Phylogenetics of mud snakes (Squamata: Serpentes: Homalopsidae): A paradox of both undescribed diversity and taxonomic inflation

Justin M. Bernstein\footnote{Corresponding author. E-mail addresses: jnbernst223@gmail.com (J.M. Bernstein), bvoris@fieldmuseum.org (H.K. Voris), rafe@ku.edu (R.M. Brown), sara.ruane@rutgers.edu (S. Ruane).}, John C. Murphy\textsuperscript{b}, Harold K. Voris\textsuperscript{b}, Rafe M. Brown\textsuperscript{c}, Sara Ruane\textsuperscript{a,d}

\textsuperscript{a} Department of Biological Sciences, 206 Boyden Hall, Rutgers University-Newark, 195 University Ave, Newark, NJ 07102, USA
\textsuperscript{b} Science and Education, Field Museum, 1400 S. Lake Shore Drive, Chicago, IL 60605, USA
\textsuperscript{c} Biodiversity Institute and Department of Ecology and Evolutionary Biology, 1345 Jayhawk Boulevard, University of Kansas, Lawrence, KS 66045, USA
\textsuperscript{d} Department of Earth and Environmental Science, Ecology & Evolution, Rutgers University-Newark, 195 University Ave, Newark, NJ 07102, USA

\textbf{ABSTRACT}

Mud snakes (Serpentes: Homalopsidae) are a family of 55 described, mainly aquatic, species primarily distributed throughout mainland Southeast Asia and the Indo-Australian Archipelago. Although they have been the focus of prior research, the basic relationships amongst genera and species remain poorly known. We used a combined mitochondrial and nuclear gene dataset to infer their phylogenetic relationships, using the highest levels of taxon and geographic sampling for any homalopsid phylogeny to date (62\% generic and 62\% species coverage; 140 individuals). Our results recover two reciprocally monophyletic groups: the fangless \textit{Brachyorrhos} and its sister clade comprised of all rear-fanged homalopsids. Most genera and interspecific relationships were monophyletic and strongly supported, but intergeneric relationships and intraspecific population structure lack support. We find evidence of both undescribed diversity as well as cases of taxonomic inflation within several species. Tree-based species delimitation approaches (mPTP) support potential new candidate species as distinct from their conspecifics and also suggest that many named taxa may not be distinct species. Divergence date estimation and lineage-through-time analyses indicate lower levels of speciation in the Eocene, with a subsequent burst in diversification in the Miocene. Homalopsids may have diversified most rapidly during the Pliocene and Pleistocene, possibly in relation to tectonic shifts and sea-level fluctuations that took place in Sundaland and the Sahul Shelf. Our analyses provide new insights on homalopsid taxonomy, a baseline phylogeny for the family, and further biogeographic implications demonstrating how dynamic tectonics and Quaternary sea level changes may have shaped a widespread, diverse family of snakes.

1. Introduction

The field of systematics strives to discover and describe the evolutionary relationships of life and integrate this knowledge to understand how species and populations interact and change over time. Phylogenies are the baseline tool to discover biodiversity and provide the historical framework for analyses that investigate taxonomic and biogeographic hypotheses \cite{Grismer2016,Li2018,Miralles2018}, hybridization \cite{Burbrink2018,Dufrenes2020}, ancestral traits \cite{Gamble2015,Takeuchi2018}, biogeography and speciation \cite{Hinckley2020,Tomasello2020}, and phenotypic trait evolution \cite{Mahler2010,Setiadi2011,Bergmann2020}. Snakes have increasingly been used as a model system to investigate evolutionary and ecological hypotheses \cite{Shine2000,Lillywhite2019}, in part due to their successful colonization of all continents except Antarctica, and having adapted to a wide variety of environments, including terrestrial, fossorial, arboreal, marine, and freshwater habitats \cite{Greene1997}. Even within families and genera, snakes are diverse in morphology, behavior, and habitat preference. Broadly construed, snakes include numerous lineages with both exceptionally wide and narrow, range-restricted geographic distributions, making them ideal for investigating biogeographic hypotheses, speciation, and adaptation, particularly in relation to body plan evolution, body size, and associated morphological traits. Although the monophyly of snakes is well-supported and among-family relationships have begun to stabilize \cite[e.g.,][]{Pyron2013,Zheng2020},
Molecular Phylogenetics and Evolution 160 (2021) 107109

J.M. Bernstein et al.

relationships within many families remain unresolved. Resolution of remaining intrafamilial systematic problems are critical challenges for the use of snakes as model study systems in evolutionary biology.

Mud snakes (Serpentes: Homalopsidae) are a group of 55 species in 29 genera, distributed as far west as the Indus River of Pakistan and ranging eastward throughout South and mainland Southeast Asia, the Philippines, the Indo-Australian archipelago, New Guinea, and northern Australia (Murphy and Voris, 2014). Homalopsids are primarily aquatic, with some species specializing in salt or brackish water systems (e.g., mangroves, coastlines), as well as freshwater systems (e.g., rivers, lakes; Murphy, 2007). Many homalopsids have morphological characters, such as semicircular valvular nostrils and dorsally located eyes, that are presumed adaptations for their largely aquatic lifestyles (Voris et al., 2002). The extraordinary morphological and ecological diversity within this group has been used to study evolutionary phenomena such as head shape and diversification (Fábreg et al., 2016), feeding performance (Jaye et al., 2018), and for testing hypotheses regarding biogeography in relation to Southeast Asia’s complex geologic history (Alfaro et al., 2008). Although these studies demonstrate the value of particular focomal and cladophyllomorph homalopsid study systems for understanding broad evolutionary themes, they are dependent on incorporation of a well-resolved phylogenetic estimate, which currently does not exist for the Homalopsidae.

Homalopsids have been long recognized as a distinct group of snakes (Bonaparte, 1845; Jan, 1863) but various studies have led to disparate hypotheses regarding their phylogenetic position, resulting in a convoluted nomenclatural history (e.g., Gray, 1849; Bouleger 1890, 1896; Smith, 1943; Gyi, 1976; Knight and Mindell, 1994; Underwood, 1999). For example, species now recognized as homalopsids have previously been a part of considered closely related to the families Colubridae, Natricidae, Viperidae, and Pareatidae (Murphy and Voris, 2014). Although the monophyly, position, and establishment of Homalopsidae as a family has consistently been supported in modern phylogenies (Voris et al., 2002; Lawson et al., 2005; Pyron et al., 2013; Burbrink et al., 2020), their intergeneric and interspecific relationships remain poorly resolved. Several fine-scale analyses have been conducted for the genera Cerberus (Alfaro et al., 2004) and Enhydris (Karns et al., 2010), but taxon and geographic sampling remain low in familial-level trees, as disparate and hypothesized to be the fangless sister group to all other species. Most taxonomically comprehensive homalopsid tree to date is based on a mitochondrial (mtDNA) gene and included 31 specimens (45% and 62% species and generic coverage, respectively; Quah et al., 2018). Prior work on the family (Alfaro et al., 2008) hypothesized divergence times for homalopsids, but with 38% of species, 48% of genera, and did not include the homalopsid genus Brachyorrhos; this genus, along with two other poorly known genera (Calamophis, Karnsophis), are ecologically disparate and hypothesized to be the fangless sister group to all other homalopsid taxa, which are rear-fanged. Thus, the inclusion of fangless homalopsids is essential for inferring accurate divergence dates and any inferences of evolutionary transitions in phenotypic or biogeographic character states.

Here, we infer a phylogeny of the Homalopsidae using a combined mitochondrial and nuclear dataset, with the highest species-, individual-, and geographic-level representation to date (62% and 62% species and generic coverage, respectively). We infer the first species tree and time-calibration analyses for the Homalopsidae to include the morphologically disparate, fangless genus Brachyorrhos. Here, we provide an initial phylogenetic framework to (1) reveal and resolve relationships within Homalopsidae with increased generic, species, individual, and geographic sampling (140 specimens across the geographic ranges of 34 species), (2) identify the position of species that have never been previously included in any phylogeny, (3) conduct time-calibration analyses to correlate homalopsid diversification with events in a geological record, and (4) provide insights into the biogeography of the family.

2. Methods and materials

2.1. Taxon sampling, DNA extraction, and sequencing

We sampled a total of 140 homalopsid species, representing 34 (62%) species and 18 (62%) genera (see Table S1 in Supplemental Material) from the rear-fanged and fangless groups; all missing species are only known from holotypes and/or formalin-preserved museum specimens and were unavailable for this project. Most publicly available sequence data for homalopsids are from mtDNA cytochrome b (cyt-b) and nuclear prolactin receptor (PRLR) regions, so we sequenced these two loci for new samples to maximize compatibility of our data with previously sampled specimens. We obtained 98 homalopsid tissues from 16 museum collections and included 48 archived sequences from Genbank (NCBI; 34 cyt-b, 14 PRLR ([Alfaro et al., 2008; Karns et al., 2010; Kumar et al., 2012; Murphy et al., 2012a; Wiens and Tiu, 2012; Alencar et al., 2016; Ukuwela et al., 2017]; see Table S1 in Supplemental Material). This sampling covers the majority of the geographic range of most species. For completeness, we also include a previously identified (Alfaro et al., 2008; Murphy and Voris, 2014), yet undescribed species of Hypisocopus from Lake Towuti, Sulawesi. Homalopsidae is consistently found as a distinct lineage from other caenophidian snakes (e.g., viperids, natricids, elapids, lamprophiids; Pyron et al., 2013; Figuerola et al., 2016; Burbrink et al., 2020), and thus we included outgroup sequences of the following from Genbank: Crotalus horridus, Bitis nasicornis, Nerodia sipedon, Bungarus fasciatus, Laticauda laticaudata, Laticauda frontalis, Boaedon fuliginosus (Alfaro and Arnold, 2001; Nagy et al., 2003; Sanders et al., 2008; Margres et al., 2015; see Table S1 in Supplemental Material, museum and institution codes follow Sabaj (2016)).

Genomic DNA was isolated from liver and muscle tissue using Qia gen® DNeasy blood and tissue kit protocols. In brief, we lysed tissues using proteinase K and ATL lysis buffer for 12 h at 56 °C, and subsequently followed the Qiagen® kit protocol. We amplified cyt-b and PRLR using the polymerase chain reaction (PCR) with primers L14910 and H16064 (cyt-b, ~1100 base pairs targeted; Burbrink et al., 2000) and PRLR_f1 and PRLR_r3 (PRLR, 532 base pairs targeted; Townsend et al., 2008). We ran 25 µL-PCR reactions with a 5-min 94 °C initial denaturation step, followed by 35 cycles of denaturing at 94 °C for 45 s, primer annealing at 52 °C, and elongation at 72 °C for 60 s, followed by a final extension step of 5 min at 72 °C. We used ExoSAP-IT™ PCR Product cleanup reagent (Thermo Fisher Scientific) to clean PCR products prior to sequencing. Genomic DNA was quantified on a Qubit 3 fluorometer (Thermo Fisher Scientific; Invitrogen) and Sanger-sequenced on an ABI 3730xl DNA analyzer at Macrogen (New York, United States). Resulting sequences were aligned for each gene using the MUSCLE alignment option under default parameters in Geneious v7.1.9 (Geneious, 2020; https://www.geneious.com) and we checked alignments by eye to ensure these protein coding loci were within the correct reading frame and to remove ambiguous base pairs at the beginning or ends of sequences. We checked PRLR sequence chromatograms for heterozygotic sites; no individuals were heterozygotic at any sites, so PRLR data were not phased.

2.2. Phylogenetic analysis

To initially explore the genetic diversity and structure amongst taxonomically and geographically distinct populations, we used maximum likelihood (ML) and Bayesian inference (BI) to reconstruct phylogenetic relationships on a concatenated dataset. All gene trees and concatenated, dual-gene trees were generated using the CIPRES Science Gateway v3.3 (Miller et al., 2010). Evolutionary models and partitioning schemes for all DNA alignments and subsequent RAxML and MrBayes analyses were selected by partitioning the concatenated dataset by gene and codon position in PartitionFinder2; the best-fit nucleotide substitution models and partitioning schemes for ML and BI analyses were
3. Results

3.1. Maximum likelihood, Bayesian inference, and pairwise distances

The final alignments for cyt-b, PRLR, and the concatenated dataset were 1053, 585, and 1638 base pairs, and the cyt-b and PRLR genealogies included 123 and 91 homalopsid respectively (See Figs. S1–S2 in Supplementary Material). Although our concatenated alignment (140 homalopsids) has 12.8% and 35.5% missing specimen data for cyt-b and PRLR, respectively, it has a higher species representation than the individual gene trees (eight species for cyt-b and one species for PRLR were only represented by that single gene in the concatenated dataset). Phylogenetic accuracy has been shown to increase if there is a higher coverage of taxon sampling compared to data coverage (Hedtke et al., 2006; Wiens and Tito, 2012). Thus, we particularly focus on the results of the concatenated and species trees in our results and discussion, which include the highest sample representation; we include the single gene trees in the supplemental materials (See Figs. S1–S2 in Supplementary Material). The complete list of specimens used in this study and their associated Genbank accession numbers are provided in Supplementary Material Table S1.
For both loci, the ML and BI analyses indicated a strongly supported, monophyletic Homalopsidae (See Fig. S1 in Supplementary Material). Most genera in the ML and BI gene trees for PRLR were monophyletic; Gylophis and Myrrophis are weakly supported within Enhydris. In contrast, all genera were monophyletic in the cyt-b gene trees. The BI cyt-b tree poorly supported Hypsicorpus as sister to all other homalopsids (including Brachyorrhos) and contained a polytomy for the placement of several genera. Bootstrap and posterior probabilities of both gene trees had comparable support values, and were similar to the concatenated tree, discussed below.

Concatenated ML and BI analyses strongly supported a monophyletic Homalopsidae (BS/PP = 1; Fig. 1), composed of two major clades: the fangless clade comprised solely of the genus Brachyorrhos and a rear-fanged clade of all other homalopsids. Most relationships were congruent between ML and BI trees and had comparable support values, with well-supported branches at the genus, species, and sister species levels, and poorly supported relationships among most genera (Fig. 1, Supplemental Figs. S1–S2). The major topological difference between the concatenated ML and BI analyses was the placement of Raclitia indica (Fig. 1).

Uncorrected pairwise cyt-b distances were calculated for specimens that were either genetically related (based on results) or geographically proximate (or both), rather than pre-existing species taxonomy, due to several cases of possible undescribed diversity or taxonomic inflation (Fig. 2). Within Brachyorrhos, there was strong support for B. albus as part of B. raffrayi, with a minimal genetic distance of 0.48%. Low levels

---

**Fig. 1.** Comparison of concatenated maximum likelihood (ML) and Bayesian inference (BI) homalopid phylogenies. Genera are color-coded on each tree. Purples lines are attached to identical tips of both trees, emphasizing areas of congruence versus topological disparity among ML and BI estimates. Unhighlighted clade in both trees are the outgroups. The red circles denote the position of Raclitia indica, the only genus recovered in different positions between both analyses. White circles at divergences indicate BS ≥ 0.70 and PP ≥ 0.95. Note: dashed lines from tips to taxon labels are for ease of visualization and are not representative of any analytical results. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
of interspecific divergence were also seen in some specimens of *Myrrophis bennettii* and *M. chinensis*. *Enhydris longicauda*, *E. innominata*, and *E. jagorii* are all recovered within a single clade with pairwise distances of 0–1.1%. Three of the seven *Homalopsis buccata* specimens, and all *H. mereljcoxi* and *H. semizonata*, were in a clade that is 0.5–1.27% in genetic distance from one another. Finally, the Palau-endemic *Cerberus dunsoni* and Philippine Lake Buhi-endemic *C. microlepis* both formed clades within the geographically widespread *C. schneiderii*. The pairwise distances between *C. dunsoni*, *C. microlepis*, and closely-related *C. schneiderii* were 1.6–3.2% and 1.3–1.7%, respectively.

Intraspecific and intragroup (for potentially synonymous taxa) distances for several species were much greater than intraspecific distances seen in other taxa. The two Australian *Pseudoferania polylepis* specimens are not sister to each other, and, with the New Guinea specimens, have an intraspecific variation of 4.2–9.9%. Two distinct clades of *Homalopsis buccata* + *Homalopsis mereljcoxi* were 3.9–5.5% divergent; one of these groups contains specimens from Malaysia and Thailand, whereas the other clade consists of Thailand, Cambodia, and Myanmar specimens.

Fig. 2. Concatenated homalopsid ML tree with mPTP species delimitation results (outgroups not shown). Black nodes indicate bootstrap values ≥ 70. Red clades and single black, terminal branches represent species from the mPTP analysis. Cytochrome b genetic distances are given for genera with high levels of diversity. Pairwise distance matrices are based by group, denoted by the adjacent colored circles. Photograph of *Cerberus australis* NTM R29883 courtesy of Museum and Art Gallery of the Northern Territory. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)


3.2. Concatenated candidate species delimitation

The mPTP analysis proposed 36 species from the 34 formally described species comprising our dataset, and, while provided quantitative support of synonymy, also yielded support for (or proposed) undescribed species diversity. Non-monophyletic species in the concatenated tree were supported as one species in the mPTP analysis. The proposal to synonymize several species was also recovered in closely-related populations of Homalopsis (H. buccata + H. mereljocoi + H. seminatona), Cerberus (C. Schneiderii + C. microlepis + C. dunsoni), Myrrophis (M. Bennetti + M. chinensis), and Enhydris (E. Longicauda + E. Innominata + E. Jagori) (Table 1).

In contrast, undescribed species diversity was proposed in Hypsiscopus, Myron, Pseudoferania, and Phyloptosis (Table 1). Our analysis putatively splits Hypsiscopus into six potential species: H. sp. (Lake Towuti, Sulawesi), H. plumbea (China), H. plumbea (Thailand + Laos), H. plumbea (Thailand + Laos + Cambodia + Malaysia), H. matannensis (South Sulawesi), and H. matannensis (Southeast Sulawesi). The New Guinea and Australian specimens of Myron richardsonii were each recovered as potentially distinct species. This was also seen with Subsessor bocourtii, one specimen each from Cambodia and Thailand as potentially different species, and to a greater extent with Pseudoferania polylepis, with all four species identified as separate candidate species. Although from the same locality, the two Phyloptosis punctata were also preliminarily delimited; this is most likely an artifact due to missing PRLR data in one of the two specimens; mPTP analyses utilize branch lengths and phylogeny and, thus, the long branch from the sample with both gene regions was proposed as a distinct species. The same may also explain preliminary delimitation of the two species of S. bocourtii as different species (but from two different localities).

3.3. Species-tree inference and divergence date estimation

For the species trees, the birth-death process of diversification had a higher likelihood score than the Yule process (-20052.08 for birth-death, -20056.17 for Yule), so the former model was used; ESS values were generally >200. In the species tree, Homalopsidae is well-supported as monophyletic, with a sister relationship between the fanged and fangless clades (Fig. 3A). The relationships in the species tree were similar to the concatenated tree, with most intergeneric relationships poorly supported. With the exception of Enhydris, all polytypic genera were recovered with strong support. The major differences between the species and concatenated trees were the placements of Dieurostus, Racilia, Erpeton + Subsessor, and Bitia + Phyloptosis (Figs. 2, 3A).

The H. plumbea from China was sister to the Thailand + Laos (northern) H. plumbea populations, and the Thailand + Laos + Cambodia + Malaysia (southern) populations were sister to H. matannensis + H. sp.-Lake Towuti (Fig. 3B). Similar to the concatenated analysis, the Australian and PNG Myron richardsonii were reciprocally monophyletic (Fig. 3C), and the West Papuan and Australian Pseudoferania polylepis were not sister taxa (Fig. 3D).

The LTT plot under the birth-death model and Pybus and Harvey’s $\gamma$ statistic indicate that homalopsids slowly diversified early in their evolution, and subsequently had a rapid increase in their diversification rate around 10 mya (HPD 8.75–11.66; $\gamma = 3.869$, $p$-value = 0.0001; Fig. 3A). Under a pure-birth model of diversification, LTT plot simulations ($n = 100$) showed a gradual increase in diversification through time (Fig. 3A). The dated analysis suggests a mid-Eocene (45.31 mya; 95% HPD 43.68–46.88 my) diversification of the crown homalopsid group (Fig. 3A). The flagless Brachyorrhos diversified during the Pleistocene (1.49 mya; 95% HPD 0.18–2.58 my), and the rear-fanged clade diversified in the upper-Miocene (10.14 mya; 95% HPD 8.75–11.66 my), with most subsequent divergences between rear-fanged genera occurring between 4.5 and 9.8 million years ago (Fig. 3A). The undescribed lineages of H. plumbea, M. richardsonii, and P. polylepis split from their sister lineages 2.47 ± 1.31, 3.83 ± 0.75, and 0.66–2.58 ± 0.59–0.91 mya respectively (Fig. 3B, C, D).

4. Discussion

Advances in phylogenomics and bioinformatics have resulted in a greater understanding of the evolution and diversity of numerous organismal groups, yet many families are still poorly known with respect
to intrageneric relationships and species limits. Our results indicate that snakes of the family Homalopsidae are in need of both broad- and fine-scale investigation with respect to species diversity, and their taxonomic placement within supra-specific higher taxa. The early branching order of our estimates and, thus, many of the intergeneric relationships of homalopsids, remain unresolved (Figs. 2, 3). However, we do find several strongly-supported sister-taxon relationships (Fig. 2, 3), which are consistent with previous studies (e.g., *Brachyorynchus*, Murphy et al., 2011; *Cerberus* relationships, Murphy et al., 2012a; a crustacean-eating clade, Alfaro et al., 2008). Additionally, results indicate many currently recognized species are likely synonymous, whereas others may harbor undescribed taxa. Although our trees are based on just two loci, mtDNA-based trees have yielded significant evolutionary and natural history findings at the species and population levels (Funk and Omland, 2003; Rubinoff and Holland, 2005; Burbrink and Ruane, 2021); despite low coverage for both genes, our additional taxonomic sampling and advanced methods, compared to previous studies, provide evolutionary insights into homalopid diversification and related taxonomy, as...
discussed below (Wiens, 2003; Wiens and Morrill, 2011; Zheng and Wiens, 2015).

4.1. Systematic implications for Homalopsidae

4.1.1. Intergeneric relationships and the inclusion of novel taxa

Our resulting trees corroborate many of the lineages and relationships found in previous homalopsid phylogenies (e.g., Alfaro et al., 2004; Alfaro et al., 2008). Novel relationships among and within genera were also recovered, when compared to studies that only used mtDNA or with lower taxonomic and geographic sampling (Voris et al., 2002; Alfaro et al., 2008). The inclusion of Cerberus dusoni, Enhydris chanardi, Homalopsis semizonata, H. nigroventralis, and Myrophis bennetti, which have never been part of any phylogenetic analysis, support the monophyly of the respective genera (Figs. 2, 3). We also find alternative topologies to previous hypotheses of the placement of particular species. For example, the placement of Raclitia indica from Malaysia is hypothesized as sister to Epeton tentaculatum from Thailand (Quah et al., 2018), but here we find R. indica as either sister to Dieurosous dussumieri or a clade containing Cerberus, Homalopsis, Bitaia, and Phylotopsis (Figs. 2, 3A).

Studies on Enhydris (e.g., Karas et al., 2005, 2010; Lukoschek et al., 2011) have focused heavily on E. jagorii, E. longicauda, E. innominata, E. enhydris, and E. subzamiana. The sixth species of the genus, E. chanardi, is a rare snake with an uncertain geographic range in Thailand (Murphy and Voris, 2014). We find support (concatenated tree BS = 98; species tree PP = 0.84) for E. chanardi as the sister to all other Enhydris. Although our specimens of E. chanardi (vouchers YPM 15033, 15037) are from the pet trade, and no other fresh tissues of this species were available for this study, an examination of the associated voucher skins renders the species confirm its identity as E. chanardi based on sculation and color pattern (JMB, unpublished data). Additionally, our mPTP analysis supports the distinctiveness of E. chanardi from the other 29 Enhydris specimens included here. In contrast, our sampling of H. mereljcoxi and H. nigroventralis indicate that although H. nigroventralis is distinct, H. mereljcoxi is within H. semizonata (Fig. 2). Similarly, recognition of Cerberus dusoni and C. microlepis renders C. schneideri polyphyletic (Fig. 2).

4.1.2. Evidence of synonymy

The species richness of homalopsid snakes has increased since Gyi’s (1970) morphology-based classification of the family (Murphy, 2007; Murphy and Voris, 2014), going from 34 species in 10 genera to the current 55 species in 29 genera. Although our study supports the recognition of many of these taxa, there are several questionable-distinct species (Table 1). We find some support from our distance calculations, concatenated tree, and mPTP species delimitation analysis, for the proposal that 13 species (among five genera) should be critically evaluated and possibly placed back in synonymy with their closely-related congeners. These results may indicate taxonomic inflation of species that have similar or overlapping diagnostic characters (See File S3 in Supplementary Material File for all comparisons, discussed below), many of which have geographically proximate, or even peripatric, distributions.

The low intraspecific genetic distances of Brachyorrhos raffrayi and B. albus indicates they may represent one species, and thorough investigation of the diagnostic characters of these understudied, semifossorial homalopsids is warranted. Brachyorrhos raffrayi and B. albus are both fangless homalopsids with allopatic distributions on the Indonesian islands of Ternate and Seram (and satellite islands, respectively). However, the known herpetofauna of island localities between B. raffrayi and B. albus are not as thoroughly investigated, and there are reports of a population of B. albus established on Pulau Bisa, off Obi Island (Murphy et al., 2012b; O’Shea, 2018), which would make the closest population B. albus 49 km away from the Bacan Islands, the northernmost which is Ternate and the known locality of B. raffrayi. Although B. raffrayi has not been reported from the rest of the Bacan Islands, undiscovered populations may exist there. Additionally, these species have similar and overlapping scale morphologies (e.g., subcaudals [males], ventrais; Murphy et al., 2012b), which are commonly used as morphological diagnostic characters for snakes. Our results suggest that the phylogenetic relationships observed may be explained by a geographically expanded B. albus, and that putting B. raffrayi in synonymy with the former may be advisable.

Taxonomic inflation is also of concern in Homalopsis, Cerberus, Myrophis, and Enhydris. The clade consisting of Homalopsis buccata, H. mereljcoxi, and H. semizonata, (Fig. 2) include Thailand H. buccata north of their known distribution (Murphy et al., 2012c; Murphy and Voris, 2014). These species may be difficult to accurately identify in the field and an examination of the H. buccata from northern Myanmar is needed to identify the species boundaries between this species and its congeners. Homalopsis mereljcoxi and H. semizonata, however, have proximate ranges in Indochina, and both are found along the Malay Peninsula, south of the Isthmus of Kra (Pauwels and Sumontha, 2016).

Meristic characters (e.g., ventrais, subcaudals) overlap between these species, and some of the diagnostic characters, such as unique three-way fragmentation of the prefrontal head plates, have been shown to be variable and similar to other Homalopsis species (Pauwels and Sumontha, 2016).

Another set of species with problematic taxonomic and biogeographic implications is the placement of Cerberus dusoni and C. microlepis within C. schneideri. Cerberus microlepis and C. dusoni differ from C. schneideri, respectively, by a higher number of dorsal scale rows and plate-like parietal scale fragments on the head (Murphy et al., 2012a; Murphy and Voris, 2014). However, much of their ranges in scale counts and their color pattern overlap with C. schneideri, which exhibits greater morphological variation than once thought (Barrera Jr. et al., 2017). The distribution of C. microlepis is restricted to Lake Buhi in the Philippines, which is geographically located within the range of C. schneideri; the endemic C. dusoni is restricted to the Palau islands of Micronesia. It has been hypothesized that C. microlepis may be a freshwater-adapted form of C. schneideri after a population became isolated to the lake when an earthquake caused the adjacent Mt. Iriga to form a natural dam a few hundred years ago (Alfaro et al., 2004).

Cerberus dusoni, which is 870 km away from its congeners, may represent a population of C. schneideri that made successive colonizations across the Kyushu-Palau Ridge, a volcanic island chain that extended north of New Guinea (Allison, 1996). Such interpretation contrasts with current taxonomy and the hypothesis that C. dusoni is independent from C. schneideri (Murphy et al., 2012a). Nevertheless, lake and island populations are often distinctly grouped in our analyses (Table 1; Figs. 2, 3), suggesting some level of isolation. A detailed study focusing on these populations and including explicit gene-flow analysis is necessary to determine the degree of separation amongst these taxa.

The two known species of Myrophis, both of which are found in southern China, are delimited by distinct numbers of several scale characters and color pattern, as well as habitat type; M. chinensis is known from freshwater habitats, whereas M. bennetti is known from brackish and marine waters (Murphy, 2007; Karas et al., 2010; Kumar et al., 2012). The grouping of these species in our analyses may represent greater variability in a single, wide-spread species and subsequent examination of the original voucher specimens would help to confirm these identifications. Similarly, the Enhydris longicauda + E. innominata + E. jagorii group is only known from Cambodia, Vietnam (and possibly Cambodia), and Thailand, respectively (Murphy and Voris, 2014). It is uncertain if these species are sympatric with one another, but their morphologies are similar in their number of dorsal scale rows and subcaudals. The relationship reported here has been recovered in other studies on Enhydris (Karas et al., 2010). We provide additional support for this potential synonymy in this genus with specific species delimitation analyses and increased sampling.

All of these cases support scenarios of synonymy for species in
Homalopsidae, which require greater sampling both for individual snakes as well as loci, ideally at the genomic level in conjunction with detailed morphological examination for the taxa in question. Thus, we remain conservative and do not suggest any taxonomic changes until future genomic analyses are conducted. The interspecific pairwise cyt-b genetic distances for most of these possibly synonymous species are much smaller (e.g., Homalopsis, Enhydris; Fig. 2) than the interspecific distances for species considered distinct in other recent snake studies that calculated mtDNA distances (e.g., Ruane et al., 2018; Keates et al., 2019). Though changes in taxonomy ideally will include increased sample sizes and more loci than presented here, these results suggest potential cases of taxonomic inflation within several homalopsid genera.

4.1.3. Undescribed diversity and intraspecific biogeography

In contrast to evidence of synonymy in some taxa, we also find several instances of undescribed homalopsid lineages. There is strong support for a north-south split between populations of Hypsicopus plumbea from northern Thailand + Laos and those from south-central Thailand + Cambodia + Malaysia. This divergence may be the result of two related geographic events in Thailand during the Quaternary. The Khorat Plateau (Khorat Basin; 180,000 km²), separating northern and southern clades of *H. plumbea*, in northeastern Thailand formed after Quaternary tectonic uplifting and tilting occurred alongside its western and southern edges (Hutchison, 1989; Rainbow, 1996). This uplift led to a river catchment event in which the Mekong River, originally flowing south towards the Gulf of Thailand (currently the Chao Phraya river plain; Carbonnel 1965; Workman, 1977), shifted to its current position running west to east, and then south through the now-Thailand-Laos border and Cambodia (Rainboth, 1966; Fontaine and Workman, 1978). During the mid-Quaternary, the Khorat Plateau had a heterogeneous landscape of river valleys and mountains that have shaped the local biodiversity (Fontaine and Workman, 1978; Hutchison, 1989). With respect to the mountains on the Khorat Plateau’s western and southern margins, shifting of major aquatic riverways (e.g., the Mekong River), and the Annamite mountain range to the east of the Khorat Plateau, our divergence dating results indicate that tectonic uplift, the heterogeneous landscape, and secondary river catchment events may have ultimately acted as barriers to gene flow between northern and southern *H. plumbea* (Fig. 3B), as has been demonstrated in gastropods, fish, and other homalopsids (Glaubrecht and Köhler, 2004; Lukoschek et al., 2011; Adamson et al., 2012). Such divisions may also be supported by the phylogenetic placement of a *H. plumbea* specimen from Bangkok (FMNH 250124) with the northern specimens (Fig. 2), despite being geographically proximate to the samples in the southern clade; this likely reflects the river catchment, where some of the population remained in the Mekong (now in northern Thailand) while other parts of the population with a similar haplotype were isolated to the Chao Phraya near Bangkok. Although some of the mountains surrounding the Khorat Plateau are not particularly high in elevation, they may still be a sufficient barrier between homalopsid populations, possibly due to their aquatic nature, as evidenced by a significant decrease in homalopsid abundance at Khorat Plateau rim sites (Karns et al., 2005). The specimen of the southern clade from Malaysian Borneo also suggests dispersal via land bridges in the Pleistocene that connected peninsular Malaysia to Borneo (Woodruff, 2010). Additionally, a potentially new species of *H. plumbea* from Guangxi, China is supported as sister to its conspecifics in northern Thailand and Laos. This locality in China is over 500 km from any of our specimens in our sampling; Guangxi, being located within one of the most biodiverse regions on the planet, may harbor undescribed species, similar to many other vertebrate groups from that region (Luo et al., 2016).

Our species delimitation analysis provides support for previous studies (Alfaró et al., 2008; Murphy and Voris, 2014) that suggest that the Sulawesi Lake Towuti Hypsicopus specimen is distinct from all other *Hypsicopus* species. The specimens of *Hypsicopus matanensis* between South and Southeast Sulawesi may also represent potentially new species, a result recovered in unrelated vertebrates that have population structure between these and other regions of the mainland (e.g., Evans et al., 2003; McGuire et al., 2007; Burton and Nietsch, 2010). Sulawesi is a composite island, in which multiple landmasses collided together, and then subsequent Pleistocene sea-level fluctuations connected and disconnected different regions (Hamilton, 1979; Hall, 1996, 1998; Moss and Wilson, 1998; Nugraha and Hall, 2018). Faunal boundaries such as the Tempe Depression in South Sulawesi and the suture (tectonic) line of Southeast Sulawesi, the latter of which is near Lake Towuti, have been proposed as significant biogeographic transition zones, which may act as barriers to gene flow (Fooden, 1969; Hall, 1998; Evans et al., 2003; Burton and Nietsch, 2010).

The trees and mPTP analysis indicate undescribed diversity in the Australasian *Myron richardsonii* and *Pseudoferania polypyle* (Table 1; Figs. 2, 3). *Myron richardsonii* specimens from localities that are (geographically) from PNG are supported as the sister taxon to the specimens from Australia. Pleistocene sea-level fluctuations have repeatedly connected and disconnected major land masses and coastlines on the Sunda and Sahul Shelves (Voris, 2000; Woodruff, 2010), resulting in distinct genetic signatures in extant taxa (Hall, 1998; Hewitt, 2000). It is likely that *M. richardsonii* was a trans-Torresian species that dispersed between Australia and PNG during periods of inter-landmass activity in the Pleistocene (Fig. 3C), which is plausible given the shallow waters of the Torres Strait and these snakes’ aquatic (estuarine/marine) nature. These patterns are not limited to our example here, but have also been shown in mammals (Aplin et al., 1993) and other snakes (e.g., elapids; Wüster et al., 2005). The phylogenetic and delimitation analyses of *Pseudoferania polypyle*, another homalopsid with an Australo-Papuan distribution, suggest four highly divergent specimens between West Papua, Indonesia, and Australia (Fig. 3D). Similar to biogeographic scenarios of the Torres Strait, the land bridge formed between Australia and PNG spanned the Arafura Sea, which is located between our *P. polypyle* specimens (Fig. 3D). Although more specimens and loci are required, multiple dispersal events may have produced the relationships observed (Fig. 3D).

These biogeographic hypotheses may explain the divergences in these taxa, though other genera included here that showed evidence of undescribed diversity (i.e., *Subsessor bocourtii* and *Phytolopis punctata*; Table 1) only included one or two specimens from each locality. More specimens and geographic sampling are needed to determine if these splits represent undescribed species or are artifacts in our analyses.

4.2. Evolution of the Homalopsidae

We present the first dating analysis specifically for the family Homalopsidae, with the inclusion of the fangless homalopsids and highest taxonomic coverage to date. These dates are likely to change with the addition of more taxa and loci, and so we suggest our biogeographic interpretations for Homalopsidae as a starting point for future studies. Our species tree indicates that the fangless *Brachyorrhors* inhabiting eastern Indonesia (Maluku Islands) and the rear-fanged group in Southeast Asia, Australia, and New Guinea diverged 45.31 ± 1.63 mya (Fig. 3). Homalopsids may have a mainland Southeast Asian origin, with subsequent expansion westward into South Asia and eastward towards the Greater and Lesser Sunda Islands, Wallacea, the Philippines, Palau, New Guinea, and Australia. An origin in mainland Southeast Asia is also supported by the observation that ~49% of homalopsid species are distributed throughout Indochina (and adjacent China), with the rest known from isolated regions or islands east or west of Indochina.

Dispersal events have likely been the result of the changing paleogeography that occurred through most of these regions, even across regions that are considered strong faunal barriers (e.g., *Hypsicopus* crossing Wallace’s Line; Alfaró et al., 2008; Esselstyn et al., 2010; Brown et al., 2013). We also find evidence for recent founder events, such as *Cerberus microlepis* of Lake Buhi, *C. dunsoni* of Palau, and the Timor population of *C. schneideri*, all of which may actually be populations of...
the phenotypically similar *C. schneideri*. Our data suggests that dispersal into isolated habitats such as lakes and islands can be followed by inhibited gene flow from the surrounding populations.

Perhaps the most striking of our results is that of the timing of the initial split between the fangless and fanged homalopsids. The time period between the diversification of crown homalopsids (~45.31 ± 1.63 mya, Fig. 3) coincides with the timing of the Australian plate moving northward through the Pacific, away from Antarctica (Hall, 2009). The Australian plate contains the landmasses of Australia, New Zealand, and New Guinea, including the Bird’s Head peninsula and eastern Indonesian Islands that harbor *Brachyorycthus*. The divergence date estimation (Fig. 3) indicates that the fangless clade was already present before any known connections to facilitate dispersal between Southeast Asia and Australia. Although these topologies are broadly consistent with previous studies inferring rear-fanged homalopsid biogeography (e.g., Alfaro et al., 2008), divergence date estimates of both fanged and fangless homalopsids greatly changes the inferred biogeographic implications on the evolution of this family.

The young ages between species of *Brachyorycthus wallacea* and *B. albus + B. raffrayi* (~10 kya–1.5 mya) may reflect the geologic history of some islands to which these species are seemingly restricted to; the Banda Arc, including Timor and Seram, are very young (~2 million years old; Hall, 2009; Lobman et al., 2011). Although some of the larger islands, such as Halmahera, are geologically dated no older than 11 my, volcanic activity of Halmahera ceased in the last 2 million years (Baker and Malahiollo, 1996; Hall, 2009). Indeed, our topology and divergence dates show a long period of ~35 million years in which there was no divergence amongst homalopsids, which may be due to missing nuclear data for *Brachyorycthus* vs. reflecting reality. The lack of a slowly-evolving nuclear gene and reliance on a fast-evolving mitochondrial gene may have resulted in the large temporal gap between the fangless and rear-fanged homalopsids in the species tree (Fig. 3). Alternatively, the temporal gap may represent missing data from real biological units (i.e., unsampled extant and/or extinct taxa). Halmahera formed due to ongoing arc collisions in the Molucca Sea, and in the next 5 million years will likely submerge below the ocean’s surface with no subsurface traces of the arc itself (Hall, 2000). If other island arcs have also disappeared from the ocean’s surface, as has been hypothesized (Hall, 2000), then the subsidence of island arcs may be correlated with the extinction of their flora and fauna. As such, if extinction has occurred in Wallacea island arcs, this can cause an artifactual gap (in our case, ~35 my long) in the chronogram and inflate the Pybus and Harvey’s γ statistic that suggested early-slow and recent-rapid diversification (i.e., the Pull of the Present; Ricklefs, 2007).

5. Conclusion

Our study provides new evolutionary insights on homalopsid snakes using the largest and most comprehensively sampled phylogeny of the family to date. Although this dataset is limited to two loci and primarily informed by mtDNA, our primary conclusions, with regard to evolutionary relationships and biogeographic patterns, are observed using either locus. Furthermore, mtDNA is a heritable unit and provides evidence of evolutionary and biogeographic phenomena (Rubinoff and Holland, 2005; Burbrink and Ruane, 2021). Previous studies on snakes (Burbrink et al., 2006, 2021), lizards (Grisser et al., 2016), salamanders (Steinfartz et al., 2000; Vences et al., 2014; Rodriguez et al., 2017), birds (Lamichhaney et al., 2015), bats (Stadelmann et al., 2007; Morales et al., 2019), and eutherian mammals (Liu and Miyamoto, 1999; Song et al., 2012) have found that initial phylogenetic reconstructions using single or few mitochondrial and/or nuclear loci often find the same lineages, sister taxa, and/or species groups when reanalyzed using more comprehensive molecular sampling. Our increased taxonomic and biogeographic sampling will be critical for accurately revisiting the evolution of organismal, geographical, and ecological traits in this group (e.g., habitat preference, salt tolerance [Kumar et al., 2012], diet and feeding behavior [Fabre et al., 2016; Jayne et al., 2018], biogeography [Alfaro et al., 2008]). Finally, increasing current knowledge of homalopsid diversity is crucial to conservation. Although mud snakes are a critical component to Southeast Asian ecosystems and contribute a large portion of resident vertebrate biomass (Murphy, 2007), populations are likely in danger of extirpation or extinction. For example, ~3.8 million homalopsid snakes harvested from the wild, yearly, in Tonle Sap, Cambodia (Brooks et al., 2007).

With the recent success of studies incorporating fluid-preserved specimens into systematic datasets (e.g., Ruane and Austin, 2017), our study and understanding of Homalopsidae can be greatly enhanced with the addition of species and populations only known from museum specimens and historical records in the literature, such as the other fangless homalopsids Karnosphis and Calamophis, as well as several rear-fanged genera. During the course of publishing this study, a new genus and species of homalopsid, endemic to Myanmar and sister to Gypophis, was described (M. thaliynensis; Köhler et al., 2021), emphasizing that the diversity and evolution of this group is still far from being recognized. The use of genomic datasets (e.g. SNPs, target capture; Leaché and Oaks, 2017; Barrow et al., 2018) and the publication of draft genomes (Köhler et al., 2021) will likely provide a more accurate interpretation of the evolution, biogeography and historic demography of this group. Additionally, increasing our geographic sampling will be crucial to elucidating the polarity of gene flow and dispersal, and provide genome-scale insights, which will improve the accuracy of species delimitation analyses (Chan et al., 2020). Although increased loci and taxon sampling will be forthcoming from both our own and other studies in progress, the analyses presented here provide new data (taxon and gene sampling), a more comprehensive phylogenetic inference, and novel evolutionary hypotheses for the continued study of this highly unique, widespread—and yet poorly known—family of aquatic caenophasid snakes.

CRediT authorship contribution statement

**Justin M. Bernstein:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Visualization.

**John C. Murphy:** Conceptualization, Validation, Resources, Supervision.

**Harold K. Voris:** Conceptualization, Validation, Resources, Supervision.

**Rafe M. Brown:** Methodology, Validation, Resources, Supervision.

**Sara Ruane:** Conceptualization, Methodology, Validation, Resources, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

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

Acknowledgements

We acknowledge the following individuals and associated institutions for providing tissues to be used for DNA sequencing in this study: Frank Burbrink, David Kizirian, and Lauren Vonnahme (American Museum of Natural History); Molly Hagemann (Bermice P. Bishop Museum); Lauren Scheinberg and Erica Ely (California Academy of Sciences); Alan Resetar and Joshua Mata (Field Museum of Natural History); Lee Grisser (La Sierra University Herpetology Collection); Christopher Austin and Jackson Roberts (Louisiana State University Museum of Natural Science); Alan Resetar and Joshua Mata (Field Museum of Natural History); Lee Grisser (La Sierra University Herpetology Collection); David Kizirian and Lauren Vonnahme (American Museum of Natural History); and Bryan Stuart (North Carolina Museum of Natural Sciences); Patrick Couper and Andrew Amey (Queensland Museum); Bob Murphy (Royal Ontario Museum); Sally South (South Australian Museum); and Addison Wynn (National Museum of Natural History); Paul Doughty and Rebecca Bray (Western Australian Museum); Gregory

10