Accepted Manuscript

Depth as a driver of evolution in the deep sea: Insights from grenadiers (Gadiformes: Macrouridae) of the genus *Coryphaenoides*

Michelle R. Gaither, Biagio Violi, Howard W.I. Gray, Francis Neat, Jeffrey C. Drazen, R. Dean Grubbs, Adela Roa-Varón, Tracey Sutton, A. Rus Hoelzel

Title:

Depth as a driver of evolution in the deep sea: insights from grenadiers (Gadiformes: Macrouridae) of the genus *Coryphaenoides*

Authors:

Michelle R. Gaither[a,b], Biagio Violi[a,c,d], Howard W. I. Gray[a], Francis Neat[e], Jeffrey C. Drazen[f], R. Dean Grubbs[g], Adela Roa-Varón[h], Tracey Sutton[i], A. Rus Hoelzel[a]

[a]School of Biological and Biomedical Sciences, Durham University, South Road, Durham DH1 3LE, UK

[b]Section of Ichthyology, California Academy of Sciences, 55 Music Concourse Drive, San Francisco, CA 94118, USA

[c]Istituto Nazionale Biostrutture e Biosistemi, Viale delle Medaglie d'Oro 305, Rome, Italy

[d]Department of Earth, Environment and life sciences, University of Genoa, Corso Europa 26, 16132, Genoa, Italy

[e]Marine Scotland-Science, 375 Victoria Rd, Aberdeen, AB11 9DB, Scotland

[f]Department of Oceanography, University of Hawaii at Manoa, 1000 Pope Rd, Honolulu, HI 96822, USA

[g]Florida State University Coastal and Marine Laboratory, 3618 Hwy 98, St. Teresa, FL 32358, USA

[h]Virginia Institute of Marine Science, PO Box 1346, Gloucester Point, VA 23062, USA

[i]Halmos College of Natural Sciences and Oceanography, Nova Southeastern University, 8000 North Ocean Drive, Dania Beach, FL 33004, USA

Corresponding authors:

Michelle R. Gaither
+44 01913341323
michellergaither@gmail.com

A. Rus Hoelzel
a.r.hoelzel@durham.ac.uk

Key words: Macrouridae, hydrostatic pressure, rattails, phylogenetics, fishes

Abstract

Here we consider the role of depth as a driver of evolution in a genus of deep-sea fishes. We provide a phylogeny for the genus *Coryphaenoides* (Gadiformes: Macrouridae) that represents the breadth of habitat use and distributions for these species. In our consensus phylogeny species found at abyssal depths (> 4000 m) form a well-supported lineage, which interestingly also includes two non-abyssal species, *C. striaturus* and *C. murrayi,* diverging from the basal node of that lineage. Biogeographic analyses suggest the genus may have originated in the Southern and Pacific Oceans where contemporary species diversity is highest. The abyssal lineage seems to have arisen secondarily and likely originated in the Southern/Pacific Oceans but diversification of this lineage occurred in the Northern Atlantic Ocean. All abyssal species are found in the North Atlantic with the exception of *C. yaquinae* in the North Pacific and *C. filicauda* in the Southern Oceans. Abyssal species tend to have broad depth ranges and wide distributions, indicating that the stability of the deep oceans and the ability to live across wide depths may promote population connectivity and facilitate large ranges. We also confirm that morphologically defined subgenera do not agree with our phylogeny and that the Giant grenadier (formally *Albatrossia pectoralis*) belongs to *Coryphaenoides,* indicating that a taxonomic revision of the genus is needed. We discuss the implications of our findings for understanding the radiation and diversification of this genus, and the likely role of adaptation to the abyss.

1. Introduction

The deep oceans (>200m depth) are vast three-dimensional habitats characterized by decreasing sunlight, low temperatures, and increasing hydrostatic pressures with depth. Life in these extreme habitats is almost entirely reliant upon the organic nutrients that rain down from the photic zone (vent and seep communities are the exception; Snelgrove and Smith, 2002). These habitats were long believed to be environmentally homogenous with limited barriers to dispersal, leading to the assumption that species in the deep sea have vast ranges with little opportunity for divergence. However, our view of this habitat has been transformed as exploration of the deep sea has revealed complex topographic features (mid-ocean ridges, seamounts, cold seeps, hydrothermal vents, hadal trenches, etc), high species richness, and species turnover rates of 45–80% over hundreds to thousands of kilometers (Brandt et al., 2005; Glover et al., 2002; Grassle and Maciolek, 1992).

With less than 1% of the deep-sea floor having been explored (McClain, 2007) and with sampling efforts concentrated in the Northern Hemisphere (Stuart et al., 2008) our knowledge of the deep sea lags behind that of shallow systems. However some broad-scale patterns concerning the distribution of biodiversity in the deep sea have emerged. Studies show that abundance (number of individuals) and biomass decreases with depth in most locations (reviewed in Rex and Etter, 2010). Furthermore, there are differences in vertical assemblages from the continental slope to the abyss (Carney, 2005; Carney et al., 1983; Haedrich et al., 1980; Lundsten et al., 2010). Other datasets indicate increasing diversity of benthic and demersal species with increasing depth from the continental shelf, reaching a peak over upper-bathyal depths, and then decreasing in deeper water (Haedrich et al., 1980; Olabarria, 2005; Pineda and Caswell, 1998; Priede et al., 2010). This unimodal pattern has been shown across a diversity of taxa in the North Atlantic and on a global scale (Priede et al., 2010; Vinogradova, 1962). However,with so few studies the generality of the pattern remains unknown (reviewed in Rex and

3

Etter, 2010). A hypothesis explaining mid-slope peaks in species richness proposes that most diversity in the deep sea originated in the heterogeneous bathyal depths (1000-2000m depth) (Etter et al., 2005; Etter and Rex, 1990).

The genus *Coryphaenoides* (family Macrouridae), known as grenadiers or rattails, is a diverse group of fishes found worldwide from tropical to polar seas. There are 65 recognized species that are found across a large depth range from the euphotic zone to the deep abyss (110 m to at least 7000m), with most species found between 700 m and 2000 m (Cohen et al., 1990; Linley et al., in review). Despite the fact that members of this genus often comprise large portions of the demersal biomass few species are commercially harvested, except *C. rupestris*, which forms a large fishery in the North Atlantic. Two species are considered circumglobal (*C. armatus* and *C. rudis*; Gaither et al., 2016) and another shows a possible anti-tropical distribution (Laptikhovsky et al., 2013). Eight species are known from abyssal depths and another two have been recorded at the edge of the deepest bathyal habitat (very near 4000 m) (Table S1).

Based on morphology, Cohen et al. (1990) tentatively divided the genus into five subgenera: *Bogoslovius* (1 sp.), *Chalinura* (10 spp.), *Coryphaenoides* (43 spp.), *Lionurus* (2 spp.), and *Nematonurus* (5 spp.). However, the generic and subgeneric designations within *Coryphaenoides* are poorly resolved based on morphological data (Iwamoto and Sazonov, 1988; Iwamoto and Stein, 1974; Marshall, 1973; Okamura, 1970a, b; Parr, 1946) or molecular data (Morita, 1999; Wilson et al., 1991; Wilson and Attia, 2003). Cohen et al. (1990) acknowledged the likely paraphyly of the largest subgenus *Coryphaenoides* and the apparent lack of diagnostic characters has left the group in 'taxonomic limbo'. Although there is uncertainty in the classification scheme put forward by Cohen et al. (1990), it provides a framework in which to test for congruence between taxonomic and molecular characters. In a phylogeny constructed for the order Gadiformes (based on mtDNA and nuclear RAG1 sequences), Roa-Varón and Ortí (2009)

4

indicated that *Coryphaenoides* is paraphyletic, with the monotypic species *Albatrossia pectoralis* nesting within the lineage, a finding which is supported by allozymes, peptide mapping and DNA sequence data (Morita, 1999; Wilson et al., 1991; Wilson, 1994; Wilson and Attia, 2003).

The genus *Coelorinchus* is the sister group to *Coryphaenoides*. The study of Roa-Varón and Ortí (2009), which was designed to resolve taxonomic relationships at the family and subfamily level among Gadiformes, included 11 species of *Coryphaenoides* but none that reside at abyssal depths. Unfortunately, other phylogenies that have focused on *Coryphaenoides* suffer from poor taxonomic sampling and are based on a single molecular marker (Fig. 1). Despite presumed low resolution, each of these phylogenies indicates that abyssal species cluster together or even form a separate clade.

Here we investigate the evolution of diversity across the bathyal-abyssal interface using the most complete phylogenetic treatment for species of *Coryphaenoides* to date. This investigation is based on sequence data from two mitochondrial (COI and 16S) and two nuclear markers (RAG1 and MYH6) from 29 of the 65 recognized species in the genus. We provide the first independent appraisal of the subgenus designations put forward by Cohen et al. (1990) and confirm the placement of *C. pectoralis* within *Coryphaenoides*. Second, we include seven of the eight abyssal *Coryphaenoides* to determine if they are monophyletic, as indicated by previous, but incomplete, phylogenies. Finally, we use biogeographic models to assess the evolutionary history of the group, and evaluate the depth distributions of species of *Coryphaenoides* to determine if there is greater species diversity at bathyal depths, as might be expected under the depth-differentiation hypothesis (Rex and Etter, 2010).

2. Materials and methods

2.1. Taxon sampling and DNA extraction

5

A total of 73 specimens across 28 of the 65 recognized species of *Coryphaenoides* were obtained for this study (Table S2). According to a previous phylogeny, the genus *Coelorinchus* is a sister taxon of *Coryphaenoides* and so we rooted our phylogenetic trees using sequences from *Coelorinchus labiatus* (Roa-Varón and Ortí, 2009).Total genomic DNA was extracted from tissues (gills or muscle) using either a phenol-chloroform protocol (after Hoelzel, 1998) or using the E.Z.N.A extraction kit (Omega Bio-Tek, USA) following the manufacturer's protocol and subsequently stored at -20 ˚C.

2.2. *PCR amplification and sequencing*

Two mitochondrial and two nuclear genes were used in this study, amounting to a total of 3086 bp. We resolved 626 bp of the mitochondrial cytochrome oxidase I gene (COI) using the primers FishF2 and FishR2 of Ward et al. (2005). We also resolved 1010 bp of the 16S mitochondrial ribosomal fragment and 715 bp of exon three of the recombination-activating protein 1 (RAG1) using the primers of Roa-Varón and Ortí (2009). Lastly, we resolved 735 bp of the MYH6 gene using two pairs of primers 459F/1325Rand 507F/1322F from Li et al. (2007) (Table S3). Both RAG1 and MYH6 required a nested PCR approach. In these cases 1 µl of a 1:20 dilution of the first PCR product was used as the template for the second PCR (Table S3). For a few low quality samples it was necessary to increase template concentration to achieve successful amplification of MYH6. Polymerase chain reactions (PCRs) were carried out in a 20 µl volume containing 1 µl of extracted DNA, 0.4 µl of each primer (10µM), 0.4 µl of dNTP mix (10mM each; Promega, USA), 2.4 µl of $MgCl_2$ (25 mM), 4 µl of 5x Green GoTaq Flexi Buffer (Promega), 0.1 µl of GoTaq Taq DNA polymerase (Promega), and deionized water to volume. PCR reactions utilized the following cycling parameters: initial denaturation at 95°C and final extension at 72°C (for 5 min each), with an intervening 35 cycles of 30 sec at 95°C, 30 sec at the annealing temperature (Table S3), and 45 sec at 72°C. Amplification products were purified using USB ExoSAP-

6

IT (Affymetrix, USA). Following the manufacturer's protocol, we incubated 5 µl of PCR product and 2 µl of ExoSAP-IT reagent at 37°C for 15 min followed by 15 min at 85°C. DNA sequencing was performed with fluorescent-labeled dideoxy terminators on an ABI 3730XL Genetic Analyzer (Applied Biosystems, USA) at the Durham University DBS genomics facility.

2.3. Sequence alignment and Phylogenetic Analysis

Sequences for each locus were edited using the DNA sequence assembly and analysis software Geneious Prov. 5.5.6 (Kearse et al., 2012). Following Morita (1999), we calculated the average rate of transversions between sequences for COI using the pairwise distance function in MEGA v. 7.0.14 (Kumar et al., 2016). Sequence alignments were conducted using the MUSCLE algorithm as implemented in Geneious Pro (see Table S2 for GenBank accession numbers). Both 16S and RAG1 alignments had gaps. The gap in RAG1 was also recovered by Roa-Varón and Ortí, 2009 and consisted of a 3 to 9 bp indel. For this gene we constructed a ML tree for a truncated alignment that omitted the gap using MEGA (as describe below). The topology of this tree was highly similar to the ML tree that resulted from the full RAG1 alignment with all major nodes retained. Therefore all subsequent analyses were conducted using the full RAG1 alignment. To improve alignments and minimize gaps in RAG1 and 16S we used the Gblocks v.0.91b web server (Castresana, 2000; Talavera and Castresana, 2007) using the codon option for RAG1 and allowing for gaps in both markers. In both cases, gaps were reduced with > 97% of positions retained (RAG1, 696 bp; 16S, 991 bp). We used MESQUITE v. 3.04 (Maddison and Maddison, 2001) to assign codon position in our protein-coding markers (COI, MYH6, and RAG1) by minimizing stop-codons and translated the sequences to ensure that no stop-codons were present. Each locus was tested for saturation using Xia's test (Xia et al., 2003) as implemented in DAMBE5 (Xia, 2013). We conducted these tests for the combined 1st and 2nd codon positions and for

the 3rd codon position. We estimated the proportion of invariable using the neighbor-joining tree algorithm as recommended by Xia and Lemey (2009). The default of 60 replicates was used and we considered only fully resolved sites, as recommended. Only COI showed signs of saturation at the 3rd codon position. Sequences were concatenated and the best-fit partitioning scheme and substitution models were investigated using PartitionFinder v. 1.1.1 (Lanfear et al., 2012). One run was performed using the "search all" algorithm with branch-lengths "linked" between partitions and a second, "greedy" search, was also performed with branch-lengths "unlinked". Best-fit schemes were identified using the Bayesian Information Criterion (BIC) (Table S4).

In order to examine the topology of phylogenetic trees for individual fragments we conducted Maximum Likelihood (ML) analyses on each marker separately using MEGA. We first selected the most appropriate model of evolution for each marker, based on the BIC, using default settings implemented in MEGA (COI:HKY G+I, 16S: TN93+G+I, RAG1:T92+G, MYH6: K2). Subsequently, we ran ML analyses invoking the appropriate model and applying 500 bootstrap replicates. Trees were rooted using the closely related Spearnose grenadier *Coelorinchus labiatus*. A Bayesian Markov Chain Monte Carlo (MCMC) analysis, as implemented in MrBayes v. 3.2 (Ronquist et al., 2012), was also conducted. Partitions were assigned according to PartitionFinder results (Table S4). Posterior probabilities were calculated using 10 million iterations of four chains (three heated) replicated in four independent runs with a sample frequency of 1000 and a burn-in fraction of 0.25. The branch length prior was set to exponential with an unconstrained molecular clock (default settings). Under these parameters standard deviations between independent runs stabilized and resulted in split frequencies below 0.01. Subsequent phylogenetic analyses of the concatenated dataset were performed on the CIPRES Science Gateway v. 3.1 (Miller et al., 2010). Maximum likelihood (ML) analyses on this larger dataset was conducted using the Randomized Accelerated Maximum Likelihood (RAxML) software v.

8

7.3.0 (Stamatakis, 2006). We set RAxML to estimate all model parameters and partitions were assigned according to PartitionFinder results (Table S4). Eight independent runs were performed and the best trees from the individual runs were compared to assess concordance in topology and ensure that the ML search was converging on the optimal area of tree space. In addition, a ML analysis with 1000 bootstrap replicates was performed to estimate support for individual clades in the tree. A Bayesian Markov Chain Monte Carlo (MCMC) analysis, as implemented in MrBayes, was performed on the concatenated dataset as described above. Under these parameters standard deviations between independent runs stabilized and resulted in split frequencies below 0.01.

A strict-consensus maximum parsimony tree was generated in PAUP* v. 4.0b10 (Swofford, 2003) from a heuristic search with TBR branch swapping, 1000 random additions and 500 bootstrap replicates. In order to test for congruence across loci, partitioned Bremer support indices (Baker and DeSalle, 1997) were calculated for each node using TreeRot v. 3 (Sorenson and Franzosa, 2007) and PAUP* by performing heuristic searches with 1000 random additions.

2.4. Biogeographic State of Ancestral Nodes

In order to reconstruct the biogeographic state of ancestral nodes, Bayesian Binary MCMC (BBM) and S-DIVA analyses were performed in RASP (Reconstruct Ancestral State Phylogenies) v. 3.2 (Yu et al., 2015). We sampled 10,000 trees at random from a Bayesian phylogenetic Markov Chain Monte Carlo (MCMC) analysis, generated from the concatenated alignment. S-DIVA analysis was run on all trees and ancestral nodes were plotted on a majority-rule consensus tree. Our objective in using biogeographic analyses was to infer the geographic origin of major lineages and therefore we assigned distributions to terminal nodes based on ocean basin: Pacific Ocean, Atlantic Ocean, and Southern Ocean. Most species are found within a single ocean basin (Fig. 3, Table S5) and only four (two of

which are considered circumglobal) are found across several ocean basins. We allowed for all three areas to be included in any single ancestral distribution. We assigned all regions to the root (*Coelorinchus labiatus*) to avoid biased reconstructions at the base of the tree. Using this same dataset, we ran the BMM for 1 million generations with a sampling frequency of 100 and 10 chains were run with a temperature of 0.1. We discarded 1000 samples (100,000 generations) as burn-in before calculating the state frequencies. The Fixed Jukes-Cantor model for state frequencies was applied with the gamma shape parameter for among-site rate variation. The analysis was run twice to ensure that estimations were converging.

Depth range estimates were based on data provided in Fishbase (http://www.fishbase.org/search.php). When possible these values were checked in the literature. However, it should be noted that obtaining accurate depth records for deep-sea species is difficult. In some cases poor sampling may underestimate the full species range while errant records can inflate depth ranges. Furthermore, the inclusion of records of larvae or very small juveniles may lead to artificially broad depth ranges as many species of fish are known to exhibit pelagic tendencies and/or ontogenetic downslope migration.

3. Results

3.1. Sequencing and partitioning of the dataset

A total of 3048 bp of DNA were resolved after editing with Gblocks. Our data matrix of 73 individuals (including *Coelorinchus labiatus*) and four loci was 97% complete (Table 1). Of the 3048 bp, 2276 were invariant (74%) and 655 were segregating (21.0%) sites. Descriptive statistics for each locus are given in Table 1.We found no evidence of saturation at any of the loci and no stop codons

10

were detected in our protein-coding markers COI, MYH6, and RAG1. All conspecific samples were monophyletic across the two mitochondrial trees. The best-fit partitioning schemes identified using the BIC from PartitionFinder are detailed in Table S4.

3.2. Phylogenetic relationships and congruence across markers

Phylogenies based on individual markers have many well-supported nodes (Fig. 2) and no strong pattern of incongruence based on the TreeRot analyses (e.g. the five deepest nodes showed Bremer index values ranging from -2.36 to 69.55 across all loci combined). The MYH6 locus showed the lowest resolution (Fig. 2d) while the tree based on mitochondrial 16S (Fig. 2b) offered the greatest resolution. In the mitochondrial phylogenies all taxonomic species were supported by high ML bootstrap values and posterior probabilities (PP). The exception was the distinction between the sister species *C. filifer* and *C. cinereus* in the 16S tree and *C. acrolepis* and *C. longifilis* in the COI tree. In all cases *C. pectoralis* nested within *Coryphaenoides* and was most closely related to *C. acrolepis* and *C. longifilis*. Only in the mitochondrial datasets were the species level designations of these three species well-resolved, with exceptions highlighted above. *Coryphaenoides dossenus* was conspicuously divergent at all loci. Of note is the presence of a well-supported and divergent abyssal clade which included all species found at depths greater than 4000 m (Fig. 2: *C. yaquinae*, *C. armatus*, *C. profundicolus*, *C. leptolepis*, *C. mediterraneus*, *C. carapinus*, and *C. brevibarbis*). Only the COI tree failed to resolve this clade (although not well supported in the MYH6 tree), but the arrangement of lineages were recovered as a polytomy and did not conflict with this grouping. Interestingly, the two non-abyssal species, *C. striaturus* and *C. murrayi*, clustered with the abyssal taxa in the mitochondrial trees (COI and 16S) while the nuclear trees showed mixed topologies. One sample of *C. striaturus* (CSI001) nested within

11

the abyssal clade at all loci, whereas the other *C. striaturus* sample (CSI002) and *C. murrayi* samples (CMY002 and CMY003) exhibited non-abyssal nuclear genotypes (Fig 2c,d).

The ML analyses based on the concatenated dataset (3048 bp) consistently returned the same tree topology across eight independent runs (Fig. 3). Similarly, our Bayesian analyses inferred the same tree topology across all runs. These two different approaches produced consensus trees of highly similar topology with strong bootstrap support (ML) and high posterior probabilities (BI) at most nodes (Fig. 3). The Bayesian 50% consensus tree is presented in Figure 3. The topology of this tree was largely consistent with the trees based on individual loci. All species-level designations were highly supported by our concatenated dataset, including the distinctions between *C. filifer* and *C. cinereus* and *C. acrolepis* and *C. longifilis*. However, some of the deeper nodes were still poorly resolved. The earliest lineage to branch from the tree included three Southern Ocean species, with the divergent *C. dossenus* branching next. A Pacific Ocean lineage that included *C. filifer*, *C. cinereus*, *C. longifilis*, *C. acrolepis* and *C. pectoralis* was well-supported and branched early. The abyssal lineage was recovered in our multi-locus dataset, with *C. striaturus* and *C. murrayi* branching from the basal node of this lineage, a pattern driven by the strong signal in the mitochondrial 16S (partitioned Bremer indices node 10; COI:-1.96, 16S:18.85, RAG2:-0.16, MYH6: -0.73). The splitting of the two *C. striaturus* samples in the concatenated dataset (Fig. 3) is driven by the divergence between these two samples at the nuclear loci with CSI001 nesting in the abyssal lineage at all loci while CSI002 nested in the abyssal lineage in mtDNA trees but in the non-abyssal lineages in the nDNA trees (Fig 2).

3.2. *RASP*

Bayesian (BBM) and event-based (S-DIVA) reconstructions of ancestral biogeography based on extant taxa produced generally consistent results (Fig. 3, Table S6). In broad terms, the analyses suggest

an origin for *Coryphaenoides* in the Southern and Pacific Oceans, with an early vicariance event (node 1) that led to a splitting of the Southern and Pacific Ocean lineages. Migrations into the Atlantic appeared later, with isolation between Pacific and Atlantic Oceans leading to further diversification (node 3). A more recent vicariant event (node 8) led to the origin of the abyssal lineage. At the base of this lineage are two non-abyssal species with reconstructions at this node indicating a Southern Ocean/Pacific Ocean origin. If we consider only the truly abyssal taxa it appears that a vicariant event isolated Southern Ocean and Atlantic Ocean lineages (node 10) with subsequent diversification of the abyssal species occurring in the Atlantic. All abyssal species reside in the Atlantic Ocean with the exception of the recently derived taxa *C. yaquinae*.

4. Discussion

Using two mitochondrial and two nuclear loci we produced a well resolved phylogeny for *Coryphaenoides*. We found that the morphological subgenera of Cohen et al. (1990) are inconsistent with the molecular data, and, as previously indicated, *C. pectoralis* nests within *Coryphaenoides* (Roa-Varón and Ortí, 2009; Wilson et al., 1991; Wilson, 1994; Wilson and Attia, 2003). Consistent with earlier studies limited by poor taxon representation and low resolution (see Fig 1), we found that species inhabiting waters below the 4000 m isobaths formed a well-supported lineage. Branching from the basal node of this divergent lineage were two species found in non-abyssal depths: *C. striaturus* and *C. murrayi*. Based on morphological evidence from Cohen et al. (1990), and the placement of these two species in the mitochondrial trees (Fig. 2a,b), these may represent abyssal species that have immigrated to a shallower habitat. The finding that three of the four individuals of *C. striaturus* and *C. murrayi* possessed nuclear alleles that nested within the non-abyssal lineage may be evidence of historical introgression between the major lineages, however the nuclear gene trees on their own are relatively

13

poorly resolved, and so poor resolution of extant diversity or incomplete lineage sorting are possible alternative interpretations..

There are several range-restricted lineages within *Coryphaenoides*. The monophyletic *C. acrolepis*, *C. longifilis*, *C. pectoralis*, *C. filifer*, and *C. cinereus* are all restricted to the North Pacific Ocean, whereas *C. serrulatus*, *C. subserrulatus* and *C. mcmillani* are found in the Southern Ocean. All of the abyssal species reside in the Atlantic Ocean with the exception of *C. yaquinae*. Furthermore, four of the eight abyssal species are also found in other ocean basins (*C. armatus*, *C. leptolepis*, *C. carapinus*, *C. filicauda*). *Coryphaenoides rudis* is the only widely distributed non-abyssal species (and is considered circumglobal; Gaither et al., 2016). In general, abyssal species have broader depth ranges than non-abyssal species (Fig. 4) and also tend to have broad horizontal distributions; a finding that may be attributed to the older and abiotically more uniform waters at depths below 1000 m. Our phylogeny is consistent with a secondary invasion into abyssal waters, and perhaps from ancestors outside of the Atlantic Ocean (as suggested by the RASP analyses; Fig 3). However, most species at abyssal depths are found in the Atlantic, suggesting that adaptations associated with living at great depth originated there (Fig. 3). Some species may have then become dependent on abyssal habitat (e.g. *C. profundicolus*), while the recent origin of *C. yaquinae* together with its isolation in Pacific Ocean suggests origination of that species after its abyssal ancestor migrated into the North Pacific.

4.1. Testing morphological hypotheses

Cohen et al. (1990) included five subgenera within *Coryphaenoides* based on morphological characters (Fig. 3). Since then there has been some minor taxonomic "reshuffling" within the genus and six new species have been described. While the authors considered this arrangement to be a putative approximation, a revision is yet to be published, and therefore it remains the current hypothesis for intra-

14

generic relationships. In general the subgenera are not well-supported by the molecular phylogeny presented here (Fig. 3). The morphological subgenus *Bogoslovius* (1 sp.) separates *C. longifilis* from all other *Coryphaenoides,* but in our treatment *C. longifilis* was part of a well-resolved lineage that includes the closely related *C. acrolepis*, *C. pectoralis*, *C. cinereus* and *C. filifer* (the latter are all in the subgenus *Coryphaenoides* except *C. pectoralis*) (Fig. 3). Similarly, the morphological subgenus *Lionurus* separates *C. carapinus*, but in our phylogeny this species nested among the other abyssal species. Moreover, the abyssal species are split among three morphological subgenera (*Chalinura, Lionurus*, and *Nematonurus*), whereas in our phylogeny they formed a monophyletic lineage. Lastly, the non-abyssal *C. serrulatus* and *C. subserrulatus* are classified among abyssal species in the morphological subgenus *Chalinura,* but in our phylogeny they formed a distinct Southern Ocean lineage (with *C. mcmillani*) at the base of the tree. While there is no clear evidence at this time that any of these species groups warrant genus-level designations it is clear that a taxonomic revision, which incorporates molecular data, is warranted.

4.2. The origin of the abyssal lineage

Our findings support earlier, less complete phylogenies, which indicated that abyssal species form a monophyletic group. We were able to sample all but one abyssal species (*C. filicauda*). Two other species have been recorded at 3900 m (*C. ferrieri* and *C. lecointei*) but are not sampled here. Due to inadequate sampling of much of the world's deep oceans it is difficult to obtain a full representation of the species named in the genus; however, we provide a broad representation and nearly complete data for the abyssal lineage.

There is no available fossil record for *Coryphaenoides* and no reliable molecular clock has been calibrated, so any estimates of divergence times need to be interpreted with caution (Morita, 1999).

15

Using COI, and assuming a rate of transversions of 0.3-0.7% per million years, Morita (1999) calculated a divergence time between abyssal and the non-abyssal *Coryphaenoides* to be between 3.2 and 7.6 My. Similar calculations based on the COI data presented here provided comparable estimates of between 3.7 and 8.7 My. These dates place the split between abyssal and non-abyssal lineages as late as the Miocene.

Species diversity for the genus is highest in the Pacific with 35 of the 66 recognized species recorded only from that ocean. Only one of these is found at abyssal depths (*C. yaquinae*). Another 13 species are restricted to the Atlantic, including three abyssal species (*C. brevibarbis*, *C. mediterraneus*, and *C. profundicolus*). If we consider only the abyssal species, a Southern /Atlantic Ocean origin is supported by the RASP results with S-DIVA favoring an Atlantic origin (Fig.3, Node 10; Table S5). However, if we consider *C. striaturus* and *C. murrayi* to be abyssal species that have secondarily moved into shallower habitat, then a Southern Ocean/Pacific Ocean origin is favored by both analyses (Fig. 3, Node 8). There is circumstantial evidence to support a Southern Ocean origin for this lineage including the restriction of the abyssal *C. filicauda* to the southern hemisphere (not sampled here). Furthermore, the Southern Ocean species *C. serrulatus*, *C. subserrulatus, C. striaturus* and *C. murrayi* are morphologically similar to several abyssal species as indicated by their inclusion in the subgenus *Chalinura,* with *C. mediterraneus*, *C. brevibarbis*, *C. profundicolus*, and *C. leptolepis* (Fig. 3).

While the origin of the abyssal lineage remains ambiguous the geographic ranges of these species (and BMM results) indicate that diversification within the lineage likely occurred in the North Atlantic. When the Atlantic Ocean first began to form around 150 Mya (Sheridan et al., 1982) it consisted of two isolated basins in the north and south. Around 80 to 65 Mya a deep-water connection between the two basins was achieved with modern circulation patterns becoming established around 35 Mya (Schopf, 1980). While the ocean floor expanded as the continental plates moved away from each

16

other at the Mid-Atlantic Ridge, the depth of the ocean floor also increased with time. As new crust formed at about 2600 m it cooled and contracted increasing depths down to over 5500 m. The current topology of the Atlantic was achieved only around 10 Mya (Priede and Froese, 2013), very near the estimated time of colonization of the North Atlantic by the abyssal lineage of *Coryphaenoides* (3-8 Mya) but we urge caution when interpreting these values. Furthermore the exterior position of the abyssal lineage in the phylogenetic tree (Fig. 3) fits the concept of recent colonization of the deep sea by species from shallower habitats (Priede and Froese, 2013).

4.4. Diversity-depth pattern

Changes in species diversity and composition with depth are often attributed to the strong bathymetric gradients such as pressure, temperature, and food availability (Wilson and Hessler, 1987); factors that are thought to influence the scale and rate of evolutionary change. Compared to abyssal depths, the bathyal environment experiences a greater influx of nutrients, a more complex current regime, and more complex topography (e.g., deep canyons), which are reflected in the more pronounced horizontal heterogeneity of the demersal fauna (Rhoads and Hecker, 1994). As evolutionary dynamics can be influenced by environmental heterogeneity and the intensity of environmental gradients (Doebeli and Dieckmann, 2003; Filin et al., 2008; Leimar et al., 2008), the depth-differentiation hypothesis predicts that rates of evolution (population differentiation and speciation) are highest where heterogeneity is the greatest and environmental gradients are most intense (Rex and Etter, 2010). Evidence has been found in deep phylogenetic-level breaks along bathyal depths in clams (Goffredi et al., 2003), gastropods (Quattro et al., 2001), amphipods (France and Kocher, 1996), polychaetes (Schüller, 2011), and hydrozoans (Lindner et al., 2008), while population-level signals have been detected in bivalves (Etter et al., 2005), fishes (White et al., 2010) and octocorals (Quattrini et al., 2015).

17

For instance, in the Atlantic bivalve *Deminucula atacellana* population differentiation was greater among individuals separated by hundreds of meters along a vertical slope (~ 3000 m) than between individuals separated by thousands of kilometers across the ocean (Chase et al., 1998; Zardus et al., 2006). If species origination is higher in the bathyal zone, as predicted under the depth-differentiation hypothesis, one might expect species diversity to be highest at these depths.

Vinogradova (1962) was the first to show a unimodal diversity-depth pattern at broad taxonomic scales. She compiled global species records by depth and recovered a strong unimodal pattern with an increase in species diversity from the continental shelf to bathyal depths but with a dramatic drop in the number of species at abyssal depths. She found a peak in species diversity at around 2000 m; a pattern she also recovered from the analysis of individual groups. Since then similar unimodal patterns have been recovered at regional scales across a diversity of taxa (Dayton and Hessler, 1972; Priede and Froese, 2013; Rex, 1981) but there are exceptions (reviewed in Rex and Etter, 2010). Interestingly, species of *Coryphaenoides* mirror this pattern with a peak in the number of taxa found at 1000 m that gradually declines to abyssal depths (Fig. 4b).

However, species diversity patterns are not solely driven by differential speciation rates. The bathyal (800-4000 m depth) habitats in the North Atlantic support larger population sizes, as well as, greater numbers of species compared to the abyssal seafloor (Rex and Etter, 2000): a pattern that is at least in part due to changes in food availability with depth. The bathyal habitat is coupled to the deep-scattering layer (DSL); a mid-water (200-1000 m) mass of small fishes, cephalopods, crustaceans and zooplankton that provides a rich source and variety of prey items (Porteiro and Sutton, 2007; Robinson et al., 2010; Trueman et al., 2014). However, beyond 1500 m depth along the continental slopes the benthic and pelagic systems become increasing decoupled resulting in a the lack of food and low particulate organic carbon (POC) influx. With biomass in the abyss below ~1g m^{-2} (biomass at the

deepest bathyal depth; Rex and Etter, 2010) it is difficult to imagine how populations at these depths are maintained. This has led some to speculate that the abyss may function as an evolutionary sink with most populations maintained by immigration from bathyal depths (Rex et al., 2005), though this is disputed for wide ocean basins such as the Pacific (Hardy et al., 2015).

4.6. Conclusions

Our phylogenetic treatment of *Coryphaenoides* indicates that the morphologically based analyses used to date are insufficient to resolve the relationships among species. Taxa inhabiting waters deeper than 4000 m form a distinct and well-supported lineage, which also includes two non-abyssal species, *C. striaturus* and *C. murrayi,* which diverge from the basal node. Examination of individual gene trees suggests that these two species may have been involved in historical introgression events between abyssal and non-abyssal taxa, as their mtDNA is abyssal in origin but most of their nuclear alleles fall within the non-abyssal lineage, though further nuclear DNA data would help confirm this. All abyssal species are found in the North Atlantic with the exception of *C. yaquinae,* thus far found only in the North Pacific, and *C. filicauda,* thus far only in the Southern Ocean. Biogeographic reconstructions indicate that the genus may have originated in the Southern/Pacific Oceans with both dispersal and vicariant events playing important roles in diversification of the group. Species' distributions support this as well, with species diversity highest in the Pacific Ocean. The abyssal lineage seems to have arisen secondarily and likely originated in the Southern/Pacific Oceans, but maximum diversification of this lineage may have occurred in the North Atlantic Ocean. Importantly, our phylogeny indicates that adaptation to the deepest oceans happened only once in this group, suggesting that movement into the abyssal realm required unique adaptations; once this novel habitat was colonized the group diversified.

19

Acknowledgements

References

Baker, R.H., DeSalle, R., 1997. Multiple sources of character information and the phylogeny of Hawaiian drosophilids. Syst. Biol. 46, 654-673.

Brandt, A., N., B., Andres, H.-G., Brix, S., Guerrero-Kommritz, J., Muhlenhardt-Siegel, U., Wagele, J.-W., 2005. Diversity of peracarid crustaceans (Malacostraca) from the abyssal plain of the Angola Basin. Organisms Diversity and Evolution 5, 105-112.

Carney, R.S., 2005. Zonation of deep biota on continental margins. Oceanogr. Mar. Biol. Annu. Rev. 43, 211-278.

Carney, R.S., Haedrich, R.L., Rowe, G.T., 1983. Zonation of fauna in the deep sea. In: Rowe, G.T. (Ed.), The Sea. Wiley, New York, USA, pp. 371-398.

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

Chase, M., Etter, R., Rex, M., Quattro, J., 1998. Bathymetric patterns of genetic variation in a deep-sea protobranch bivalve, Deminucula atacellana. Mar. Biol. 131, 301-308.

Cohen, D., Inada, T., Iwamoto, T., Scialabba, N., 1990. FAO species catalogue. Vol. 10. Gadiform fishes of the world (Order Gadiformes). An annotated and illustrated catalogue of cods, hakes, grenadiers and other gadiform fishes known to data. FAO Fish. Synop., 452.

Dayton, P., Hessler, R., 1972. The role of disturbance in the maintenance of deep-sea diversity. Deep-Sea Res 19, 199-208.

Doebeli, M., Dieckmann, U., 2003. Speciation along environmental gradients. Nature 421, 259-264.

Etter, R., Rex, M.A., Chase, M.R., Quattro, J.M., 2005. Population differentiation with depth in deep-sea bivalves. Evolution 59, 1479-1491.

Etter, R.J., Rex, M.A., 1990. Population differentiation decreases with depth in deep-sea gastropods. Deep Sea Research Part A. Oceanographic Research Papers 37, 1251-1261.

Filin, I., Holt, R.D., Barfield, M., 2008. The relation of density regulation to habitat specialization, evolution of a species' range, and the dynamics of biological invasions. The American Naturalist 172, 233-247.

France, S.C., Kocher, T., 1996. Geographic and bathymetric patterns of mitochondrial 16S rRNA sequence divergence among deep-sea amphipods, Eurythenes gryllus. Mar. Biol. 126, 633-643.

Gaither, M.R., Bowen, B.W., Rocha, L.A., Briggs, J.C., 2016. Fishes that rule the world: circumtropical distributions revisited. Fish Fisheries. Online Early.

Glover, A.G., Smith, C.R., Paterson, G.L.J., Wilson, G.D.F., Hawkins, L., Sheader, M., 2002. Polychaete species diversity in the central Pacific abyss: local and regional patterns, and relationships with productivity. Mar. Ecol. Prog. Ser. 240, 157-170.

Goffredi, S., Hurtado, L., Hallam, S., Vrijenhoek, R., 2003. Evolutionary relationships of deep-sea vent and cold seep clams (Mollusca: Vesicomyidae) of the "pacifica/lepta" species complex. Mar. Biol. 142, 311-320.

Grassle, J.F., Maciolek, N.J., 1992. Deep-sea species richness: regional and local diversity estimates from quantitative bottom samples. The American Naturalist 139, 313-341.

Haedrich, R.L., Rowe, G.T., Polloni, P.T., 1980. The megabenthic fauna in the deep sea south of New England, USA. Mar. Biol. 57, 165-179.

Hardy, S.M., Smith, C.R., Thurnherr, A.M., 2015. Can the source–sink hypothesis explain macrofaunal abundance patterns in the abyss? A modelling test. Proc. R. Soc. B. The Royal Society, p. 20150193.

Hoelzel, A.R., 1998. Molecular genetic analysis of populations: a practical apporach. Oxford University Press, UK.

Iwamoto, T., Sazonov, Y., 1988. A review of the southeastern Pacific Coryphaenoides (sensu lato)(Pisces, Gadiformes, Macrouridae). Revisión de los Coryphaenoides (sensu lato)(Pisces, Gadiformes, Macrouridae) del sureste Pacífico. Proceedings of the California Academy of Sciences. 45, 35-82.

Iwamoto, T., Stein, D.L., 1974. Systematic review of the rattail fishes (Macrouridae: Gadiformes) from Oregon and adjacent waters. California Academy of Sciences, San Francisco (USA); Oregon State Univ., Corvallis (USA). School of Oceanography.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious

Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647-1649.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. accepted.

Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695-1701.

Laptikhovsky, V., Gaither, M.R., Black, A., 2013. A Pacific grenadier *Coryphaenoides acrolepis* in the south-west Atlantic and environmental changes in the Falkland deep seas. Marine Biodiversity Records 6.

Leimar, O., Doebeli, M., Dieckmann, U., 2008. Evolution of phenotypic clusters through competition and local adaptation along an environmental gradient. Evolution 62, 807-822.

Li, C., Orti, 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, 44.

Lindner, A., Cairns, S.D., Cunningham, C.W., 2008. From offshore to onshore: multiple origins of shallow-water corals from deep-sea ancestors.

Linley, T.D., Gerringer, M.E., Yancy, P.H., Drazen, J.C., Weinstock, C.L., Jamieson, A.J., in review. Fishes of the hadal zone: new observations and updated assessment of hadal fishes and a new snailfish (Liparidae) seen at 8145 m in the Mariana Trench Deep Sea Res. (I Oceanogr. Res. Pap.) submitted.

Lundsten, L., Schlining, K.L., Frasier, K., Johnson, S.B., Kuhnz, L.A., Harvey, J.B.J., Clague, G., Vrijenhoek, R.C., 2010. Time-series analysis of six whale-fall communities in Monterey Canyon, California, USA. Deep Sea Research Part I: Oceanographic Research Papers 57, 1573-1584.

Maddison, W.P., Maddison, D.R., 2001. Mesquite: a modular system for evolutionary analysis. Version 3.04 http://mesquiteproject.org.

Marshall, N., 1973. Subfamily Macrourinae. Fishes of the western North Atlantic. DM Cohen (ed.). Memoir. Sears Found. Mar. Res. Part 6, 530-533.

McClain, C.R., 2007. Seamounts: identity crisis or split personality? J. Biogeogr. 34, 2001-2008.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, LA, pp. 1-8.

Morita, T., 1999. Molecular phylogenetic relationships of the deep-sea fish genus Coryphaenoides (Gadiformes: Macrouridae) based on mitochondrial DNA. Mol. Phylogenet. Evol. 13, 447-454.

Okamura, O., 1970a. Fauna Japonica: Macrourina (Pisces). Acad. Press of Japan.

Okamura, O., 1970b. Studies on the macrouroid fishes of Japan: morphology, ecology and phylogeny. Rep. Usa Mar. Biol Stat. 17, 1-179.

Olabarria, C., 2005. Patterns of bathymetric zonation of bivalves in the Porcupine Seabight and adjacent Abyssal plain, NE Atlantic. Deep Sea Research Part I: Oceanographic Research Papers 52, 15-31.

Parr, A.E., 1946. The Macrouridae of the western North Atlantic and central American seas. The Bulletin of the Bingham Oceanographic Collection 10, 1-99.

Pineda, J., Caswell, H., 1998. Bathymetric species-diversity patterns and boundary constraints on vertical range distributions. Deep Sea Research Part II: Topical Studies in Oceanography 45, 83-101.

Porteiro, F.M., Sutton, T., 2007. Midwater fish assemblages and seamounts. Seamounts: Ecology, Fisheries, and Conservation 12, 101-116.

Priede, I., Froese, R., 2013. Colonization of the deep sea by fishes. J. Fish Biol. 83, 1528-1550.

24

Priede, I.G., Godbold, J.A., King, N.J., Collins, M.A., Bailey, D.M., Gordon, J.D., 2010. Deep-sea demersal fish species richness in the Porcupine Seabight, NE Atlantic Ocean: global and regional patterns. Mar. Ecol. 31, 247-260.

Quattrini, A.M., Baums, I.B., Shank, T.M., Morrison, C.L., Cordes, E.E., 2015. Testing the depth-differentiation hypothesis in a deepwater octocoral. Proc Biol Sci 282, 20150008.

Quattro, J., Chase, M.R., Rex, M.A., Greig, T.W., Etter, R.J., 2001. Extreme mitochondrial DNA divergence within populations of the deep-sea gastropod *Frigidoalvania brychia*. Mar. Biol. 139, 1107-1113.

Rex, M.A., 1981. Community structure in the deep-sea benthos. Annu. Rev. Ecol. Syst. 12, 331-353.

Rex, M.A., Etter, R.J., 2010. Deep-sea biodiversity: pattern and scale. Harvard University Press, Cambridge, MA.

Rex, M.A., McClain, C.R., Johnson, N.A., Etter, R.J., Allen, J.A., Bouchet, P., Waren, A., 2005. A source-sink hypothesis for abyssal biodiversity. The American Naturalist 165, 163-178.

Rhoads, D.C., Hecker, B., 1994. Processes on the continental slope off North Carolina with special reference to the Cape Hatteras region. Deep-Sea Research II 41, 965-980.

Roa-Varón, A., Ortí, G., 2009. Phylogenetic relationships among families of Gadiformes (Teleostei, Paracanthopterygii) based on nuclear and mitochondrial data. Mol. Phylogenet. Evol. 52, 688-704.

Robinson, C., Steinberg, D.K., Anderson, T.R., Arístegui, J., Carlson, C.A., Frost, J.R., Ghiglione, J.-F., Hernández-León, S., Jackson, G.A., Koppelmann, R., 2010. Mesopelagic zone ecology and biogeochemistry–a synthesis. Deep Sea Research Part II: Topical Studies in Oceanography 57, 1504-1518.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539-542.

Schopf, T.J.M., 1980. Paleoceanography. Harvard University Press, Cambridge, MA.

Schüller, M., 2011. Evidence for a role of bathymetry and emergence in speciation in the genus Glycera (Glyceridae, Polychaeta) from the deep Eastern Weddell Sea. Polar Biol. 34, 549-564.

Sheridan, R.E., Gradstein, F.M., Barnard, L.A., Bliefnick, D.M., Habib, D., Jenden, P.D., Kagami, H., Keenan, E.M., Kostecki, J., Kvenvolden, K.A., 1982. Early history of the Atlantic Ocean and gas hydrates on the Blake Outer Ridge: results of the Deep Sea Drilling Project Leg 76. Geological Society of America Bulletin 93, 876-885.

Snelgrove, P.V.R., Smith, C.R., 2002. A riot of species in an environmental calm,the paradox of species richness. Oceanogr. Mar. Biol. Annu. Rev. 40, 311-342.

Sorenson, M., Franzosa, E., 2007. TreeRot, version 3. Boston University, Boston, MA.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688-2690.

Stuart, C.T., Martinez Arbizu, P., Smith, C.R., Molodtsova, T., Brandt, A., Etter, R.J., Escobar-Briones, E., Fabri, M.C., Rex, M.A., 2008. CeDAMar global database of abyssal biological sampling. Aquatic Biology 4, 143-145.

Swofford, D.L., 2003. PAUP*: phylogenetic analysis using parsimony, version 4.0 b10.

Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56, 564-577.

Trueman, C., Johnston, G., O'Hea, B., MacKenzie, K., 2014. Trophic interactions of fish communities at midwater depths enhance long-term carbon storage and benthic production on continental slopes. Proceedings of the Royal Society of London B: Biological Sciences 281, 20140669.

Vinogradova, N.G., 1962. Vertical zonation of deep-sea bottom fauna in the ocean. Deep-Sea Research 8, 245-250.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D.N., 2005. DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B: Biological Sciences 360, 1847-1857.

White, T.A., Fotherby, H.A., Hoelzel, A., 2011. Comparative assessment of population genetics and demographic history of two congeneric deep sea fish species living at different depths. Mar. Ecol. Prog. Ser. 434, 155-164.

White, T.A., Stamford, J., Rus Hoelzel, A., 2010. Local selection and population structure in a deep-sea fish, the roundnose grenadier (Coryphaenoides rupestris). Mol. Ecol. 19, 216-226.

Wilson, G.D.F., Hessler, R.R., 1987. Speciation in the deep sea. Annu. Rev. Ecol. Syst. 18, 185-207.

Wilson, R., Siebenaller, J.F., Davis, B.J., 1991. Phylogenetic analysis of species of three subgenera of *Coryphaenoides* (Teleostei: Macrouridae) by peptide mapping of homologs of LDH-A4 Biochemical Systematics and Ecology 19, 277-287.

Wilson, R.R., 1994. Interrelationships of the subgenera of *Coryphaenoides* (Gadiformes: Macrouridae): comparison of protein electrophoresis and peptide mapping. Copeia, 42-50.

Wilson, R.R., Attia, P., 2003. Interrelationships of the subgenera of *Coryphaenoides* (Teleostei: Gadiformes: Macrouridae): synthesis of allozyme, peptide mapping, and DNA sequence data. Mol. Phylogenet. Evol. 27, 343-347.

Xia, X., 2013. DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Mol. Biol. Evol. 30, 1720-1728.

Xia, X., Lemey, P., 2009. Assessing substitution saturation with DAMBE. The phylogenetic handbook: a practical approach to DNA and protein phylogeny 2, 615-630.

Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. 26, 1-7.

Yu, Y., Harris, A.J., Blair, C., He, X., 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. Mol. Phylogenet. Evol. 87, 46-49.

Zardus, J.D., Etter, R.J., Chase, M.R., Rex, M.A., Boyle, E.E., 2006. Bathymetric and geographic population structure in the pan-Atlantic deep-sea bivalve *Deminucula atacellana* (Schenck, 1939). Mol. Ecol. 15, 639-651.

Figure legends

Fig. 1. Proposed phylogenetic hypotheses for the relationships of species of *Coryphaenoides* based on a) peptide mapping of homologus of LDH-A$_4$ (Least-Squares analyses of dissimilarity matrix; Wilson et al., 1991), b) 12S rRNA gene (Morita, 1999; percent bootstrap support for Maximum Parsimony analysis shown), and c) COI sequences (White et al., 2011; percent bootstrap support for Neighbor-Joining analysis shown). Brackets indicate abyssal species. Also see Wilson and Attia (2003). Photo of the abyssal grenadier *C. armatus* made available through Wikimedia Commons.

Fig. 2. Individual gene trees for species of *Coryphaenoides* inferred using Maximum Likelihood (ML) and Bayesian (BA) analyses. Trees are rooted with *Coelorinchus labiatus*. The topology shows the 50% Bayesian consensus tree with only percent bootstrap (BS) support ≥ 70% and posterior probabilities (PP) ≥ 0.70 shown. Asterisks indicate nodes with both BS values ≥ 95% and PP ≥ 0.95.For clarity duplicate names for samples from the same taxa that cluster together have been omitted. Taxa names in blue are those species found at abyssal depths while those in red indicate non-abyssal taxa that cluster with abyssal species in the mitochondrial trees (COI and 16S).

Fig. 3. Phylogeny of *Coryphaenoides* inferred using Maximum Likelihood (ML) and Bayesian (BA) analyses of the concatenated dataset (3048 bp; CO1, 16S, RAG2, and MYH6). The topology shows the 50% Bayesian consensus tree with onlypercent bootstrap (BS) support ≥ 70% and posterior probabilities (PP) ≥ 0.70 shown. Asterisks indicate nodes with both BS values ≥ 95% and PP ≥ 0.95.Subgenera based on Cohen et al. (1990) are shown in parentheses (*Bogoslovius*, Bog; *Chalinura*, Cha; *Coryphaenoides*, Cor; *Lionurus*, Lio; *Nematonurus*, Nem). Taxa names in blue are those species found at abyssal depths while those in red indicate non-abyssal taxa that cluster with abyssal species in the mitochondrial trees

29

(COI and 16S). Bars next to species names indicate geographic range according to key. Pie charts represent biogeographic states of ancestral nodes as determined using RASP (Reconstruct Ancestral State Phylogenies; see key).

Fig. 4. Depth range for species of *Coryphaenoides* (Table S1). Only species with both a minimum and maximum depth recorded in FishBase are included (53 of 65 species). A) Each bar represents the depth range for a single species: black bars species found at non-abyssal depths, blue bars are species found at abyssal depths, red bars are non-abyssal taxa that cluster with abyssal species in the consensus tree (see Fig. 3), and grey bars are species not sampled in this study. B) Bar graph representing the total number of species found at each depth (at intervals of 500 m). C) Bar graph showing the size of depth range for abyssal and non-abyssal species. Data is grouped into 500m bins with the proportion of species falling into each bin indicated.

Table 1. Contribution of each locus to the total dataset is described including number of characters per locus (after editing with Gblocks for 16S and RAG1), number of invariable, and number of segregating sites. Number of species and total number of individual represented in the dataset are listed.

gene/dataset	# sites	# invariable sites	# segregating sites	# species/individuals[a]
COI	626	367	201	27/71
16S	991	731	230	30/76
RAG1	702	529	141	30/73
MYH6	735	649	83	30/74
mtDNA (COI, 16S)	1617	1098	431	N/A
nDNA (RAG1, MYH6)	1440	1178	224	N/A

[a]Out of 30 species and 76 individuals (including three *Coelorinchus labiatus* individuals)

Table S1. List of depth ranges in meters for species of *Coryphaenoides*. Data were retrieved from Fishbase (http://www.fishbase.org/search.php). References listed are additional citations found in the literature and sources often conflicted. Depths used in Fig. 4 are underlined and represent the best estimates based on the available data.

Table S2. Tissues used in this study including species, common name, subgenus (according to Cohen et al. 1990), tissue source (listed as collector or repository and unique identifier or musuem catalog number), collection location, and GenBank accession numbers. JNMNS, Japan National Museum of Nature and Science; KU, the University of Kansas Biodiveristy Institute; NIWA, National Institute of Water and Atmospheric Research, New Zealand. Tissues provided by Adela Roa-Varón were originally collected by Tomio Iwamoto.

Table S3. List of genes and the forward (F) and reverse (R) primers used for amplification (A) or sequencing (S). Annealing temperatures for each primer combination is shown. For the RAG1 and

31

MYH6 genes two PCRs were required. For these both sets of primers are listed for each step (PCR steps).

Table S4. Partitioning schemes and substitution models used for Bayesian and RaxML analyses.

Table S5. Species distributions used for S-DIVA and BMM analyses in RASP. See Table S1 for sample information. The distributions are A=Pacific Ocean, B = Atlantic Ocean, and C = Southern Ocean.

Table S6. Results from S-DIVA and BMM analyses ran in RASP. Probabilities of ancestral range are listed for each node. The distributions are A=Pacific Ocean, B = Atlantic Ocean, and C = Southern Ocean. * is unclassified according to RASP.

(a) Wilson et al., 1991
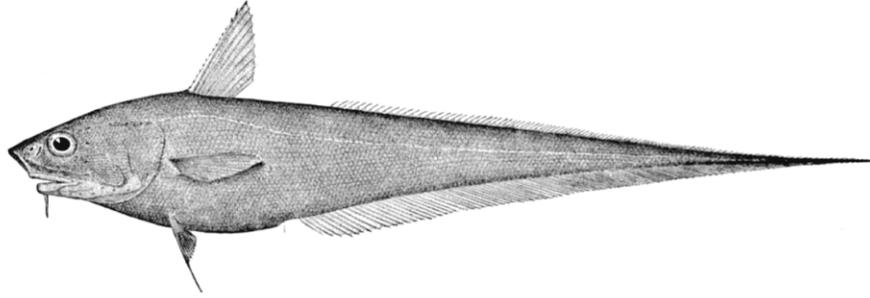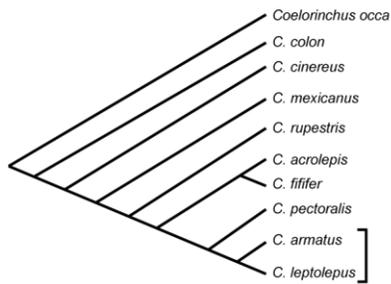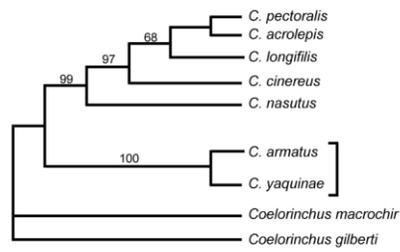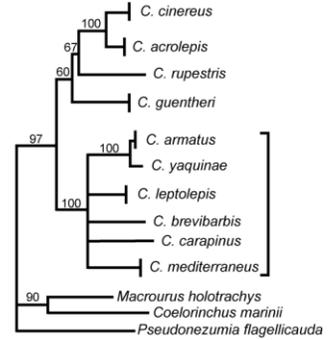
Coelorinchus occa
C. colon
C. cinereus
C. mexicanus
C. rupestris
C. acrolepis
C. fififer
C. pectoralis
C. armatus
C. leptolepus

(b) Morita, 1999

C. pectoralis
C. acrolepis
C. longifilis
C. cinereus
C. nasutus
C. armatus
C. yaquinae
Coelorinchus macrochir
Coelorinchus gilberti

68
97
99
100

(c) White et al., 2011

C. cinereus
C. acrolepis
C. rupestris
C. guentheri
C. armatus
C. yaquinae
C. leptolepis
C. brevibarbis
C. carapinus
C. mediterraneus
Macrourus holotrachys
Coelorinchus marinii
Pseudonezumia flagellicauda

100
67
60
97
100
100
90

33

(a) COI
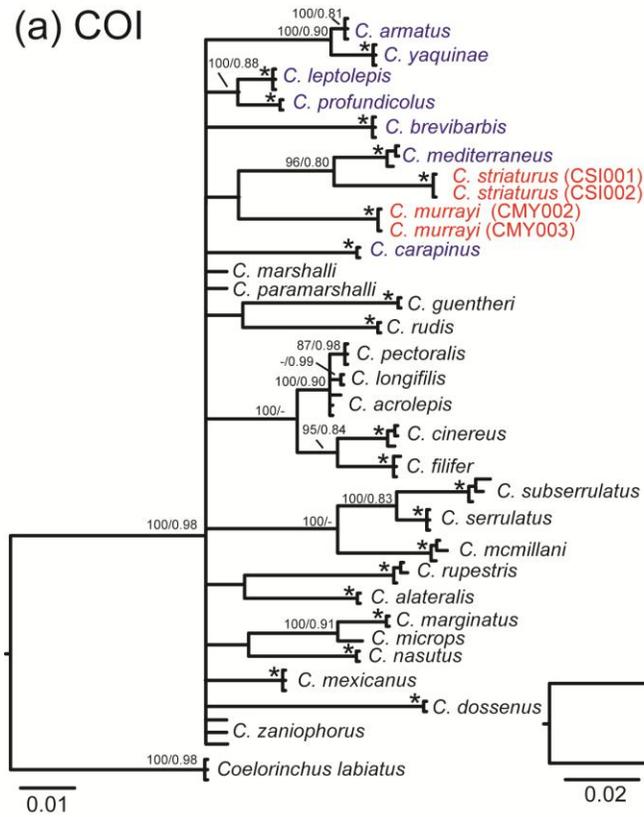
(b) 16S
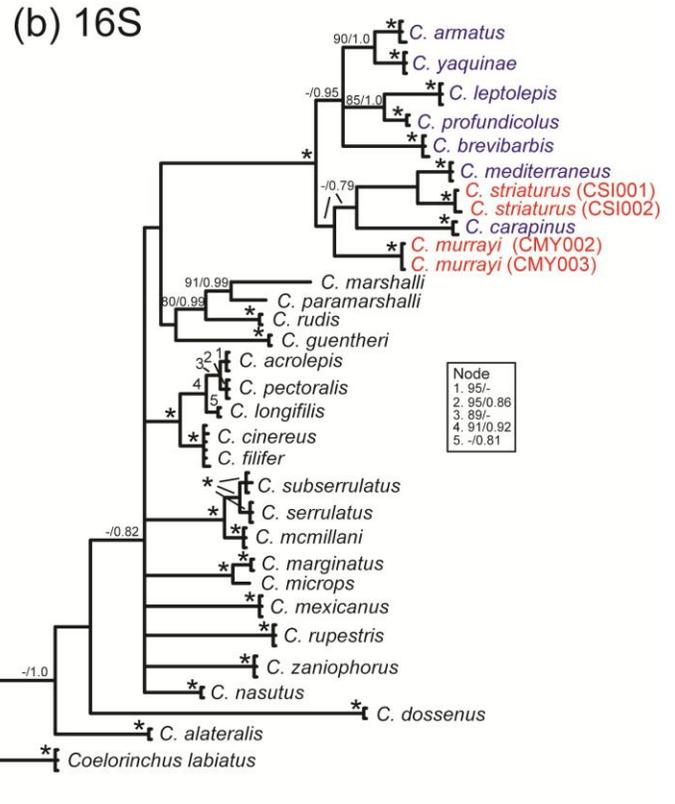
(c) RAG1
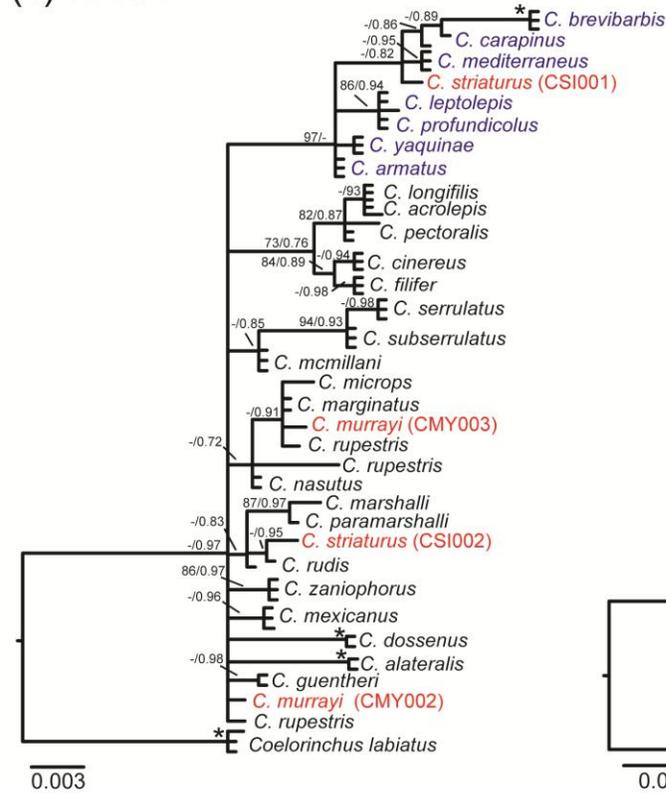
(d) MYH6
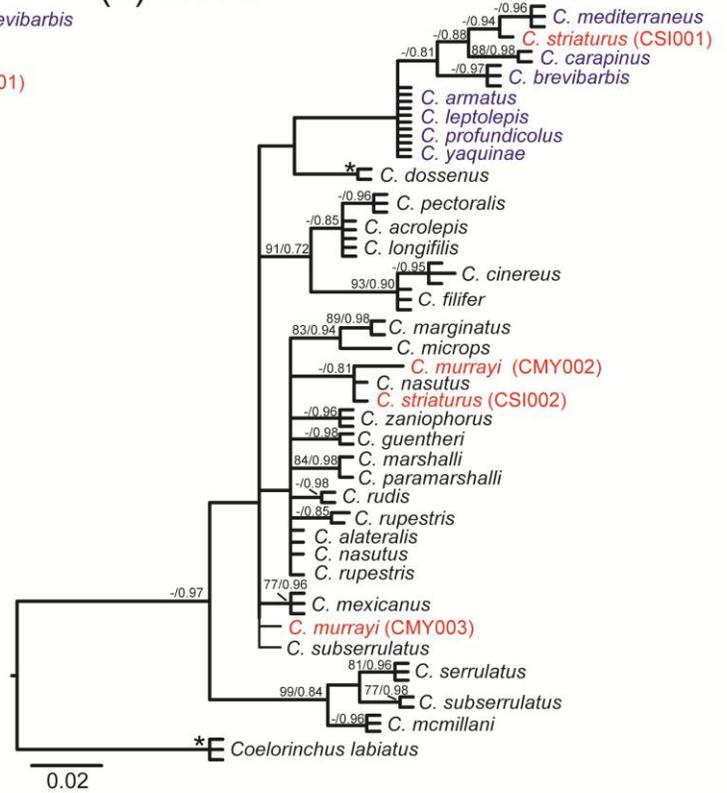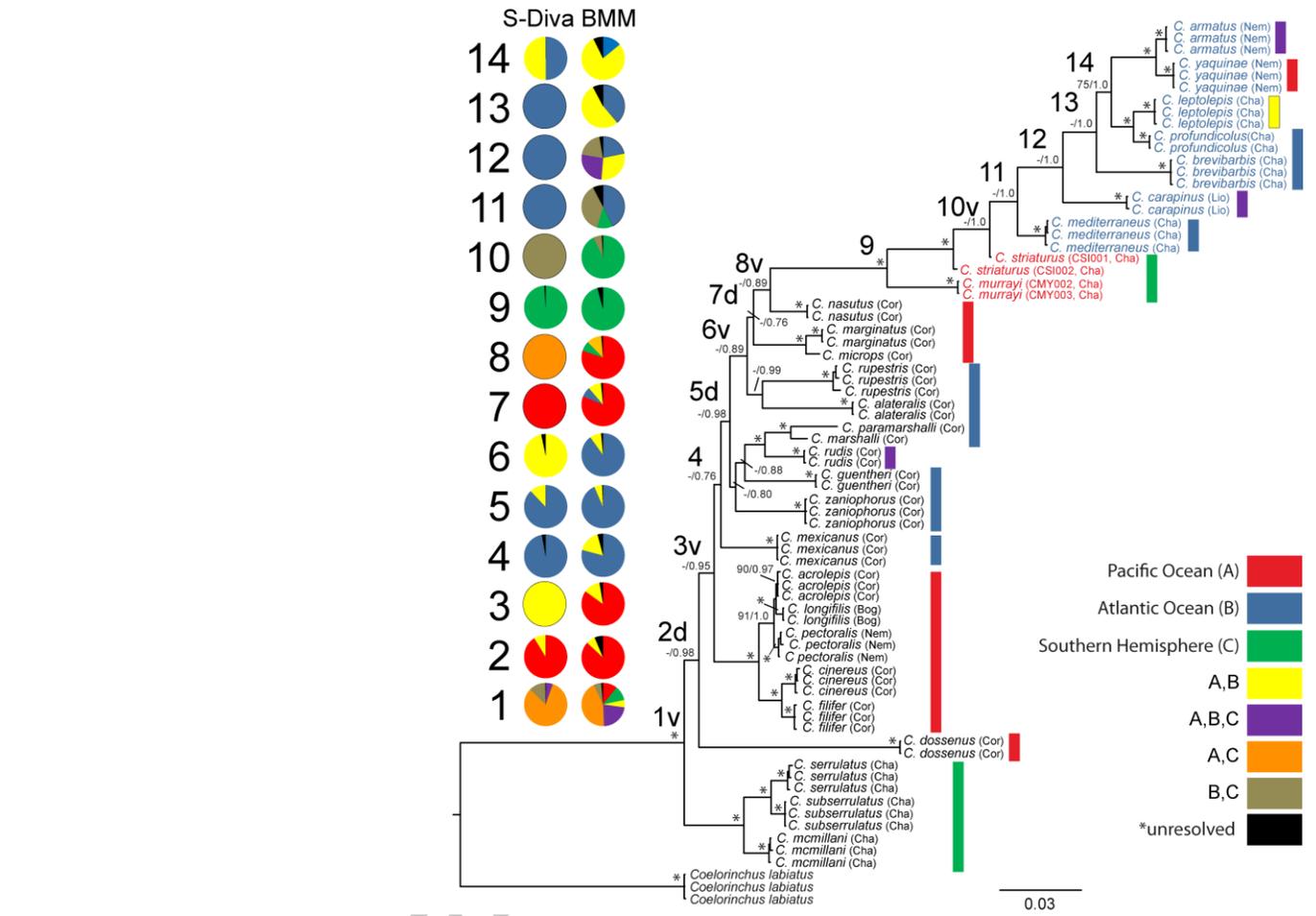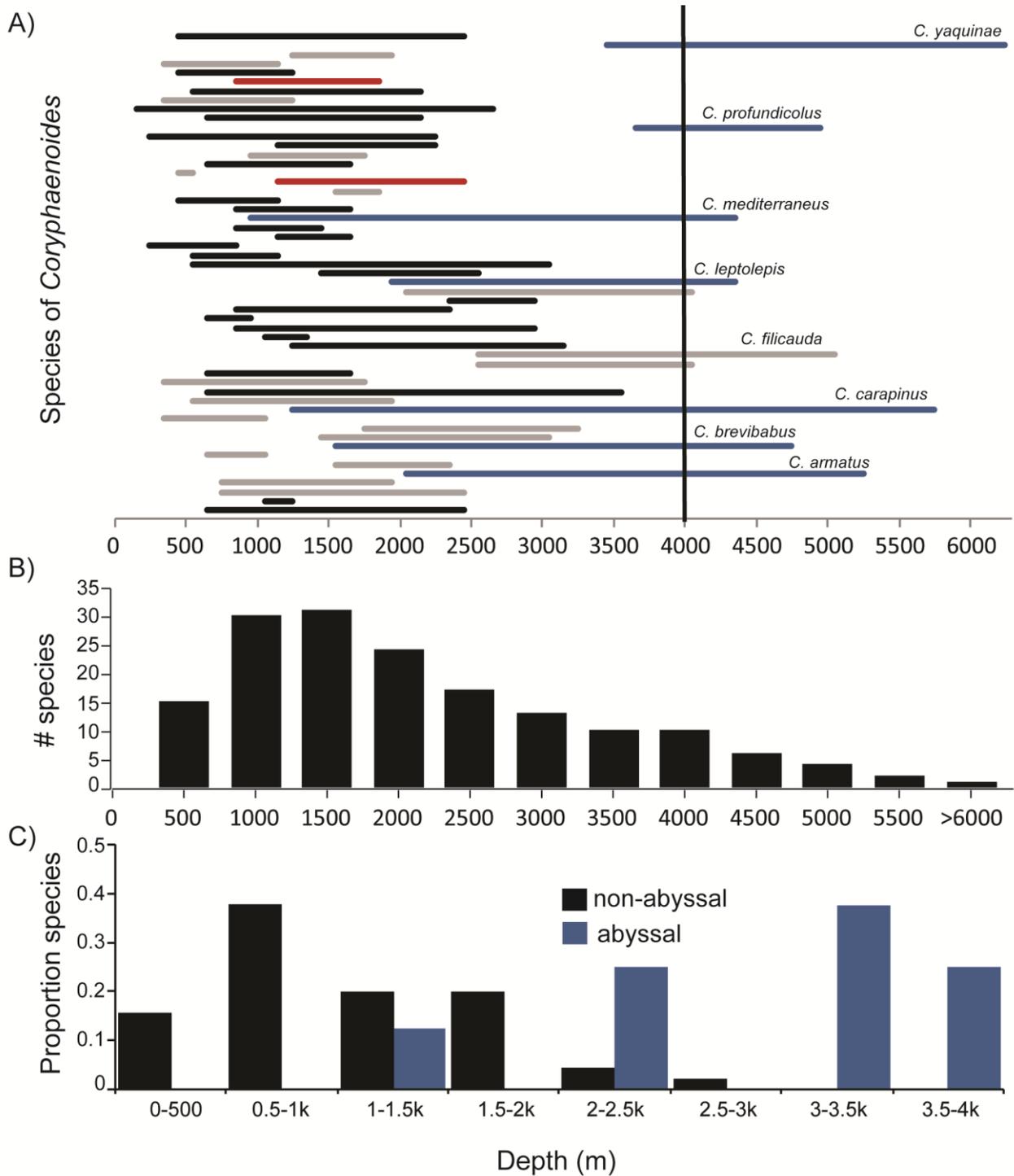
Graphical abstract
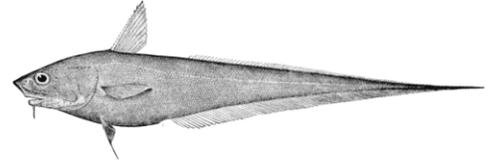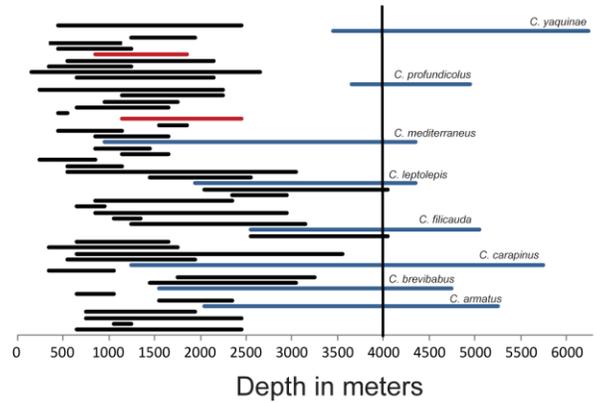
Highlights

- Phylogeny for the genus *Coryphaenoides* (Gadiformes: Macrouridae) that represents the breadth of habitat use and distribution.

- Abyssal species form a well-supported lineage that includes two non-abyssal species diverging from the basal node.

- The genus may have originated in the Southern and Pacific Oceans where contemporary species diversity is highest.

- The abyssal lineage likely arose secondarily and in the South/Pacific Oceans but diversified in the North Atlantic.