeneichthys stotti*, *Solegnathus spinosissimus*, *Phycodurus eques*, *Phyllopteryx taeniolatus*; Hamilton et al. 2017). While the 8-area scheme also accounts for TA species, the additional area coded with this scheme (TNP) only includes a small proportion of endemics (1.1%).

Our analyses identified the ancient Tethys Sea as the center of origin for syngnatharians during the Late Cretaceous (86.8 Ma, HPD 84.4–94.4 Ma), followed by the origination of suborder-level lineages (∼70–83 Ma) before the Cretaceous-Palaeogene (K-Pg) mass extinction event (∼66 Ma). All families had an ancestor that was present in the Tethys Sea before their widespread colonization of the IP (WIO, CIP, CP, and/or TA). Syngnathidae originated in the late Cretaceous at 63.1 Ma (HPD 57.5–69.5 Ma) and started colonizing the New World in the Late Eocene. Subsequent to the origin of Centriscidae (62.6 Ma, HPD 43.9–77.1 Ma; Palaeocene/Eocene), the genera *Centriscus* and *Aeoliscus* persisted in the IP, while *Notopogon* and *Macroramphosus* dispersed into the Atlantic and EP during the Pliocene. Crown Callionymidae originated and diversified in the IP during the Eocene (46.2 Ma, HPD 40.6–51.8 Ma) while lineages in the genera *Callionymus* and *Synchiropus* colonized the Atlantic later in the Miocene. The family Dactylopteridae, which originated in the Tethys Sea/IP, also colonized the Atlantic during the Middle Miocene (genus *Dactylopterus*; 14.6 Ma, HPD 10.0–23.5 Ma). Among members of the family Mullidae (origin dated at 18.0 Ma, HPD 14.9–21.8 Ma), only the genus *Upeneus* remained restricted to the ancestral IP range, while the rest of the genera in the family colonized the Atlantic and the EP during the Miocene. Whereas the total group origin for the families Fistulariidae and Aulostomidae dates back to 51.6 Ma (HPD 48.6–58.9 Ma), their crown members diversified more recently at 6.1 Ma (HPD 4.3–8.6Ma; Miocene/Pliocene) and 3.3 Ma (HPD 2.8–4.5 Ma; Pliocene/Pleistocene), respectively. Our biogeographic analyses indicate that these two families dispersed into the three major basins (from a Tethys Sea ancestor of the total group Fistulariidae + Aulostomidae) around the Pliocene. Draconettidae (79.3 Ma; HPD 73.0–86.3 Ma; Late Cretaceous), Pegasidae (36.9 Ma; HPD 27.3–47.5 Ma; Eocene/Oligocene), and Solenostomidae (9.5 Ma; HPD 5.7–14.1 Ma; Neogene) are the only families that did not disperse outside their center of origin. Most of the genera also are inferred to have a Tethys Sea/IP (WIO, CIP, CP, or TA) origin except for two genera in the family Syngnathidae, *Enneacampus* and *Pseudophallus*, which probably had a WA origin.

Our ancestral range estimates combined with stochastic mapping suggest different routes of colonization to the EP and the Atlantic (Fig. 3). Except for the widespread species with circumglobal or

semi-circumglobal distributions, the EP was colonized 6–8 times, with at least one event taking place eastwards across the EP Barrier (*Mulloidichthys*) and the rest occurring via the WA through the Central American Seaway before the closure of the Isthmus of Panama (*Synchiropus*, *Hippocampus*, *Pseudophallus*; Fig. 3). Similarly, the Atlantic was colonized 6–14 times through three different routes: (i) 6–8 lineages were either Tethyan relicts or crossed the Tethys Seaway before its closure *ca.* 12–18 Ma (Steininger and Rögl 1979; Adams et al. 1983; Rögl 1998; e.g., *Dactylopterus volitans*, *Entelurus aequoreus* + *Nerophis ophidion*), (ii) 1–4 lineages colonized the Atlantic via southern Africa (e.g., *Mullus*), (iii) and at least one lineage passed from the EP to the Atlantic prior to the emergence of the Isthmus of Panama >2.8 Ma (e.g., *Mulloidichthys*). Finally, four different syngnatharian lineages (*Pseudupeneus prayensis*, *Synchiropus phaeton*, *Enneacampus ansorgii*, and some species in *Syngnathus*) colonized the EA via a west-to-east Atlantic route (Fig. 3).

Discussion

We investigated the evolutionary and biogeographic history of marine fish species in Syngnatharia by combining genomic (UCEs), paleontological, geologic, and geographic data layers. Although the biogeographic history of a few families or genera in this group have been examined in detail—-for example, Aulostomidae (Bowen et al. 2001), Syngnathidae (Hamilton et al. 2017), *Hippocampus* (Teske et al. 2004, 2007; Boehm et al. 2013; Li et al. 2021), and *Mulloidichthys* (Lessios and Robertson 2013) this is the first biogeographic study conducted for the entire clade. Our analyses accounting for topological,

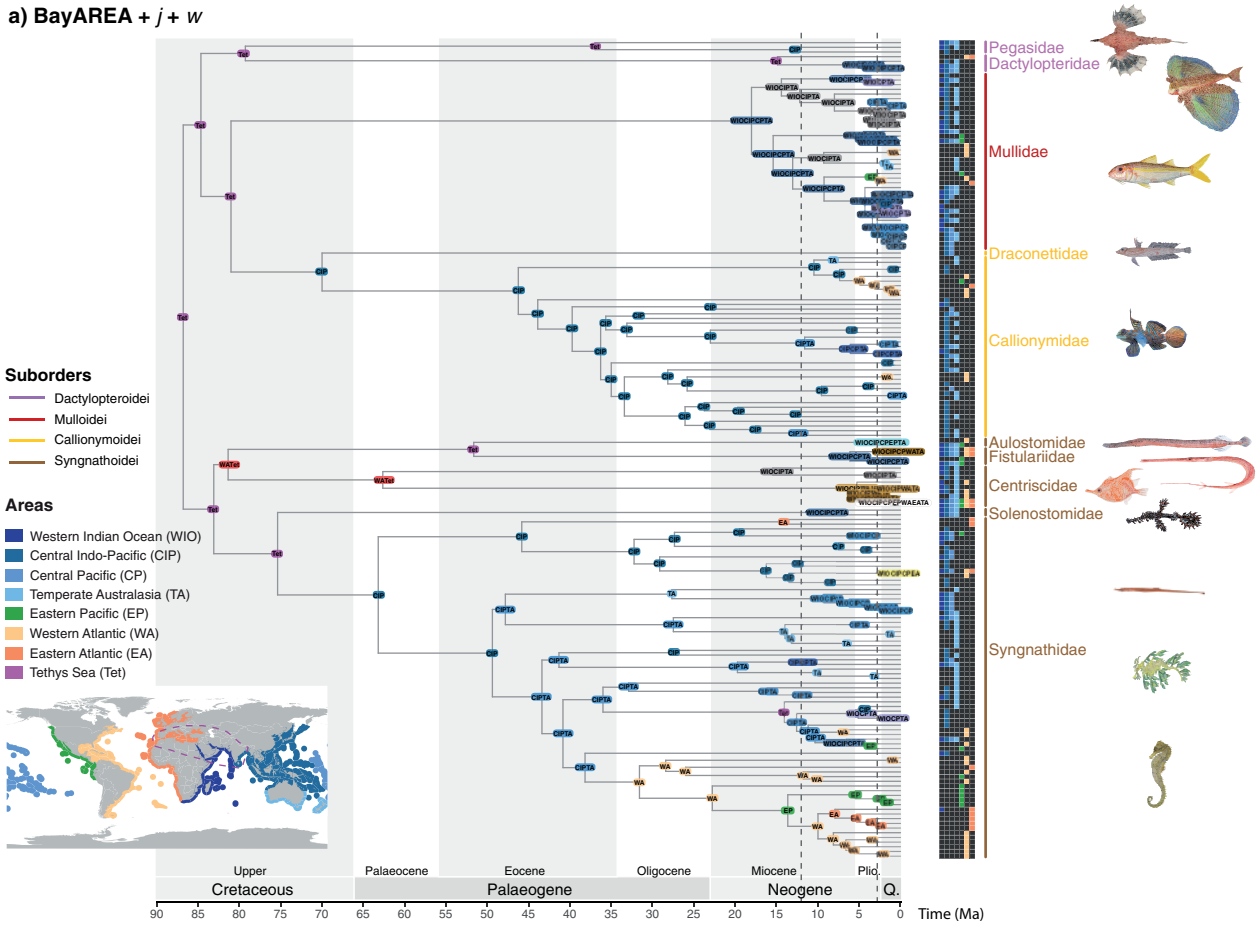

FIGURE 2. Ancestral range estimations for Syngnatharia based on the seven-area scheme applied to the 28 trees using the "master tree" as fixed topology in BioGeoBEARS. a) Best-fit biogeographic model based on the BayAREA+*j*+*w* model. b) Given recent criticisms around the use of the *j* parameter (Ree and Sanmartín 2018) the BayAREA model is also reported here (see also Supplementary Table S4 available on Dryad). Note that similar results were obtained with and without the *w* parameter, suggesting that this parameter alone is not a confounding factor in these comparisons. Size of boxes at the nodes are proportional to the number of areas in the estimated ancestral ranges. The map shows the seven marine biogeographic regions used to code the geographic distribution of extant species (based on Spalding et al. 2007; Kulbicki et al. 2013) and the ancient Tethys Sea. Families are color-coded by suborder. Dotted lines represent the time constraints corresponding to two major biogeographic events: the Tethys Seaway closure (12–18 Ma) and the undisputed minimum age for the closure of the Isthmus of Panama (2.8 Ma; see comments under divergence-time calibrations). Plio. = Pliocene; Q. = Quaternary; Ma = millions of years.

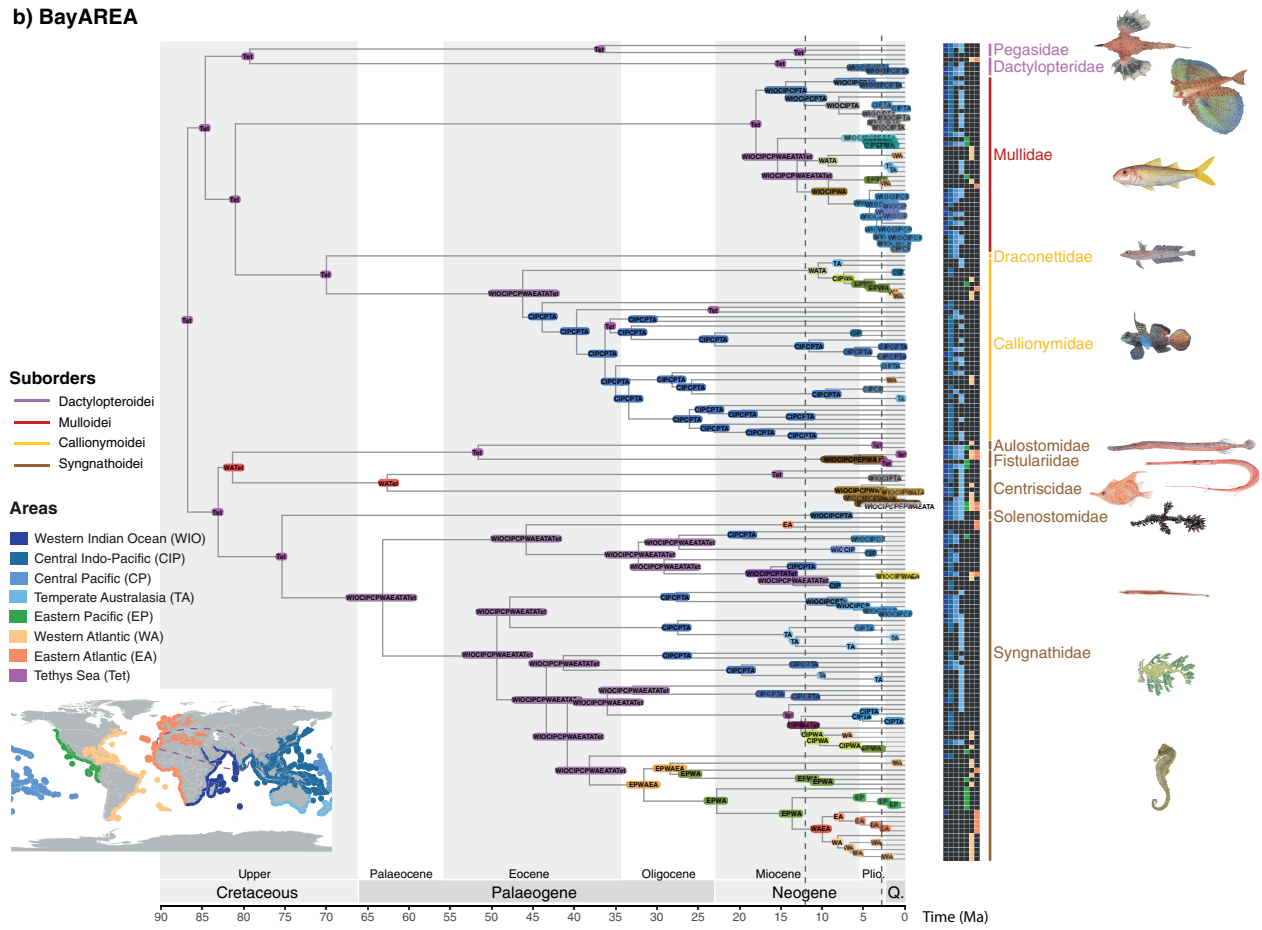

FIGURE 2. (Continued)

temporal, and biogeographic uncertainty support a Late Cretaceous origin of syngnatharians in the Tethys Sea, with subsequent dispersal into the central Indo-Pacific and independent colonizations of the eastern Pacific and the Atlantic by most families through alternative routes.

Evolutionary Relationships and Divergence Times.—The phylogenetic relationships among major clades differ slightly depending on the methodological approach used. In agreement with other recent studies (Longo et al. 2017; Alfaro et al. 2018; Fig. 1 and Supplementary Figs. S3 and S5 available on Dryad), concatenation-based analyses resolved an early split that supports the reciprocal monophyly of the long-snouted Syngnathoidei and a clade including the bottom-dwelling suborders Mulloidei, Callionymoidei, and Dactylopteroidei, most of which are short-snouted. The exception to this is Pegasidae, which like Syngnathoidei has an elongated snout due to enlargement of specialized bones of the neurocranium and suspensorium (Bergert and Wainwright 1997). The mouth in pegasids, however, is oriented ventrally (vs. terminal in syngnathoids; Pietsch 1978; Bergert and Wainwright 1997). Coalescent-based analyses, in contrast, resolved the suborder Syngnathoidei nested within a paraphyletic group comprising the bottom-dwelling suborders (Supplementary Figs. S4 and S6 available on Dryad). Remarkably, irrespective of the reconstruction method, all analyses resolved a monophyletic Syngnathoidei, which is consistent with results from many previous studies (e.g., Near et al. 2013; Betancur-R et al. 2017, based on exonic data; Longo et al. 2017; Alfaro et al. 2018, based on UCE data), but contrasts with others that examined fewer genes (Near et al. 2012; Betancur-R et al. 2013; Song et al. 2014) or taxa (Hughes et al. 2018) rejecting the monophyly of the suborder. The relationships among major lineages of Syngnathidae, the most diverse syngnatharian family, are also largely congruent with previous studies, showing an early divergence between trunk- and tail-brooders (Wilson and Orr 2011; Hamilton et al. 2017). Finally, while this and other previous phylogenetic studies provide support for the monophyly of all genera in Mullidae (Kim 2002; Longo et al. 2017), the resolution of other intrafamilial relationships is more elusive, including the nonmonophyly of genera in Callionymidae (e.g., *Synchiropus* and *Callionymus*) and Syngnathidae (e.g., *Microphis* and *Cosmocampus*).

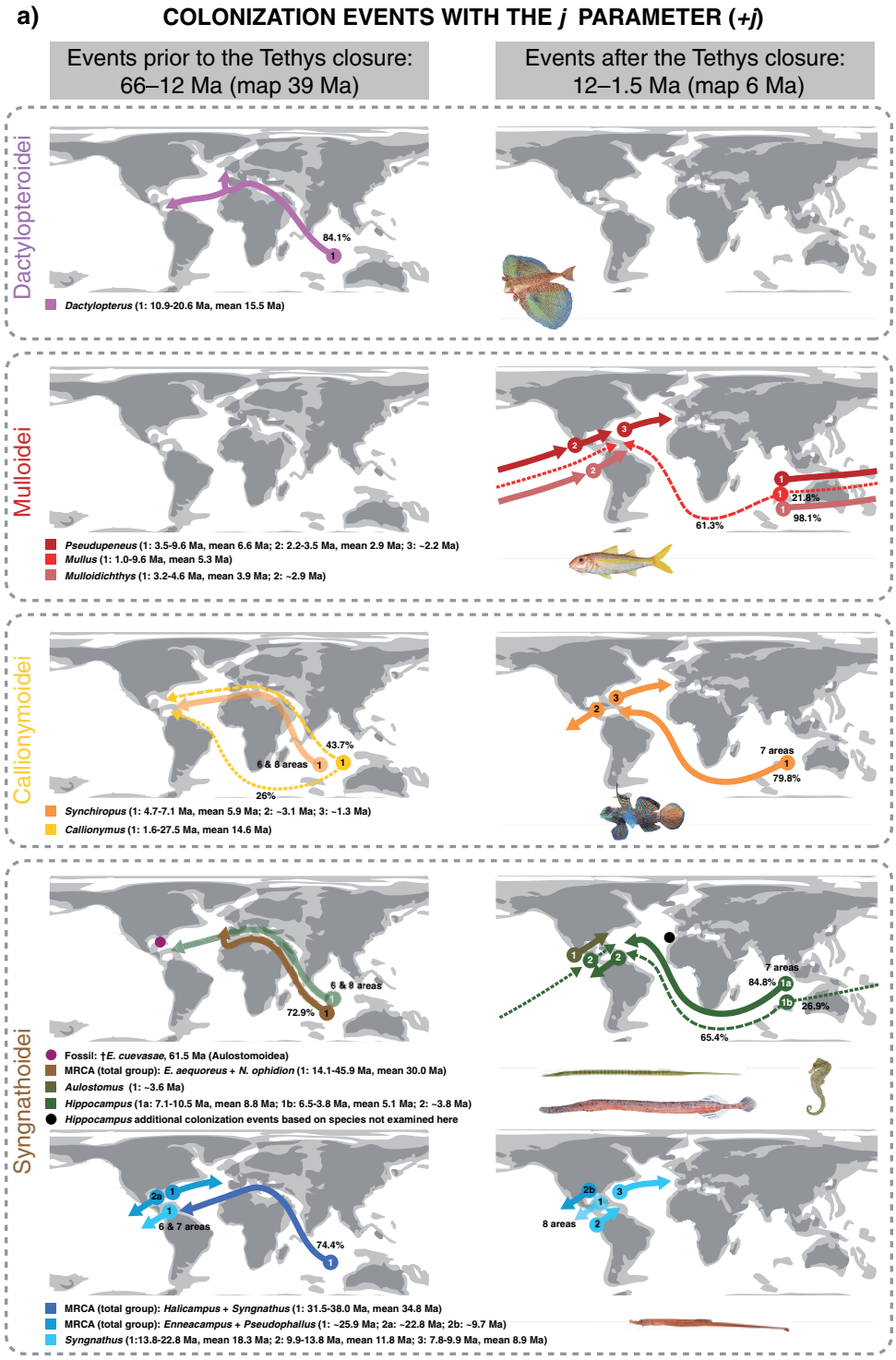

FIGURE 3. Possible dispersal routes that led to the historical colonization of the Atlantic and eastern Pacific basins (from a Tethys Sea or Indo-Pacific basin) in different clades of syngnatharian fishes as inferred using two alternative models. Ancestral ranges estimated using the BayAREA model: a) including the long-distance dispersal parameter (*j*), and b) without the *j* parameter. Solid arrows indicate ancestral range estimations using the favored seven-area scheme (see main text), and differences obtained with alternative area schemes (6 or 8) are denoted with transparency. The timing of dispersal events indicated are based on the seven-area scheme alone (Fig. 2; but see also Supplementary Figs. S13 and S14 available on Dryad for ages inferred with 6 and 8 areas). Dispersal routes mapped are macroevolutionary in scope, involving vicariant speciation events leading to the origin of at least one lineage restricted to one of the major basins. Clades including multiple widespread species with circumtropical and/or circum-temperate distributions (e.g., *Aulostomus chinensis*, *Fistularia* spp., *Doryrhamphus excisus*, *Syngnathus acus*, Centriscidae spp., and *Mulloidichthys vanicolensis*) are better examined using phylogeographic analyses (e.g., Bowen et al. 2001; Lessios and Robertson 2006) and are thus not mapped here. In cases where multiple routes are possible (e.g., *Dactylopterus volitans*), route probability is estimated based on biogeographic stochastic mapping using the "master tree" (indicated with dashed lines;

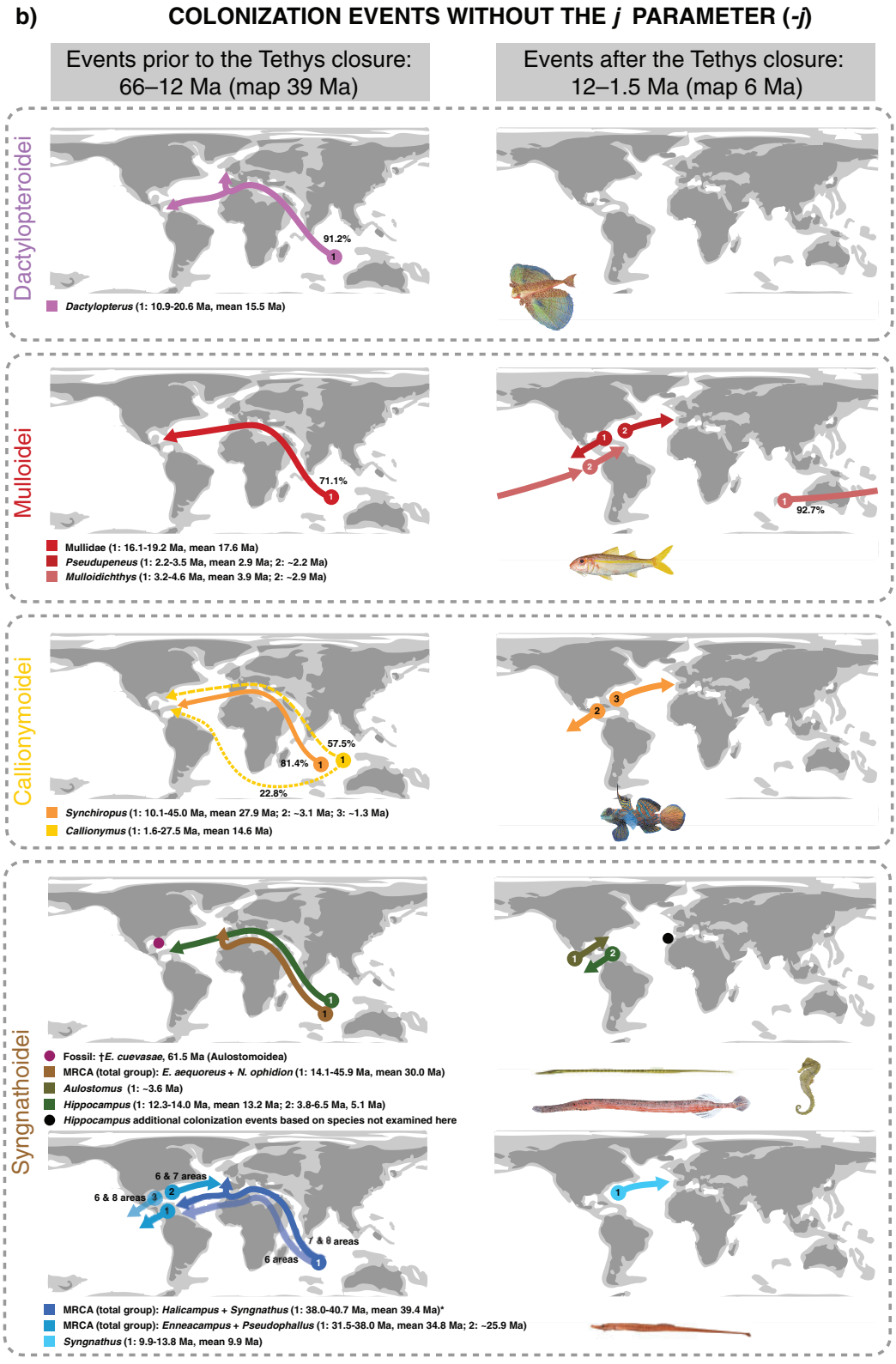

FIGURE 3. (Continued)
size of dashes are proportional to the probability). MRCAs indicate events for total groups (crown and stem lineages) given by the two taxa indicated in each case. Age ranges indicated per event denote the minimum and maximum ages for crown vs. stem clades obtained from the 28 inferred trees. Due to age uncertainty and overlapping, some events could be depicted on both maps; thus, selection of maps for event depiction is based on mean ages. Fish drawings are shown only for clades involved in mapped events. Maps modified from GPlate (Müller et al. 2018) represent the mean age from the following time slices: 12–66 Ma and 0–12 Ma. Ma = millions of years.

Our results indicate a Late Cretaceous (~86.8 Ma) origin of crown Syngnatharia, which is roughly 10 million years older than the evolutionary timescales inferred by recent studies (Near et al. 2012; Betancur-R et al. 2017; Alfaro et al. 2018; Hughes et al. 2018). These remarkable differences are likely the result of new interpretations for the age of the fossil †*Gasterorhamphosus zuppichinii* (Sorbini 1981), which implies that all previous studies underestimated the group's crown age. This interpretation is based on a recent stratigraphic analysis of the Calcari di Melissano, showing a lower Campanian (83.6 Ma; Schlüter et al. 2008) instead of a Maastrichtian (72.1 Ma; Sorbini 1981) age for this formation (see additional details in the Supplementary material available on Dryad). Although most family-level diversification events happened during the Cenozoic, the origin of suborders and most family-level total group predates the end of the Cretaceous (Fig. 1). This result runs counter to the notion that the divergence of major syngnatharian lineages is associated with the K-Pg mass extinction (Alfaro et al. 2018). While the split between Draconettidae and Callionymidae as well as the crown ages for Centriscidae and Syngnathidae are estimated to be around the K-Pg, our time-calibrated trees show no signs of diversification bursts associated with this extinction event. Instead, they reveal that the early Eocene was a period of exceptional diversification within Syngnathidae and Callionymidae. This period coincides with the early expansion of other reef-associated families (e.g., Apogonidae, Labridae, Pomacentridae; Cowman and Bellwood 2011; Bellwood et al. 2017; Fig. 1). The origin of crown Mulloidei, the youngest among syngnatharians suborders, has been linked to a rapid diversification process associated with extensive coral reef rearrangements as a result of the origin of the Indo-Australian-Archipelago (IAA) marine biodiversity hotspot in the Miocene (Renema et al. 2008; Bellwood et al. 2017; Siqueira et al. 2019, 2020; Fig. 1).

Uncertainties in Biogeographic Analyses.—Ancestral range estimations are typically inferred using a single tree and a predefined area scheme, resolving the most probable history based on a single best-fit model (e.g., Feng et al. 2017; Tea et al. 2019; Varela et al. 2019; Batista et al. 2020). In some cases, however, a set of trees are sampled from the Bayesian posterior distribution and used to gauge the effect of alternative phylogenetic resolutions (e.g., Berger et al. 2016). Here, we inferred the biogeographic history of syngnatharians based on comprehensive approaches designed to better capture uncertainties in estimated ancestral ranges. Given that the implementation of the founder-event speciation or jump-dispersal (*j*) parameter has been suggested to favor an unparsimonious numbers of long-distance dispersal events (Ree and Sanmartín 2018), we examined the results of our best-fit biogeographic model (BayAREA), both with and without the *j* parameter. As expected, the addition of the *j* parameter resulted in a better-fit model in all cases, increasing the probability of long-distance and more recent colonization events (Figs. 2 and 3). Overall, the inclusion/exclusion of the *j* parameter had a stronger effect on our biogeographic inferences relative to the number of areas considered or the alternative topologies used, despite considerable topological discordance. Furthermore, most discrepancies were observed among different area schemes with models that incorporate the *j* parameter, indicating a confounding interaction between these two variables (Fig. 3).

After the concerns raised by Ree and Sanmartín (2018) a number of studies using BioGeoBEARS have omitted the *j* parameter (e.g., Dong et al. 2018; Vargas and Dick 2020), including a recent investigation of the biogeography of marine angelfishes (Baraf et al. 2019). We believe that *j* can be informative for modeling the biogeography of marine fishes in general and reef-associated fishes in particular (like most syngnatharians), which can feature long-distance dispersal during pelagic larval stages or through rafting (e.g., sargassum-associated species) aided by oceanic currents (Luiz et al. 2012). Noteworthy, the critique of Ree and Sanmartín (2018) regarding the implementation of *j* in a model-fitting framework (e.g., by comparing DEC and DEC+*j*) was more recently challenged by Klaus and Matzke (2020) on the basis of previously conducted simulations (Matzke 2014), a review of a number of empirical studies that do not seem to show inflated likelihood scores in favor of *j*, and Ree and Sanmartín's (2018) use of a small hypothetical data set to emphasize their points (Klaus and Matzke 2020). Given these ongoing debates, we opt to focus on the similarities obtained between the two different estimations (with and without *j*), rather than their differences, to investigate the biogeographic history of syngnatharians (see next section).

Another factor of uncertainty relates to the use of alternative phylogenies to conduct ancestral range estimations. In this case, tree uncertainty appears to have a much smaller effect in this study possibly because early branching lineages that show a higher degree of topological discordance (e.g., full data set vs. subsets; concatenation vs. multispecies coalescent analyses; Supplementary Fig. S7 available on Dryad) are invariably estimated as having a Tethys/Indo-Pacific origin (Supplementary Tables S5 and S6 available on Dryad). Therefore, relatively lower sensitivity to phylogenetic uncertainty is probably a factor that is case-specific to syngnatharians and should not be generalized to other groups.

Finally, to identify the alternative colonization routes that different lineages followed, we calculated their relative probabilities using biogeographic stochastic mapping or BSM (Fig. 3). In some cases, the probability of a specific route was high. For example, *Mulloidichthys* dispersed through the EP Barrier with ~94% probability, and *Dactylopterus* colonized the Atlantic via the Tethys Seaway with ~87% probability. In other cases, BSM resulted in greater uncertainties, such as in *Callionymus*

that colonized the western Atlantic through two possible routes (~50% via Tethys Seaway and ~25% via southern Africa). More alternative dispersal routes were available before the closure of the Tethys Sea than after its closure, increasing the challenge to infer the pathways lineages followed. However, our analyses suggest that the most likely colonization route to the Atlantic for older dispersal events occurred via the Tethys Seaway (Fig. 3).

Several previous studies have examined the biogeographic history of Syngnathidae (Hamilton et al. 2017), particularly seahorses (genus *Hippocampus*, Teske et al. 2004, 2007; Li et al. 2021). These previous studies, however, did not consider biogeographic or phylogenetic variance, identifying two independent colonizations of the Atlantic by seahorse lineages—an ancient event (14.2–15.12 Ma Teske et al. 2007; 13.6–15.6 Ma Li et al. 2021) via the Tethys Seaway (MRCA *H. zosterae* + *H. erectus*) and a younger event through either The Cape of Good Hope (southern Africa; 3.1–4.6 Ma Teske et al. 2007; 3.6–4.9 Ma Li et al. 2021) or the EP Barrier (3.1–4.6 Ma; MRCA *H. algiricus* + *H. ingens*, Teske et al. 2007). By contrast, although our results and those from previous studies concur in identifying an Indo-Pacific (CIP + TA) origin of seahorses (Teske et al. 2004, 2007; Hamilton et al. 2017; Li et al. 2021), we found two possible biogeographic histories for the colonization of the Atlantic (Figs. 2 and 3). These include a single colonization event (12.3–14.0 Ma) through the Tethys Seaway in the MRCA of *H. subelongatus* + *H. ingens*, and two colonization events taking place after the closure of the Tethys Seaway via southern Africa (6.8–10.3 Ma and 3.7–6.5 Ma, respectively). It is also possible that the first colonization event of the Atlantic in *Hippocampus* suggested by these previous studies (or the single colonization proposed here) happened southwestwards via The Cape of Good Hope rather than northwestwards through the Tethys Seaway, as the seaway started to close at 18 Ma (Steininger and Rögl 1979; Adams et al. 1983; Rögl 1998).

The disagreements regarding the alternative colonization routes of the Atlantic in *Hippocampus* appear to stem from conflicts associated with divergence time estimations. The timing of the second Atlantic colonization was similar among the three studies, likely an indication of the similar use of a geological calibration based on the Isthmus of Panama for the MRCA of the geminate species pair *H. ingens* and *H. reidi*. (Note that Teske et al. 2007 and Hamilton et al. 2017 did not date their trees). The major difference concerns the age of the first colonization event (see above) and the age of crown *Hippocampus* (~24 Ma in Li et al. 2021 vs. ~14 Ma in this study; note that Teske et al. 2007 did not infer an age for this node), both of which are substantially older than our estimates (Supplementary Table S7 available on Dryad). These studies either applied a fossil calibration to the crown *Hippocampus* where it should have instead been placed in the stem lineage (see Žalohar and Hitij 2012; Li et al. 2021) or used geological calibrations only (Teske et al. 2007). In contrast, our divergence time estimates are based on both primary fossil and geological calibrations available for the entire Syngnatharia, as well as a secondary root calibration based on multiple global ray-finned fish time-calibrated trees that used dozens of primary fossil calibrations (Supplementary Tables S1 and S2 available on Dryad).

Biogeographic History of Syngnatharia.—Irrespective of the model, number of areas or tree used, all ancestral range estimations invariably reveal that the center of origin for syngnatharians was the ancient Tethys Sea. Ancestral lineages subsequently occupied the Indo-Pacific Ocean, followed by multiple independent colonization events of the eastern Pacific and the Atlantic basins, which took place via different routes both before and after the closure of the Tethys Seaway (Figs. 2 and 3, Supplementary Figs. S10–S14 available on Dryad). Like most marine fishes with long-distance dispersal capabilities, these patterns show that syngnatharians are successful at colonizing different oceanic realms. In fact, some species have circumglobal distributions or occur in at least two major basins (e.g., *Aulostomus chinensis*, *Fistularia* spp., *Doryrhamphus excisus*, *Syngnathus acus*, Centriscidae spp., and *Mulloidichthys vanicolensis*).

The western Tethys Sea was one of the richest regions for fossil teleost species, being a hotspot of marine biodiversity in the Eocene (Renema et al. 2008; Friedman and Carnevale 2018). In fact, the majority of syngnatharian fossils are currently described from different time horizons in that region (Sorbini 1981; Žalohar et al. 2009; Carnevale et al. 2014), suggesting that the Tethys Sea was the center of origin for the group, as it has been shown for other reef-fish families and reef-associated taxa (Renema et al. 2008; Cowman and Bellwood 2013a; Siqueira et al. 2019). The Terminal Tethyan Event (TTE), which opened and closed intermittently between 12 and 18 Ma, divided the western and the eastern Tethys Sea (Steininger and Rögl 1979; Adams et al. 1983; Rögl 1998) and created an important oceanic barrier that shaped the dispersal of marine species (Bellwood and Wainwright 2002; Barber and Bellwood 2005; Cowman et al. 2009, 2017; Cowman and Bellwood 2013b; Hou and Li 2017). Before its final closure (12 Ma), some lineages dispersed eastwards to the Indo-Pacific where diversification of most of the families occurred (e.g., Syngnathidae and Callionymidae). Rather than dispersing via The Cape of Good Hope (southern Africa), most other lineages colonized the Atlantic from the western Tethys Sea or crossed the Tethys Seaway from the Tethys Sea/Indo-Pacific (e.g., *Synchiropus*, *Callionymus*, *Dactylopterus volitans*, *Entelurus aequoreus* + *Nerophis ophidion*). During the Miocene, the syngnatharian biodiversity hotspot moved to the IAA where most reef fish clades also originated and diversified (e.g., Pomacanthidae, Baraf et al. 2019; Lutjanidae, Rincon-Sandoval et al. 2020). After the closure of the Tethys Seaway, lineages took the two remaining routes available for dispersal towards

the Atlantic and the eastern Pacific: the EP Barrier or southern Africa. Like other reef fishes (Lessios and Robertson 2006), goatfish (*Mulloidichthys*) and pipefish (*Doryrhamphus*) genera crossed the EP Barrier to colonize the eastern Pacific. Other genera in these groups (e.g., *Mullus* and *Microphis*), however, most probably dispersed into the western Atlantic via southern Africa through the warm Agulhas rings that occasionally penetrates into the Atlantic (see also Rocha et al. 2005; Floeter et al. 2008). See above regarding uncertainties in dispersal routes to the Atlantic in *Hippocampus*.

During the Neogene, the rising of the Isthmus of Panama interrupted gene flow between the western Atlantic and the eastern Pacific, ultimately producing many geminate species pair (see Lessios 2008 for a review). Before its final closure at some point before 2.8 Ma (Montes et al. 2015; O'Dea et al. 2016), many lineages dispersed across the Central American Seaway in either direction (i.e., goatfishes, dragonets, trumpetfishes, seahorses, and pipefishes), but predominantly through a WA-to-EP route (at least five events vs. two events from eastern Pacific to western Atlantic). Similar asymmetric dispersal routes between these two basins have been reported for other groups—for example, Lutjanidae (Rincon-Sandoval et al. 2020), Gobiidae (Thacker 2015), and *Apogon* (Thacker 2017). Finally, lineages that crossed the Mid-Atlantic Barrier to colonize both sides of the Atlantic either took a westwards route before the Tethys Seaways closure—probably via the North Equatorial Current (e.g., flying gurnards, pipefishes)—or dispersed eastwards no later than ~10 Ma, most likely using the Gulf Stream current (e.g., West African goatfish, Phaeton dragonet, pipefishes; see also Floeter et al. 2008; Luiz et al. 2012). While the predominant dispersal mode for these and other reef fish groups is via planktonic larvae, pipefishes and seahorses can also disperse by rafting on pelagic *Sargassum* and other macroalgae (Teske et al. 2005; Casazza and Ross 2008; Woodall 2009; Luiz et al. 2012; Boehm et al. 2013; Hamilton et al. 2017).

Aside from the Tethys Sea/Indo-Pacific region, the temperate Australasia and the Atlantic have also served as a center of origin for some genera (Fig. 2; Hamilton et al. 2017). The temperate Australasia region harbors significant endemicity and biodiversity of syngnatharians (particularly syngnathids) due to its extensive coastal seagrass habitats and its unique oceanographic conditions (Poore 1995; Shepherd and Edgar 2013). In fact, our analyses suggest that seadragons (*Phycodurus* and *Phyllopteryx*) and pipehorses (*Solegnathus*) originated in temperate Australasia. In the western Atlantic, the genus *Pseudophallus* is restricted to freshwater and brackish waters in Central and South America, whereas *Enneacampus* is currently distributed in the eastern Atlantic, suggesting that its ancestral lineage crossed the Mid-Atlantic Barrier eastwards (Floeter et al. 2008). A caveat of our study is that incomplete biogeographic sampling for some taxa may have affected our ancestral range estimation analyses, leading to area misplacement of lineage origin or underestimation of the number of colonization events. For instance, the genera *Bryx* and *Cosmocampus* have circumtropical distributions, but our sampling only includes species from the western Atlantic and the eastern Pacific/western Atlantic, respectively. Likewise, we lack representatives for the genus *Hippocampus* distributed in the eastern Atlantic, an area that *H. hippocampus* is thought to have recolonized from the western Atlantic by crossing the Mid-Atlantic Barrier eastwards via the Gulf Stream Current (Teske et al. 2007; Boehm et al. 2013; Li et al. 2021).

In conclusion, our study uses an integrative approach in a robust phylogenomic framework to account for a number of uncertainties in phylogenetic comparative inferences to trace the biogeographic history of Syngnatharia. We identified multiple independent colonizations of the Atlantic and the eastern Pacific and inferred possible dispersal routes from the Indo-Pacific and their center of origin, the Tethys Sea. While for some lineages the biogeographic history did not change using different area schemes or including/excluding the jump-dispersal (*j*) parameter, for other clades we identified a number of alternative colonization timings and routes, particularly when different area schemes are implemented in combination with the *j* parameter. Contrary to the common practice of estimating the biogeographic history using a single tree, a predefined set of areas, and a single biogeographic model, our study highlights the importance of embracing uncertainty in ancestral range estimations. We show that the common practice can be overly simplistic, failing to capture intrinsic complexities in historical biogeographic inferences. Our results ultimately provide a robust framework to address future questions on the evolutionary history of syngnatharians, such as understanding the factors driving their evolutionary radiation and explaining the uneven richness and morphological disparity across globally distributed clades.

Supplementary Material

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.4xgxd2580.

Acknowledgments

We are thankful to K.P. Maslenikov (Burke Museum of Natural History and Culture, University of Washington), L. Smith and A. Bentley (University of Kansas), and J.M.D. Astarloa (Universidad Nacional de Mar del Plata) for providing tissue samples. A. Graham (CSIRO Australian National Fish Collection, Hobart), Huddleston and D. Pitassy (Smithsonian Institution National Museum of Natural History), R. Peterson and V. Rodriguez (The George Washington University), M. Delpiani (Universidad Nacional de Mar del Plata) helped prepare and ship tissue samples and/or conducted DNA extractions. Bioinformatic analyses were facilitated

by the High-Performance Computing facility of UPR-RP (funded by INBRE Grant P20GM103475) and the University of Oklahoma Supercomputing Center for Education & Research (OSCER). We also thank J. Enk (Arbor BioSciences) for assistance in designing the probes and library preparation. We are grateful to K. Marske, D. Moen, G. Tolentino-Ramos, and M. Wersebe for helpful comments on the manuscript. Finally, we thank E. Santaquiteria for providing artistic illustrations.

Funding

This project was supported by National Science Foundation (NSF) [DEB-1932759 and DEB-1929248 to R.B.R., DEB-1457426 and DEB-1541554 to G.O., and DEB-1541552 to C.C.B.].

References

Adams C.G., Gentry A.W., Whybrow P.J. 1983. Dating the terminal Tethyan event. Utrecht Micropaleontological Bulletins. 30:273–298.

Alfaro M.E., Faircloth B.C., Harrington R.C., Sorenson L., Friedman M., Thacker C.E., Oliveros C.H., Ěerný D., Near T.J. 2018. Explosive diversification of marine fishes at the Cretaceous-Palaeogene boundary. Nat. Ecol. Evol. 2:688–696.

Atz J.W. 1962. Does the shrimpfish swim head up or head down? Anim Kingdom. 66:175–179.

Bannikov A.F. 2014. The systematic composition of the Eocene actinopterygian fish fauna from Monte Bolca, northern Italy, as known to date. Misc. Paleontol. 12:22–34.

Baraf L.M., Pratchett M.S., Cowman P.F. 2019. Ancestral biogeography and ecology of marine angelfishes (F: Pomacanthidae). Mol. Phylogenet. Evol. 140:106596.

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.

Batista R., Olsson U., Andermann T., Aleixo A., Ribas C.C., Antonelli A. 2020. Phylogenomics and biogeography of the world ' s thrushes (Aves, *Turdus*): new evidence for a more parsimonious evolutionary history. Proc. R. Soc. B Biol. Sci. 287:20192400.

Bellwood D.R., Goatley C.H.R., Bellwood O. 2017. The evolution of fishes and corals on reefs: Form, function and interdependence. Biol. Rev. 92:878–901.

Bellwood D.R., Wainwright P.C. 2002. The history and biogeography of fishes on coral reefs. Coral Reef Fishes. In: Sale P.F., editor. Dynamics and Diversity in a Complex Ecosystem. San Diego: Academic Press. p. 5–32.

Berger B.A., Kriebel R., Spalink D., Sytsma K.J. 2016. Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. Mol. Phylogenet. Evol. 95:116–136.

Bergert B.A., Wainwright P.C. 1997. Morphology and kinematics of prey capture in the syngnathid fishes *Hippocampus erectus* and *Syngnathus floridae.* Mar. Biol. 127:563–570.

Berglund A., Rosenqvist G., Svensson I. 1986. Reversed sex roles and parental energy investment in zygotes of two pipefish (Syngnathidae) species. Mar. Ecol. Prog. Ser. 29:209–215.

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

Betancur-R. R., Arcila D., Ballesteros J.A., Roa-Varon A., Ortí G., Broughton R.E., Cureton II J.C., Zhang F., Hough D.J., Wiley E.O., Arratia G., Carpenter K., Sanciangco M., López J.A., Campbell M.A., Li C., Holcroft N.I., Willis S., Borden W.C., Rowley T., Reneau P.C., Lu G., Buser T., Grande T. 2013. The tree of life and a new classification of bony fishes. PLoS Curr. 5:0–45.

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.

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. Evolution. 55:1029.

Cantalice K., Alvarado-Ortega J. 2016. *Eekaulostomus cuevasae* gen. and sp. nov., an ancient armored trumpetfish (Aulostomoidea) from Danian (Paleocene) marine deposits of Belisario Domínguez, Chiapas, southeastern Mexico. Palaeontol. Electron. 18:1–24.

Carnevale G., Bannikov A.F. 2019. A dragonet (Teleostei, callionymoidei) from the eocene of Monte Bolca, Italy. Boll. Soc. Paleontol. Ital. 58:295–307.

Carnevale G., Bannikov A.F., Landini W., Sorbini V. 2006. Volhynian (Early Sarmatian Sensu Lato) fishes from Tsurevsky, North Caucasus, Russia. J. Paleontol. 80:684–699.

Carnevale G., Bannikov A.F., Marramà G., Tyler J.C., Zorzin R. 2014. The Pesciara-Monte Postale Fossil-Lagerstätte: 2. Fishes and other vertebrates. Rend. Soc. Paleontol. Ital. 4:37–63.

Casazza T.L., Ross S.W. 2008. Fishes associated with pelagic Sargassum and open water lacking Sargassum in the Gulf Stream off North Carolina. Fish. Bull. 106:348–363.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17:540–552.

Chen W.J., Bonillo C., Lecointre G. 2003. Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. Mol. Phylogenet. Evol. 26:262–288.

Cowman P.F., Bellwood D.R. 2011. Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. J. Evol. Biol. 24: 2543–2562.

Cowman P.F., Bellwood D.R. 2013a. The historical biogeography of coral reef fishes: Global patterns of origination and dispersal. J. Biogeogr. 40:209–224.

Cowman P.F., Bellwood D.R. 2013b. Vicariance across major marine biogeographic barriers: temporal concordance and the relative intensity of hard versus soft barriers. Proc. R. Soc. B Biol. Sci. 280:20131541.

Cowman P.F., Bellwood D.R., van Herwerden L. 2009. Dating the evolutionary origins of wrasse lineages (Labridae) and the rise of trophic novelty on coral reefs. Mol. Phylogenet. Evol. 52:621–631.

Cowman P.F., Parravicini V., Kulbicki M., Floeter S.R. 2017. The biogeography of tropical reef fishes: endemism and provinciality through time. Biol. Rev. 92:2112–2130.

Dong J., Kergoat G.J., Vicente N., Rahmadi C., Xu S., Robillard T. 2018. Biogeographic patterns and diversification dynamics of the genus *Cardiodactylus* Saussure (Orthoptera, Grylloidea, Eneopterinae) in Southeast Asia. Mol. Phylogenet. Evol. 129:1–14.

dos Reis M., Yang Z. 2019. Bayesian molecular clock dating using genome-scale datasets. Methods Mol. Biol. 1910:309–330.

Dupin J., Matzke N.J., Särkinen T., Knapp S., Olmstead R.G., Bohs L., Smith S.D. 2017. Bayesian estimation of the global biogeographical history of the Solanaceae. J. Biogeogr. 44:887–899.

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

Feng Y.-J., Blackburn D.C., Liang D., Hillis D.M., Wake D.B., Cannatella D.C., Zhang P. 2017. Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. Proc. Natl. Acad. Sci. USA 114:E5864–E5870.

Fish F.E., Holzman R. 2019. Swimming turned on its head: stability and maneuverability of the shrimpfish (*Aeoliscus punctulatus*). Integr. Org. Biol. 1:1–14.

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.

Fricke R., Eschmeyer W.N., Fong J.D. 2020. Eschmeyer's catalog of fishes. species by family/subfamily. Available from: http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp.

Friedman M., Carnevale G. 2018. The Bolca Lagerstätten: shallow marine life in the Eocene. J. Geol. Soc. Lond. 175:569–579.
Froese R., Pauly D. 2019. FishBase. Available from: www.fishbase.org.
Gosline W.A. 1984. Structure, function, and ecology in the goatfishes (Family Mullidae). Pacific Sci. 38:312–323.
Hamilton H., Saarman N., Short G., Sellas A.B., Moore B., Hoang T., Grace C.L., Gomon M., Crow K., Brian Simison W. 2017. Molecular phylogeny and patterns of diversification in syngnathid fishes. Mol. Phylogenet. Evol. 107:388–403.
Hou Z., Li S. 2017. Tethyan changes shaped aquatic diversification. Biol. Rev. 93:874–896.
Huelsenbeck J.P., Rannala B., Masly J.P., Huelsenbeck J.P., Rannala B., Masly J.P. 2000. Accommodating phylogenetic uncertainty in evolutionary studies. Published by: American Association for the Advancement of Science Stable. Available from: http://www.jstor.org/stable/3075584.
Hughes L.C., Ortí 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. Proc. Natl. Acad. Sci. USA 115:6249–6254.
Hughes L.C., Ortí G., Saad H., Li C., White W.T., Baldwin C.C., Crandall K.A., Arcila D., Betancur-R. R. 2020. Exon probe sets and bioinformatics pipelines for all levels of fish phylogenomics. Mol. Ecol. Resour. 21:816–833.
IUCN. 2019. The IUCN Red List of Threatened Species. Available from: https://www.iucnredlist.org.
Johnson G.D. 1993. Percomorph phylogeny: progress and problems. Bull. Mar. Sci. 52:3–28.
Johnson G.D., Patterson C. 1993. Percomorph phylogeny: a survey of acanthomorphs and a new proposal. Bull. Mar. Sci. 52:554–626.
Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772–780.
Kawahara R., Miya M., Mabuchi K., Lavoué S., Inoue J.G., Satoh T.P., Kawaguchi A., Nishida M. 2008. Interrelationships of the 11 gasterosteiform families (sticklebacks, pipefishes, and their relatives): a new perspective based on whole mitogenome sequences from 75 higher teleosts. Mol. Phylogenet. Evol. 46:224–236.
Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton, B., Mentjies P., Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28:1647–1649.
Kim B. 2002. Comparative anatomy and phylogeny of the family Mullidae (Teleostei: Perciformes). Mem. Grad. Sch. Fish. Sci. Hokkaido Univ. 49:1–75.
Klaus K. V., Matzke N.J. 2020. Statistical comparison of trait-dependent biogeographical models indicates that Podocarpaceae dispersal is influenced by both seed cone traits and geographical distance. Syst. Biol. 69:61–75.
Kulbicki M., Parravicini V., Bellwood D.R., Arias-Gonzàlez E., Chabanet P., Floeter S.R., Friedlander A., McPherson J., Myers R.E., Vigliola L., Mouillot D. 2013. Global biogeography of reef fishes: a hierarchical quantitative delineation of regions. PLoS One. 8:e81847.
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.
Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2017. Partitionfinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34:772–773.
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., Robertson D.R. 2006. Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. Proc. R. Soc. B Biol. Sci. 273:2201–2208.
Lessios H.A., Robertson D.R. 2013. Speciation on a round planet: phylogeography of the goatfish genus *Mulloidichthys*. J. Biogeogr. 40:2373–2384.
Li C., Olave M., Hou Y., Qin G., Schneider R.F., Gao Z., Tu X., Wang X., Qi F., Nater A., Kautt A.F., Wan S., Zhang Y., Liu Y., Zhang H., Zhang B., Zhang H., Qu M., Liu S., Chen Z., Zhong J., Zhang H., Meng L., Wang K., Yin J., Huang L., Venkatesh B., Meyer A., Lu X. 2021. Genome sequences reveal global dispersal routes and convergent developmental mechanism in seahorse evolution. Nat. Commun. 12:1094.
Lin Q., Fan S., Zhang Y., Xu M., Zhang H., Yang Y., Lee A.P., Woltering J.M., Ravi V., Gunter H.M., Luo W., Gao Z., Lim Z.W., Qin G., Schneider R.F., Wang X., Xiong P., Li G., Wang K., Min J., Zhang C., Qiu Y., Bai J., He W., Bian C., Zhang X., Shan D., Qu H., Sun Y., Gao Q., Huang L., Shi Q., Meyer A., Venkatesh B. 2016. The seahorse genome and the evolution of its specialized morphology. Nature 540:395–399.
Longo S.J., Faircloth B.C., Meyer A., Westneat M.W., Alfaro M.E., Wainwright P.C. 2017. Phylogenomic analysis of a rapid radiation of misfit fishes (Syngnathiformes) using ultraconserved elements. Mol. Phylogenet. Evol. 113:33–48.
Luiz O.J., Madin J.S., Ross Robertson D., Rocha L.A., Wirtz P., Floeter S.R. 2012. Ecological traits influencing range expansion across large oceanic dispersal barriers: Insights from tropical Atlantic reef fishes. Proc. R. Soc. B Biol. Sci. 279:1033–1040.
Maddison W.P. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.
Matzke N.J. 2013. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R Scripts. R Packag. version 0.2.
Matzke N.J. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol. 63:951–970.
Matzke N.J. 2019. BioGeoBEARS - Run BioGeoBEARS on multiple trees. Available from: https://github.com/nmatzke/BioGeoBEARS/blob/master/R/BioGeoBEARS_on_multiple_trees_v1.R.
McGee M.D., Faircloth B.C., Borstein S.R., Zheng J., Hulsey C.D., Wainwright P.C., Alfaro M.E. 2016. Replicated divergence in cichlid radiations mirrors a major vertebrate innovation. Proc. R. Soc. B Biol. Sci. 283:20151413.
Mirarab S., Warnow T. 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics. 31:i44–i52.
Miya M., Friedman M., Satoh T.P., Takeshima H., Sado T., Iwasaki W., Yamanoue Y., Nakatani M., Mabuchi K., Inoue J.G., Poulsen J.Y., Fukunaga T., Sato Y., Nishida M. 2013. Evolutionary origin of the scombridae (tunas and mackerels): members of a paleogene adaptive radiation with 14 other pelagic fish families. PLoS One :e73535.
Montes C., Cardona A., Jaramillo C., Pardo A., Silva C., Valencia V., Ayala C., Pérez-Angel L., Rodriguez-Parra L., Ramirez V., Niño H. 2015. Middle Miocene closure of the Central American Seaway. Science. 348:226–229.
Müller R.D., Cannon J., Qin X., Watson R.J., Gurnis M., Williams S., Pfaffelmoser T., Seton M., Russell S.H.J., Zahirovic S. 2018. GPlates: building a virtual earth through deep time. Geochem. Geophys. Geosyst. 19:2243–2261.
Near T.J., Dornburg A., Eytan R.I., Keck B.P., Smith W.L., Kuhn K.L., Moore J.A., Price S.A., Burbrink F.T., Friedman M., Wainwright P.C. 2013. Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. Proc. Natl. Acad. Sci. USA. 110:12738–12743.
Near T.J., Eytan R.I., Dornburg A., Kuhn K.L., Moore J.A., Davis M.P., Wainwright P.C., Friedman M., Smith W.L. 2012. Resolution of ray-finned fish phylogeny and timing of diversification. Proc. Natl. Acad. Sci. USA. 109:13698–13703.
Nelson G. 1989. Phylogeny of major fish groups. hierarchy life. Mol. Morphol. phylogenetic Anal. Proc. from Nobel Symp. 70. ICS824.
Nelson J.S., Grande T.C., Wilson M.V.H. 2016. Fishes of the world. Hoboken (NJ): John Wiley & Sons.
Neutens C., Adriaens D., Christiaens J., De Kegel B., Dierick M., Boistel R., Van Hoorebeke L. 2014. Grasping convergent evolution in syngnathids: a unique tale of tails. J. Anat. 224:710–723.
O'Dea, Aaron, 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. 2:1–12.
OBIS. 2019. Data from the Ocean Biogeographic information system. Intergovernmental Oceanographic Commission of UNESCO.
Pietsch T. 1978. Evolutionary Relationships of the Sea Moths (Teleostei: Pegasidae) with a Classification of Gasterosteiform Families. Copeia. 1978:517–529.
Poore G.C.B. 1995. State of the marine environment report for Australia: the marine environment - technical annex: 1. Canberra, Australia.
Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67:901–904.
Ree R.H., Sanmartín I. 2018. Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. J. Biogeogr. 45:741–749.
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.
Renema W., Bellwood D.R., Braga J.C., Bromfield K., Hall R., Johnson K.G., Lunt P., Meyer C.P., McMonagle L.B., Morley R.J., O'Dea A., Todd J.A., Wesselingh F.P., Wilson M.E.J., Pandolfi J.M. 2008. Hopping hotspots: global shifts in marine biodiversity. Science 321:654–657.
Revell L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217–223.
Ribeiro E., Davis A.M., Rivero-Vega R.A., Ortí G., Betancur-R. R. 2018. Post-Cretaceous bursts of evolution along the benthic-pelagic axis in marine fishes. Proc. R. Soc. B Biol. Sci. 285:20182010.
Richardson J.E., Chatrou L.W., Mols J.B., Erkens R.H.J., Pirie M.D. 2004. Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. Philos. Trans. R. Soc. B Biol. Sci. 359:1495–1508.
Rincon-Sandoval M., Duarte-Ribeiro E., Davis A.M., Santaquiteria A., Hughes L.C., Baldwin C.C., Soto-Torres L., Acero A., Walker Jr. H.J., Carpenter K.E., Sheaves M., Ortí G., Arcila D., Betancur-R. R. 2020. Evolutionary determinism and convergence associated with water-column transitions in marine fishes. Proc. Natl. Acad. Sci. USA. 117:33396–33403.
Rocha L.A., Robertson D.R., Rocha C.R., Van Tassell J.L., Craig M.T., Bowen B.W. 2005. Recent invasion of the tropical Atlantic by an Indo-Pacific coral reef fish. Mol. Ecol. 14:3921–3928.
Rögl F. 1998. Palaeogeographic considerations for Mediterranean and Paratethys seaways. Ann. Naturhist. Mus. Wien. 99A:279–310.
Rokas A., Carroll S.B. 2006. Bushes in the tree of life. PLoS Biol. 4:1899–1904.
Ronquist F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Biol. 46:195–203.
Sanciangco M.D., Carpenter K.E., Betancur-R R., Betancur-R. R. 2016. Phylogenetic placement of enigmatic percomorph families (Teleostei: Percomorphaceae). Mol. Phylogenet. Evol. 94:565–576.
Schlüter M., Steuber T., Parente M. 2008. Chronostratigraphy of Campanian-Maastrichtian platform carbonates and rudist associations of Salento (Apulia, Italy). Cretac. Res. 29:100–114.
Shepherd S., Edgar G. (Eds) 2013. Ecology of Australian Temperate reefs: the unique south. CSIRO Publishing, Melbourne, Australia.
Siqueira A.C., Bellwood D.R., Cowman P.F. 2019. Historical biogeography of herbivorous coral reef fishes: the formation of an Atlantic fauna. J. Biogeogr. 46:1611–1624.
Siqueira A.C., Morais R.A., Bellwood D.R., Cowman P.F. 2020. Trophic innovations fuel reef fish diversification. Nat. Commun. 11:2669.
Song H.Y., Mabuchi K., Satoh T.P., Moore J.A., Yamanoue Y., Miya M., Nishida M. 2014. Mitogenomic circumscription of a novel percomorph fish clade mainly comprising "Syngnathoidei" (Teleostei). Gene 542:146–155.
Sorbini L. 1981. The Cretaceous fishes of Nardò. I. Order Gasterosteiformes (Pisces). Boll Mus. Civ. Stor. Nat. Verona. 8:1–27.
Spalding M.D., Fox H.E., Allen G.R., Davidson N., Ferdaña Z.A., Finlayson M., Halpern B.S., Jorge M.A., Lombana A., Lourie S.A., Martin K.D., McManus E., Molnar J., Recchia C.A., Robertson J. 2007. Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. Bioscience 57:573–583.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30:1312–1313.
Steininger F., Rögl F. 1979. The paratethys history. A contribution towards the Neogene geodynamics of the alpine orogene. Ann. Geol. des Pays Hell. 3:1153–1165.
Tagliacollo V.A., Lanfear R. 2018. Estimating improved partitioning schemes for ultraconserved elements. Mol. Biol. Evol. 35:1798–1811.
Tea Y.K., Van Der Wal C., Ludt W.B., Gill A.C., Lo N., Ho S.Y.W. 2019. Boomeranging around Australia: historical biogeography and population genomics of the anti-equatorial fish *Microcanthus strigatus* (Teleostei: Microcanthidae). Mol. Ecol. 28:3771–3785.
Teske P.R., Cherry M.I., Matthee C.A. 2004. The evolutionary history of seahorses (Syngnathidae: *Hippocampus*): molecular data suggest a West Pacific origin and two invasions of the Atlantic Ocean. Mol. Phylogenet. Evol. 30:273–286.
Teske P.R., Hamilton H., Matthee C.A., Barker N.P. 2007. Signatures of seaway closures and founder dispersal in the phylogeny of a circumglobally distributed seahorse lineage. BMC Evol. Biol. 7:138.
Teske P.R., Hamilton H., Palsbøll P.J., Choo C.K., Gabr H., Lourie S.A., Santos M., Sreepada A., Cherry M.I., Matthee C.A. 2005. Molecular evidence for long-distance colonization in an Indo-Pacific seahorse lineage. Mar. Ecol. Prog. Ser. 286:249–260.
Thacker C.E. 2015. Biogeography of goby lineages (Gobiiformes: Gobioidei): origin, invasions and extinction throughout the Cenozoic. J. Biogeogr. 42:1615–1625.
Thacker .E. 2017. Patterns of divergence in fish species separated by the Isthmus of Panama. BMC Evol. Biol. 17:111.
Townsend J.P., Su Z., Tekle Y.I. 2012. Phylogenetic signal and noise: predicting the power of a data set to resolve phylogeny. Syst. Biol. 61:835–849.
Varela L., Tambusso P.S., Mcdonald H.G., Fariña R.A. 2019. Phylogeny, macroevolutionary trends and historical biogeography of sloths: insights from a bayesian morphological clock analysis. Syst. Biol. 68:204–218.
Vargas O.M., Dick C.W. 2020. Diversification history of neotropical lecythidaceae, an ecologically dominant tree family of amazon rain forest. In: Rull V., Carnaval A., editors. Neotropical Diversification: Patterns and Processes. Basel, Switzerland: Springer. p. 791–809.
Whitfield A.K. 1999. Ichthyofaunal assemblages in estuaries: a South African case study. Rev. Fish Biol. Fish. 9:151–186.
Wiley E., Johnson G. 2010. A teleost classification based on monophyletic groups. In: Nelson J.S., Schultze H.P, MVH W., editors. Orig. Phylogenetic, Interrelat. Teleosts. (Munich Verlag Dr. Friedrich Pfeil). 23–182.
Wilson A.B., Orr J.W. 2011. The evolutionary origins of Syngnathidae: pipefishes and seahorses. J. Fish Biol. 78:1603–1623.
Woodall L. 2009. Population genetics and mating systems of the European seahorses *Hippocampus guttulatus* and *Hippocampus hippocampus* [PhD Thesis]. Royal Holloway, UK: University of London.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.
York P.H., Booth D.J., Glasby T.M., Pease B.C. 2006. Fish assemblages in habitats dominated by *Caulerpa taxifolia* and native seagrasses in south-eastern Australia. Mar. Ecol. Prog. Ser. 312:223–234.
Žalohar J., Hitij T. 2012. The first known fossil record of pygmy pipehorses (Teleostei: Syngnathidae: Hippocampinae) from the Miocene Coprolitic Horizon, Tunjice Hills, Slovenia. Ann. Paleontol. 98:131–151.
Žalohar J., Hitij T., Križnar M. 2009. Two new species of seahorses (Syngnathidae, *Hippocampus*) from the Middle Miocene (Sarmatian) Coprolitic Horizon in Tunjice Hills, Slovenia: the oldest fossil record of seahorses. Ann. Paleontol. 95:71–96.