Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

Phylogenomic analysis of a rapid radiation of misfit fishes (Syngnathiformes) using ultraconserved elements

S.J. Longo [a,*], B.C. Faircloth [b], A. Meyer [c], M.W. Westneat [d], M.E. Alfaro [e], P.C. Wainwright [a]

[a] Department of Evolution and Ecology, University of California, Davis, CA 95616, USA
[b] Department of Biological Sciences and Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803, USA
[c] Department of Biology, University of Konstanz, D-78457 Konstanz, Germany
[d] Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL 60637, USA
[e] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA

ARTICLE INFO

ABSTRACT

Phylogenetics is undergoing a revolution as large-scale molecular datasets reveal unexpected but repeatable rearrangements of clades that were previously thought to be disparate lineages. One of the most unusual clades of fishes that has been found using large-scale molecular datasets is an expanded Syngnathiformes including traditional long-snouted syngnathiform lineages (Aulostomidae, Centriscidae, Fistulariidae, Solenostomidae, Syngnathidae), as well as a diverse set of largely benthic-associated fishes (Callionymoidei, Dactylopteridae, Mullidae, Pegasidae) that were previously dispersed across three orders. The monophyly of this surprising clade of fishes has been upheld by recent studies utilizing both nuclear and mitogenomic data, but the relationships among major lineages within Syngnathiformes remain ambiguous; previous analyses have inconsistent topologies and are plagued by low support at deep divergences between the major lineages. In this study, we use a dataset of ultra-conserved elements (UCEs) to conduct the first phylogenomic study of Syngnathiformes. UCEs have been effective markers for resolving deep phylogenetic relationships in fishes and, combined with increased taxon sampling, we expected UCEs to resolve problematic syngnathiform relationships. Overall, UCEs were effective at resolving relationships within Syngnathiformes at a range of evolutionary timescales. We find consistent support for the monophyly of traditional long-snouted syngnathiform lineages (Aulostomidae, Centriscidae, Fistulariidae, Solenostomidae, Syngnathidae), which better agrees with morphological hypotheses than previously published topologies from molecular data. This result was supported by all Bayesian and maximum likelihood analyses, was robust to differences in matrix completeness and potential sources of bias, and was highly supported in coalescent-based analyses in ASTRAL when matrices were filtered to contain the most phylogenetically informative loci. While Bayesian and maximum likelihood analyses found support for a benthic-associated clade (Callionymidae, Dactylopteridae, Mullidae, and Pegasidae) as sister to the long-snouted clade, this result was not replicated in the ASTRAL analyses. The base of our phylogeny is characterized by short internodes separating major syngnathiform lineages and is consistent with the hypothesis of an ancient rapid radiation at the base of Syngnathiformes. Syngnathiformes therefore present an exciting opportunity to study patterns of morphological variation and functional innovation arising from rapid but ancient radiation.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Syngnathiformes (seahorses, pipefishes, and their relatives) are a morphologically distinctive group of fishes with novel modes of locomotion, reproduction, and feeding. Lineages within Syngnathiformes have evolved vertical swimming postures (head-down in

shrimpfish: Centriscidae; head-up in seahorses: Syngnathidae), prehensile tails (seahorses, pipehorses: Syngnathidae), power-amplified feeding (Syngnathidae), female egg-brooding in specialized pelvic fins (ghost-pipefishes: Solenostomidae), and, of course, male pregnancy (Syngnathidae). The Syngnathidae is by far the most studied family within Syngnathiformes, and they have become a model system for the study of sexual selection and sex-role reversal (Berglund and Rosenqvist, 2003; Mobley and Jones, 2013; Stolting and Wilson, 2007; Wilson et al., 2001, 2003)

* Corresponding author.
E-mail address: sjlongo@ucdavis.edu (S.J. Longo).

and exceptional adaptation, including feeding behaviors and tail morphology (Neutens et al., 2014; Porter et al., 2015; Van Wassenbergh et al., 2008, 2009). However, the study of the origin and evolution of traits within syngnathids, and within syngnathiforms as a whole, has been hampered by the lack of well-sampled, and well-supported, phylogenetic hypotheses of the relationships between syngnathiform families and their association with other acanthomorph fishes.

The order Syngnathiformes encompasses a number of families characterized by an elongated tube-shaped snout with small jaws at the end and traditionally includes at least five families: Aulostomidae (trumpetfish), Centriscidae (shrimpfish and snipefish), Fistulariidae (cornetfish), Solenostomidae (ghost pipefish), and Syngnathidae (seahorses, pipefish, and seadragons) (Nelson, 2006). The Pegasidae (seamoths) have often been considered close relatives of the Syngnathiformes based on a number of characters, including their gill filament structure and the possession of bony armor (Johnson and Patterson, 1993; Nelson, 2006; Pietsch, 1978). Morphological traits, including the presence of armor and absence of pleural ribs in some taxa, has also led to the association of Syngnathiformes with the Gasterosteiformes, which historically included the families Aulorhynchidae (tubesnouts), Gasterosteidae (sticklebacks), Hypoptychidae (sand eels), and Indostomidae (armored sticklebacks) (Nelson, 2006). The history of phylogenetics and taxonomic work on this group is rich, with many of the possible combinations of relationships among syngnathiform and gasterosteiform lineages proposed previously (Wilson and Orr, 2011). However, an alternative to a close syngnathiform-gasterosteiform relationship had never been suggested before the use of molecular phylogenetics.

In contrast with morphological studies, large-scale molecular datasets with broad taxonomic sampling have consistently rejected a close relationship between Syngnathiformes and Gasterosteiformes (Betancur-R et al., 2013a; Chen et al., 2003; Dettai and Lecointre, 2005; Kawahara et al., 2008; Leo and Craig, 2007; Li et al., 2009; Near et al., 2012, 2013; Sanciangco et al., 2016; Smith and Wheeler, 2004, 2006; Song et al., 2014). Instead, traditional long-snouted syngnathiform lineages fall out among a surprising array of largely benthic-associated fishes, including the familiar pegasid seamoths, Dactylopteridae (flying gurnards), Callionymoidei (dragonets), and Mullidae (goatfishes) (Fig. 1). Each of these groups has been historically problematic in their own right, with much debate about their relationships to other fishes arising based on morphological studies. Although Pietsch (1978) noted similarities linking Dactylopteridae to Pegasidae and/or Syngnathiformes, more recent morphological classifications have placed Dactylopteridae as a suborder within Scorpaeniformes (Eschmeyer et al., 2017; Nelson, 1994, 2006; Washington et al., 1984), its own order closely related to Scorpaeniformes (Johnson and Patterson, 1993), or a perciform family closely related to the Malacanthidae (Imamura, 2000). The Callionymoidei is composed of two families of dragonets, the Callionymidae and Draconettidae; this order has generally been considered a notothenioid derivative with a strong affinity with gobiesocid clingfishes (Allen, 1984; Gosline, 1970; Nelson, 1994, 2006; Springer and Johnson, 2004), although this affinity has been challenged based on key osteological differences (Allen, 1984; Houde, 1984). Mullidae, or goatfishes, are largely characterized by a number of unique morphological characteristics, many of them relating to the specialization of the hyoid apparatus into barbels, which "have no counterpart, rudimentary or well-developed, elsewhere among percoids" (Gosline, 1984). Based on vertebral and cranial similarities, goatfish have been linked within Perciformes to the Lutjanidae or Sparidae (Gosline, 1984; Gregory, 1933), though Gosline notes that many such characters could be the result of convergent adaptation to benthic feeding. The surprising but replicated rearrangement of

families previously thought to be associated with disparate lineages across the Scorpaeniformes, Perciformes, and Gasterosteiformes into this new syngnathiform clade has been one of the major upheavals in modern molecular phylogenetics for fishes, akin to the finding of a close relationship between Lophiiformes and Tetraodontiformes (Miya et al., 2003; Song et al., 2014).

Studies incorporating broad sampling across the fish tree of life with a variety of genetic loci and phylogenetic reconstruction methods provide strong evidence for a monophyletic Syngnathiformes comprised of the Synganthidae, Solenostomidae, Aulostomidae, Fistulariidae, Centriscidae, Pegasidae, Dactylopteridae, Callionymidae, Draconnetidae, and Mullidae (Betancur-R et al., 2013a; Kawahara et al., 2008; Near et al., 2012, 2013; Sanciangco et al., 2016; Song et al., 2014). Song et al. (2014) recently used whole mitogenomes and a likelihood approach to confirm the monophyly of this clade; based on their results, Song et al. suggested this clade be referred to as 'Syngnathiformes' over other alternatives (e.g., Clade D, Syngnathoidei, Syngnatharia), and we use this more inclusive definition of Syngnathiformes here. While the affinities of these diverse fishes are strongly supported by molecular data, the interrelationships of syngnathiform families remain ambiguous because previous molecular phylogenies recover conflicting topologies (Fig. 1; also see review in Song et al., 2014). For instance, traditional long-snouted syngnathiforms (Aulostomidae, Centriscidae, Fistulariidae, Solenostomidae, Synganthidae) may be monophyletic and nested within the benthic-associated families (Pegasidae, Dactylopteridae, Callionymidae, Draconettidae, and Mullidae) similar to Near et al. (2013). Alternatively, the benthic lineages may be interspersed among traditional syngnathiforms similar to Betancur-R et al. (2013a) or Song et al. (2014). In addition to topological inconsistency, previous analyses found low support at multiple nodes separating the deep divergences among the major lineages within Syngnathiformes (Fig. 1).

One factor that may account for the inconsistency among previous studies is taxonomic coverage. Of the previous molecular datasets, only Song et al. (2014) included a representative from all eleven revised syngnathiform families. In addition, because these studies were focused on broad-scale relationships among the fish tree of life, most included a single species or few representatives per lineage. Sanciangco et al. (2016) included the most syngnathiform taxa to date, but their overall species representation is still low for the most speciose families (e.g. they included only eight syngnathid species, which is about 2.5% of the 316 valid species in this family according to Eschmeyer et al. (2017)). Although there is some debate about the effectiveness of increasing taxon sampling in light of the computational demands of increasingly large matrices, it is generally agreed that higher species coverage should increase phylogenetic accuracy (Heath et al., 2008; Lecointre et al., 1993; Zwickl and Hillis, 2002). The addition of taxa is also vitally important when lineages have deep divergences and are therefore separated by long branches (Hillis et al., 2003). We expect such challenges apply to the accurate recovery of the sygnanthiform tree, given that prior studies have found that major lineages sit on relatively long branches that in turn stem from a series of rapid divergences with short internodes (e.g., Betancur-R et al., 2013a; Near et al., 2012; Song et al., 2014). Some of the longest branches, such as those leading to aulostomids and solenostomids, will be impossible to break given that these families circumscribe few recently-diverged species (e.g., Bowen et al., 2001). However, other families like the Syngnathidae, Callionymidae, and Mullidae are much richer in species than their representation in previous large-scale analyses, and better sampling from these clades may help alleviate long-branch attraction artifacts (Heath et al., 2008; Hillis et al., 2003).
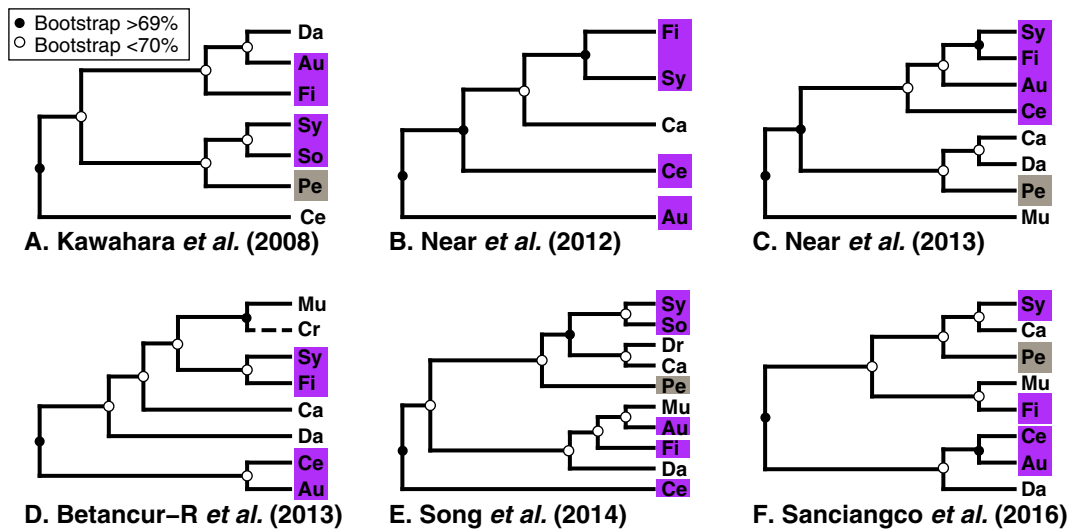
Fig. 1. Recently published relationships for syngnathiform lineages based on large-scale molecular datasets. Although sampling varies along studies, traditional long-snouted syngnathiform lineages (purple) and pegasids (gray) fall out among an array of benthic-associated lineages. However, among-family relationships are highly variable and often poorly supported. Cladograms are redrawn from the following sources: (A) Kawahara et al. (2008; Fig. 3), (B) Near et al. (2012; Fig. S1), (C) Near et al. (2013; Fig. S1C), (D) Betancur-R et al. (2013a; Fig. 4), (E) Song et al. (2014; Fig. 2), (F) Sanciangco et al. (2016; Fig. S1). Nodes are labeled by bootstrap support: greater than or equal to 70% (black) or less than 70% (white). Abbreviations: Au = Aulostomidae, Ce = Centriscidae (which includes 'Macroramphosidae'), Ca = Callionymidae, Cr = Creediidae. Da = Dactylopteridae, Dr = Draconettidae, Fi = Fistulariidae, Mu = Mullidae, Pe = Pegasidae, Sy = Syngnathidae, So = Solenostomidae. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In this study, we investigate the evolutionary history of Syngnathiformes by inferring several phylogenies using ultraconserved elements (UCEs). Our goal is to estimate the relationships among syngnathiform fishes by including members of all major lineages, increasing the sampled species diversity, and leveraging the power of a phylogenomic dataset. We use an actinopterygian-specific set of probes to target approximately 1340 UCEs that have been effective markers for resolving deep phylogenetic relationships, particularly for acanthomorph fishes (Faircloth et al., 2013; Gilbert et al., 2015). Hybrid capture of UCEs renders sequence data that have a dramatic substitution-rate profile—ranging from sites with very low substitution rates near the UCE core to sites with substantially higher rates in the flanking regions (Faircloth et al., 2012). Both the core and flanking regions of UCEs have been shown to have considerably higher net phylogenetic informativeness compared to protein coding genes over a range of divergence times (Gilbert et al., 2015), and UCE sequence data has been used to resolve phylogenetic relationships spanning a wide range of divergence dates in a number of clades across the tree of life (Blaimer et al., 2015; Crawford et al., 2012; Faircloth et al., 2014; Harvey et al., 2016; Manthey et al., 2016; McCormack et al., 2012; Smith et al., 2014). Therefore our UCE dataset, combined with increased taxonomic coverage, should provide increased resolution of the problematic and deep branching patterns within Syngnathiformes spanning a 60–90 million-year interval (Betancur-R et al., 2013a; Near et al., 2012; Sanciangco et al., 2016).

2. Materials and methods

2.1. Taxon sampling

Our dataset is composed of 112 species representing 10 out of the 11 syngnathiform families as follows: Syngnathidae (59 species, out of 316), Solenostomidae (2 species, out of 6), Centriscidae (4 species out of 13, including species of *Macroramphosus* which were previously grouped in the family Macroramphosidae), Aulostomidae (2 species, out of 3), Fistulariidae (4 species, out of 4), Pegasidae (2 species, out of 6), Dactylopteridae (3 species, out of 7), Callionymidae (22 species, out of 198), and Mullidae (14 spe-

cies, out of 87). Of all syngnathiforms lineages, only a representative of the callionymoid family Draconnetidae (15 species) is missing due to insufficient sample quality. We used the scombroid *Taractichthys longipinnis* as an outgroup, because recent molecular studies using nuclear datasets have identified the Scombroidei (or Scombroidei plus other Perciformes) as the closest relatives of Syngnathiformes (Betancur-R et al., 2013a; Dettai and Lecointre, 2005; Near et al., 2012, 2013; Sanciangco et al., 2016). Although Betancur-R et al. (2013a) reported the inclusion of Creediidae within Syngnathiformes, this result was not replicated by a follow-up study by the same authors, where they noted that such spurious placement was likely the result of low sample quality or contamination (Sanciangco et al., 2016). Therefore this family is not included here. Details regarding sampled taxa can be found in Supplementary Table S1.

2.2. Library preparation, targeted enrichment of UCEs, sequencing, and assembly

Aside from samples loaned as DNA extracts (Supplementary Table S1), we extracted DNA from ethanol-preserved tissues following a modified Qiagen DNeasy protocols (McGee et al., 2016). We quantified all DNA extracts with a Qubit Fluorometer, assessed each extract for quality by electrophoresis, and then sheared extracts by sonication to a target size of 600 bp on a Bioruptor (Diagenode, Inc.). Library preparation and targeted enrichment for UCEs followed protocols described in detail by McGee et al. (2016), which are also available online (http://ultraconserved.org). In brief, we prepared libraries from sheared DNA using commercially available library preparation reagents (Kapa Biosystems), a generic SPRI substitute (Glenn et al., 2016) and custom adapters (Glenn et al., 2016). Following amplification, we grouped libraries into pools of eight based on taxonomic similarity, and we enriched pooled libraries for UCE targets using Mycoarray MYbaits UCE Capture Kits designed to target approximately 1340 highly conserved loci across acanthomorph fishes (McGee et al., 2016). We determined the size of enriched and purified pools with a Bioanalyzer (Agilent Technologies, Inc.) and we quantified enriched libraries by qPCR (Kapa Biosystems) prior to sequencing. We sequenced

libraries using approximately 2/3 of an Illumina NextSeq 500 PE150 run (Georgia Genomics Facility).

We demultiplexed and converted reads for downstream processing using bcl2fastq2 conversion software (v02.14.01.07. Illumina Inc.) and cleaned the reads with illumiprocessor (https://github.com/faircloth-lab/illumiprocessor) (Supplementary Table S2). Then we assembled the reads into contigs with Trinity within the PHYLUCE package (Supplementary Table S3) and matched the assembled contigs to UCE probes (uce-fish-1k-prob e-set/fish-uce-1k-probes.fasta) using the PHYLUCE (Faircloth, 2015) function match_contigs_to_probes.py (Supplementary Table S4). We extracted those contigs that represented UCE loci and prepared these for alignment using PHYLUCE (Faircloth, 2015). UCE raw read and contig data are available from NCBI and NCBI GenBank (BioProject PRJNA378844, SRA SRP103699). Sequence assemblies, along with sequence alignments and phylogenetic trees are available from Dryad http://dx.doi.org/10.5061/dryad.v7f76.

2.3. Alignment, trimming, and matrices

We performed alignments for each UCE using MAFFT with a maximum sequence divergence of 0.2. We trimmed the resulting alignments in gblocks (Castresana, 2000) to remove ambiguous and difficult to align portions of the flanking regions. After alignment and trimming, our dataset consisted of 1180 UCE loci ranging from 93 to 697 bp (mean 289.9, 95% CI 5.47) and a total concatenated alignment length of 342,080 bp (alignment length prior to trimming was 2,056,192 bp). We created two matrices varying in the amount of missing data: 95% and 75% complete matrices (Table 1). The 95% complete matrix was composed of 448 UCEs, each of which contained sequence data for at least 107 of our 113 taxa. The 75% complete matrix was composed of 973 UCEs, each of which contained sequence data for at least 84 taxa. These two matrices differed in their completeness in terms of the percentage of taxa with sequence data for each UCE.

2.4. Phylogenetic inference

2.4.1. Data partitioning

We estimated best-fit partitioning schemes using Partition Finder v1.1.1, which implements search algorithms designed for large phylogenomic datasets (Lanfear et al., 2014). For each matrix, we compared separate analyses using both the strict hierarchical clustering algorithm (hcluster method) and the relaxed hierarchical clustering algorithm (rcluster method), and we kept the best scheme based on the Bayesian Information Criterion (BIC). Only the GTR + G model of molecular evolution was considered for each subset within possible schemes. Due to computational demands

when performing the relaxed clustering method, we checked the top 10% of schemes for the 95% compete matrix and the top 1% of schemes for the 75% compete matrix. It was not feasible to test all possible weighting schemes, so preliminary analyses were run on the 95% and 75% matrices using the hcluster method and one of two weighting schemes: an equal weighting scheme or the default scheme that only weights the rate of the subset. In all cases, the default-weighting scheme was a better fit (lower BIC) and was used thereafter. The best scheme for our 95% matrix had 28 subsets with 1–90 loci per subset (rcluster p = 0.1 BIC = 2,887,106 chosen over hcluster BIC = 2,887,892). The best scheme for our 75% matrix had 12 subsets with 1–191 loci per subset (hcluster BIC = 5,630,897 chosen over rcluster p = 0.01 BIC = 5,633,988).

2.4.2. Bayesian analyses

We used the multi-threaded MPI hybrid variant of ExaBayes v1.4.2 (Aberer et al., 2014) to estimate phylogenies in a Bayesian inference (BI) framework for both our 95% and 75% partitioned alignments. For each matrix, we ran 4 independent runs with 4 coupled chains (3 heated chains) for 1–2 million generations. For each run, we sampled every 500 generations. Convergence was assessed by the average deviation of split frequencies (ADSF), potential scale reduction factor (PSRF), and by visually inspecting the traces and estimated sample sizes (ESS) for all parameters in Tracer v1.6 (Rambaut et al., 2014) for parameters estimated in the individual and combined runs. The 95% analysis converged after 1 million generations (ASDF 0.001%, average PRSF 1.00 ± 0.00, average ESS 3898 ± 896, minimum combined ESS 1902) and the 75% analysis converged after 2 M generations (ASDF 0.00%, average PRSF 1.00 ± 0.00, average ESS 7732 ± 1848, minimum combined ESS 3869). The lowest ESS value for any parameter in any independent run was 401. Consensus trees are presented with posterior probability support summarized from the marginal distribution of trees with a 25% burnin.

2.4.3. Maximum likelihood analyses

Maximum likelihood (ML) phylogenetic analyses for our concatenated matrices were performed with RAxML v8.2.4 (Stamatakis, 2014). We specified a GTRGAMMA substitution model and set *Taractichthys longipinnis* as the outgroup. The best tree was selected from among 20 ML searches, and we separately performed non-parametric bootstrap replicates using the autoMRE criterion, which stopped the bootstrap searches after 60 replicates for all analyses. For both the 95% and 75% matrices, we ran a fully concatenated (single-partition) analysis, as well as a partitioned analysis using the best-fitting partitioning scheme for that matrix, as detailed above.

Because the Pegasidae and Solenostomidae were resolved on long branches, we also reran partitioned ML analyses after removing all pegasid or all solenostomid species to evaluate the effect of long branch artifacts on the topology obtained.

2.4.4. Gene-tree species-tree analyses

Individual gene trees were estimated for each UCE locus in our 95% and 75% complete matrices with RAxML v8.0.19 (Stamatakis, 2014). We specified a GTRGAMMA substitution model and the best-fit tree for each UCE was selected from among 20 ML searches. 400 bootstrap replicates were performed using RAxML for each UCE in the 95% complete matrix and 300 bootstraps were performed for each UCE in the 75% complete matrix. We carried out species tree analyses with ASTRAL v4.7.8 using the heuristic algorithm due to the large number of taxa in our dataset (Mirarab et al., 2014). We performed site + gene multi-locus bootstrapping in ASTRAL (300 ASTRAL replicates for the 95% complete matrix and 200 replicates for the 75% complete matrix). We report the greedy

Table 1
Information about the two alignment matrices used in this study. Percentage (95% and 75%) refers to the taxon coverage for each UCE in the matrix. See text for further explanation.

	95% Matrix	75% Matrix
Min. taxa per locus	107	84
Alignment length (bp)	149,710	285,228
Variable sites (%)	48.7	50.2
Parsimony informative sites (%)	37.5	38.5
Missing data (%)	6.04	9.73
CG content (%)	46.1	46.6
Loci (#)	448	973
Min. length of loci (bp)	149	102
Max. length of loci (bp)	636	697
Mean length of loci (95% CI)	334.17 (8.70)	293.14 (6.09)
Subsets	28	12

majority rule consensus tree calculated from among the bootstrap-replicate species trees output by ASTRAL.

ASTRAL uses a quartet-based approach to estimate a species tree from a set of gene trees under the assumption that all gene trees are correct. However, because of the conserved core regions, individual UCE loci can have relatively low phylogenetic informativeness (Gilbert et al., 2015), which may result in high gene tree error. To partially address this issue, we repeated ASTRAL analyses on subsets of our datasests with the most phylogenetically informative loci, using the number of parsimony informative sites as a proxy for phylogenetic informativeness (Meiklejohn et al., 2016). For each level of matrix coverage (95% and 75%), we calculated the number of parsimony phylogenetically informative sites for each locus, and then we selected all loci in the upper quartile (upper 25%) using the AMAS software package (Borowiec, 2016). For the 95% dataset, this reduced the number of loci from 448 to 116. For the 75% dataset, this reduced the number of loci from 973 to 250. We then estimated consensus gene tree species trees again in ASTRAL for each dataset of informative loci using the same methods as for the full datasets.

For each consensus topology, we used ASTRAL to calculate the normalized quartet score from the best-fit gene trees, which is interpreted as the percent of quartet trees induced by the best-fit gene trees that are present in the topology of interest (Mirarab et al., 2014). ASTRAL runs on unrooted trees, so we manually rooted trees with *Taractichthys longipinnis* for visualization.

2.4.5. Compositional bias

Base compositional bias can produce inconsistent phylogenetic results and phylogenomic data sets may be especially vulnerable to this source of bias (Dávalos and Perkins, 2008; Foster and Hickey, 1999; Jermiin et al., 2004; Phillips et al., 2004; Rodríguez-Ezpeleta et al., 2007). We performed a Chi-squared test of homogeneity of state frequencies across taxa in Paup* v4.0a147 for Macintosh (Swofford, 2002). GC content for each species was calculated as a percentage of the total alignment length (not including missing data or indels) for each species (Supplementary Table S5). Average base counts were also calculated for each of the nine syngnathiform families and tested for homogeneity using the chisq.test function in R (Supplementary Fig. S4; Supplementary Table S6).

To determine if heterogeneous base frequencies could be leading to the artificial grouping of sequences based on similar base frequencies, we performed ML analyses of "RY-coded" matrices (Woese et al., 1991). While this approach does not explicitly model base heterogeneity across the tree, it normalizes base frequencies by coding bases as either purine or pyrimidine and has been shown to outperform ML analyses with homogenous approaches in the presence of parallel (non-homologous) compositional heterogeneity (Ishikawa et al., 2012).

We converted the concatenated 95% and 75% complete alignments into binary matrices composed of 0's (purines) and 1's (pyrimidines). We then repeated ML analyses of these RY-coded matrices in RAxML v8.2.4 (Stamatakis, 2014) using the BINGAMMA model for binary characters. As before, we performed 20 independent searches for the best-fit tree and nonparametric bootstrapping with the autoMRE cutoff, which stopped bootstrapping after 60 iterations.

3. Results

3.1. Bayesian and maximum likelihood phylogenetic analyses

The consensus tree from the Bayesian analysis of the partitioned 95% matrix (Fig. 2) was fully supported at all nodes (posterior probability = 1) and topologically identical to the best-fit topology for both unpartitioned (not shown) and partitioned ML analyses using the same matrix (Supplementary Fig. S1). In the partitioned ML analysis for this dataset (95%), we resolved most relationships with high bootstrap support (≥92), although there were exceptions (Fig. 2). The clade containing Fistulariidae, Aulostomidae, and Centriscidae had low support (bootstrap = 53), and relationships between the callionymid genera *Repomucenus*, *Draculo, Diplogrammus*, and *Calliurichthys* were not all fully supported (bootstraps ranging from 77 to 100). Within syngnathids, the clade containing seadragaons (Fig. 2, red clade) and messmate pipefishes (Fig. 2, blue clade) was only resolved with moderate support (bootstrap = 72).

The Bayesian consensus and best-fit ML trees estimated from the analyses of the partitioned 75% complete matrix were fully congruent with one another and generally recovered syngnathiform relationships with high support (Supplementary Figs. S2–3). This topology was also identical to that estimated from ML analysis of the concatenated unpartitioned 75% matrix (not shown). The topology from the 75% analyses was also congruent with those from the 95% analyses, except for the placement of the seadragon clade (Fig. 2 insert, red clade). Relative support for nodes across the 75% phylogenetic analyses mirrored the 95% analyses. For instance, the bootstrap support for the placement of the seadragon clade was only moderate (bootstrap = 67) for the 75% partitioned ML tree, but was fully supported in the 75% BI consensus tree. The callionymid genera *Repomucenus*, *Draculo, Diplogrammus*, and *Calliurichthys* were resolved with better support in analyses of the 75% complete matrix (e.g., bootstraps 90–100 for the 75% partitioned ML analysis) compared to the 95% matrix, although these genera were still polyphyletic.

Because the Pegasidae and Solenostomidae were on long branches, we reran partitioned ML analyses after removing both pegasid species or both solenostomid species to evaluate the effect of long branch artifacts on the estimated topology. While the removal of the Pegasidae did not change the topology depicted in Fig. 2 (not shown), omitting the Solenostomidae did change the topological arrangement of the families Centriscidae, Aulostomidae, Fistulariidae, and Syngnathidae (Fig. 3). Fistulariidae and Aulostomidae were inferred as the closest relative of the Syngnathidae instead of the Centriscidae but with very low support at a node subtended by an extremely short branch (bootstrap = 55).

3.2. Gene-tree species tree

For both the 95% and 75% complete matrices, we used ASTRAL to estimate consensus species trees from bootstrapped sets of gene trees (Fig. 4A, Supplementary Fig. S4A). These represent estimates of the true species tree topology that take into account individual locus histories and gene tree uncertainty. The consensus topology using all loci in the 95% matrix supported a clade of long-snouted syngnathiforms as in the ML/BI topologies (e.g., compare Fig. 4A to Fig. 2), but this clade had lower support in the consensus species tree (MLBS = 58). While the benthic lineages Callionymidae, Pegasidae, Mullidae, and Dactylopteridae form a clade sister to traditional long-snouted syngnathiform families in all BI/ML topologies, the consensus species tree for the full 95% matrix nests the traditional syngnathiforms within the benthic lineages. However, the exact arrangement of lineages is unclear given the low support for the sister-group relationship between Callionymidae and the long-snouted clade (MLBS = 47).

The lack of support for the divergences among major syngnathiform lineages was exacerbated when we analyzed the larger, less complete dataset, the 75% complete matrix with all loci included (Supplementary Fig. S4A). Callionymidae, Pegasidae, the clade including Centriscidae, Aulostomidae, Fistulariidae, and the clade
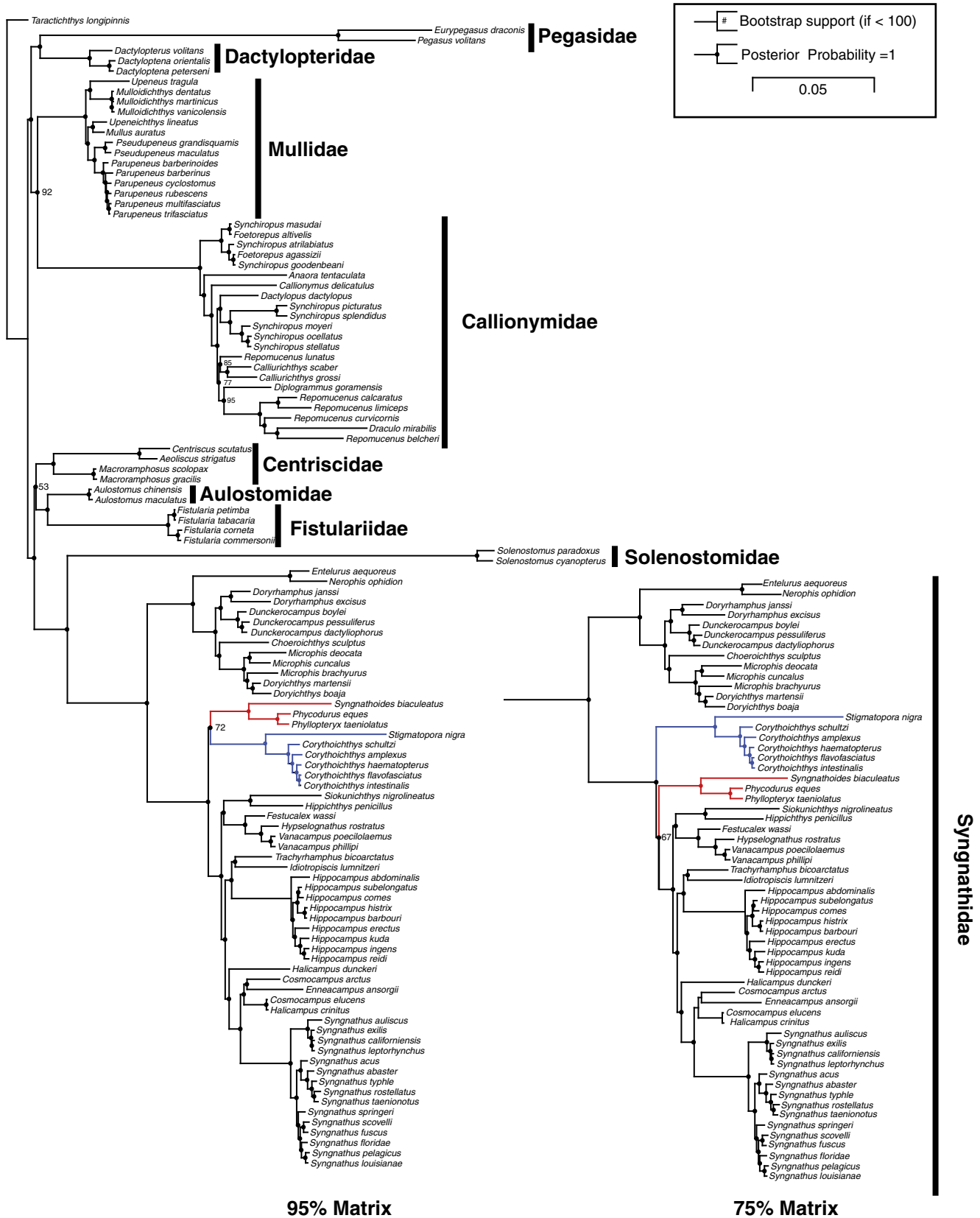
Fig. 2. Consensus phylogeny from the Bayesian analysis of the partitioned 95% matrix. All nodes were resolved with a posterior probability of 1 (black circles). Maximum likelihood analysis recovered the same topology; only bootstrap support values less than 100% are denoted. Insert: The Bayesian and maximum likelihood analyses for the 75% matrix recovered different relationships among the clades including seadragons (red) and messmate pipefishes (blue) and other tail-brooding pipefishes compared to the 95% matrix.
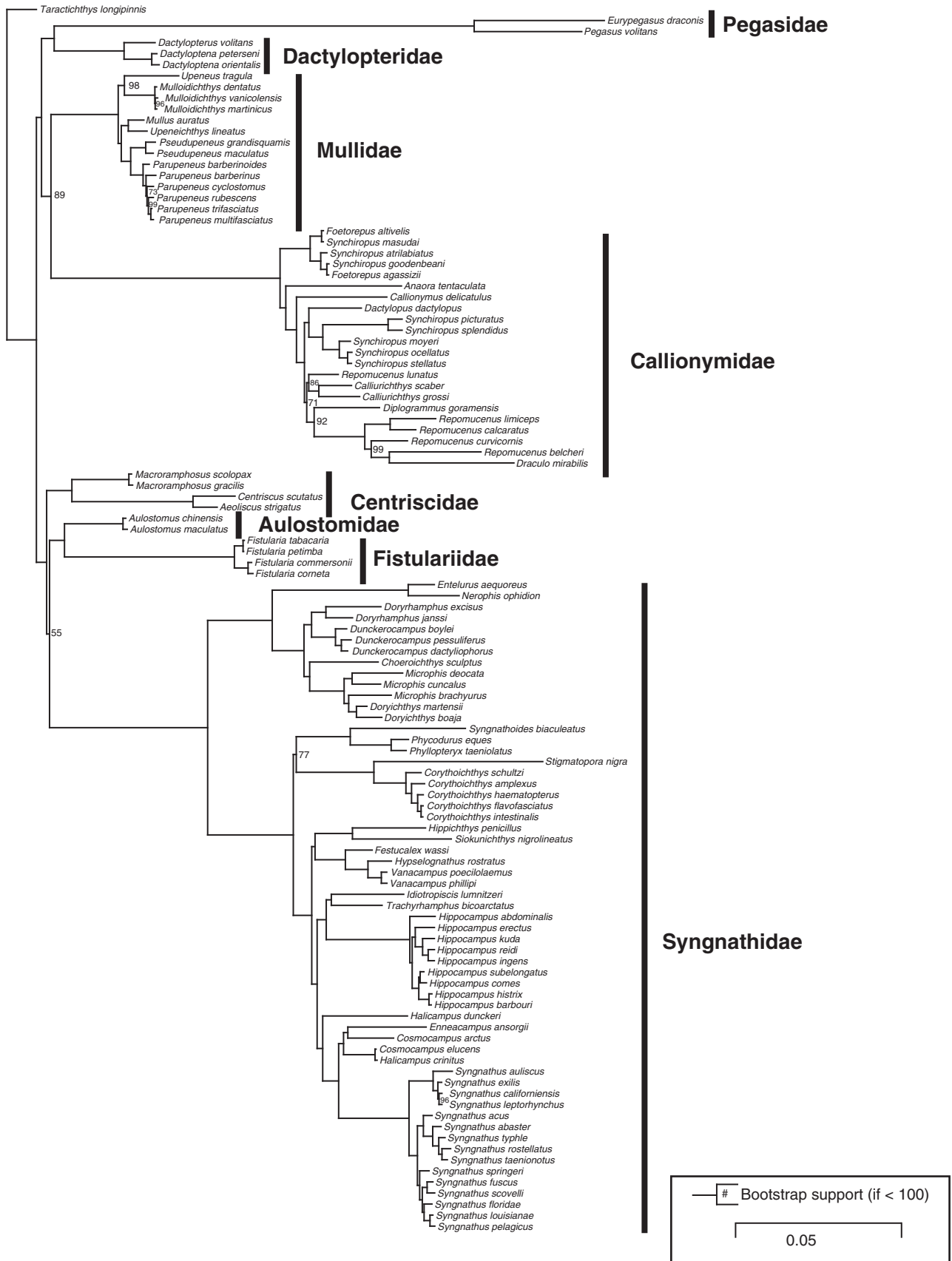
Fig. 3. Best-fit phylogeny from the RAxML analysis of the partitioned 95% matrix with Solenostomidae excluded. Bootstrap support values less than 100% are denoted. Only relationships among Syngnathidae, Centriscidae, and Aulostomidae plus Fistulariidae were altered (compare with Fig. 2). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
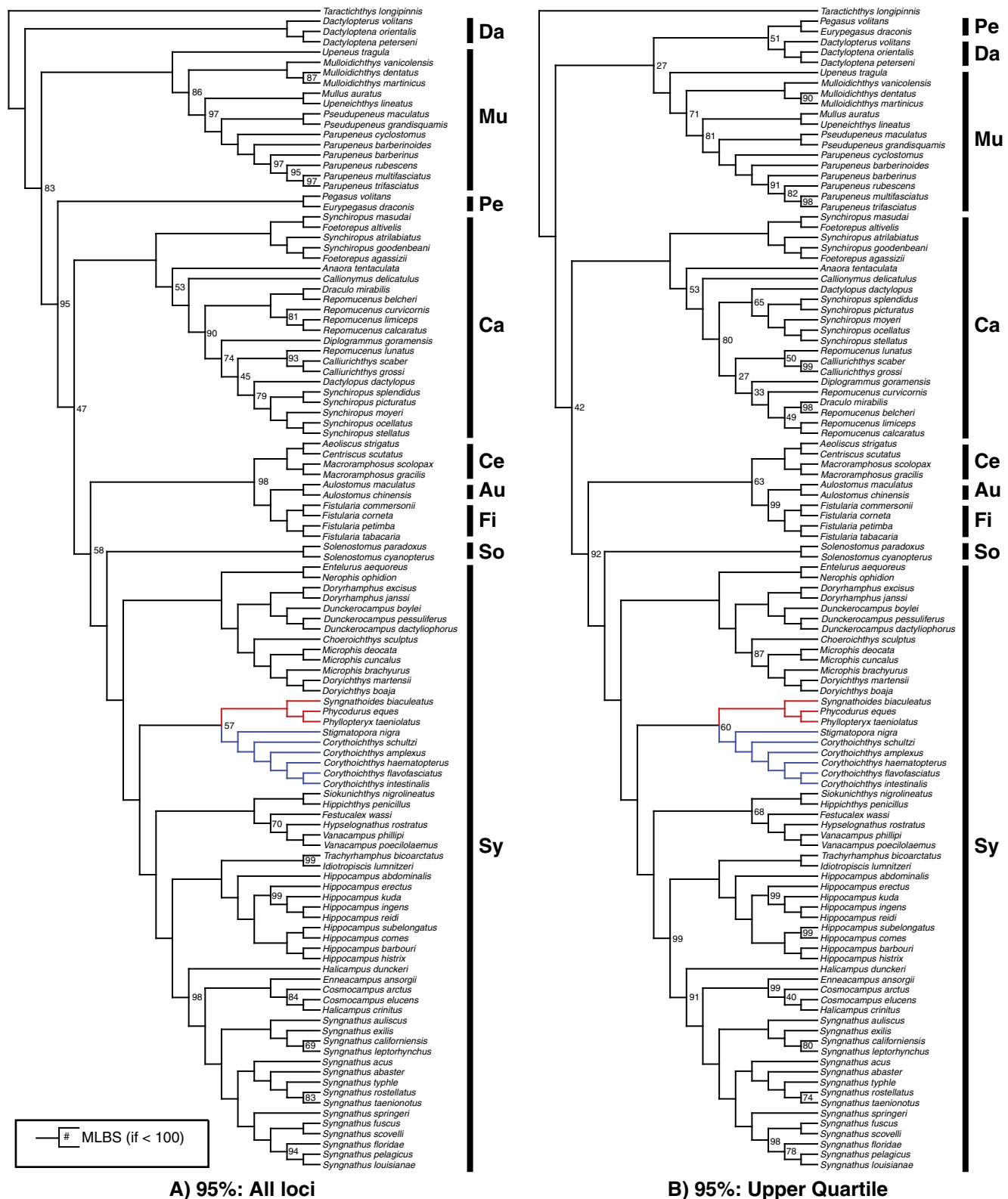
Fig. 4. Cladograms representing the greedy consensus topologies from the ASTRAL gene-tree species-tree analyses for the 95% complete matrix. We obtained different topologies when the analysis was run on all loci (left) or only the most informative loci (loci in the upper quartile for parsimony informative sites; right). Nodes are labeled when multi-locus bootstrap support (MLBS) was less than 100%. The placement of the seadragon (red) and messmate pipefish (blue) clades are highlighted for comparison with other topologies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

including Solenostomidae and Syngnathidae are essentially a polytomy in the consensus species tree for the full 75% dataset (MLBS = 31–46). Overall, global support for the consensus species trees from our ASTRAL analysis on the full 75% dataset (normalized quartet score = 0.849) was lower than that for the full 95% dataset (normalized quartet score = 0.861).

Because trimmed UCEs are relatively short loci with portions of highly conserved sequence, many of the input gene trees into ASTRAL analyses are poorly resolved. For instance, average bootstrap support for an individual gene trees in the 95% dataset were as low as 19.2 (range 19.2–68.3; mean 43.9 ± 9.5). This may be problematic, because ASTRAL assumes that all input gene trees are known without error (Mirarab et al., 2014). To reduce the amount of noise in our input loci, we also ran ASTRAL analyses with matrices filtered to contain informative loci, or only loci in the upper quartile for number of parsimony informative sites (Fig. 4B, Supplementary Fig. S4B).

Overall, using datasets of informative loci resulted in an increase in global support: the normalized quartet score for the consensus topology for the 95% analysis using informative loci was 0.887 and 0.883 for the 75% analysis using informative loci. Both analyses filtered for informative loci were resolved with high support for the monophyly of the long-snouted syngnathiform families (95% MLBS = 92, Fig. 4B; 75% MLBS = 100, Supplementary Fig. S4B), and relationships among the long-snouted families agreed with those from the ML/BI analyses. In particular, Centriscidae, Aulostomidae, and Fistulariidae formed a clade sister to Solenostomidae plus Syngnathidae. Callionymidae was resolved as the sister group to the long-snouted syngnathiforms in both analyses using informative loci, but with low support (95% MLBS = 42, 75% MLBS = 54), and the arrangement of Dactylopteridae, Mullidae, and Pegasidae were not consistent between analyses.

Overall, within-family relationships between species were well-resolved in the ASTRAL trees and were largely congruent with the BI/ML topologies. Still, some species-level relationships within Mullidae, Callionymidae, and Syngnathidae differ from or have lower replicate support compared to the BI/ML analyses. Often these discrepancies correspond to areas with extremely short internode branch lengths in the BI/ML topologies (e.g., relationships among *Mulloidichthys* species). Although MLBS support is not high in the ASTRAL topologies (ranging from 46–60), the placement of the seadragon clade (Fig. 4, red clade) consistently agrees with the partitioned analyses of the 95% complete matrix (Fig. 2, red clade).

3.3. Compositional bias

Chi-square tests indicated that bases deviated from homogeneity across taxa (95% concatenated alignment: $\chi^2 = 2193$, df = 336, p = 0.00; 75% concatenated alignment: $\chi^2 = 3827$, df = 336, p = 0.00). Average GC composition for the 95% matrix across species was 46.1 ± 61.2% with a skew of 1.43 (75%: 46.6 ± 60.4%, 0.824 skew). Average GC content in the 95% matrix at the family level was 46.18 ± 1.15% skew 1.64 (Supplementary Table S6). Base composition within Fistulariidae, Solenostomidae, Pegasidae, and Dactylopteridae did not deviate from homogeneity, while other families did at the 0.05 significance level (Supplementary Table S6; Supplementary Fig. S5). Homogeneity across families was also rejected (95% concatenated alignment: $\chi^2 = 571.6$, df = 24, p = 0.00), even with the removal of the Aulostomidae and Solenostomidae (95% concatenated alignment: $\chi^2 = 64.85$, df = 18, p = 0.00).

We compared the topologies from ML analyses of the RY-coded concatenated alignments to the ML topologies on the original concatenated alignments to assess whether nonhomogeneous base compositions were potentially influencing relationships. For both the 95% and 75% matrices, the RY-coded ML topologies were largely consistent with those obtained from the analyses with all four bases. Focusing first on deeper relationships, the best-fit trees from the RY-coded analyses placed Centriscidae sister to all other long-snouted syngnathiforms (Fig. 5) instead of sister to Aulostomidae

+ Fistulariidae. Within families, very few species level relationships differed between RY-coded analyses and analyses on the original alignments. Overall, the RY-coded topologies had lower bootstrap support at many nodes compared to topologies from the original alignments, and the bootstrap support for the placement of benthic associated lineages (Callionymidae, Mullidae, and Pegasidae + Dactylopteridae) was particularly low in the 95% complete RY-coded analysis (40–53 bootstrap support). This could indicate that similarities in base frequencies are artificially inflating support values at some nodes. Conversely, lower support values across the tree could be a byproduct of the decreased informativeness of the RY-coded matrices compared to those with all four bases represented.

4. Discussion

4.1. Evaluation of topological support and potential sources of error

In this study, we use a phylogenomic dataset of ultraconserved elements (UCEs) from a large sample of taxa to reevaluate historically problematic relationships among a morphologically diverse group of fishes, Syngnathiformes, which diverged from a common ancestor approximately 80 million years ago (Betancur-R et al., 2013a; Near et al., 2012; Sanciangco et al., 2016). From the posterior probabilities in Fig. 2, it would appear that the power of phylogenomic-scale data has confidently resolved syngnathiform relationships spanning from recent to deep divergences. These results are particularly striking, considering that many other studies using phylogenomic data still fail to resolve some relationships among deep and rapid divergences, including studies of fishes spanning similar clade ages (e.g., Eytan et al., 2015) and using mitogenomes for Syngnathiformes (Song et al., 2014). It is possible that the extremely high support in our study is due to the superior phylogenetic information contained within UCE loci compared to coding loci (Gilbert et al., 2015). However, large datasets can potentially lead to the inference of highly supported, but incorrect, topologies due to nonrandom or systematic sources of error (Jeffroy et al., 2006; Kumar et al., 2012; Lemmon and Lemmon, 2013; Nishihara et al., 2007; Philippe et al., 2011; Swofford et al., 1996). Systematic error is not a new issue for phylogenetic analysis, but has become more prominent with the advent of genome-scale studies, because systematic biases can compound with increasing dataset size (Felsenstein, 1978). It is therefore common practice to assess potential sources of systematic error in phylogenomic datasets such as compositional heterogeneity, rate differences among genes, long-branch attraction, and gene-tree discordance; and to look for topological convergence across analyses (Cummings and Meyer, 2005; Eytan et al., 2015; Jeffroy et al., 2006; Kumar et al., 2012; Meiklejohn et al., 2016; Nishihara et al., 2007; Phillips et al., 2004; Steinke et al., 2006).

We do not find strong evidence that either long branches or base compositional heterogeneity are inducing topological artifacts at the deep divergences among syngnathiform families. While we rejected chi-square tests of homogeneity of base composition across species (Supplementary Table S6), and families in our analyses have relatively different GC composition means and ranges (Supplementary Fig. S5), compositional heterogeneity is explained in part by the known shared evolutionary history within each lineage and stark variation in species diversity across the syngnathiform families. Both are factors ignored by standard homogeneity tests. It is also not surprising that homogeneity is easily rejected even when differences in base proportions among lineages may not seem large, given the taxonomic scope and extremely great number of sites in our analysis (149,610–285,228 sites). Fortunately, there is no obvious pattern between base composition
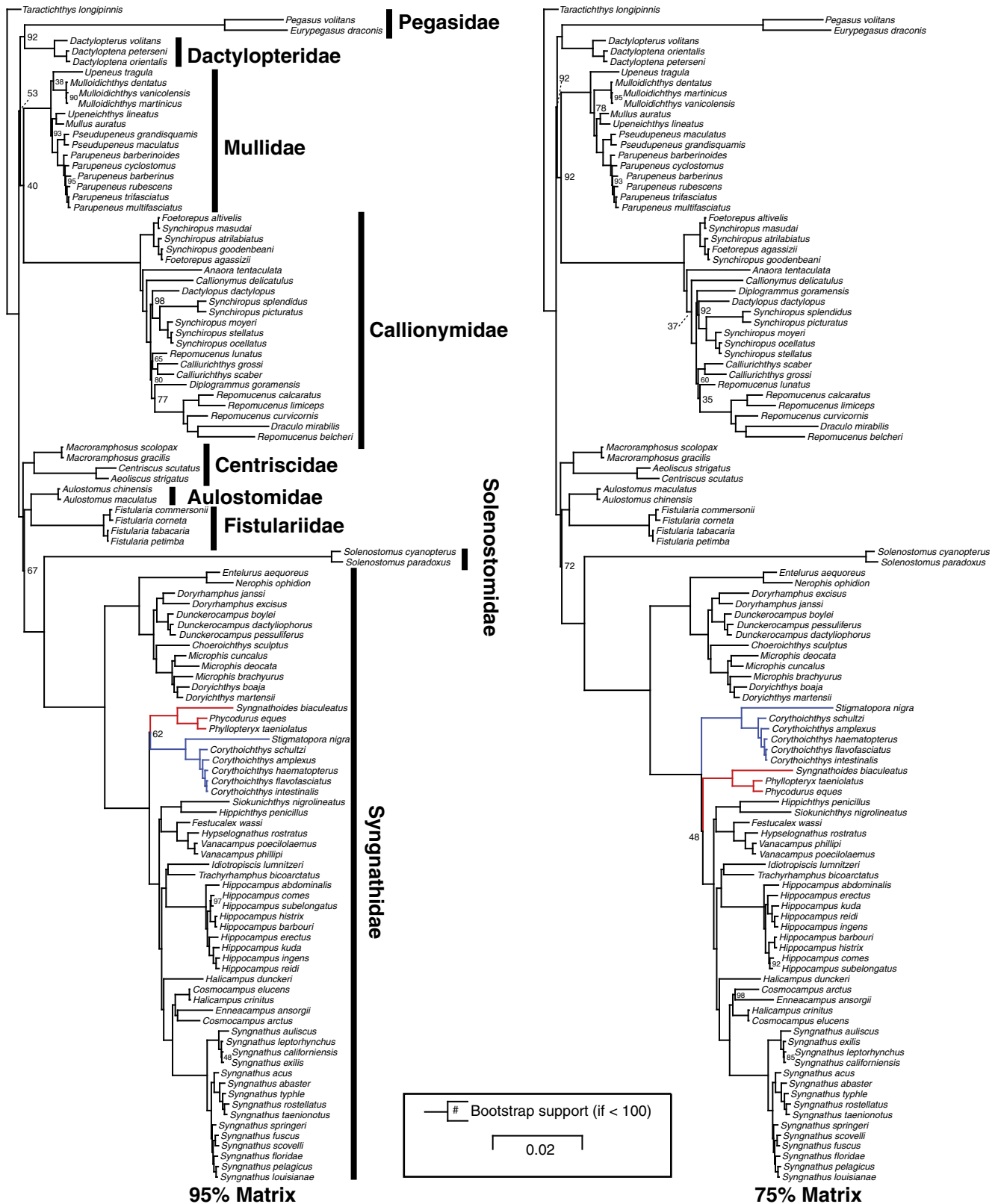
Fig. 5. Best-fit phylogenies from the RAxML analyses of the RY-coded matrices (95% on the left, 75% on the right). Bootstrap support values less than 100% are denoted. Only relationships among Syngnathidae, Centriscidae, and Aulostomidae plus Fistulariidae were altered (compare with Fig. 2).

and phylogenetic placement of the major syngnathiform lineages. For instance, three out of the four pairs of sister families in our BI/ML analyses (Pegasidae and Dactylopteridae, Solenostomidae and Syngnathidae, Fistulariidae and Aulostomidae) do not have overlapping GC composition ranges (Supplementary Fig. S5). Corroborating this observation, RY-coded analyses did not dramatically alter the best-fit topology compared to analyses with all four bases represented (Fig. 5).

Instead of biasing topology, violation of base heterogeneity assumed by our models of sequence evolution may be affecting branch lengths or bootstrap support. In some cases, species with extreme GC proportions compared to the rest of their family corresponded to tips with unusually long branches (e.g., *Stigmatopora nigra* in Syngnathidae and *Draculo mirabilis* in Callionymidae; Fig. 1 and Supplementary Fig. S5). Likewise, the comparatively high GC composition of solenostomids relative to other taxa may account for the long branch inferred for this lineage (Fig. 2, Supplementary Fig. S5). However, *P. cyclostomus* appears to be an outlier compared other goatfishes but has a very short branch (Fig. 2). We note that the RY-coded analyses showed relatively lower bootstrap support at many nodes compared to topologies from the original alignments. This could indicate that similarities in base frequencies are artificially inflating support values at some nodes, but it is also possible that lower support values are simply due to decreased informativeness of the RY-coded matrices compared to unaltered matrices.

Long branch attraction is a well-documented phenomenon that can produce incorrect topologies (Felsenstein, 1978). Despite the presence of several long branches at the base of major lineages such as Pegasidae and Solenostomidae, these clades did not tend to group together in the Bayesian or likelihood analyses (Fig. 2, Supplementary Fig. S1–3). Removing the Pegasidae had no effect on topology, but the exclusion of Solenostomidae changed relationships among Centriscidae, Fistulariidae, Aulostomidae, Solenostomidae and Syngnathidae (Fig. 2). Nevertheless, changes in topology after the removal of Solenostomidae do not indicate that its long subtending branch was causing long-branch attraction artifacts: no particularly long branches shifted away from this portion of the tree after removal, and bootstrap support for the shifted placement of Aulostomidae plus Fistulariidae was not improved. Instead, changes in topology after the exclusion of certain lineages could indicate that relationships are sensitive to taxonomic sampling (Heath et al., 2008). The removal of Solenostomidae rearranged nearby relationships separated by extremely short internal branch lengths, which are likely to be most sensitive to changes in the data matrix. As we describe in more detail below, the inclusion of Solenostomidae may be important for informing relationships among what was an apparently old and rapid radiation. While there is not a straightforward case for long-branch attraction, long branches can be symptomatic of and coincident with other phylogenetic challenges, including compositional heterogeneity, lineage specific rate heterogeneity and heterotachy, all of which can mislead phylogenetic inference (Kumar et al., 2012; Rothfels et al., 2012; Whitfield and Kjer, 2008).

Concatenated datasets can result in the inference of incorrect topologies by ignoring conflicting phylogenetic signal among loci. One common source of discordance among loci in vertebrate taxa is incomplete lineage sorting or ILS (Maddison, 1997). ILS results in conflict between some gene trees and the true species tree and a coalescent approach should more accurately recover the true species tree when there is a moderate to large degree of ILS (Mirarab and Warnow, 2015; Mirarab et al., 2014). Given the large number of UCE loci in our datasets and the relatively short internal branch lengths between speciation events leading to almost all families in our dataset (Fig. 1), we used a coalescent-consistent method (ASTRAL) to account for the potential negative impact(s) of ILS.

The consensus species trees we obtained from our ASTRAL analyses differed from our BI/ML topology in a number of ways, including major rearrangements of the relationships among the benthic-associated lineages with respect to each other and to the clade of traditional syngnathiform lineages (Fig. 4, Supplementary Fig. S4). For instance, while the BI/ML analyses resolved long-snouted syngnathiforms as sister to a monophyletic group of all the benthic-associated families (Callionymidae, Dactylopteridae, Mullidae, and Pegasidae), this arrangement was not recovered in any of the ASTRAL trees. Instead, the long-snouted syngnathiforms were usually nested as a clade within the benthic associated lineages in the ASTRAL topologies, although the actual ordering of benthic clades was not consistent and poorly supported.

Both shallow and deep topological discrepancies between the ASTRAL topology and BI/ML topology often correspond to areas with extremely short internode branch lengths in the BI/ML topologies (e.g., relationships among *Mulloidichthys* species). All else being equal, we would expect these areas of rapid diversification to have higher rates of ILS, and therefore be more likely to differ from concatenated analyses. While discordance (low nodal support) can be due to actual historical events such as ILS, it can be falsely inflated by violations of the assumptions of the method, such as assuming that input gene trees are correct. We believe this to be an important issue for our UCE dataset because gene trees from short, highly conserved loci often produced poorly-resolved gene trees. Additionally, poorly resolved gene trees may be more sensitive to sources of systematic error, such as base compositional bias, compared to multi-locus alignments (Betancur-R et al., 2013b). We therefore ran additional ASTRAL analyses that were filtered to include more informative loci (loci in the upper quartile for number of parsimony informative sites). This approach does not solve the issue of low gene tree accuracy, but it does remove the most spurious gene trees from the dataset being analyzed. Using more informative loci increased support for some relationships and decreased support at others (Fig. 4, Supplementary Fig. S4). Most notably, the analyses with informative loci resulted in higher support for monophyletic long-snouted syngnathiforms in both matrices. Callionymids also emerged as sister clade to this group in both analyses, though with low support. Overall, focusing only on the most informative loci (loci in the upper quartile of parsimony informative sites versus all loci) increased global topological support, although relationships among benthic associated lineages generally remained inconclusive, as summarized in Fig. 6. Low support at these nodes could reflect true discordance due to historical events (e.g., ILS) or insufficient phylogenetic information in our loci, but it is reassuring that incongruent relationships between analyses on informative loci are reflected by low node support (Fig. 4B, Supplementary Fig. 4B).

4.2. Evolutionary relationships among long-snouted Syngnathiformes

While there are some striking differences between the topologies we present, they agree on a number of fundamental points relating to controversial relationships among syngnathiform lineages. Fig. 4 summarizes the two topological arrangements of major syngnathiform lineages we have described and their various measures of support. Across all analyses, we resolved traditional, long-snouted syngnathiform fishes (Centriscidae, Aulostomidae, Fistulariidae, Solenostomidae, and Syngnathidae) as a monophyletic group, sharing a more recent common ancestor with each other than they do with any of the benthic syngnathiform lineages (Dactylopteridae, Pegasidae, Mullidae, Callionymidae). This node was fully supported in the BI/ML analyses (posterior probability = 1, bootstrap = 100) and the gene-tree species-tree analyses filtered for informative loci (95% MLSB = 92, 75% MLBS = 100). With the exception of Near et al. (2013), previous large-scale molecular analyses nested benthic-associated lineages among traditional syngnathiforms (Betancur-R et al., 2013a; Dettai and Lecointre, 2005; Kawahara et al., 2008; Smith and Craig, 2007; Near et al., 2012; Sanciangco et al., 2016; Song et al., 2014). Our finding of a monophyletic long-snouted syngnathiform group is easier to reconcile with the morphological literature where a number of key features, including aspects of the snout (e.g., Keivany and Nelson,
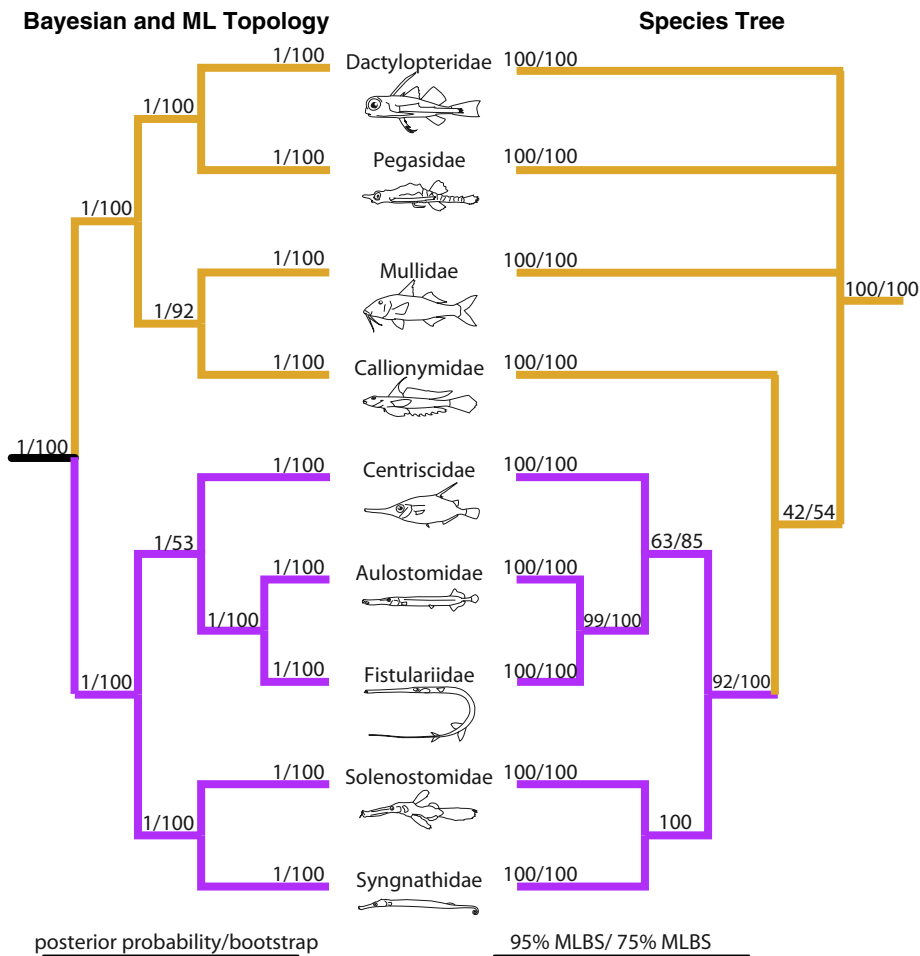
Fig. 6. These cladograms compare the two main topologies of evolutionary relationships among major syngnathiform lineages produced in this study. On the left are relationships recovered by all Bayesian (BI) and maximum likelihood (ML) analyses that included all lineages. The BI posterior probability followed by the ML bootstrap support is labeled above each branch. The topology on the right summarizes relationships from the ASTRAL gene-tree species-tree analyses using datasets of informative loci. Relationships that were inconsistent between analyses (and had low support) were collapsed into polytomies. Site-only multi-locus bootstrap (MLBS) support is included above each branch. Relationships among traditional long-snouted syngnathiform families are highlighted by purple branches while the benthic-associated lineages are in orange.

2006), suggest that these lineages should remain monophyletic despite the shift in their global position among all fishes. Nevertheless, Pegasidae was previously grouped with Syngnathiformes based on morphological characteristics (Nelson, 2006; Pietsch, 1978), but we do not find evidence that they are more closely related to traditional syngnathiforms than the other benthic lineages (Fig. 5). Instead, Pegasidae may be most closely related to the Dactylopteridae, which has been proposed based on morphology (Pietsch, 1978), although this relationship is not reflected in our consensus species trees (Fig. 4, Supplementary Fig. 4).

Relationships among the long-snouted syngnathiform families Aulostomidae, Fistulariidae, Solenostomidae, and Syngnathidae were also generally consistent across analyses using UCEs, although the relative placement of Centriscidae shifted depending on matrix and inferential approach. Aulostomidae plus Fistulariidae and Solenostomidae plus Syngnathidae were each fully supported clades (posterior probability = 1, bootstrap = 100, MLBS = 100). Despite short internal branch lengths in this area of the tree (as small as 0.0007 units for the 95% partitioned ExaBayes Tree; Fig. 2), the Bayesian analyses placed the Centriscidae sister to Aulostimidae plus Fistuliidae with high confidence (posterior probability = 1). While the best trees from our maximum likelihood analyses also recovered this clade with Centriscidae, Fistulariidae, and Aulostomidae, there was only moderate support at

this node (bootstrap 53–65 depending on matrix completeness). Likewise, ASTRAL analyses consistently yielded a clade composed of Centriscidae, Fistulariidae, and Aulostomidae, but the support at this node was only moderate for matrices enriched for informative loci (MLBS = 63–85). There are cases in which Bayesian posterior probabilities can be inflated and topological uncertainty may be better represented by the more conservative nonparametric bootstraps (Alfaro et al., 2003; Cummings et al., 2003; Kumar et al., 2012; Simmons et al., 2004), so the RAxML and ASTRAL trees may better reflect uncertainty in relationships among Centriscidae, Fistulariidae, and Aulostomidae. It is also important to note that Centriscidae, Aulostomidae, and Fistulariidae did not form a clade in the RY-coded analyses (Fig. 5), indicating that some inflated support for this clade could be caused by similarity in base composition, although the placement of Aulostomidae and Fistulariidae relative to other syngnathiforms was only moderately supported (bootstrap = 67–72).

Placement of Centriscidae, Aulostomidae, and Fistulariidae relative to other syngnathiforms has been historically inconsistent and poorly supported (Betancur-R et al., 2013a; Near et al., 2012; Sanciangco et al., 2016; Song et al., 2014). The short internal branch lengths in this region of the tree (e.g., Fig. 2) suggest that these lineages rapidly diverged from one another in deep time, which is an evolutionary scenario that can be difficult to

reconstruct. One way to diagnose an ancient rapid radiation is to look for topological incongruence with consistently low support among analyses and sensitivity to such factors as changes in taxonomic representation (Whitfield and Kjer, 2008; Whitfield and Lockhart, 2007). Here, removal of a long-branch (Solenostomidae) changed the relationships among the long-snouted families, even though short branch lengths and low bootstrap support remain a problem in this area of the phylogeny (Fig. 2). Changes in topology after the removal of certain lineages could indicate that syngnathiform relationships are sensitive to taxonomic sampling artifacts, such as the inclusion of Solenostomidae. Indeed, the removal of Solenostomidae caused the topology of long-snouted syngnathiform taxa to converge on that published by Near et al. (2013), which did not sample a representative of this family.

4.3. Uncertain placement of benthic-associated lineages

While we argue that the overarching signal across our analyses supports that long-snouted syngnathiforms form a clade, the relationship between this clade and the other syngnathiform families (Callionymidae, Dactylopteridae, Mullidae, and Pegasidae) remains ambiguous (Fig. 6). Note that in the BI/ML topologies benthic-associated lineages always formed a monophyletic group, while the ASTRAL analyses tended to nest the long-snouted clade within the benthic associated lineages (albeit with low support). The discrepancies between these two scenarios are significant and have important implications for our understanding of syngnathiform morphological evolution and diversification. Future UCE studies that sample a larger diversity of percomorph outgroup taxa may help alleviate some uncertainty in this area of the syngnathiform tree.

4.4. Resolution of species-level relationships

Within-lineage relationships were highly stable across analyses and remarkably well-resolved. As shown in Fig. 6, all nine syngnathiform families were fully supported monophyletic groups (posterior probability = 1, bootstrap = 100, MLBS = 100). Most genera were also monophyletic, with the exception of those in Callionymidae where many were polyphyletic (Figs. 2–5). Issues within callionymids may reflect differing opinions about generic definitions among morphological taxonomists (Nelson, 2006), and these problems highlight the need for additional work on this phylogenetically understudied group of fishes. Within the Syngnathidae, the genus *Microphis* was paraphyletic with respect to the closely related *Dorichthys* (Dawson, 1981, 1984). In agreement with a recently published syngnathid phylogeny (Hamilton et al., 2016), *Cosmocampus* and *Halicampus* were both polyphyletic. It appears that a number of syngnathid genera will require revisiting as more species are included in future phylogenetic analyses with increased taxonomic sampling (Hamilton et al., 2016).

The evolutionary relationships among syngnathid species reported in this study largely agree with recently published phylogenetic analyses (Hamilton et al., 2016; Mwale et al., 2013; Teske and Beheregaray, 2009; Wilson and Rouse, 2010). One of the major dichotomies within syngnathids is male brood pouch location: on the trunk region (trunk-brooder) or on the tail region (tail-brooder) (Hamilton et al., 2016; Wilson and Orr, 2011; Wilson et al., 2001). It has been previously shown that each type of brooder forms a monophyletic group, with the exception of the trunk-brooder, *Syngnathoides biaculeatus*, which shares a more recent common ancestor with tail-brooding pipefishes (Hamilton et al., 2016; Wilson and Rouse, 2010). In this study, we sampled an additional genus of trunk brooding pipefish (*Doryichthys*) that has not been included in prior molecular syngnathid phylogenies. *Doryichthys* was resolved, as expected, among the trunk brooding

clade (e.g., Fig. 2). Overall, we verify that all trunk-brooding pipefish species (except *S. biaculeatus*) are monophyletic and sister to all tail-brooding pipefishes, which was consistent across every analysis in this study with high support. These two brooding strategies are apparently highly conserved within syngnathids (Hamilton et al., 2016).

The evolutionary relationships of the unusual and charismatic seadragons (*Phycodurus eques* and *Phyllopteryx taeniolatus*) have received specific attention in the past (Wilson and Rouse, 2010). We find high support for a close relationship between *S. biaculeatus* and seadragons (red clade in Figs. 2–5). This does not agree with the previously published phylogeny using mitochondrial data (Wilson and Rouse, 2010), but is consistent with the results using nuclear loci in Hamilton et al. (2016). However, the placement of the clade including seadragons and *S. biaculeatus* within Syngnathidae was problematic. Their relationship to other tail-brooding pipefishes was the only topological rearrangement that occurred when matrix completeness was altered (Fig. 2, red and blue clades) and this node had low bootstrap and MLSB support across analyses (Figs. 2–5). The addition of more taxa in future UCE analyses may help stabilize the placement of this early-branching lineage within tail-brooding syngnathids. For instance, in a recent study that sampled many more genera, including close relatives of seadragons (*Solegnathus* pipe horses; Hamilton et al. (2016)), the messmate pipefish clade (*Corythoichthys* and *Stigmatopora*) was the earliest branching tail-brooding lineage as in the 75% analysis in Fig. 2 (blue clade) .

Because of their specialized brood pouches, upright posture, and extreme tail-prehension, the placement of seahorses within syngnathids has garnered special attention. Until recently, it was generally accepted that seahorses are sister to pygmy pipehorses, which also possess tail prehension and lack a caudal fin (Casey et al., 2004; Kuiter, 2009; Neutens et al., 2014; Teske and Beheregaray, 2009; Teske et al., 2004). However, we find that the sister group to seashores consists of our pygmy pipehorse representative, *Idiotropiscus lumnitzeri*, and the pipefish *Trachyrhamphus biocoarctatus*. This result was repeated across all datasets and methods with high support, and replicates the findings for a smaller dataset of nuclear loci (Hamilton et al., 2016). A close relationship between *Trachyrhamphus* and either pygmy pipehorses or seahorses has not been suggested based on morphology. Superficially, *Trachyrhamphus* species appear quite different from seahorses due to their often-drab coloration and relatively large size (Kuiter, 2009). However, on closer inspection, *Trachyrhamphus* pipefishes possess a number of key traits resembling seahorses and pipehorses. For instance, they have a relatively bent-head posture, may exhibit dermal flaps, and the caudal fin is lost or rudimentary in adults (Dawson, 1985). If these relationships are born out, study of *Trachyrhamphus* armor and caudal development may help inform the evolution of the seahorse's striking anatomy.

4.5. An ancient rapid radiation of misfit fishes

This is the largest study focused on syngnathiform relationships and the first to use a targeted-enrichment approach. The results presented here have given significant insights into the evolutionary relationships among syngnathiform lineages, especially among the long-snouted families. Although we did not perform analyses that infer ultrametric or dated topologies, the extremely short branch lengths estimated in our Bayesian and maximum likelihood analyses are consistent with a rapid radiation at the base of syngnathiforms. From previous work, Syngnathiformes are estimated to be 60–90 million years old (Betancur-R et al., 2013a; Near et al., 2012; Sanciangco et al., 2016). Such ancient rapid radiations are notoriously difficult phylogenetic problems and can be compounded by a suite of associated concerns including long-branch

attraction and rate heterogeneity (Whitfield and Lockhart, 2007). A rapid burst of speciation followed by long periods of independent diversification and radiation within lineages helps explain why traditional approaches before molecular systematics failed to correctly identify syngnathiform outgroups and why the resolution of branching relationships among currently recognized families has been inconsistent across large-scale molecular studies to date. While we anticipate that future studies focusing on this group of "misfit fishes" will reveal previously overlooked morphological characters uniting them with traditional syngnathiforms, it is also possible that rapid radiation followed by extended periods of divergence has eroded phylogenetic signal in morphology. This may be especially true for lineages such as the goatfish (Mullidae), which have many apomorphic traits associated with adaptations for benthic foraging (Gosline, 1984) and which have been historically difficult to place in the fish tree of life based on morphological characters.

Our study adds further evidence to the case that UCEs have the potential to resolve difficult issues within rapid radiations (McGee et al., 2016; Meiklejohn et al., 2016). However, sources of bias must be investigated, especially because systematic error is expected to be particularly problematic for instances where basal internode distances are closely spaced (Whitfield and Lockhart, 2007). In our case, we show that incomplete sampling of certain lineages (e.g., exclusion of the solenostomids) can lead to the inference of highly supported alternative topologies. In agreement with recent studies using UCEs (e.g., Meiklejohn et al., 2016), we find that the species tree summary method ASTRAL resulted in high discordance and a shift to unsymmetrical topologies when used on a UCE dataset. The situation may also become worse when applied to rapid radiations that are also ancient, since most loci will be uninformative for very short branches (Townsend et al., 2011). Filtering datasets to include informative loci may help alleviate some discordance due to poorly resolved gene trees, although deep relationships in our phylogeny remained poorly supported and inconsistent across matrices after filtering.

Syngnathiform lineages have acquired a range of functional, morphological, and behavioral innovations including power-amplified head rotation, independently mobile sensory barbels, pectoral-fin subdivision, extreme jaw protrusion, segmented armor, male and female brooding, and hovering locomotion. There are also substantial differences in morphological disparity and lineage diversity between families, ranging from three nearly-morphologically identical species of Aulostomidae to Syngnathidae, which includes more than 316 species having exceptionally high body shape disparity (the highest among 53 families of coral reef fishes in Price et al., (2015)). It could be that features of the rapid radiation itself, such as mosaicking of genetic backgrounds via ancestral polymorphism and hybridization, or periods of very high genetic variation within populations caused by these processes, led to such disparate lineages and facilitated the evolution of extremely unusual forms. Recent rapid radiations have received much attention for the unusual circumstances that allow early divergence, especially in fishes (Henning and Meyer, 2014; Martin et al., 2015; Meyer et al., 1990; Pickrell and Pritchard, 2012; Seehausen, 2006), but the long-term impact of these events on diversification patterns are not often apparent because of the focus on young radiations. The nature of the syngnathiform radiation suggests the possibility that there were long-term effects of this event that were key to the rapid diversification that founded this interesting group of fishes. It is unmistakable that this ancient radiation has had a lasting impact, encompassing some of the oddest and most successful groups of fishes. Syngnathiformes therefore present an exciting opportunity to study patterns of novelty and disparity arising from a fairly ancient radiation. We anticipate future work focused on these ideas to benefit from improved methodologies, models, and computational innovations designed to accommodate phylogenomic datasets.

Acknowledgements

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2017.05.002.

References

Aberer, A.J., Kobert, K., Stamatakis, A., 2014. ExaBayes: Massively Parallel Bayesian tree inference for the Whole-Genome Era. Mol. Biol. Evol. 31, 2553–2556.

Alfaro, M.E., Zoller, S., Lutzoni, F., 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol. Biol. Evol. 20, 255–266.

Allen, L.G., 1984. Gobiesociformes: development and relationships. Ontogen. System. Fishes, 629–636.

Berglund, A., Rosenqvist, G., 2003. Sex role reversal in pipefish. Adv. Study Behav. 32, 131–168.

Betancur-R, R., Broughton, R.E., Wiley, E.O., Carpenter, K., López, J.A., Li, C., Holcroft, N.I., Arcila, D., Sanciangco, M., Ii, J.C.C., et al., 2013a. The tree of life and a new classification of bony fishes. PLoS Curr. 5, 1–45.

Betancur-R, R., Li, C., Munroe, T.A., Ballesteros, J.E., Ortí, G., 2013b. Addressing gene tree discordance and non-stationarity to resolve a multi-locus phylogeny of Flatfishes (Teleostei: Pleuronectiformes). Syt. Biol. 62, 763–785.

Blaimer, B.B., Brady, S.G., Schultz, T.R., Lloyd, M.W., Fisher, B.L., Ward, P.S., 2015. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. BMC Evol. Biol. 15, 271.

Bowen, B.W., Bass, A.L., Rocha, L.A., Grant, W.S., Robertson, R., 2001. Phylogeography of the trumpetfishes (Aulostomus): ring species complex on a global scale. Evolution (N. Y) 55, 1029–1039.

Borowiec, M.L., 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4, e1660.

Casey, S.P., Hall, H.J., Stanley, H.F., Vincent, A.C.J., 2004. The origin and evolution of seahorses (genus Hippocampus): a phylogenetic study using the cytochrome b gene of mitochondrial DNA. Mol. Phylogenet. Evol. 30, 261–272.

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.

Crawford, N.G., Faircloth, B.C., McCormack, J.E., Brumfield, R.T., Winker, K., Glenn, T. C., 2012. More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs. Biol. Lett. 8, 783–786.

Cummings, M.P., Handley, S.A., Myers, D.S., Reed, D.L., Rokas, A., Winka, K., 2003. Comparing bootstrap and posterior probability values in the four-taxon case. Syst. Biol. 52, 477–487.

Cummings, M.P., Meyer, A., 2005. Magic bullets and golden rules: data sampling in molecular phylogenetics. Zoology 108, 329–336.

Dawson, C.E., 1981. Review of the indo-pacific doryrhamphine pipefish genus doryichthys. J. Ichthyol. 28, 1–18.

Dawson, C.E., 1984. Revision of the Genus Microphis Kaup (Pisces: Syngnathidae). Bull. Mar. Sci. 35, 117–181.

Dawson, C.E., 1985. Indo-Pacific Pipefishes. Ocean Springs. The Gulf Coast Research Laboratory, Mississippi.

Dávalos, L.M., Perkins, S.L., 2008. aturation and base composition bias explain phylogenomic conflict in Plasmodium. Genomics 91, 433–442.

Dettai, A., Lecointre, G., 2005. Further support for the clades obtained by multiple molecular phylogenies in the acanthomorph bush. Comptes Rendus – Biol. 328, 674–689.

Eschmeyer, W.N., Fricke, R., van der Laan, R., 2017. Catalog of Fishes: Genera, Species, References. <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>. Electronic version accessed 14 04 2017.

Eytan, R.I., Evans, B.R., Dornburg, A., Lemmon, A.R., Lemmon, E.M., Wainwright, P.C., Near, T.J., 2015. Are 100 enough? Inferring acanthomorph teleost phylogeny using Anchored Hybrid Enrichment. BMC Evol. Biol. 15, 113.

Faircloth, B.C., 2015. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32, 786–788.

Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61, 717–726.

Faircloth, B.C., Sorenson, L., Santini, F., Alfaro, M.E., 2013. A phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of Ultraconserved Elements (UCEs). PLoS ONE 8, e65923.

Faircloth, B.C., Branstetter, M.G., White, N.D., Brady, S.G., 2014. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. Mol. Ecol. Resour. n/a–n/a.

Felsenstein, J., 1978. Cases in which parsimony or compatibility methods will be positively misleading. Soc. Syst. Biol. 27, 401–410.

Foster, P.G., Hickey, D.A., 1999. Compositional Bias May Affect Both DNA-Based and Protein-Based Phylogenetic Reconstructions. pp. 284–290.

Gilbert, P.S., Chang, J., Pan, C., Sobel, E.M., Sinsheimer, J.S., Faircloth, B.C., Alfaro, M.E., 2015. Genome-wide ultraconserved elements exhibit higher phylogenetic informativeness than traditional gene markers in percomorph fishes. Mol. Phylogenet. Evol. 92, 140–146.

Glenn, T.C., Nilsen, R.A., Kieran, T.J., Finger Jr., J.W., Pierson, T.W., Bentley, K.E., Hoffberg, S.L., Louha, S., García-De León, F.J., Portilla, M.A.D.R., et al., 2016. Adapterama I: Universal stubs and primers for thousands of dual-indexed Illumina libraries (iTru & iNext). BioRxiv, 049114.

Gosline, W.A., 1970. A reinterpretation of the teleostean fish order Gobiesociformes. Proc. Calif. Acad. Sci. 38, 363–381.

Gosline, W.A., 1984. Structure, Function, and Ecology in the Goatfishes (Family Mullidae). Pacific Sci. 38, 312–323.

Gregory, W.K., 1933. Fish Skulls: A Study of the Evolution of Natural Mechanisms. Trans. Amer. Phil. Soc. 23, 75–481.

Hamilton, H., Saarman, N., Short, G., Sellas, A.B., Moore, B., Hoang, T., Grace, C.L., Gomon, M., Crow, K., Brian, W., 2016. Molecular phylogenetics and evolution molecular phylogeny and patterns of diversification in syngnathid fishes. Mol. Phylogenet. Evol. 107, 388–403.

Harvey, M.G., Smith, B.T., Glenn, T.C., Faircloth, B.C., Brumfield, T., Rouge, B., Rouge, B., Harvey, M.G., 2016. Sequence capture versus rad-seq for shallow systematics. Syst. Biol. 65, 910–924.

Heath, T.A., Hedtke, S.M., Hillis, D.M., 2008. Taxon sampling and the accuracy of phylogenetic analyses. J. Syst. Evol. 46, 239–257.

Henning, F., Meyer, A., 2014. The evolutionary genomics of cichlid fishes: explosive speciation and adaptation in the postgenomic era. Annu. Rev. Genom. Hum. Genet. 15, 417–441.

Hillis, D.M., Pollock, D.D., McGuire, J.A., Zwickl, D.J., 2003. Is sparse taxon sampling a problem for phylogentic inference? Syst. Biol. 52, 124–126.

Houde, E.D., 1984. Callionymidae: development and relationships. In: Ontogeny and Systematics of Fishes. American Society of Ichthyologists and Herpetologists Special Publication, pp. 637–640.

Imamura, H., 2000. An alternative hypothesis on the phylogenetic position of the family Dactylopteridae (pisces: Teleostei), with a proposed new classification. Ichthyol. Res. 47, 203–222.

Ishikawa, S.A., Inagaki, Y., Hashimoto, T., 2012. RY-coding and non-homogeneous models can ameliorate the maximum-likelihood inferences from nucleotide sequence data with parallel compositional heterogeneity. Evol. Bioinforma. 8, 357–371.

Jeffroy, O., Brinkmann, H., Delsuc, F., Philippe, H., 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22, 225–231.

Jermiin, L., Ho, S.Y., Ababneh, F., Robinson, J., Larkum, A.W., 2004. The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. Syst. Biol. 53, 638–643.

Johnson, G.D., Patterson, C., 1993. Percomorph phylogeny: a survey of acanthomorphs and a new proposal. Bull. Mar. Sci. 52, 554–626.

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.

Keivany, Y., Nelson, J.S., 2006. Interrelationships of Gasterosteiformes (Actinopterygii, Percomorpha). J. Ichthyol. 46, S84–S96.

Kuiter, R., 2009. Seahorses and Their Relatives. Aquatic Photographics, Seaford, Australia, p. 334.

Kumar, S., Filipski, A.J., Battistuzzi, F.U., Kosakovsky Pond, S.L., Tamura, K., 2012. Statistics and truth in phylogenomics. Mol. Biol. Evol. 29, 457–472.

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., Stamatakis, A., 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evolut. Biol. 14, 1–14.

Lecointre, G., Philippe, H., Lanh van Le, H., le Guyader, H., 1993. Species sampling has a major impact on phylogenetic inference. Mol. Phylogenet. Evol. 2, 205–224.

Lemmon, E.M., Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics. Annu. Rev. Ecol. Evol. Syst. 44, 99–121.

Leo, W., Craig, M.T., 2007. Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. Copeia 1, 35–55.

Li, B., Dettaï, A., Cruaud, C., Couloux, A., Desoutter-Meniger, M., Lecointre, G., 2009. RNF213, a new nuclear marker for acanthomorph phylogeny. Mol. Phylogenet. Evol. 50, 345–363.

Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.

Manthey, J.D., Campillo, L.C., Burns, K.J., Moyle, R.G., 2016. Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: a test in cardinalid tanagers (Aves, Genus: Piranga). Syst. Biol. syw005.

Martin, C.H., Cutler, J.S., Friel, J.P., Dening Touokong, C., Coop, G., Wainwright, P.C., 2015. Complex histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on one of the clearest examples of sympatric speciation. Evolution 69, 1406–1422.

McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome Res. 22, 746–754.

McGee, M.D., Faircloth, B.C., Borstein, S.R., Zheng, J., Darrin Hulsey, C., Wainwright, P.C., Alfaro, M.E., 2016. Replicated divergence in cichlid radiations mirrors a major vertebrate innovation. Proc. Roy. Soc. B 283, 20151413.

Meiklejohn, K.A., Faircloth, B.C., Glenn, T.C., Kimball, R.T., Braun, E.L., 2016. Analysis of a rapid evolutionary radiation using ultraconserved elements (UCEs): evidence for a bias in some multispecies coalescent methods. Syst. Biol. 0, syw014.

Meyer, A., Kocher, T.D., Basasibwaki, P., Wilson, A., 1990. Monophyletic origin of Lake Victoria cichlid Fishes suggested by mitochondrial DNA sequences. Nature 347, 550–553.

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.

Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30, i541–i548.

Miya, M., Takeshima, H., Endo, H., Ishiguro, N.B., Inoue, J.G., Mukai, T., Satoh, T.P., Yamaguchi, M., Kawaguchi, A., Mabuchi, K., et al., 2003. Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. Mol. Phylogenet. Evol. 26, 121–138.

Mobley, K.B., Jones, A.G., 2013. Overcoming statistical bias to estimate genetic mating systems in open populations: a comparison of Bateman's principles between the sexes in a sex-role-reversed pipefish. Evolution 67, 646–660.

Mwale, M., Kaiser, H., Barker, N.P., Wilson, a.B., Teske, P.R., 2013. Identification of a uniquely southern African clade of coastal pipefishes Syngnathus spp. J. Fish Biol. 82, 2045–2062.

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.

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., et al., 2013. Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. PNAS 110, 12738–12743.

Nelson, J.S., 1994. Fishes of the World. John Wiley & Sons, New York. 600p.

Nelson, J.S., 2006. Fishes of the World. John Wiley & Sons, Hoboken, NJ. 600p.

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.

Nishihara, H., Okada, N., Hasegawa, M., 2007. Rooting the eutherian tree: the power and pitfalls of phylogenomics. Genome Biol. 8, R199.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9, e1000602.

Phillips, M.J., Delsuc, F., Penny, D., 2004. Genome-scale phylogeny and the detection of systematic biases. Mol. Biol. Evol. 21, 1455–1458.

Pickrell, J.K., Pritchard, J.K., 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. 8, e1002967.

Pietsch, T.W., 1978. Evolutionary relationships of the sea moths (Teleostei: Pegasidae) with a classification of gasterosteiform families. Copeia 1978, 529.

Porter, M.M., Adriaens, D., Hatton, R.L., Meyers, M.a., McKittrick, J., 2015. BIOMECHANICS. Why the seahorse tail is square. Science, vol. 349, pp. aaa6683.

Price, S.A., Claverie, T., Near, T.J., Wainwright, P.C., 2015. Phylogenetic insights into the history and diversification of fishes on reefs. Coral Reefs 34, 997–1009.

Rambaut, A., Suchard, M. A., Xie, D., Drummond, A.J., 2014. Tracer v1.6, Available from <http://beast.bio.ed.ac.uk/Tracer>.

Rodríguez-Ezpeleta, N., Brinkmann, H., Roure, B., Lartillot, N., Lang, B.F., Philippe, H., 2007. Detecting and overcoming systematic errors in genome-scale phylogenies. Syst. Biol. 56, 389–399.

Rothfels, C.J., Larsson, A., Kuo, L.Y., Korall, P., Chiou, W.L., Pryer, K.M., 2012. Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. Syst. Biol. 61, 490–509.

Sanciangco, M.D., Carpenter, K.E., Betancur-R, R., 2016. Phylogenetic placement of enigmatic percomorph families (Teleostei: Percomorphaceae). Mol. Phylogenet. Evol. 94, 565–576.

Seehausen, O., 2006. African cichlid fish: a model system in adaptive radiation research. Proc. Roy. Soc. B 273, 1987–1998.

Simmons, M.P., Pickett, K.M., Miya, M., 2004. How meaningful are bayesian support values? Mol. Biol. Evol. 21, 188–199.

Smith, B.T., Harvey, M.G., Faircloth, B.C., Glenn, T.C., Brumfield, R.T., 2014. Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. Syst. Biol. 63, 83–95.

Smith, W.L., Craig, M.T., 2007. Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. Copeia, 35–55.

Smith, W.L., Wheeler, W.C., 2004. Polyphyly of the mail-cheeked fishes (Teleostei: Scorpaeniformes): evidence from mitochondrial and nuclear sequence data. Mol. Phylogenet. Evol. 32, 627–646.

Smith, W.L., Wheeler, W.C., 2006. Venom evolution widespread in fishes: a phylogenetic road map for the bioprospecting of piscine venoms. J. Hered. 97, 206–217.

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.

Springer, V.G., Johnson, G.D., 2004. Study of the dorsal gill-arch musculature of the teleostome fishes, with special reference to the actinopterygii. Bull. Biol. Soc. Washingt., 237–260

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

Steinke, D., Salzburger, W., Meyer, A., 2006. Novel relationships among ten fish model species revealed based on a phylogenomic analysis using ESTs. J. Mol. Evol. 62, 772–784.

Stolting, K.N., Wilson, A.B., 2007. Male pregnancy in seahorses and pipefish: Beyond the mammalian model. BioEssays 29, 884–896.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Swofford, D.L., Olsen, G.J., Waddell, P.J., D.M.H., 1996. Phylogenetic inference. In: Hllis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sunderland: Sinauer Associates, pp. 407–514.

Teske, P.R., Beheregaray, L.B., 2009. Evolution of seahorses' upright posture was linked to Oligocene expansion of seagrass habitats. Biol. Lett. 5, 521–523.

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.

Townsend, T.M., Mulcahy, D.G., Noonan, B.P., Sites, J.W., Kuczynski, C.A., Wiens, J.J., Reeder, T.W., 2011. Phylogeny of iguanian lizards inferred from 29 nuclear loci, and a comparison of concatenated and species-tree approaches for an ancient, rapid radiation. Mol. Phylogenet. Evol. 61, 363–380.

Van Wassenbergh, S., Strother, J.A., Flammang, B.E., Ferry-Graham, L.A., Aerts, P., 2008. Extremely fast prey capture in pipefish is powered by elastic recoil. J. Roy. Soc. Interface 5, 285–296.

Van Wassenbergh, S., Roos, G., Genbrugge, A., Leysen, H., Aerts, P., Herrel, A., 2009. Suction is kid's play: extremely fast suction in newborn seahorses. Biol. Lett. 5, 200–203.

Washington, B.B., Eschmeyer, W.N., Howe, K., 1984. Scorpaeniformes: relationships. Ontogen. System. Fishes, 438–447.

Whitfield, J.B., Kjer, K.M., 2008. Ancient rapid radiations of insects: challenges for phylogenetic analysis. Annu. Rev. Entomol. 53, 449–472.

Whitfield, J.B., Lockhart, P.J., 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22, 258–265.

Wilson, A.B., Orr, J.W., 2011. The evolutionary origins of Syngnathidae: pipefishes and seahorses. J. Fish Biol. 78, 1603–1623.

Wilson, N.G., Rouse, G.W., 2010. Convergent camouflage and the non-monophyly of "seadragons" (Syngnathidae): suggestions for a revised taxonomy of syngnathids. Zool. Scr. 39, 551–558.

Wilson, A.B., Vincent, A., Ahnesjö, I., Meyer, a., 2001. Male pregnancy in seahorses and pipefishes (family Syngnathidae): rapid diversification of paternal brood pouch morphology inferred from a molecular phylogeny. J. Hered. 92, 159–166.

Wilson, A.B., Ahnesjö, I., Vincent, A.C.J., Meyer, A., 2003. The dynamics of male brooding, mating patterns, and sex roles in pipefishes and seahorses (family Syngnathidae). Evolution 57, 1374–1386.

Woese, C.R., Achenbach, L., Rouviere, P., Mandelco, L., 1991. Archaeal phylogeny: reexamination of the phylogenetic position of Archaeoglohus fulgidus in light of certain composition-induced artifacts. Syst. Appl. Microbiol. 14, 364–371.

Zwickl, D.J., Hillis, D.M., 2002. Increased taxon sampling greatly reduces phylogenetic error. Syst. Biol. 51, 588–598.