Molecular Phylogenetics and Evolution 189 (2023) 107933

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylogenetic review of the comb-tooth blenny genus *Hypleurochilus* in the northwest Atlantic and Gulf of Mexico

Joshua E. Carter [*,1], Megan A. Sporre [1], Ron I. Eytan

Department of Marine Biology, Texas A&M University at Galveston, 1001 Texas Clipper Road, Galveston, TX 77554, United States

ARTICLE INFO

Keywords:
Blenniidae
Species delimitation
Sequence mining
Phylogeny
Biogeography
Systematics

ABSTRACT

As some of the smallest vertebrates, yet largest producers of consumed reef biomass, cryptobenthic reef fishes serve a disproportionate role in reef ecosystems and are one of the most poorly understood groups of fish. The blenny genera *Hypleurochilus* and *Parablennius* are currently considered paraphyletic and the interrelationships of *Parablennius* have been the focus of recent phylogenetic studies. However, the interrelationships of *Hypleurochilus* remain understudied. This genus is transatlantically distributed and comprises 11 species with a convoluted taxonomic history. In this study, relationships for ten *Hypleurochilus* species are resolved using multi-locus nuclear and mtDNA sequence data, morphological data, and mined COI barcode data. Mitochondrial and nuclear sequence data from 61 individuals collected from the western Atlantic and northern Gulf of Mexico (N. GoM) delimit seven species into a temperate clade, a tropical clade, and a third distinct lineage. This lineage, herein referred to as *H.* cf. *aequipinnis,* may represent a species of *Hypleurochilus* whose range has expanded into the N. GoM. Inclusion of publicly available *COI* sequence for an additional three species provides further phylogenetic resolution. *H. bananensis* forms a new eastern Atlantic clade with *H.* cf. *aequipinnis*, providing further evidence for a western Atlantic range expansion. Single marker *COI* delimitation was unable to elucidate the relationships between *H. springeri/H. pseudoaequipinnis* and between *H. multifilis/H. caudovittatus* due to incomplete lineage sorting. Mitochondrial data are also unable to accurately resolve the placement of *H. bermudensis*. However, a comprehensive approach using multi-locus phylogenetic and species delimitation methods was able to resolve these relationships. While mining publicly available sequence data allowed for the inclusion of an increased number of species in the analysis and a more comprehensive phylogeny, it was not without drawbacks, as a handful of sequences are potentially mis-identified. Overall, we find that the recent divergence of some species within this genus and potential introgression events confound the results of single locus delimitation methods, yet a combination of single and multi-locus analyses has allowed for insights into the biogeography of this genus and uncovered a potential transatlantic range expansion.

1. Introduction

In recent decades the advent of molecular systematics has prompted a resurgence of interest in combtooth blennies (Blenniidae Rafinesque 1810; herein, blennies), a taxonomically and ecologically diverse clade of cryptobenthic fishes with ~58 genera and >400 species worldwide (Hastings & Springer 2009). Several blenny phylogenetic review studies have emerged in the last decade, providing the information for greater resolution and understanding of taxonomic relationships within the clade (Hundt et al., 2014; Araujo et al., 2020; Vecchioni et al., 2022). Presently nested within the Almadablennius tribe, the blenny genera *Hypleurochilus* Gill, 1861 and *Parablennius* are considered paraphyletic based on modern systematics (Almada et al., 2005). The interrelationships of *Parablennius* have been the focus of a handful of recent phylogenetic studies (Almada et al., 2005; Levy et al., 2013; Vecchioni et al., 2022). However, those of *Hypleurochilus* remain understudied and phylogenetic analyses have yet to be performed on more than four *Hypleurochilus* species (Levy et al., 2013), leaving relationships within this genus largely undefined.

The genus *Hypleurochilus* is comprised of small scaleless fishes, typically less than 5 cm in length, inhabiting shallow marine environments with a demersal spawning mode of reproduction and limited adult

* Corresponding author.
E-mail address: carterjoshe.1990@gmail.com (J.E. Carter).
[1] Equal Contribution.

https://doi.org/10.1016/j.ympev.2023.107933
Received 10 April 2023; Received in revised form 25 September 2023; Accepted 25 September 2023
Available online 26 September 2023

mobility (Greenfield & Johnson, 1981; Springer, 1993). These species are confined to temperate and tropical waters of the Atlantic Ocean and Mediterranean Sea. In the eastern Atlantic they range from the Mediterranean Sea, along the coast of Africa, to the mouth of the Congo River and in the western Atlantic they range from New Jersey, USA, throughout the Caribbean and south to Brazil (Fig. 1). There are currently 11 valid species broadly distributed across the western and eastern Atlantic tropical and temperate marine waters (Pinheiro et al., 2013). Eight of these species are reported in the western Atlantic, with *H. pseudoaequipinnis* being the only species within the genus reported on both sides of the Atlantic (Levy et al., 2013). These species inhabit several ecosystems including coral and rocky reef areas. Some of these species have been found inhabiting oil platforms (Topolski & Szedlmayer, 2004).

Systematics of this group are largely based on morphological characters, but taxonomic uncertainty persists due to a lack of easily observable diagnostic traits and cryptic coloration (Araujo et al., 2020; Levy et al., 2011). *H. pseudoaequipinnis* and its eastern counterpart *H. aequipinnis* were classified as sister taxa by Randall (1966) and Bath (1994) based on morphological data. Genetic analysis by Levy et al. (2013) also recovered this sister relationship. They reported a second sister relationship between *H. fissicornis* and *H. n sp*, later identified as *H. brasil* by Pinheiro et al. (2013). More recently, Vecchioni et al. (2022) reported a sister relationship between western Atlantic species, *H. geminatus* and *H. fissicornis*, with the eastern Atlantic *H. bananensis* being sister to the western Atlantic lineages. These two phylogenetic studies include the largest number of *Hypleurochilus* species and present two competing hypotheses about the relationships of western and eastern Atlantic species. Levy et al. (2013) report paraphyletic western Atlantic relationships, where *H. pseudoaequipinnis* (western) is more closely related to *H. aequipinnis* (eastern) than the two other western species studied (*H. fissicornis* and *H. brasil*), while Vecchioni et al. (2022) report that two western Atlantic species (*H. geminatus* and *H. fissicornis*) are monophyletic compared to the eastern Atlantic species (*H. bananensis*). These and other previous studies lack an in-depth review of the genus and instead have include a limited number of representatives in the context of larger blenniid phylogenetic reconstructions (Javonillo & Harold, 2010; Levy et al., 2013; Hundt et al., 2014; Vecchioni et al., 2019; Attaran-Fariman et al. 2016).

Overall, the relationships of western Atlantic *Hypleurochilus* species remain understudied and rely on morphological comparisons to define them. *Hypleurochilus geminatus* (Wood 1825), *H. multifilis*, and *H. caudovittatus* were previously assigned to *H. geminatus* before being described as different species by Bath (1994). In this species description, *H. geminatus* and *H. multifilis* share more overlap in characters than *H. caudovittatus* with the exception of constant or variable pore counts and pelvic fin rays. Bath (1994) also describes *H. fissicornis* as the most closely related species to the sister species grouping of *H. geminatus/H. multifilis*. Randall (1966) describes *H. bermudensis, H. springeri* and *H. pseudoaequipinnis* (née *H. aequipinnis*) as being morphologically similar. There are no molecular studies to date that define the relationships between these species occurring in the western Atlantic. To provide a better understanding of western Atlantic *Hypleurochilus* systematics, we conducted an integrated study using a multi-locus dataset, morphological data, and molecular species delimitation methods with specimens representing six species from the western Atlantic, along with those of previously unreported *H.* cf. *aequipinnis* in the northern Gulf of Mexico (N. GoM). In addition, we performed phylogenetic analysis that included genetic data from three additional species from GenBank and the Barcode of Life Database (BOLD). This study aimed to 1) examine the phylogenetic and biogeographic relationships among species in the northern Gulf of Mexico (N. GoM) and western Atlantic with multilocus mitochondrial and nuclear data, 2) provide molecular support for *H. caudovittatus* as a putative sister taxon to *H. multifilis*, and 3) evaluate the identification of *Hypleurochilus* specimens in BOLD. In addition, we provide the first genetic records for a previously unreported *Hypleurochilus* sp. in the N. GoM, here identified as *H.* cf. *aequipinnis* based on the species description by Bath (1994).

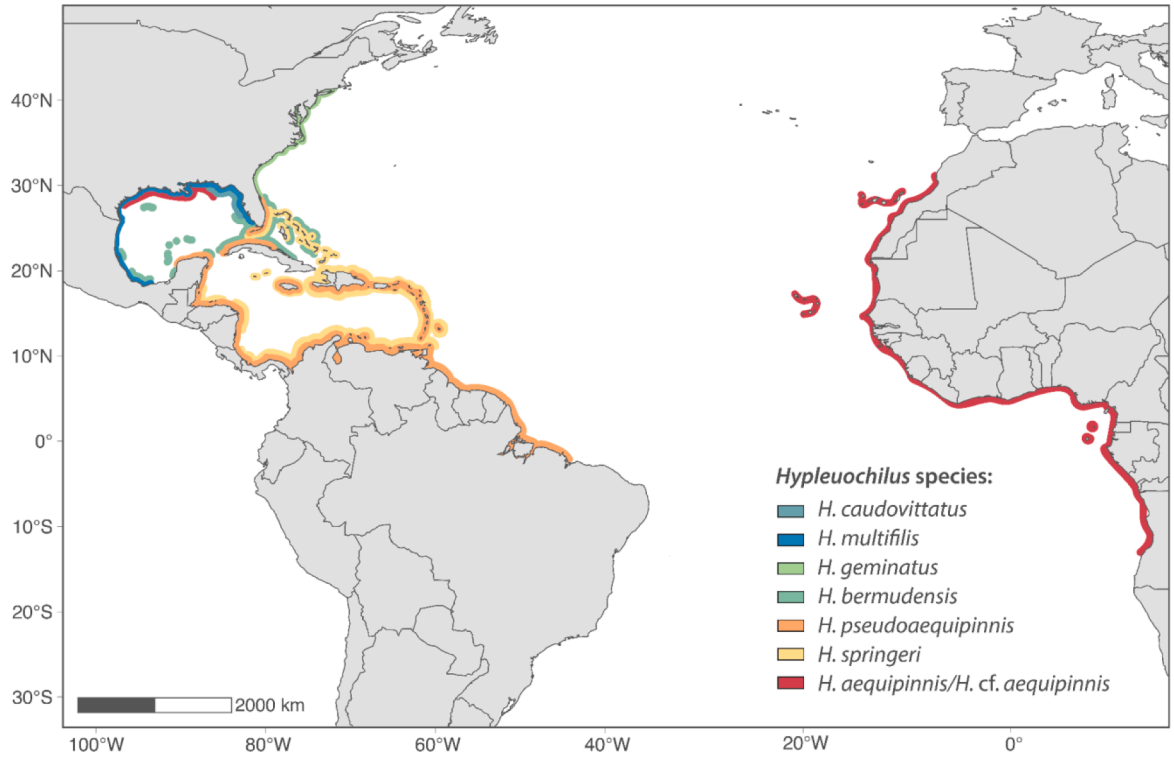

Fig. 1. Known distributions of ten species of *Hypleurochilus* and an additional range expansion for *H. aequipinnis* (herein referred to as *H.* cf. *aequipinnis*) in the northern Gulf of Mexico. (FishBase, 2001; STRI).

2. Materials and methods

2.1. Taxon sampling

A total of 83 specimens belonging to eight species of *Hypleurochilus* were collected in the field, obtained from museum collections, or sequences were obtained from BOLD and GenBank databases (Table S1-S3). Also included are specimens of a species not reported in the N. GoM designated here as *H.* cf. *aequipinnis* based on description of *H. aequipinnis* by Bath (1994). Specimens were collected from the N. GoM, east coast of the United States, and Curacao using dipnets or on SCUBA using slurp guns and quinaldine sulfate anesthetic. Specimens were photographed immediately after euthanasia to record fresh coloration followed by fixation in 10 % formalin for 24 h before being transferred to 70 % ethanol for long-term storage. Frozen specimens were thawed, photographed, and sampled for lateral muscle tissue and stored in 96 % EtOH at −80 °C. All specimens were identified using the most relevant morphological identification keys (FAO 2002) and species descriptions. Tissues for DNA extraction and comparative material for morphological analysis were acquired from the Academy of National Sciences, Philadelphia (ANSP), Smithsonian Institution National Museum of Natural History (USNM), and University of Florida (UF) (Table S1-S2). Tissue from the right lateral muscle block was sampled for DNA extraction and stored in 96 % EtOH at −80 °C.

2.2. DNA extraction and sequencing

Total genomic DNA was extracted from muscle or fin clip using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA, USA) according to the manufacturer recommended protocol. DNA sequence data were generated from the mitochondrial gene *Cytochrome Oxidase subunit I* (*COI*) and four protein-encoding nuclear genes (*RAG1, PLAGL2, MYH6* and *SH3PX3*) using polymerase chain reaction (PCR) with Promega GoTAQ (Madison, WI). PCR conditions followed Li et al. (2007). Amplicon products were purified using a polyethylene glycol (PEG) incubation and EtOH wash (Glenn, 2019). Automated Sanger sequencing was performed using ABI 3730xl capillary instrument (Applied Biosystems, Foster City, CA, USA) at the Keck DNA Sequencing Facility at Yale University, CT. Novel sequences were aligned using the MAFFT method (Katoh and Standley, 2013) in Geneious v 9.1.8 (Kearse et al., 2012) and deposited in GenBank and Dyrad databases (See Table S1 for Accession Numbers). Alignments of novel sequences and those downloaded from BOLD or GenBank were trimmed to fragments of 597 bp (*COI*), 1122 bp (*RAG1*), 564 bp (*PLAG12*), 615 bp (*MYH6*), and 660 bp (*SH3PX3*), respectively. Heterozygous loci were confirmed visually and resolved using PHASE v.2.1 using a probability threshold of 0.7 (Stephens et al., 2001). Input files were generated by the online software package SeqPhase (Flot, 2010). Multilocus datasets were generated for 7 species of *Hypleurochilus* (*H. caudovittatus*, *H. multifilis*, *H. geminatus*, *H. bermudensis*, *H. pseudoaequipinnis*, *H. springeri*, and *H.* cf. *aequipinnis*). Sequence divergence for COI was calculated in Geneious v 9.1.8 using the sequence matrix created from the MAFFT alignment.

2.3. Species tree estimation

Species trees were inferred in BEAST 2.7.4 (Bouckaert et al., 2019; Ogilvie et al., 2017) using StarBeast3 (Douglas et al. 2022). Four sets of trees were created: one based on nuclear genes only, a second set with combined mitochondrial and nuclear genes, a third set with our mitochondrial COI DNA only, and a fourth with our COI data plus the COI data from BOLD. Nucleotide alignments were partitioned by codon position, site models were unlinked, and trees and clocks were linked. Substitution models were inferred using bModelTest (Bouckaert and Drummond 2017). A species tree relaxed clock was used with units in substitutions/site/my. The MCMC was run twice for 50,000,000 generations, sampling every 1000. Convergence was confirmed using Tracer v1.7.2 (Rambaut et al. 2018) based on ESS values > 200. Maximum clade credibility trees were summarized in TreeAnnotator (Drummond and Rambaut, 2007) after discarding 25 % of trees as burn-in. Trees to be used in BioGeoBears analyses were rescaled in R using the ape package (Paradis et al. 2004). Tree outputs were visualized in FigTree 1.4.4 (Rambaut and Drummond, 2016) and final trees were created with the *ggtree* R-package (Yu et al., 2017). A DensiTree plot was generated to visualize all trees in each set and the consensus tree topology using DensiTree v 2.7.4 (Bouckaert, 2010). Trees to be used in BioGeoBears analyses were rescaled in R using the ape package (Paradis et al. 2004).

A third set of trees using 61 *COI* sequences was created in BEAST using the best model found by JModelTest (Posada, 2008). A relaxed lognormal clock model was applied under a birth–death tree prior, and data was partitioned by codon position. MCMC chain lengths were set to 50,000,000 generations, sampling every 1,000 generations. The first 10 % of trees produced were discarded as burn-in and BEAST log files were analyzed using Tracer. An MCC tree with 95 % highest posterior probability density and 10 % burn-in was generated. This tree was used for GMYC analysis (see below).

2.4. Species delimitation

Two approaches were used to delimit species – a discovery method (GMYC) and a validation method (BPP). Generalized mixed Yule coalescent (GMYC) is a single-locus method that uses a likelihood approach to identify the boundary between a Yule speciation process and intraspecific coalescence (Pons et al., 2006). The GMYC model was applied to ultrametric COI trees generated in BEAST. Two replicate GMYC runs were performed for two independent analyses implementing a Yule model (pure birth) using a constant clock model in the 'splits' R-package. The Bayesian Phylogenetics and Phylogeography (BPP version 3.4a) program employs MCMC for inferring species delimitations using DNA sequences under the multi-species coalescent (MSC) model (Yang, 2015). Given the sensitivity of BPP to selected parameters, four runs of both 'A10′ and 'A11′ analyses each with different θ (ancestral population sizes) and τ_0 (divergence time among species) priors were performed (Leaché & Fujita, 2010). Each of these four independent analyses were run twice using a phylogeny estimated by replicate StarBEAST2 v2.6 analyses as the guide tree in the 'A10′ analysis. BPP 'A10′ analysis performs species delimitation of a given species tree while the 'A11′ analysis performs both species tree estimation and species delimitation. Each analysis was set to run 100,000 MCMC generations from different starting seeds with a burn-in period of 8,000 and a sampling interval of 2. In total, sixteen BPP runs were carried out using the four phased nuclear alignments. A species probability value of 0.95 was considered strong support of BPP estimated tree nodes representing lineage splitting or speciation event estimated by BPP analyses.

2.5. Morphometric analysis

Due to low resolution of phylogenetic trees reconstructed from mitochondrial markers, morphological measurements were taken on the same 7 species used in genetic analysis. This data was also used to compare the morphology of the previously unreported species in the Gulf of Mexico, *H. cf aequipinnis,* to other species found in the Gulf. Meristic (discrete counts of segmented fin rays, supraorbital cirri count, cephalic pore count, and lateral line pore condition type) and morphometric data (continuous linear measurements) generally followed Hubbs & Lagler (1958) and Bath (1994) with some additional measurements. Methods for counts followed Springer and Gomon (1975). Arrangement of cephalic sensory pores and laterosensory conditions are denoted in Figure S1. All measurements were acquired from preserved specimens to avoid effects on dimensions from potential shrinkage. A total of 74 individuals were measured for 13 morphometric characters (Table S2). Point-to-point measurements and pore counts were taken from the left side of each specimen using digital Vernier calipers to the nearest 0.1

mm. The following characters were used: supraorbital cirri count (SC), cephalic sensory pore count (CP), dorsal notch as ratio between length of last dorsal spine and 1st dorsal soft ray (DN), depth at pelvic (DAP), depth at anus (DAN), head length (HL), eye diameter (ED), anal length (AL), peduncle length (PL), maxillary length (ML), upper lip thickness (UL), snout length (SNL), interorbit distance (ID), and lateral line system type.

Continuous variable measurements were scaled proportional to SL and HL and to unit variance to reduce variation due to allometric difference using R 3.6.1 (R Core Team, 2017). Sex effects were not considered in this study; therefore, all statistical analyses were performed on a dataset containing both males and females since all measurements were scaled and collected samples were predominantly male in all species. Box plots and QQ plots were generated for each variable to visually assess normality using the R package *ggplot2* (Wickham, 2016). Multiple characters, both standard and log-transformed, did not meet normality assumption, thus nonparametric methods were applied to test character significance between species. Levene's F-Test was used to assess homogeneity of variance, Kruskal Wallis Tests and Dunn Tests were used to evaluate character significance between species and to assess variance estimates of the Kruskal Wallis tests.

Morphometric and meristic data were combined and used to perform a principal component analysis (PCA). The most informative principal components (PC1 and PC2) and characters contributing most to PC variation were identified using their loading values and visualized in a biplot using the R package *FactoMineR* (Lê et al., 2008). PCA loading scores are reported in Table S4.

2.6. BOLD database mining and phylogenetic reconstruction

Forty-nine COI sequences of *Hypleurochilus* species were mined from GeneBank and BOLD (accession numbers found in Table S3). The 49 mined sequences and 61 sequences generated for this study were aligned using MAFFT (Katoh and Standley, 2013) in Geneious v 9.1.8 (Kearse et al., 2012). This dataset contained sequences from 10 species of *Hypleurochilus*. A Bayesian phylogenetic tree was reconstructed in BEAST2 v 2.6.0 (Bouckaert et al., 2014). Site and clock models were unlinked. Site models were chosen using BModelTest. A strict clock model was applied under a Yule tree prior. The MCMC was run for 10,000,000 generations, sampling every 5,000. The maximum clade credibility tree with median node heights after burn-in of 10 % of the trees was used as the final BEAST tree and was visualized in FigTree 1.4.3 (Rambaut and Drummond, 2016). Node support was determined using posterior probability (PP).

2.7. Historical biogeography

Ancestral ranges were inferred using BioGeoBears (Sidje 1998, Matzke, 2013, Matzke 2014, Matzke, 2018a,b, Matzke, 2023). Two species trees were used: the nDNA + mtDNA tree, and the BOLD COI tree. Three ancestral areas were taken from Robertson and Cramer (2014): 1) Northern Caribbean; 2) Central Caribbean; and 3) Southern Caribbean. Two additional areas were taken from Briggs and Bowen (2013): 4) Brazil Coast; and 5) Eastern Atlantic. *Hypleurochilus* range information came from the literature (Bath 1994; Pinheiro et al., 2013) the STRI shorefishes database, and personal collections. The DEC model of range evolution was used to infer ancestral ranges (Ree and Smith 2008). The BOLD species tree was used to infer ancestral ranges so as to have all *Hypleurochilus* species present in a phylogeny.

3. Results

3.1. Molecular data

Sequences for the mitochondrial gene COI (597 bp) and nuclear genes MYH6 (615 bp), PLAG12 (564 bp), SH3PX3 (660 bp) and RAG1 (1122 bp) were generated and deposited in GenBank and Dryad databases (GenBank Accession numbers can be found in Appendix S1). COI sequence divergence the between *H.* cf. *aequipinnis* and congeners ranged from 5.2 % to 5.8 %. Sixty-one *Hypleurochilus* specimens representing seven species were used to generate three different species trees (mtDNA, nDNA, and mtDNA + nDNA). For species delimitation 61 specimens were used to generate a COI gene tree and 60 individuals from six species of *Hypleurochilus* were used to generate multilocus species trees.

3.2. Phylogenetic relationships and species trees

All multilocus and mtDNA COI species trees support *H.* cf. *aequipinnis* as a monophyletic lineage (PP > 0.98) (Figs. 2 & 3). Posterior effective sample sizes (ESSs) were high (ESS > 250) for all StarBEAST2 and BEAST runs. All species trees recovered two sister relationships, one between *H. caudovittatus/ H. multifilis* (PP = 1) and another between *H. pseudoaequipinnis/H.springeri* (PP between 0.64 and 0.98) (Fig. 2). There is lower posterior support and incongruencies in the placement of internal nodes of the species trees. All three datasets recover *H. geminatus* as sister to *H. caudovittatus/ H. multifilis* with moderate support (PP between 0.38 and 0.74) (Fig. 2). The mtDNA species tree (Fig. 2A) and mtDNA gene tree (used in GMYC analysis; Fig. 4A) recover *H. bermudensis* as sister to *H. pseudoaequipinnis/H.springeri* (PP > 0.94), however nDNA and nDNA + mtDNA species trees (Fig. 2B & 2C) recover *H. bermudensis* as sister to *H. caudovittatus/ H. multifilis* (PP between 0.21 and 0.44). For clarity we use the topology of the mtDNA + nDNA species tree when referring to relationships within this genus, more temperate species are assigned into clade I (*H. multifilis, H. caudovittatus, H. geminatus,* and *H. bermudensis*), more tropical species into clade II (*H. pseudoaequipinnis* and *H. springeri*) and *H* cf. *aequipinnis* assigned to lineage/clade III.. All species trees recovered this split between the putative clades (PP between 0.30 and 0.35) (Fig. 2). Low posterior probabilities for internal nodes can be ascribed to having only a single representative of *H. bermudensis* and the discordance of assignment to clade I or II for this species based on dataset used. The uncertainty in the mtDNA + nDNA species tree is visualized in Fig. 3.

Multiple independent runs of mtDNA COI in BEAST recovered trees with similar topology to that of the mtDNA species trees; however, the placement of clade II shifted between a sister relationship with clade I or that of *H.* cf. *aequipinnis* (Fig. 4A). This tree was used for single locus delimitation. The species tree used for multilocus delimitation did not include *H. bermudensis*, but recovered the same general topology as species trees containing *H. bermudensis*; namely sister relationships between *H. caudovittatus/ H. multifilis* (PP = 1.0) and *H. pseudoaequipinnis/ H.springeri* (PP = 0.62) and the recovery of the two clades (PP = 0.56).

3.3. Species delimitation

Discovery and validation species delimitation analyses both support *H.* cf. *aequipinnis* as a distinct lineage. GMYC analysis recovers 5 distinct lineages between the 7 species included. All *H.* cf. *aequipinnis* and *H. geminatus* specimens, as well as the single *H. bermudensis* specimen, are assigned to monophyletic lineages (Fig. 4A). The *H. multifilis/ H. caudovittatus* lineages and *H. springeri/H. pseudoaequipinnis* are collapsed into two lineages (Fig. 4A). Both GMYC models performed well and provided a significantly better fit to the data than the null models hypothesis of the entire sample being derived from a single species with uniform branching. The Yule prior model analyses are represented by 4 ML clusters (4–4) and 5 entities (5–5) and Constant Coalescent prior model by 4 ML clusters (4–9) and 5 entities (5–10). *Hypleurochilus bermudensis* is represented by only one sequence, thus the number of entities is greater than the number of clusters for both models. Delimited GMYC clusters (*H. multifilis/H. caudovittatus, H. geminatus, H. springeri/H. pseudoaequipinnis,* and *H.* cf. *aequipinnis*) were largely congruent with clades defined by the COI BEAST tree (Fig. 4A) but were

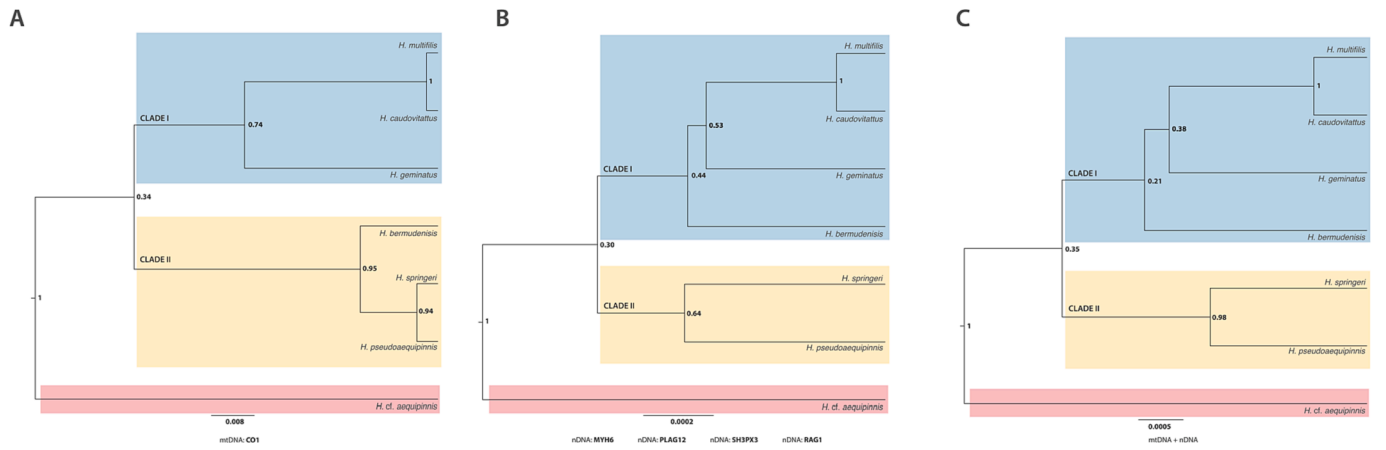

Fig. 2. Species trees reconstructed in StarBeast2 using data from 61 individuals across 7 species. Panels represent species trees reconstructed with mitochondrial data (A), nuclear data (B) and combined mitochondrial and nuclear data (C). Node values represent posterior probabilities.

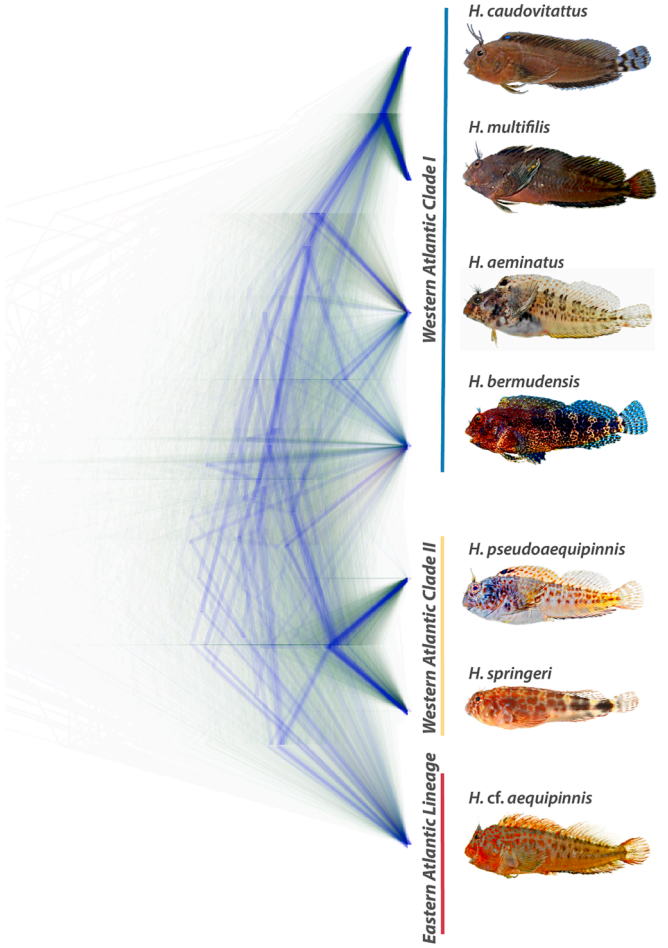

Fig. 3. DensiTree visualization of species trees reconstructed from the nDNA + mtDNA sequence dataset in StarBEAST2. Areas of the tree in topological and branch length agreement are highly colored while areas with little agreement are more diffuse. The consensus topology is denoted in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

less resolved than lineages determined by the mtDNA species tree (Fig. 2A). GMYC analysis failed to support the splits between H. *multifilis*/H. *caudovittatus* and *H. springeri*/*H. pseudoaequipinnis,* instead collapsing those four species into two clade.

BPP consistently found very high posterior probabilities (1.0) for delimitations of all six species included across multiple analyses with different algorithms and prior distributions (Fig. 4B). Unlike the single-locus delimitation, the multi-locus delimitation recovers the sister groupings of *H. multifilis*/H. *caudovittatus* and *H. springeri*/*H. pseudoaequipinnis* as four distinct lineages.

3.4. Morphology and diagnostic traits

Principal component analysis (PCA) of the seven *Hypleurochilus* species resulted in PC1 and PC2 explaining 23.0 % and 16.4 % of the total variance, respectively (Fig. 5). PCA loadings (Table S4) indicate that PC1 was most influenced by DAP, DAN, HL, ML and SNL (>±0.30) and PC2 loadings were most influenced by DN, CP, PL and SNL (>±0.30). PCA biplot shows a clear partition between *H.* cf. *aequipinnis* and the six other species (Fig. 5). There is a slightly larger partitioning between species of clade I when compared to species of clade II. *Hypleurochilus pseudoaequipinnis* and *H. springeri* almost completely overlap in morphospace. *Hypleurochilus bermudensis* is centrally located in morphospace between species from both clades. There is a large amount of overlap between *H. bermudensis*, *H. caudovittatus,* and *H. multifilis* in clade I morphospace, while *H. geminatus* appears to be largely partitioned from the rest of clade I (Fig. 6). Furthermore, PCA analysis indicates a clear morphological differentiation between *H.* cf. *aequipinnis* and *H. pseudoaequipinnis*. Genetic and morphological data support a distinction between these two species. Morphological data show similar patterns of differentiation to that of COI delimitation; the overlap in morphospace of *H. multifilis*/*H. geminatus* and *H. springeri*/*H. pseudoaequipinnis* mirrors the lack of resolution in COI data for these two groups of species.

3.5. BOLD sequence phylogeny

A Bayesian phylogenetic tree was reconstructed with a total of 110 COI sequences (Fig. 6). The *H. bananensis* sequence (MG837129) was recovered as sister to all *H. cf aequipinnis* specimens with 0.99 posterior probability support, putatively forming an eastern Atlantic clade. This eastern Atlantic clade groups as sister to the western Atlantic clade I with low posterior probability (0.40). The western Atlantic clade I included no sequences from BOLD or GenBank, and recovered *H. geminatus* as a monophyletic lineage, sister (PP = 0.99) to a paraphyletic lineage containing *H. multifilis* and *H. caudovittatus*. The western Atlantic clade II was recovered as a sister clade to the eastern Atlantic/western Atlantic clade I grouping with a posterior probability of 1.0. This is the same topology recovered by the mtDNA BEAST tree used in GMYC analysis (Fig. 4A).

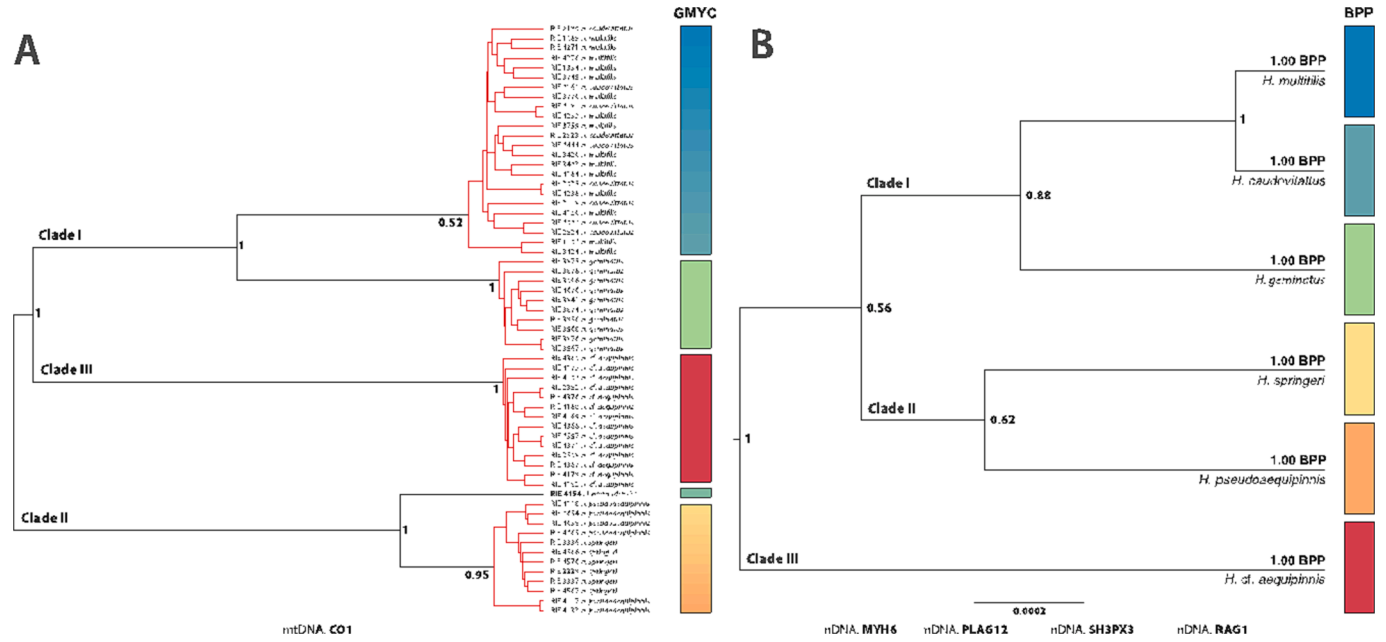

Fig. 4. A) Gene-tree based on cytochrome *c* oxidase subunit I (COI) used as input for GMYC model analysis of 61 *Hypleurochilus* specimens collected by the authors. White squares indicate sequence clusters corresponding to single GMYC species. Node support values are Bayesian posterior probabilities. B) Species-tree based on four nuclear markers used as input for BPP species delimitation of 60 *Hypleurochilus* species collected by the authors. Bayesian posterior probabilities support values indicate support at internal nodes and terminal species nodes. Note: single specimen of *H. bermudensis* is included in GMYC gene-tree, but not in species-tree as more than one specimen is required for BPP species delimitation analysis.

Within clade II, two specimens identified as *H. bermudensis* and a third unidentified *Hypleurochilus* specimen (MXV140) formed a monophyletic lineage (PP = 1.0). The placement of this *H. bermudensis* clade as more closely related to other clade II individuals compared to clade I is in disagreement with the species tree topologies recovered using nDNA and mtDNA + nDNA and a single *H. bermudensis* specimen (Fig. 2B & C, Fig. 3). A paraphyletic clade with low posterior probability (PP = 0.08) contained 11 specimens identified as *H. fissicornis*, seven specimens identified as *H. pseudoaequipinnis*, three specimens identified as *H. geminatus* (likely mis-identified), and one individual identified as *H. springeri*. A clade containing 12 specimens identified as *H. springeri* also contained two specimens identified as *H. pseudoaequipinnis* (PP = 0.90) and was sister to the majority *H. pseudoaequipinnis/H. fissicornis* clade. Ten specimens identified as *H. brasil* formed a monophyletic clade (PP = 0.99) that was sister to the *H. pseudoaequipinnis/H. fissicornis/H. springeri* lineages. Additionally, 11 specimens identified as *H. fissicornis* and one specimen identified as *H. geminatus* grouped outside of the *Hypleurochilus* lineages. BLAST results show each of these sequences most closely identifying as either *Auchenionchus microcirrus*, *Calliclinus geniguttu*, or *Scartichthys viridis* with > 90 % percent identity.

3.6. Biogeographic reconstruction

The topologies of both species trees, as implemented by BioGeoBears, identified the combined regions of the Eastern Atlantic and the Northern Caribbean as the most probable ancestral area of origination for *Hypleurochilus*. This large area may be due to low resolution in nodes at the base of the tree and from lack of more eastern Atlantic species included in the analysis. The wide ranging area may represent transoceanic dispersal from either western Atlantic to eastern Atlantic or eastern Atlantic to western Atlantic. Reconstruction based on the mtDNA + nDNA species tree suggests that *Hypleurochilus* had a widespread common ancestor that broke apart and evolved to inhabit the eastern Atlantic region and the northern Caribbean region (Fig. 7). The clade II ancestor dispersed to the central and southern Caribbean regions from the northern Caribbean. *Hypleurochilus springeri* and *H. pseudoaequipinnis* have current known ranges in both the central and southern Caribbean. Ancestors of the western Atlantic clade I persisted in the northern Caribbean before there was a secondary dispersal event of *H. bermudensis*. The three other species in Clade I (*H. geminatus, H. multifilis*, and *H. caudovittatus*) have persisted in the northern Caribbean. *Hypleurochilus* cf. *aequipinnis* has persisted in the eastern Atlantic until its possible introduction to the western Atlantic (this reconstruction used only the native range of *H. aequipinnis* and not the potentially introduced range). Reconstruction based on the species tree produced using the COI dataset containing mined sequences suggests that *Hypleurochilus* had a widespread common ancestral area that broke apart and lineages evolved to inhabit the eastern Atlantic region and the northern Caribbean region (Figure S2). The ancestor of western Atlantic Clade II dispersed to the northern Caribbean and then south into the central Caribbean, southern Caribbean, and Brazilian regions. *H. bermudensis* persisted in the northern and central Caribbean, *H. brasil* dispersed to and persisted in only the Brazilian region, *H. springeri* and *H. pseudoaequipinnis* persisted in the northern, central and southern Caribbean regions, and there was a secondary dispersal event of *H. fissicornis* to the Brazilian region. Western Atlantic Clade I and eastern Atlantic clades persisted in the northern Caribbean and eastern Atlantic before breaking apart. *Hypleurochilus bananensis* and *H.* cf *aequipinnis* persisted in the eastern Atlantic (possible introduced range not included in model), while *H. geminatus, H. multifilis,* and *H. caudovittatus* have persisted in the northern Caribbean.

4. Discussion

4.1. Phylogeography & delimitation

The phylogenetic analysis and species delimitation results presented here clarify the relationships of the western Atlantic species of *Hypleurochilus*. Based on single and multilocus species delimitations and species tree inference there are two distinct clades in the western Atlantic and a third distinct lineage that may represent the range expansion of a species from the eastern Atlantic or a species previously

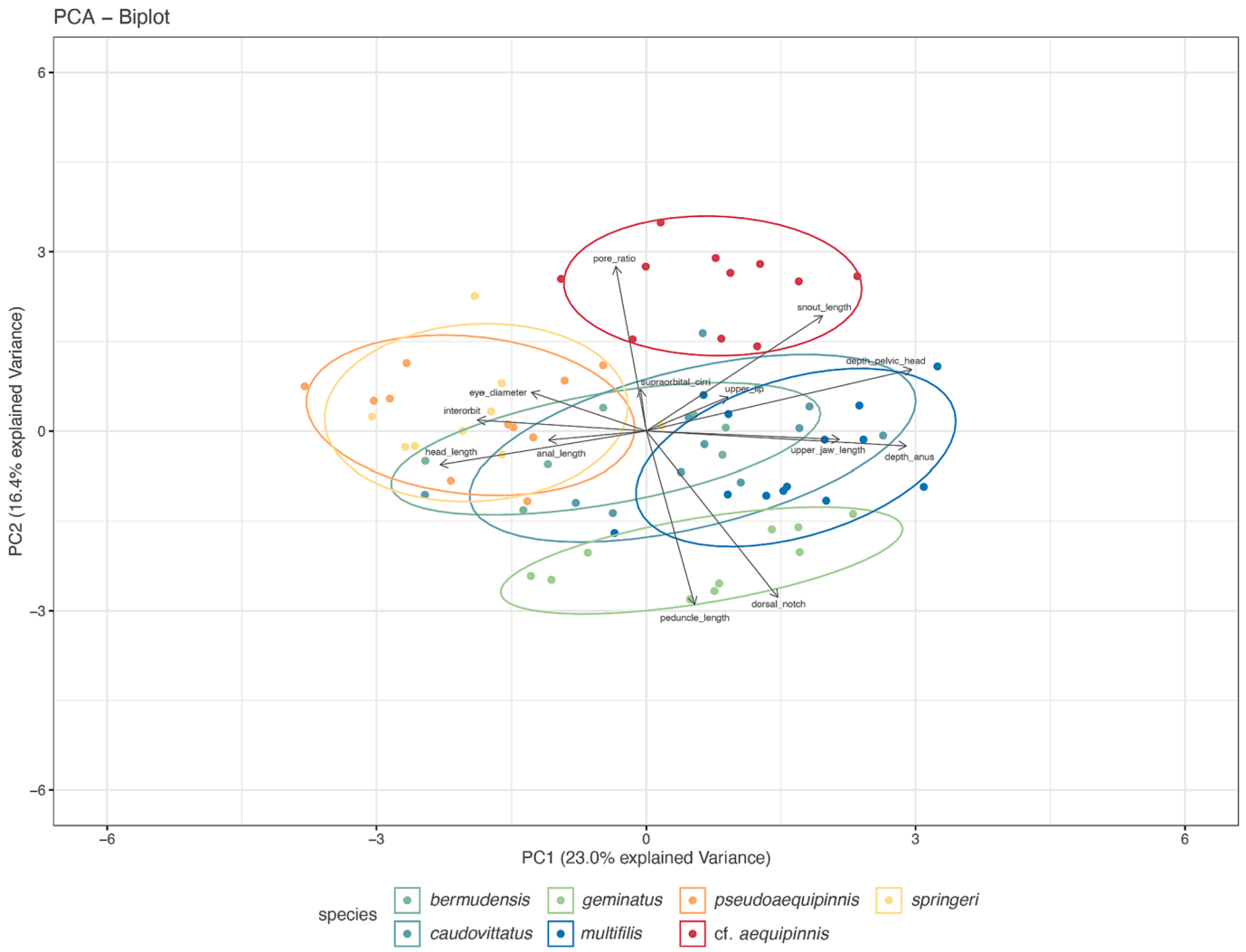

Fig. 5. PCA of morphometric analysis for *Hypleurochilus* Clades I and II and *H.* cf. *aequipinnis*.

unknown to occur in the Gulf of Mexico. Clade I consists of temperate water species that are distributed from the northern Caribbean and Gulf of Mexico to New Jersey, USA and includes *H. caudovittatus, H. multifilis, H. geminatus,* and *H. bermudensis.* Clade II consists of more tropical water species that are distributed from Brazil to Florida in the northern Caribbean and includes *H. pseudoaequipinnis* and *H. springeri*. Single locus delimitation uses a tree topology inconsistent with both nDNA and combined nDNA + mtDNA species tree topologies concerning the positioning of *H. bermudensis*; however, it effectively identifies this species as a distinct lineage. With only a single representative of *H. bermudensis*, it is difficult to determine its exact relationship to other species within the genus. Overall, multi-locus delimitation results support the classification of all seven *Hypleurochilus* species studied as distinct lineages. While not used in delimitation analysis, by mining sequence data from the BOLD database and GenBank we were able to obtain sequence data for an additional three species of *Hypleurochilus*. *H. fissicornis* and *H. brasil* grouped within Clade II and a sister relationship between *H.* cf *aequipinnis* and an eastern Atlantic species, *H. bananensis*, provided more support to the hypothesis of a potential range expansion for a species from the eastern Atlantic.

Previous, more extensive blenniid phylogenetic studies, which utilized a limited array of *Hypleurochilus* samples, have produced varying hypotheses regarding the relationships between species located in the eastern and western Atlantic. The first hypothesis suggests the western Atlantic species are paraphyletic, with *H. pseudoaequipinnis* and *H. aequipinnis* forming a sister clade to *H. fissicornis* and *H. brasil* (Levy et al. 2013). Our mtDNA gene trees (Figs. 4 & 6) did recover a topology in which the western Atlantic species are paraphyletic, however it is species in Clade I (*H. multifilis*, *H. caudovittatus*, and *H. geminatus*) that group as sister to a putative eastern Atlantic clade (*H.* cf. *aequipinnis* and *H. bananensis*) and not *H. pseudoaequipinnis*. Contrary to the relationships reported by Levy et al. (2013), our mtDNA phylogeny recovered using mined sequences found *H. pseudoaequipinnis* and *H. fissicornis* to be a paraphyletic lineage. It should be noted that the *H. pseudoaequipinnis* individual used in Levy et al. (2013) was collected from the Gulf of Guinea in the eastern Atlantic. Further investigation using the same molecular markers for direct comparison of *H.* cf *aequipinnis* to this sequence data would provide more clarification. Another blenniid with a similar eastern and western Atlantic distribution is the genus *Ophioblennius*. Using only a single mitochondrial marker, Muss et al. (2001) recovered a sister relationship between Caribbean species and eastern Atlantic species, similar to the sister relationship recovered in the *Hypleurochilus* phylogeny reconstructed with mitochondrial COI, where the species with the most northern distributions in the western Atlantic are the most closely related to the species in the eastern Atlantic. Relationships within the tropical western Atlantic Clade II, also appear to follow strong phylogeographic breaks. While their relationship remains unresolved using a single marker phylogeny, *H. fissicornis* and *H. pseudoaequipinnis* are geographically separated along the Brazilian coast via the discharge of freshwater by the Amazon and Orinoco rivers.

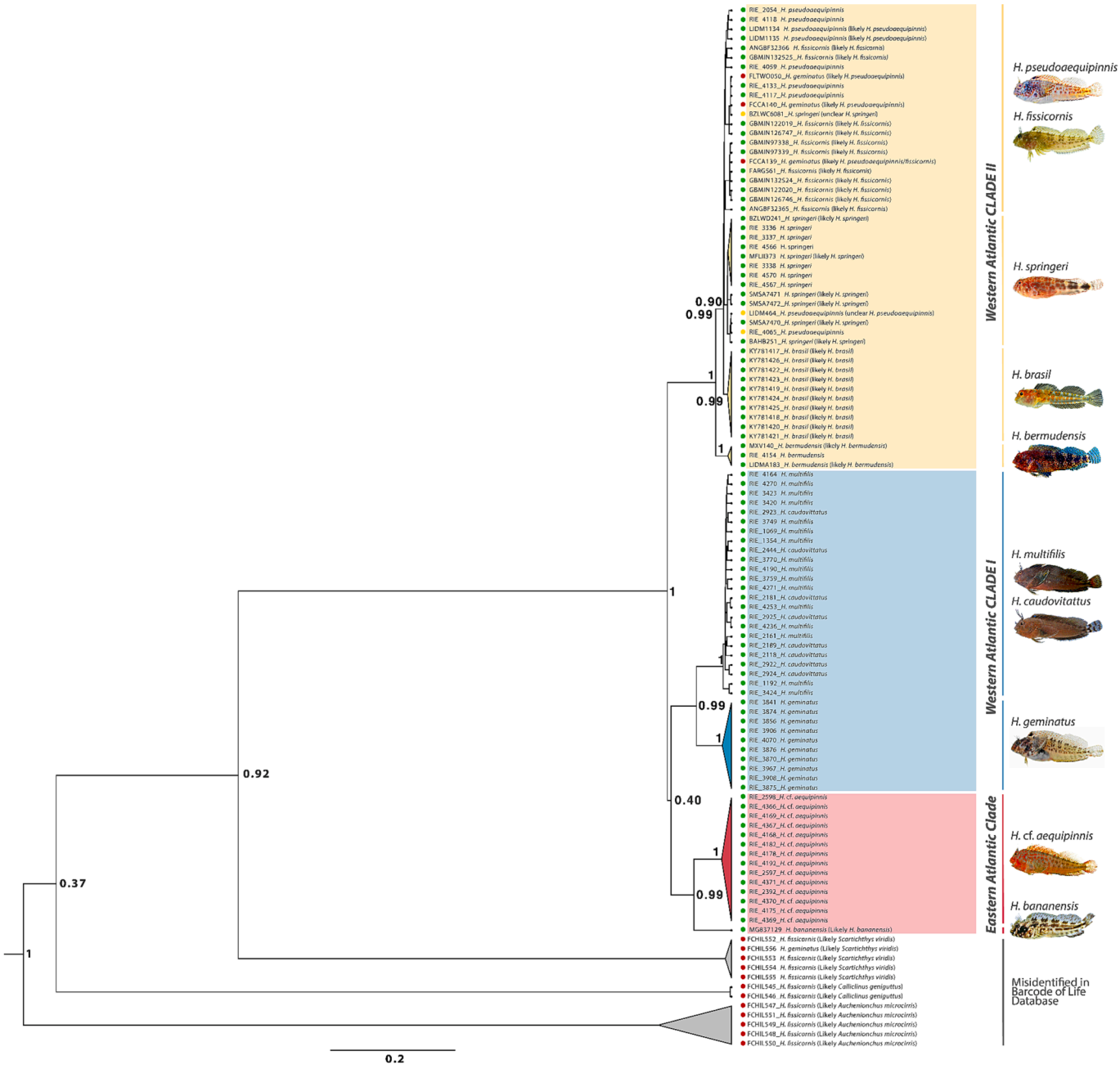

Fig. 6. Phylogenetic tree reconstructed from *COI* sequencing including sequences mined from GenBank and BOLD databases. Node labels indicate posterior probability, red circles denote a mis-identified specimen, yellow circles denote a likely mis-identification, green circles denote a correctly identified specimen. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We were unable to fully resolve this relationship, but it may represent allopatric speciation, as the discharge of freshwater has been shown to be an effective barrier to gene flow (Muss et al., 2001; Floeter et al., 2008; Araujo et al., 2020).

A more recent blenniid phylogeny using *Hypleurochilus* species from temperate (*H. geminatus*) and tropical (*H. fissicornis*) western Atlantic locations and the eastern Atlantic (*H. bananensis*) suggests a monophyletic relationship between the western Atlantic species compared the eastern Atlantic species (Vecchioni et al. 2022). This is the same relationship recovered by three different species trees reported here with high posterior probability (PP = 1.0) (Fig. 2). Additionally, this relationship is supported by morphological data, as the eastern Atlantic species (*H.* cf *aequipinnis*) is morphologically distinct from all of the western Atlantic species. An eastern Atlantic/western Atlantic phylogenetic break is also seen within the blenniid genus *Scartella* (Araujo et al., 2020). There is also a tropical/temperate split within western Atlantic species of this genus. The recovery of a temperate and tropical clade in the western Atlantic is not surprising, as the Atlantic-Gulf disjunction, where a shift between warm tropical and cooler temperate water occurs) is a well-established phylogeographic barrier (Avise, 2000; Lee and Ó Foighil, 2004). Additionally, the Florida peninsula may cause a biogeographic break explaining the sister relationship between *H. geminatus* and *H. caudovittaus/H. multifilis* (Avise, et al., 1987; Bowen and Avise, 1990; Avise, 1992).

Support for the transatlantic sister relationship observed in the COI gene trees (EA and temperate WA) has been reported in other fish species (Casey et al., 2004; Boehm et al., 2013; Araujo et al., 2020). In this scenario, *Hypleurochilus*, originating in the northern Caribbean and

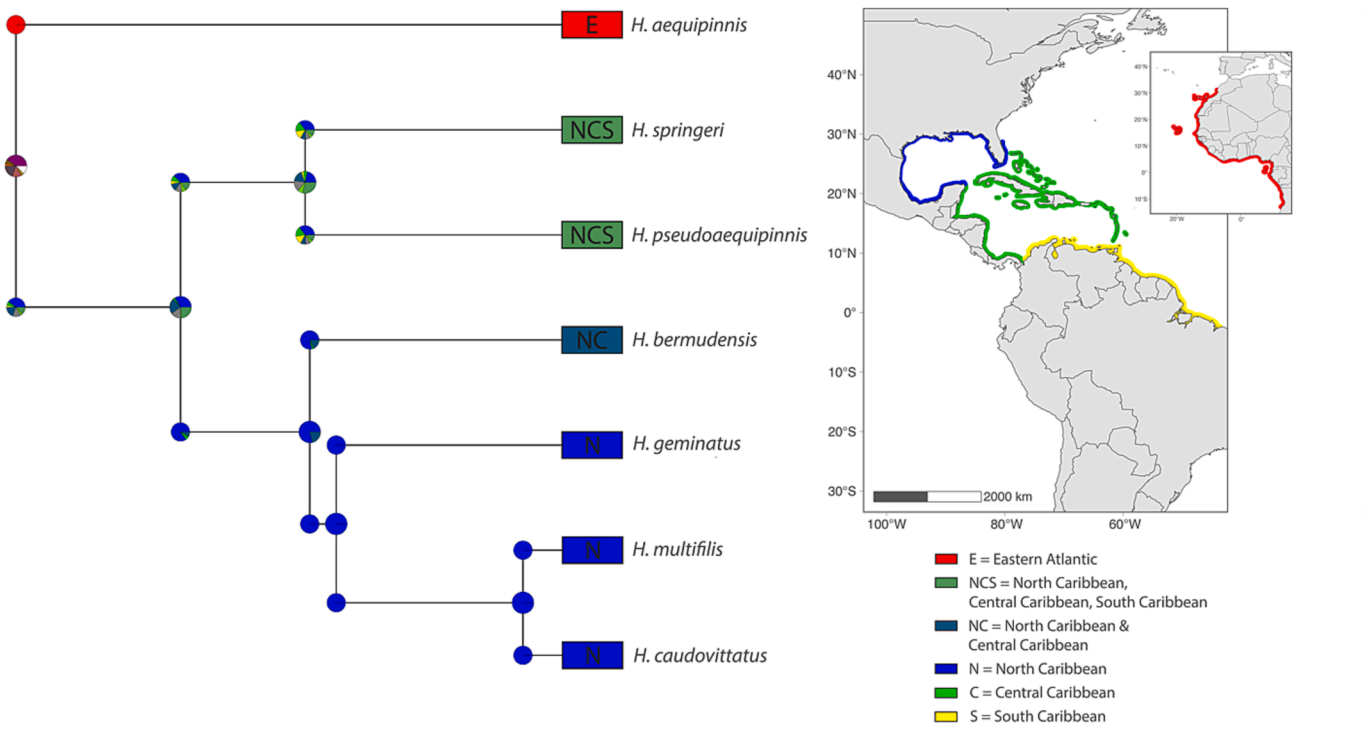

Fig. 7. BioGeoBears Ancestral range reconstruction based on the mtDNA + nDNA species tree reconstructed in StarBeast2. Pie charts on nodes indicate proportions of ancestral range reconstruction for the common ancestor.

eastern Atlantic, would have diverged into a tropical clade in the south and central Caribbean. A common ancestor persisting in the northern Caribbean and eastern Atlantic would have split, leading to the emergence of two distinct *Hypleurochilus* groups: one dispersing and diversifying in the temperate western Atlantic/northern Caribbean and another in the eastern Atlantic. In the second hypothesis of a monophyletic western Atlantic clade, *Hypleurochilus* would again have originated from a broadly dispersed ancestor in the eastern Atlantic and temperate western Atlantic/northern Caribbean before breaking apart and diverging from the temperate western Atlantic/northern Caribbean to the tropical western Atlantic/southern and central Caribbean regions. This transatlantic relationship is similar to distributions of *Sparisoma* (Robertson et al., 2006; Floeter et al., 2008). In both scenarios, low resolution in nodes at the base of the tree and a lack of other eastern Atlantic species produce an unclear result that there was a wide ranging or broadly dispersed common ancestor. It is also possible that the common ancestor originated in either the eastern Atlantic or temperate western Atlantic and a transatlantic dispersal event occurred. Previously, eastern Atlantic to western Atlantic range expansion in the goby genus, *Gnatholepis*, has been linked to global warming and the potential for islands around the mid-Atlantic ridge to aid in dispersal (Rocha et al., 2005). Overall, multilocus species trees and morphological data provide more support for a monophyletic western Atlantic clade than for the sister relationship between the temperate western Atlantic and potential eastern Atlantic clades, leaving the historic biogeography of this genus unresolved.

4.2. Single locus delimitation fails to delimit *H. multifilis* and *H. caudovittatus* or accurately reconstruct the relationship of *H. bermudensis*

Mitochondrial barcoding has been broadly used in species discovery and single-locus delimitation analyses due to its rapid mutation rate and relative ease of amplification and has been used for the discovery of many blenny species (Victor, 2013; 2017; Araujo et al., 2020). Compared to nuclear DNA, mitochondrial markers have a reduced effective population size and allow for rapid sorting of recently diverged populations (Rand et al., 2004; Rubinoff & Holland, 2005; Sunnucks et al., 2017). Yet, within both the western Atlantic clades, gene tree reconstruction using the mitochondrial COI marker was unable to differentiate *H. multifilis/H. caudovittatus* and *H. pseudoaequipinnis/H. springeri*. Single locus species delimitation was also unable to differentiate *H. multifilis/H. caudovittatus*. Species trees created with mtDNA, nDNA and mtDNA + nDNA as well as multilocus delimitation were able to differentiate the *H. multifilis* and *H. caudovittatus* lineages. The discrepancy may be caused by mitochondrial introgression, a phenomenon that has been reported in other blenniid species with mito-nuclear discordance (Belaiba et al., 2019). However, mitochondrial introgression is typically unidirectional, from the local species to an invading species (Currat et al., 2008) and this pattern is not seen in the mitochondrial gene tree topologies reported here. It is more plausible that the discordance between the mitochondrial gene tree and the nuclear species tree stems from incomplete lineage sorting (ILS), a phenomenon relatively common among species that have diverged recently and rapidly (Galtier and Daubin, 2008). Incipient species should first form a polyphyletic group that will evolve to become a paraphyletic group and eventually form a monophyly (Avise, 2000). The paraphyletic nature of both *H. multifilis/H. caudovittatus* and *H. pseudoaequipinnis/H. springeri* suggests that the divergence of these species is relatively recent and ongoing, and that the monophyletic resolution of these two species is limited by the retention of ancestral polymorphism (Tang et al., 2012).

The retention of ancestral polymorphism may also explain the morphological similarity between both *H. multifilis/H. caudovittatus* and *H. pseudoaequipinnis/H. springeri*. These species share significant overlap in morphospace (Fig. 5). In all phylogenetic analyses, *H. geminatus* is sister to the *H. caudovittatus/H. multifilis* group. This is surprising as Bath (1994) describes *H. geminatus* and *H. caudovittatus* as being more morphologically similar to each other than to *H. multifilis*. However, the use of a multilocus species delimitation clearly defines *H. multifilis* and *H. caudovittatus* as genetically distinct sister species. The inclusion of *H. geminiatus* as sister to *H. multifilis/H. caudovittatus* confirms the findings of Bath (1994) that the previously known *H. geminatus* complex is comprised of three distinct species. While unavailable for this study,

the inclusion of nuclear markers and morphological data for *H. fissicornis* would potentially allow for higher resolution of the western Atlantic Clade II species.

There was discordance between the mtDNA species tree and nDNA and mtDNA + nDNA species trees regarding the placement of *H. bermudensis*. Mitochondrial species and gene trees recovered *H. bermudensis* as the earliest branching lineage of western Atlantic Clade II while nDNA and mtDNA + nDNA species trees recovered *H. bermudensis* as the earliest branching lineage of western Atlantic Clade I. This discordance may be caused by ILS or ancient introgression events but a larger sample size of *H. bermudensis* individuals is needed to confirm. Morphological data provides little resolution as this species is clustered intermediately between the two clades in the PCA analysis. Based on ancestral reconstruction, *H. bermudensis* does appear to represent a lineage that has dispersed southward into the central Caribbean region compared to the rest of the Clade I species. Based on current geographic distributions, *H. bermudensis* does have the most range overlap with species in Clade II.

Increased marker sampling through next-generation sequencing techniques may be able to provide higher resolution for internal species tree nodes that had low posterior probability and provide clearer support for the inclusion of *H. bermudensis* the western Atlantic Clade I. Overall, we find that a multilocus approach to species delimitation can more precisely reconstruct the phylogenetic relationships of the *Hypleurochilus* species found in the western Atlantic compared to single marker reconstruction.

4.3. The double-edged sword of publicly available sequence data

Through the addition of publicly available sequence data to our dataset, we can identify two sequences previously identified only to the genus level. The *Hypleurochilus* sp. (GenBank: GU224877; BOLD: MFLII373) first reported in Valdez-Moreno et al. (2010) and subsequently used in phylogenies reconstructed by Attaran-Fariman et al. (2016) and Vecchioni et al. (2019) should be recognized as *H. springeri*. BOLD specimen MXV140 should be recognized as *H. bermudensis*.

Mining sequence data has its challenges; of the 49 sequences mined, 12 were misidentified, with BLAST searches linking to non-*Hypleurochilus* species with >90% ID. Additionally, three sequences mined were incorrect identifications within the *Hypleurochilus* genus. These three specimens identified as *H. geminatus* (GenBank Accession Numbers: KP255118, FCCA139, and FCCA140) group with the western Atlantic clade II, *H. pseudoaequipinnis/H. fissicornis* individuals, all other *H. geminatus* species group as a monophyletic lineage within western Atlantic clade I. Three additional sequences mined are possible misidentifications. Two specimens identified as *H. pseudoaequipinnis* (RIE_4065 and LIDM464) group within the *H. springeri* lineage and an individual identified as *H. springeri* (BZLWC6081) groups within the *H. pseudoaequipinnis/H. fissicornis* lineage. In these instances, a multilocus approach should be implemented to positively identify the species.

Publicly available sequence data and museum specimens can greatly aid in research efforts, but care should be taken to ensure that specimens and sequences are correctly identified and labeled, as mis-identified specimens can confound taxonomic studies (Kitchener et al., 2020). We recommend the routine inclusion of specimen photographs in data repositories for each sequence. This practice could potentially reduce confusion and increase confidence in utilizing mined data.

4.4. Potential range expansion of H. aequipinnis in western Atlantic waters

The unidentified species, listed here as *H. cf aequipinnis* is genetically distinct from all of the species in Clades I and II, and represents a unique lineage that is sister to the two Western Atlantic clades. Initial observations of the unknown *H. cf aequipinnis* species indicated a superficial resemblance to the Atlantic oyster blenny, *H. pseudoaequipinnis*. However, comparisons of morphologic traits also support the unknown *H. cf aequipinnis* as morphologically distinct from six other w. Atlantic *Hypleurochilus* species, including *H. pseudoaequipinnis*. Specimens of *H. cf aequipinnis* are identified based on the species description of *H. aequipinnis* by Bath (1994). Pairwise multiple comparison tests revealed significant variation in morphometric traits with *H. cf aequipinnis* featuring a higher cephalic sensory pore count relative to size than any other species and a longer snout length and shorter peduncle length compared to species in Clades I and II.

The morphological distinction between *H. cf aequipinnis* and *H. pseudoaequipinnis* is important, as these two fish have a convoluted history. Originally thought to be the same species (*H. aequipinnis*) occurring on both sides of the Atlantic (Randall, 1966), the western Atlantic species was renamed *H. pseudoaequipinnis* after a more detailed description of the eastern Atlantic *H. aequipinnis* was provided (Bath & Wirtz, 1981; Bath, 1994). *Hypleurochilus aequipinnis* is formally described as not occurring on both sides of the Atlantic (Bath, 1994), while in 2007 *H. pseudoaequipinnis* is reported in the Gulf of Guinea (Wirtz et al., 2007). The first molecular data from these two species confirm *H. aequipinnis* and *H. pseudoaequipinnis* as genetically distinct sister species in a multilocus phylogeny (Levy et al. 2013). However we report them occupying different clades. While material of *H. aequipinnis* was unobtainable for this study, mining the sequence databases for additional *Hypleurochilus* sequences provided sequence data for another eastern Atlantic species, *H. bananensis*. Within the phylogeny reconstructed using the GenBank/BOLD sequences, *H. cf aequipinnis* grouped as sister to *H. bananensis*, providing additional support for the hypothesis that the unknown species collected from the Gulf of Mexico may represent the range expansion of a species from the eastern Atlantic.

If this species reported from the GoM is indeed *H. aequipinnis*, it would be the second species from the genus to be reported on both sides of the Atlantic Ocean and *H. aequipinnis* would not be the first reef fish to expand their range into the Caribbean and Gulf of Mexico. Multiple species of invasive blenniids have been reported along the western Atlantic, including *Omobranchus punctatus* (Gerhardinger et al., 2006; Lasso-Alcalá et el., 2011) and *O. sewali* (Cabezas et al., 2022). Human-mediated dispersal may be the most significant cause for these invasions. Many collection locations for *O. punctatus* in the southwestern Atlantic are proximal to ports and zones of heavy ship traffic support the hypothesis that this species may have been initially introduced via ships from India (Lasso-Alcalá et el., 2011). Transportation of species through ship ballast water is common (Wonham et al., 2000) and because of the benthic association of cryptobenthic fishes, hull-fouling transport may be another mode of invasion (Ferreira et al., 2006). A third mode of transport for invasives is the transport of oil rigs (Wanless et al., 2010). The heavy biofouling on these structures allow for the transport of species in the same way as the biofouling on ship hulls. The transport of oil rigs from the Indian Sea, around Africa and into the Gulf of Mexico has been suggested as the cause for recent invasions of the Indo-Pacific damselfish, *Neopomacentrus cyanomos* (Robertson et al., 2016, 2018).

5. Conclusions

The interrelationships among nine out of the eleven described species within the *Hypleurochilus* genus are resolved using a combination of single locus and multilocus species delimitation methods. Mining sequence databases allowed for the inclusion of more species in a single marker species delimitation than that of the multilocus delimitation. The single marker delimitation using mitochondrial *COI* includes all species of *Hypleurochilus* found in the western Atlantic and one species, *H. bananensis,* found in the eastern Atlantic. Single marker delimitation failed to uncover the relationships between *H. springeri/H. pseudoaequipinnis* and between *H. multifilis/H. caudovittatus*, though it successfully recovered two distinct clades in the western Atlantic and a potential eastern Atlantic clade, indicating a possible range expansion of *Hypleurochilus* species in the western Atlantic. Morphological data shows

significant overlap in morphospace between *H. multifilis/H. caudovittatus* and *H. springeri/H. pseudoaequipinnis*, but marked distinction between *H. cf aequipinnis* and western Atlantic species. Moreover, the multilocus species delimitation and phylogenetic reconstruction efficiently delineate the two western Atlantic clades and a third distinct lineage while also revealing a sister relationship between *H. caudovittatus* and *H. multifilis*.

CRediT authorship contribution statement

Joshua E. Carter: Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Data curation, Writing – original draft, Writing – review & editing. **Megan A. Sporre:** Investigation, Formal analysis, Visualization, Data curation, Writing – original draft, Writing – review & editing. **Ron I. Eytan:** Supervision, Funding acquisition, Methodology, Resources, Writing – review & editing.

Declaration of Competing Interest

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

Acknowledgements

We would like to thank K. St Clair and the Texas A&M University at Galveston Sea Life Facility staff; The Academy of National Sciences, Philadelphia (ANSP), Smithsonian Institution National Museum of Natural History (USNM), and University of Florida (UF) for providing tissues and voucher specimens; and the reviewers for their insightful comments that helped to improve this manuscript.

Funding Statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2023.107933.

References

Almada, F., Almada, V.C., Guillemaud, T., Wirtz, P., 2005. Phylogenetic relationships of the north-eastern Atlantic and Mediterranean blenniids. Biol. J. Linn. Soc. 86 (3), 283–295. https://doi.org/10.1111/j.1095-8312.2005.00519.x.

Araujo, G.S., Vilasboa, A., Britto, M.R., Bernardi, G., Von der Heyden, S., Levy, A., Floeter, S.R., 2020. Phylogeny of the comb-tooth blenny genus Scartella (Blenniiformes: Blenniidae) reveals several cryptic lineages and a trans-Atlantic relationship. Zool. J. Linn. Soc. 190 (1), 54–64. https://doi.org/10.1093/zoolinnean/zlz142.

Attaran-Fariman, G., Estekani, S. and Ghasemzadeh, J., 2016. The first study of the phylogenetic relationships of three rare species of Blenniids (Fishes: Blenniidae) from Makoran Coast based on the gene sequencing of Cytochrome Oxidase 1. 10.22092/ISFJ.2017.110229.

Avise, J.C., 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. Oikos 62–76.

Avise, J.C., 2000. Phylogeography, the History and Formation of Species. Harvard University, England.

Avise, J.C., Reeb, C.A., Saunders, N.C., 1987. Geographic population structure and species differences in mitochondrial DNA of mouthbrooding marine catfishes (Ariidae) and demersal spawning toadfishes (Batrachoididae). Evolution 41 (5), 991–1002. https://doi.org/10.1111/j.1558-5646.1987.tb05872.x.

Bath, H., 1994. Untersuchung der Arten Hypleurochilus geminatus (Wood 1825), H. fissicornis (Quoy & Gaimard 1824) und H. aequipinnis (Günther 1861). *mit Revalidation von Hypleurochilus multifilis (Girard 1858) und Beschreibung von zwei neuen Arten*. Senckenberg. Biol. 74, 59–85.

Bath, H., Wirtz, P., 1981. Rediscovery of Hypleurochilus aequipinnis (Günther 1861) in West Africa (Pisces: Blenniidae). Senckenb. Biol. 61 (5–6), 349–355.

Belaiba, E., Marrone, F., Vecchioni, L., Bahri-Sfar, L., Arculeo, M., 2019. An exhaustive phylogeny of the combtooth blenny genus Salaria (Pisces, Blenniidae) shows introgressive hybridization and lack of reciprocal mtDNA monophyly between the marine species Salaria basilisca and Salaria pavo. Mol. Phylogenet. Evol. 135, 210–221.

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 (10), 1839–1849.

Bouckaert, R.R., 2010. DensiTree: making sense of sets of phylogenetic trees. Bioinformatics 26 (10), 1372–1373. https://doi.org/10.1093/bioinformatics/btq110.

Bouckaert, R.R., Drummond, A.J., 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. BMC Evol. Biol. 17, 42.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15, e1006650.

Bowen, B.W., Avise, J.C., 1990. Genetic structure of Atlantic and Gulf of Mexico populations of sea bass, menhaden, and sturgeon: influence of zoogeographic factors and life-history patterns. Mar. Biol. 107 (3), 371–381.

Briggs, J.C., Bowen, B.W., 2013. Marine shelf habitat: biogeography and evolution. J. Biogeogr. 40, 1023–1035.

Cabezas, M.P., Lasso-Alcalá, O.M., Quintero-T, E., Xavier, R., Giarrizzo, T., Nunes, J.L., Machado, F.S., Gómez, J., Silva Pedroza, W., Jowers, M.J., 2022. Clarifying the taxonomy of some cryptic blennies (Blenniidae) in their native and introduced range. Sci. Rep. 12 (1), 1–14.

Carpenter, K.E. (ed.) The living marine resources of the Western Central Atlantic. Volume 2: Bony fishes part 1 (Acipenseridae to Grammatidae). FAO Species Identification Guide for Fishery Purposes and American Society of Ichthyologists and Herpetologists Special Publication No. 5. Rome, FAO. 2002. pp. 601-1374.

Casey, S.P., Hall, H.J., Stanley, H.F., Vincent, A.C., 2004 Feb 1. The origin and evolution of seahorses (genus Hippocampus): a phylogenetic study using the cytochrome b gene of mitochondrial DNA. Mol. Phylogenet. Evol. 30 (2), 261–272. https://doi.org/10.1016/j.ympev.2003.08.018.

Currat, M., Ruedi, M., Petit, R.J., Excoffier, L., 2008 Aug. The hidden side of invasions: massive introgression by local genes. Evolution: International Journal of Organic. Evolution 62 (8), 1908–1920. https://doi.org/10.1111/j.1558-5646.2008.00413.x.

Douglas, J., Jiménez-Silva, C.L., Bouckaert, R., 2022. StarBeast3: Adaptive Parallelized Bayesian Inference under the Multispecies Coalescent. Syst. Biol. 71, 901–916.

Drummond, A.J., Rambaut, A., 2007. Tree Annotator version 2.6.

Ferreira, C.E.L., Gonçalves, J.E.A., Coutinho, R., 2006. Ship hulls and oil platforms as potential vectors to marine species introduction. J. Coastal Res. 1340–1345.

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., Gasparini, J.L., Brito, A., 2008. Atlantic reef fish biogeography and evolution. J. Biogeogr. 35 (1), 22–47.

Flot, J.F., 2010. SeqPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. Mol. Ecol. Resour. 10 (1), 162–166. https://doi.org/10.1111/j.1755-0998.2009.02732.x.

Galtier, N., Daubin, V., 2008 Dec 27. Dealing with incongruence in phylogenomic analyses. Philos. Trans. R. Soc., B 363 (1512), 4023–4029. https://doi.org/10.1098/rstb.2008.0144.

Gerhardinger, L.C., Freitas, M.O., Andrade, Á.B., Rangel, C.A., 2006. *Omobranchus punctatus* (Teleostei: Blenniidae), an exotic blenny in the Southwestern Atlantic. Biol. Invasions 8 (4), 941–946.

Glenn, T. PEG, 2019. Precipitation of PCR products. Protocolo online.

Greenfield, D.W., Johnson, R.R., 1981. The Blennioid fishes of Belize and Honduras, Central America, with comments on their sytematics, ecolgy, and distribution (Blenniidae, Chaenopsidae, Labrisomidae, Tripterygiidae). *Fieldiana*. Zoology 8, 105.

Hastings, P.A., Springer, V.G., 2009. Systematics of the Blenniidae (combtooth blennies). Biol. Blennies 69, 91.

Hubbs, C.L., Lagler, K.F., 1958. Fishes of the Great Lakes region. Univ. Mich. Press, Ann Arbor, p. 213.

Hundt, P.J., Iglésias, S.P., Hoey, A.S., Simons, A.M., 2014. A multilocus molecular phylogeny of combtooth blennies (Percomorpha: Blennioidei: Blenniidae): multiple invasions of intertidal habitats. Mol. Phylogenet. Evol. 70, 47–56. https://doi.org/10.1016/j.ympev.2013.09.001.

Javonillo, R., Harold, A.S., 2010. A systematic review of the genus Chasmodes (Teleostei: Perciformes: Blenniidae). Zootaxa 2558 (1), 1–16. https://doi.org/10.11646/zootaxa.2558.1.1.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30 (4), 772–780. https://doi.org/10.1093/molbev/mst010.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28 (12), 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.

Kitchener, A.C., Machado, F.A., Hayssen, V., Moehlman, P.D., Viranta, S., 2020. Consequences of the misidentification of museum specimens: the taxonomic status of *Canis lupaster soudanicus*. J. Mammal. 101 (4), 1148–1150. https://doi.org/10.1093/jmammal/gyaa080.

Lasso-Alcalá, O., Nunes, J.L., Lasso, C., Posada, J., Robertson, R., Piorski, N.M., Tassell, J. V., Giarrizzo, T., Gondolo, G., 2011. Invasion of the indo-pacific blenny Omobranchus punctatus (Perciformes: Blenniidae) on the Atlantic coast of central and south America. Neotropical Ichthyol. 9, 571–578.

Lê, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. J. Stat. Softw. 25, 1–18. https://doi.org/10.18637/jss.v025.i01.

Leaché, A.D., Fujita, M.K., 2010. Bayesian species delimitation in West African forest geckos (Hemidactylus fasciatus). Proc. R. Soc. B Biol. Sci. 277 (1697), 3071–3077. https://doi.org/10.1098/rspb.2010.0662.

Levy, A., Wirtz, P., Floeter, S.R., Almada, V.C., 2011. The Lusitania Province as a center of diversification: The phylogeny of the genus *Microlipophrys* (Pisces: Blenniidae). Mol. Phylogenet. Evol. 58 (2), 409–413. https://doi.org/10.1016/j.ympev.2010.12.008.

Lee, T., Ó Foighil, D., 2004. Hidden Floridian biodiversity: mitochondrial and nuclear gene trees reveal four cryptic species within the scorched mussel, Brachidontes exustus, species complex. Mol. Ecol. 13 (11), 3527–3542.

Levy, A., Von der Heyden, S., Floeter, S.R., Bernardi, G., Almada, V.C., 2013. Phylogeny of Parablennius Miranda Ribeiro, 1915 reveals a paraphyletic genus and recent Indo-Pacific diversification from an Atlantic ancestor. Mol. Phylogenet. Evol. 67 (1), 1–8. https://doi.org/10.1016/j.ympev.2012.12.010.

Li, C., Ortí, G., Zhang, G., Lu, G., 2007. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. BMC Evol. Biol. 7 (1), 1–11. https://doi.org/10.1186/1471-2148-7-44.

Matzke, N.J., 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. University of California, Berkeley.

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., 2018a. _cladoRcpp v0.15.1: C++ Implementations of Phylogenetic Cladogenesis Calculations. University of Auckland, New Zealand.

Matzke, N.S.R., Schmidt, D., 2023. _rexpokit v0.26.6.8: R wrappers for EXPOKIT; other matrix functions. School of Biological Sciences, University of Auckland, New Zealand.

Matzke, N. J. 2013b. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Frontiers of Biogeography **5**.

Matzke, N. 2018. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis with R scripts. version 1.1. 1. San Francisco, CA: GitHub. Available online at: https://github. com/nmatzke/BioGeoBEARS.

Muss, A., Robertson, D.R., Stepien, C.A., Wirtz, P., Bowen, B.W., 2001. Phylogeography of Ophioblennius: the role of ocean currents and geography in reef fish evolution. Evolution 55 (3), 561–572. https://doi.org/10.1111/j.0014- 3820.2001.tb00789.x.

Ogilvie, H.A., Bouckaert, R.R., Drummond, A.J., 2017. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. Mol. Biol. Evol. 34 (8), 2101–2114. https://doi.org/10.1093/molbev/msx126.

Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20, 289–290.

Pinheiro, H.T., Gasparini, J.L., Rangel, C.A., 2013. A new species of the genus Hypleurochilus (Teleostei: Blenniidae) from Trindade Island and Martin Vaz archipelago, Brazil. Zootaxa 3709 (1), 095–100. https://doi.org/10.11646/zootaxa.3709.1.5.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55 (4), 595–609. https://doi.org/10.1080/10635150600852011.

Posada, D., 2008 Jul 1. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25 (7), 1253–1256.

Rambaut, A., Drummond, A., 2016. Figtree version 1, 3.

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 (5), 901–904. https://doi.org/10.1093/sysbio/syy032.

Rand, D.M., Haney, R.A., Fry, A.J., 2004. Cytonuclear coevolution: The genomics of cooperation. Trends Ecol. Evol. 19, 645–653.

Randall, J.E., 1966. The West Indian blenniid fishes of the genus hypleurochilus, with the description of a new species.

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.

Robertson, D. R., and K. L. Cramer. 2014. Defining and dividing the Greater Caribbean: Insights From the biogeography of shorefishes. PLoS ONE 9:e102918.

Robertson, D.R., Karg, F., de Moura, R.L., Victor, B.C., Bernardi, G., 2006. Mechanisms of speciation and faunal enrichment in Atlantic parrotfishes. Mol. Phylogenet. Evol. 40 (3), 795–807. https://doi.org/10.1016/j.ympev.2006.04.011.

Robertson, D.R., Simoes, N., Gutierrez Rodriguez, C., Pineros, V.J., Perez-Espana, H., 2016. An Indo-Pacific damselfish well established in the southern Gulf of Mexico: prospects for a wider, adverse invasion. J. Ocean Sci. Found.

Robertson, D.R., Dominguez-Dominguez, O., Victor, B., Simoes, N., 2018. An Indo-Pacific damselfish (*Neopomacentrus cyanomos*) in the Gulf of Mexico: origin and mode of introduction. PeerJ 6, e4328.

Rocha, L.A., Robertson, D.R., Rocha, C.R., Van Tassell, J.L., Craig, M.T., Bowen, B.W., 2005 Nov. Recent invasion of the tropical Atlantic by an Indo-Pacific coral reef fish. Mol. Ecol. 14 (13), 3921–3928.

Rubinoff, D., Holland, B.S., 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst. Biol. 54 (6), 952–961.

Sidje, R.B., 1998. A Software Package for Computing Matrix Exponentials. ACM-Trans. Math. Software 24, 130–156.

Springer, V.G., 1993. Definition of the suborder Blennioidei and its included families (Pisces: Perciformes). Bull. Mar. Sci. 52 (1), 472–495.

Springer, V.G., Gomon, M.F., 1975. Revision of the blenniid fish genus Omobranchus, with descriptions of three new species and notes on other species of the tribe Omobranchini. Smithsonian Contributions to Zoology.

Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68 (4), 978–989. https://doi.org/10.1086/319501.

Sunnucks, P., Morales, H.E., Lamb, A.M., Pavlova, A., Greening, C., 2017. Integrative approaches for studying mitochondrial and nuclear genome co-evolution in oxidative phosphorylation. Front. Genet. 8, 25.

Tang, Q.Y., Liu, S.Q., Yu, D., Liu, H.Z., Danley, P.D., 2012 May. Mitochondrial capture and incomplete lineage sorting in the diversification of balitorine loaches (Cypriniformes, Balitoridae) revealed by mitochondrial and nuclear genes. Zool. Scr. 41 (3), 233–247. https://doi.org/10.1371/journal.pone.0080641.

Topolski, M.F., Szedlmayer, S.T., 2004. Vertical Distribution, Size Structure, and Habitat Associations of Four Blenniidae Species on Gas Platforms in the Northcentral Gulf of Mexico. Environ. Biol. Fishes 70, 193–201. https://doi.org/10.1023/B:EBFI.0000029364.23532.94.

Valdez-Moreno, M., Vásquez-Yeomans, L., Elías-Gutiérrez, M., Ivanova, N.V., Hebert, P. D., 2010. Using DNA barcodes to connect adults and early life stages of marine fishes from the Yucatan Peninsula, Mexico: potential in fisheries management. Mar. Freshw. Res. 61 (6), 655–671. https://doi.org/10.1071/MF09222.

Vecchioni, L., Ching, A.C., Marrone, F., Arculeo, M., Hundt, P.J., Simons, A.M., 2022. Multi-locus phylogenetic analyses of the Almadablennius clade reveals inconsistencies with the present taxonomy of blenniid fishes. Diversity 14 (1), 53. https://doi.org/10.3390/d14010053.

Vecchioni, L., Marrone, F., Belaiba, E., Tiralongo, F., Bahri-Sfar, L., Arculeo, M., 2019 Feb. The DNA barcoding of Mediterranean combtooth blennies suggests the paraphyly of some taxa (Perciformes, Blenniidae). J. Fish Biol. 94 (2), 339–344.

Victor, B.C., 2013. The Caribbean Roughhead Triplefin (Enneanectes boehlkei): DNA barcoding reveals a complex of four West Indian sympatric cryptic species (Teleostei: Blennioidei: Tripterygiidae). J. Ocean Sci. Found. 7, 44–73.

Victor, B.C., 2017. Review of the Indo-Pacific Pseudojuloides cerasinus species complex with a description of two new species (Teleostei: Labridae). J. Ocean Sci. Found. 29, 11–31.

Wanless, R.M., Scott, S., Sauer, W.H., Andrew, T.G., Glass, J.P., Godfrey, B., Griffiths, C., Yeld, E., 2010. Semi-submersible rigs: a vector transporting entire marine communities around the world. Biol. Invasions 12 (8), 2573–2583.

Wirtz, P., Ferreira, C.E.L., Floeter, S.R., Fricke, R., Gasparini, J.L., Iwamoto, T., Rocha, L., Sampaio, C.L., Schliewen, U.K., 2007. Coastal fishes of São Tomé and Príncipe islands, Gulf of Guinea (eastern Atlantic Ocean)–an update. Zootaxa 1523 (1), 1–48.

Wonham, M.J., Carlton, J.T., Ruiz, G.M., Smith, L.D., 2000. Fish and ships: relating dispersal frequency to success in biological invasions. Mar. Biol. 136 (6), 1111–1121.

Yang, Z., 2015. The BPP program for species tree estimation and species delimitation. Curr. Zool. 61 (5), 854–865. https://doi.org/10.1093/czoolo/61.5.854.

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.Y., 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8 (1), 28–36. https://doi.org/10.1111/2041-210X.12628.