Molecular phylogenetics of moray eels (Muraenidae) demonstrates multiple origins of a shell-crushing jaw (Gymnomuraena, Echidna) and multiple colonizations of the Atlantic Ocean

Joshua S. Reece, Brian W. Bowen, David G. Smith, Allan Larson

1. Introduction

Moray eels (Muraenidae) are apex predators on coral reefs around the world, but they are not well studied because their cryptic habitats and occasionally aggressive behaviors make them difficult to collect. We provide a molecular phylogeny of moray eels including 44 species representing two subfamilies, eight genera, and all tropical ocean basins. Phylogenetic relationships among these taxa are estimated from portions of mitochondrial loci cytochrome b (632 bp) and cytochrome oxidase subunit 1 (596 bp), and portions of the nuclear loci RAG-1 (421 bp) and RAG-2 (754 bp). We test four sets of contrasting phylogenetic hypotheses using Bayes Factors, Shimodaira–Hasegawa tests, and Templeton tests. First, our results support the subfamily-level taxonomic distinction among true morays (Muraeninae) and snakemorays (Uropterygiinae), statistically rejecting hypotheses of non-monophyly for each subfamily. Second, we reject a monophyletic grouping of the genera Gymnomuraena and Echidna, which share a durophagous (shell-crushing) cranial morphology and dentition, indicating that the durophagous characters are not homologous. Third, we demonstrate that durophagous feeding habits and associated morphological characters have evolved in parallel in an ancestor of Gymnomuraena and at least three additional times within the genus Echidna. Finally, the tree topology indicates multiple invasions of the Atlantic from the Indo-Pacific, one of these occurring immediately prior to formation of the Isthmus of Panama approximately 2.8 MYA (million years ago) and one or two others occurring in the early to mid Miocene. Cladogenesis occurring within the Atlantic during the mid Miocene and Pliocene also contributed to moray species diversity. These data include a pair of sister species separated by the Isthmus of Panama, allowing a time-calibrated tree with an estimated crown age for Muraenidae at between 41 and 60 MYA, consistent with fossil evidence. Most lineage accumulation within morays occurred from the late Oligocene (24–27 MYA) through the Miocene (5–23 MYA) to the late Pliocene (~2.5 MYA).

© 2010 Elsevier Inc. All rights reserved.
often occur sympatrically and can be diagnosed with varying levels of resolution by the position of dorsal and anal fins, tooth placement and morphology, larval pigmentation, adult color pattern, and vertebral numbers (Böhleke et al., 1989).

Mehta (2008) described two morphotypes within Muraenidae corresponding to piscivorous (fish eating) and durophagous (shell-crushing) jaws. These morphotypes differ in their cranial and pharyngeal jaw structures and most noticeably in their dentition. Piscivorous species have elongate oral jaws with numerous long, recurved fang-like teeth. The oral jaws of durophagous morays are short, recurved, and covered with numerous molariform teeth for crushing shelled prey. Dietary studies confirm that species classified by their cranial morphology, oral jaws and dentition (as described above) as durophagous feed primarily on crustaceans, whereas those with piscivorous morphology feed primarily or exclusively on fishes (literature reviewed in Table 1 of Mehta, 2008). Taxonomically, the piscivorous morphology characterizes the largest genus of Muraeninae (Gymnothorax) and most of the remaining genera, and it resembles the jaws of uropterygines, which lack the durophagous form. The durophagous morphology occurs in the genus Echidna (with 11 recognized species) and the monotypic genus Gymnomuraena. It is unknown whether the durophagous morphology is homologous among these two genera and 11 species, or whether they include parallel origins of a durophagous jaw from a (presumed) piscivorous ancestor.

Muraenidae includes approximately 150 Indo-Pacific species and 50 Atlantic species (Böhleke et al., 1989; Böhleke and Smith, 2002). Most reef fishes have a center of diversity in the Indo-Pacific (Briggs, 1999), which potentially includes the sites of origin of many taxa. Butterflyfishes (Fessler and Westneat, 2007) and wrasses of the family Labridae (Westneat and Alfaro, 2005) show multiple invasions from the Indo-Pacific into the Atlantic. Conversely, Atlantic gobies of the genus Gnatholepis (Rocha et al., 2005), pygmy angelfishes of the genus Centropyge (Bowen et al., 2006) and the wrasse genera Halichoeres (Barber and Bellwood, 2005) and Thalassoma (Bernardi et al., 2004), represent single colonization events from the Indo-Pacific via southern Africa. At least three well-surveyed groups of reef fishes (damsel, some wrasses, and parrotfishes) show multiple invasions of and local diversification within the western Atlantic (Barber and Bellwood, 2005; Robertson et al., 2006; Rocha et al., 2008). Multiple genera of moray eels occur in the Atlantic, and the widespread genus Gymnothorax occurs in every ocean basin. It is unclear whether the occurrence of Gymnothorax species in the Atlantic represents multiple invasions from the Indo-Pacific or a single invasion and subsequent speciation in the Atlantic.

We generate a phylogeny of 44 moray eel species to test four major hypotheses about moray eel evolutionary history and biogeography. First, we test hypotheses of monophyly for the morphologically diagnosable subfamilies Muraeninae and Uropterygiinae. Second, we test the hypothesis that the durophagous genera Gymnomuraena and Echidna form a monophyletic group and that their evolutionarily derived feeding morphologies are homologous. Prior taxonomic grouping of the monotypic genus Gymnomuraena with Echidna is based primarily on shared durophagy and not well supported by other morphological characters. Third, we test monophyly of the genus Echidna and the associated hypothesis that durophagy is homologous within the genus. Finally, we test the hypothesis that
Atlantic members of Gymnothorax form a monophyletic group and represent a single invasion from the Indo-Pacific. Fig. 1 summarizes expected topologies corresponding to each of the four hypotheses. We use fossil and biogeographic data to produce a time-calibrated phylogeny to test these hypotheses and to estimate a crown age for the extant moray lineages, and approximate timing of interoceanic exchanges between the Atlantic and Indo-Pacific.

2. Materials and methods

2.1. Sequence generation

We acquired specimens from tissue banks, personal collections, collaborators, and the pet trade when voucher and locality information were available. Author DGS at the United States National Museum confirmed the identity of a voucher specimen for each DNA sequence in this study. We extracted DNA using Viogene DNA Kits (www.viogene.com) and manufacturer’s protocols. Polymerase chain reactions featured a total volume of 25 μL including 5 μL of Promega (www.promega.com) 5× buffer, 2.5 μL of 25 mM MgCl₂, 2.5 μL of 0.2 μM dNTPs, 2.5 μL of 0.2 μM of each primer, 0.125 μL (1 unit) of Promega GoTaq DNA polymerase, and 2 μL of template DNA at approximately 5 ng/μL. A 632-bp fragment of cytochrome b (CYB) was amplified using the primers LI14725 (5′-GTG ACT TGA AAA ACC ACC GTT G-3′) and FishF2 (5′-TCG ACT AAT CAT AAA GAT ATC GGC AC-3′) (Taberlet et al., 1992) and an annealing temperature of 50 °C. A 596-bp fragment of cytochrome oxidase subunit 1 (COI) was amplified using primers FishF2 (5′-TCG ACT AAT CAT AAA GAT ATC GGC AC-3′) and FishR2 (5′-ACT TCA GGA TGA CCG AAG AAT CAG AA-3′) (Ward et al., 2005) and annealing temperature of 50 °C. A 421-bp fragment of the nuclear recombination activation gene RAG-1 was amplified using primers RAG1-F3 (5′-GCC TCA GAA AAC ATG GTG CTG-3′) and RAG1-R3 (5′-CCA CAC AGG TTT CAT CTG GA-3′) (Reece et al., 2010) with an annealing temperature of 50 °C. A 754-bp fragment of the nuclear recombination activation gene RAG-2 was amplified using primers RAG2-F3 (5′-AGG TGA CCC TTC GTG AGT-3′) and RAG2-R3 (5′-ATG AGG CTC CCT TCC AAA GT-3′) (Reece et al., 2010) at an annealing temperature of 52 °C. The thermal profiles for PCR were 95 °C for 3 min, followed by 35 cycles of 95 °C for 1 min, annealing temperature for 40 s, and 72 °C for 45 s, with a final elongation at 72 °C for 7 min.

PCR products were visualized through 1.5% agarose gel electrophoresis and purified using Exo-Sap or Viogene Gel Purification Kits using manufacturer’s protocols. Sequences were generated on ABI 3130 and ABI 3130 Automated DNA sequencers at the Washington University Genome Sequencing Center and the Smithsonian Museum Support Center using PCR primers listed above. DNA sequences were manually edited using Sequencher v.4.8, and aligned by hand. Author JSR performed all sequencing and editing. For nuclear markers, heterozygous positions were identified by a secondary peak in the electropherograms reaching at least 25% of the intensity of the primary peak. Genetic phases of nuclear sequences with more than a single heterozygous site were estimated using a Bayesian approach implemented in the software program Phase v.2.1 (Stephens and Donnelly, 2003; Stephens et al., 2001). All Phase analyses were run through five iterations with different random-number seeds and run for 1000 iterations with a single thinning interval and 100 burn-in iterations. Consistency of results was determined by examining allele frequencies and coalescent goodness-of-fit measures estimated for each of the four runs. If haplotypes could not be estimated with 90% posterior probability, each ambiguous site was coded as missing data. Less than 3% of all nucleotide characters were coded as missing data under this criterion.

2.2. Phylogenetic analyses

We calculated Tajima’s D (Tajima, 1989) test for selection in DNAps v.5.1.0 (Rozas et al., 2003) to ensure that the loci included in this study conformed to neutral expectations. Phylogenetic trees were constructed in the programs MrBayes v.3.1 (Ronquist and Huelsenbeck, 2003) and BEAST v.1.5.4 (Drummond and Rambaut, 2007). We used four anguilliform outgroups (Fig. 2) for combined portions of the mitochondrial genes CYB and COI, and for the nuclear genes RAG-1 and RAG-2. A model of evolution was computed for three partitions within each gene region, corresponding to 1st, 2nd and 3rd codon positions, using jModelTest v.0.1.1 (Posada, 2008). Each of three MrBayes runs, corresponding to combined mtDNA, RAG-1, and RAG-2, consisted of 3,000,000 iterations of four chains replicated in two independent runs with a sampling interval of 100 iterations and burn-in of 7500 runs. Each run reached stationarity with these values; effective sample-size values were over 200 for each parameter estimated, and the standard deviations between independent runs had stabilized and were below 0.001. Stationarity of all runs was interpreted using Tracer v.1.5.0 (Rambaut and Drummond, 2007). To demonstrate that a concatenated dataset was appropriate for these gene regions, we conducted pairwise Shimodaira–Hasegawa (S-H) tests (Shimodaira and Hasegawa, 1999) among the gene-tree topologies (mtDNA, RAG-1, RAG-2). After all gene topologies were identified as congruent by S-H tests, we used a concatenated dataset (maintaining the partitioning strategy described above for a total of 4 genes × 3 partitions within each gene = 12 partitions) to reconstruct evolutionary relationships using the same run parameters as above but for 10,000,000 generations. The phylogenetic reconstruction executed in BEAST differed only in the use of time calibrations, and those methods are discussed below.

We tested four phylogenetic hypotheses that correspond to (1) the two subfamilies of Muraeninae and Uropterygidae are monophyletic, (2) the durophagous genera Echidna and Gymnomuraena form a monophyletic group with respect to the genus Gymnothorax, (3) genus Echidna is monophyletic, and (4) Atlantic species of Gymnothorax form a monophyletic group. The taxonomic sampling scheme used to address each of these tests is described below. Each of these hypotheses was tested by three metrics (in order of increasing stringency) to determine whether the data statistically discriminate hypotheses 1–4 from contrasting alternatives: Bayes Factors (Kass and Raftery, 1995), S-H tests (Shimodaira and Hasegawa, 1999), and Templeton tests (Templeton, 1983). Bayes Factors were calculated by comparing the logarithms of overall likelihood scores of the favored topology to the best topology compatible with the contrasting hypothesis. This value was then doubled and interpreted as the Bayes Factor (2logB10) according to Table 2 in Kass and Raftery (1995). A Bayes Factor exceeding 10 is considered definitive rejection of the less likely topology (Kass and Raftery, 1995). S-H and Templeton tests were executed in PAUP v.4.0 (Swoford, 2003). S-H tests with 1000 bootstrap replicates checked statistical significance of differences in likelihood values of the favored tree and those of the 100 most likely trees sampled from the Bayesian posterior probability distribution of the contrasting hypothesis. We considered a hypothesis rejected if all of the 100 best trees compatible with it were statistically rejected as less compatible with the data than the favored tree. For Templeton tests, we constrained a parsimony analysis to find the shortest tree compatible with hypotheses 1–4 or their converse as appropriate, and then evaluated statistical significance of the differences in length of the contrasting pairs of topologies.

All four phylogenetic hypotheses were tested using the concatenated gene trees comprising the three gene regions deemed congruent by pairwise topological tests. We also evaluated support for each hypothesis among each of the three gene trees by the same
topological tests. We report only the results from the concatenated gene tree because the full concatenated dataset never contradicted support from any of the constituent gene trees (see Section 3).

All species sampled in this study were used to test each hypothesis, and the pertinent taxonomic coverage is as follows: To test for monophyly of Muraeninae and Uropethygiinae, we used samples from three species of two genera in Uropethygiinae: Uropethygius macrocephalus, U. fuscoguttatus, and Scuticaria tigrina and 41 Muraeninae species representing the genera Gymnomuraena, Gymnotherax, Echidna, Enchelynassa, Rhinomuraena, and Enchelycore. We tested monophyly of each subfamily and evaluated the possibility that either family is monophyletic and nested within a paraphyletic group comprising species from the other subfamily. Our second hypothesis was that the durophagous species Gymnomuraena zebra, Echidna nebula, E. polyzona, E. rhodochilus, and E. leucotaenia form a monophyletic group relative to the 36 piscivorous species sampled. Monophyly of Echidna was tested using the four Echidna species relative to all other samples. Lastly, members of Indo-Pacific Muraeninae were compared to five Atlantic Gymnotherax species (G. saxicola, G. vicinus, G. acellatus, G. moringa, and G. miliaris) to test the hypothesis of a single invasion into the Atlantic from the Indo-Pacific.

The program BEAST was used to estimate both a phylogenetic topology and the timing of cladogenesis events based on two independent and cross-validated calibrations. The BEAST analysis included the same partitioning strategy and models of evolution as did the MrBayes run. A joint tree was estimated from the four gene trees, and molecular clock and mutation models were unlinked across all 12 partitions. A relaxed uncorrelated lognormal clock prior was used, along with a Yule speciation process, per recommendations for interspecific phylogenies (Drummond and Rambaut, 2007) and produced a time-calibrated phylogenetic tree because the full concatenated dataset...
for the sister taxa *G. miliaris* and *G. meleagris*, we calculate a mutation rate of 2.1% divergence per MY (between lineages) for portions of the CYB gene and 1.2% for portions of the COI gene; both values mirror those reported in the literature for marine fishes (see Section 4). Calibrations for RAG-1 and RAG-2 based on combined mtDNA divergences yield estimated rates of 0.45% and 0.35% divergence per MY (between lineages), respectively.

### 4. Discussion

Despite the potential difficulties in identifying and classifying moray species (Randall, 2007), our results support the deepest morphology-based taxonomy within Muraenidae; subfamily classifications of Muraeninae and Uropterygiinae. These subfamilies are distinguished by a suite of diagnostic characteristics, including restriction of the anal and dorsal fins to the posterior tip of the tail in Uropterygiinae and their extension beyond this region in Muraeninae. Based on our current sampling of Muraenidae, the two subfamilies are monophyletic and sister taxa. A previous study with more thorough sampling of Uropterygiinae (Loh et al., 2008) is compatible with this conclusion. Our molecular data support the inferences that morphological synapomorphies of the Muraeninae (loss of hypobranchial elements from the pharyngeal jaws) and Uropterygiinae (dorsal and anal fins restricted to tip of tail) are robust, thus we doubt that additional sampling of species will render either family non-monophyletic.

The zebra moray, *Gymnomuraena zebra*, represents a monotypic genus characterized by a strongly ossified skeleton, enlarged dermal bones, and posterior placement of the anus. *Gymnomuraena zebra* and members of the genus *Echidna* feed almost exclusively on crustaceans, and all share common morphological features associated with this diet, including rounded, molariform teeth, and short, recurved jaws. *Gymnomuraena zebra* appears to have diverged from a common ancestor of all other Muraeninae approximately 43 MYA, and based on our phylogenetic inferences, its durophagous feeding and associated morphological characters evolved separately from those of other durophagous muraenines. Nonmonophyly of the durophagous genus *Echidna* indicates that durophagous feeding and associated jaw morphology evolved at least three additional times in Muraeninae; six unsampled species of *Echidna*, including one in the Atlantic, might include additional origins of durophagous feeding. Morphology-based taxonomy in this case is confounded by parallel evolution of durophagy and associated jaw morphology in Muraeninae. Remarkably, the nominal *Echidna* species in our survey are separated by about 20 MY. This finding will warrant further morphological investigations and revision of the genus *Echidna*. The genus *Echidna* is formally recognized by a type specimen of *E. nebulosa*, which in our sampling represents the sister taxon to a clade comprising *Gymnothorax*, *Enchelycore*, *Enchelynassa*, *Rhynomuraena*, and the remaining sampled species of *Echidna*. The genus *Echidna* is perhaps best retained for *E. nebulosa* and any unsampled species later found to be closer to *E. nebulosa* than to any species of *Gymnothorax*. Because *Gymnothorax* is rendered paraphyletic by the remaining sampled species of *Echidna* and by the genera *Enchelycore*, *Enchelynassa*, and *Rhynomuraena*, to reflect a monophyletic taxonomic grouping, all of these species are best placed in genus *Gymnothorax*.

Taxonomic families and subfamilies of reef fishes typically have broad or even cosmopolitan distributions with a center of species
diversity in an area called the Coral Triangle (Allen, 2008; Briggs, 2009; Veron et al., 2009). Species diversity decreases with distance from the Coral Triangle, and is almost universally lower in the Atlantic (e.g., Briggs, 1995; Myers, 1991; Springer, 1982; Veron, 1995). Some lineages of reef fishes have invaded the Atlantic multiple times (e.g., Flessier and Westneat, 2007), and others have diversified following a single invasion (Muss et al., 2001). These invasions could occur from the proto-eastern Pacific prior to closure of the Isthmus of Panama 2.8–3.1 MYA (Duque-Caro, 1990; Lessios, 2008), from the Indian Ocean via southern Africa (Gordon, 2003), or through the now closed (15–20 MYA) Tethys Sea (Smith et al., 2004). The southern African pathway effectively closed to tropical fauna about 2.5 MYA with the onset of cold-water (Benguela) upwelling, but opened intermittently at the end of each Pleistocene glacial cycle (Peeters et al., 2004; Shannon, 1985). The five Atlantic Gymnothorax species sampled represent at least two and probably three invasions from the Indo-Pacific. Cladogenesis associated with formation of the Isthmus of Panama explains the separation of the Atlantic G. miliaris from its Pacific sister species G. melanoleucus 2.9 MY. Two other pairs of Atlantic sister species are G. morgina and G. vicinus (14.2 MY; Fig. 2) and G. ocellatus and G. saxicolor (3.2 MYA; Fig. 2). Our phylogenetic analysis indicates that the ancestral lineages of these two pairs of species each separated from its closest Indo-Pacific relative in the early to mid Miocene (16–21 MYA), coincident with the closing of the Tethys Sea. Although our favored topology indicates two separate invasions of the Atlantic in the early to mid Miocene, branch support is not sufficient to reject the alternative hypothesis of a single Atlantic invasion followed by return of a descendant lineage to the Indo-Pacific. The Atlantic contains most members of the unsampled genus Muraena, which is hypothesized to be a relic formerly distributed across the Tethys Sea 15–20 MYA (Almada et al., 2009; Smith et al., 1994). This time frame is compatible with our inferences of the oldest invasions of the Atlantic by Gymnothorax. We conclude that recurring invasions from the Indo-Pacific and in situ speciation both contribute to Atlantic moray diversity (Fig. 2).

Although our taxonomic sampling is too limited to estimate lineage-accumulation rates for moray eels, our results indicate that lineage accumulation was greatest in the Miocene (approximately 5–23 MYA), with relatively few lineages predating a late Oligocene date of ~25 MYA and few origination occurring in the Pliocene (approximately 2.5–5 MYA). These dates are consistent with limited fossil evidence for the first appearance of moray eels at 34–54 MYA (Sels, 1953), and of morphologically/Gymnothorax-like eels in the Mediterranean as recently as 5.3 MYA (Arambourg, 1927; Gaudant, 2002). Our prior work documents geographic genetic continuity within species of Gymnothorax (Reece, 2010; Reece et al., 2010), indicating that widespread discovery of cryptic species in moray eels is unlikely, in contrast to results obtained for many terrestrial and freshwater vertebrate taxa. Our estimate that extant moray species trace their cladogenesis primarily to events occurring between the late Oligocene and Pliocene is unlikely to be biased by overlooking large numbers of younger, cryptic species, and is consistent with a broad taxonomic review of marine fish diversification (Rocha and Bowen, 2008). Our prior results also indicate that some Gymnothorax species have likely maintained geographically widespread distributions throughout the Indo-Pacific for much of their evolutionary history. Although climatic cycles produce recurring isolation of marine fish populations across the Sunda Shelf on a scale of ~100,000 years, geographically isolated populations appear rarely to evolve reproductive isolation on this timescale, and they subsequently merge genetically when sea levels rise again to permit gene flow across the Sunda Shelf. The estimated mitochondrial-haplotype coalescence time within G. tundulus is late Miocene (~5.9 MYA), and comparable phylogeographic studies of three additional moray species indicate that mitochondrial-haplotype coalescence exceeds 1.7 million years in each case (Reece, 2010; Reece et al., 2010). If these results are typical for moray species, young cryptic species are probably uncommon; therefore, our finding that most lineage accumulation dates from the late Oligocene through the Miocene to the early Pliocene should be robust to further species sampling.

Acknowledgments

Funding for this project was provided by National Science Foundation grants DDDG-0909756 to A.L. and J.S.R., OCE-0453167 and OCE-0929301 to B.W.B. Additional funding was provided to J.S.R. through the PADI Foundation, The Society of Systematic Biologists, the Fish Barcode of Life Initiative, the DeepFin Project, and the National Geographic Young Scientists Award. We thank Jeffrey Hunt and Lee Weigt of the Smithsonian Institution, Arnold Suzumoto and Jack Randall of the Bishop Museum, Randall Kosaki and the crew of the NOAA ship H/Alalak, the Stock Assessment Program at the NOAA Pacific Islands Fisheries Science Center, and the Hawaii Institute of Marine Biology staff. Tissues and vouchers were provided by the University of Kansas Biodiversity Institute, the Smithsonian Institution, Australian Museum, the South Australian Institute for Aquatic Biodiversity, and the California Aquarium. We thank the Papahānaumokuākea Marine National Monument, US Fish and Wildlife Services, and Hawaii Division of Aquatic Resources (DAR) for coordinating research activities and permits for the NW Hawaiian Islands. Thanks to Jeff Eble, Zoltan Szabo, Rita Mehta, Rob Toonen, Michelle Gaither, Robert Moffitt, Joe O’Malley, Toby Daly-Engel, and Mathew Furtoad of Koolau Pets for assistance collecting specimens. Vadim Goz and Kavita Joshi contributed to the genetic data used in this study. Matt Craig, Nicolas Kooyers, Jose Ramirez, Luiz Rocha, Alan Templeton, Vitas Wagner, and anonymous reviewers and provided helpful comments on this manuscript.

References
