Cenozoic colonisation of the Indian Ocean region by the Australian freshwater-originating glassperch family Ambassidae (Teleostei)

Siti Zafirah Ghazali [a], Sébastien Lavoué [b,*], Tedjo Sukmono [c], Ahasan Habib [d,e], Min Pau Tan [a], Siti Azizah Mohd Nor [a,b]

[a] Institute of Marine Biotechnology, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia
[b] School of Biological Sciences, Universiti Sains Malaysia, 11800 Pulau Pinang, Malaysia
[c] Department of Biology, Universitas Jambi, Jalan Lintas Jambi - Muara Bulian Km15, 36122 Jambi, Sumatra, Indonesia
[d] Faculty of Fisheries and Food Science, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia
[e] Department of Fisheries and Marine Science, Noakhali Science and Technology University, Noakhali 3814, Bangladesh

ARTICLE INFO

ABSTRACT

We examined the phylogeny and biogeography of the glassperch family Ambassidae (Teleostei), which is widely distributed in the freshwater, brackish and marine coastal habitats across the Indo-West Pacific region. We first built a comprehensive time-calibrated phylogeny of Ambassidae using five genes. We then used this tree to reconstruct the evolution of the salinity preference and ancestral areas. Our results indicate that the two largest genera of Ambassidae, *Ambassis* and *Parambassis*, are each not monophyletic. The most recent common ancestor of Ambassidae was freshwater adapted and lived in Australia about 56 million years ago. Three independent freshwater-to-marine transitions are inferred, but no marine-to-freshwater ones. To explain the distribution of ambassids, we hypothesise two long-distance marine dispersal events from Australia. A first event was towards Southeast Asia during the early Cenozoic, followed by a second one towards Africa during mid-Cenozoic. The phylogenetic signal associated with the salinity adaptation of these events was not detected, possibly because of the selective extinction of intermediate marine lineages. The Ambassidae shares two characteristics with other freshwater fish groups distributed in continental regions surrounding the Indian Ocean: They are too young to support the hypothesis that their distribution is the result of the fragmentation of Gondwana, but they did not retain the phylogenetic signal of their marine dispersal.

1. Introduction

The glassperch family Ambassidae (Teleostei) includes freshwater and marine/brackish species distributed in the coastal, brackish and continental ecosystems across the Indo-West Pacific (IWP) region (Fig. 1). It comprises 54 valid species that are currently classified into seven valid genera: the species-rich genera *Ambassis* (21 species) and *Parambassis* (21 species) and the species-poor genera *Gymnochanda* (five species), *Tetracentrum* (three species), *Denariusa* (two species), *Chanda* (one species) and *Paradoxodacna* (one species) (Fricke et al., 2023). The genus *Ambassis* contains freshwater and marine/brackish species while the other six genera are freshwater. Several species of *Ambassis* are found along coastal areas and often in brackish water in estuarine and mangrove regions where they form an important community of these fragile ecosystems. Several of these marine/brackish species are widely distributed along the coasts of the Indo-West Pacific (IWP) region (e.g., *Ambassis interrupta* and *Ambassis dussumieri*) whereas the distribution of each freshwater species is restricted to one of these four continental regions: Madagascar, South Asia (=Indian subcontinent), Southeast Asia and Papua-New Guinea-Australia.

The biogeographical causes explaining the complex distribution of Ambassidae within the IWP region (with freshwater species having narrow distributions and marine species having widespread distributions), have not been studied. The general physical configuration of the IWP region was shaped by the fragmentation of East Gondwana (Antarctica, Madagascar, India, and Papua-New Guinea-Australia) that ended about 90 million years ago (Ma), followed by the northward drifting of the India plate until its collision with Eurasia, about 35 Ma and the closure of the Tethys Sea after Africa and Eurasia collided, about 25 Ma (Aitchison et al., 2007; Ali & Aitchison, 2008; Feng et al., 2017;

Karanth, 2021; Salles et al., 2023). In this context, the trans-IWP distribution of taxa is hypothesized to be the consequence of either freshwater and marine vicariant events caused by the East Gondwanan break up or more recent marine dispersal events. To study the biogeography of Ambassidae and determine the role of vicariant and dispersal mechanisms as well as the evolution of salinity preference, a time-calibrated phylogenetic tree of Ambassidae is needed.

The monophyly of Ambassidae is supported by several morphological synapomorphies (Roberts and Jumnongthai, 1999) but the inter- and intra-familial relationships are poorly studied. Recent molecular and morphological studies found that the Ambassidae belongs to a large clade of spiny-fish, the Ovalentaria (e.g. Hughes et al., 2018; Thieme et al., 2022). This clade comprises 48 families and>6,000 species (Hughes et al., 2018; Thieme et al., 2022). Within the Ovalentaria, the family Mugilidae is the possible sister group of Ambassidae (Hughes et al., 2018; Ghezelayagh et al., 2022). The relationships within the Ambassidae have not yet been comprehensively investigated. Whereas the species-poor genera are diagnosed based on synapomorphies, the two largest genera, *Ambassis* and *Parambassis*, lack synapomorphy-based definition and their respective monophyly has never been comprehensively tested (Roberts, 1995). Verma et al. (2019), who focused their molecular study on Indian freshwater ambassids, found *Chanda nama* nested within *Parambassis*.

In this study, we aim to examine the biogeography of the family Ambassidae by reconstructing a molecular time-calibrated phylogenetic tree comprising 33 species (representing about 60 % of the 54 valid ambassid species) and six genera (out of seven). Using this tree, we first comment on the genus-level classification of this family; we then infer the salinity preference evolution within Ambassidae, the geographical origin of the ancestors, and the biogeographical events (vicariant versus dispersal) leading to distribution shifts and their chronology.

2. Material and methods

2.1. Taxonomic and molecular sampling

We collected 32 specimens representing 17 species from five out of the seven currently valid ambassid genera (Fricke et al., 2023). After capture, specimens were handled according to the ethical prescriptions proposed by Blessing et al. (2010). Pectoral fin clips or dorsal muscles were taken and preserved in 99 % ethanol. All specimens were preserved either in 10 % formaldehyde or 95 % ethanol before being deposited in the ichthyological collections of the Makmal Rujukan Zoologi (Zoological Reference Laboratory), Universiti Sains Malaysia (collection code: USMFC), and South China Sea Repository and Reference Centre, Universiti Malaysia Terengganu (UMTF) (Museum accession numbers given in Table S1). We then combined our 17 species with 16 species (one of them from the genus *Denariusa*) examined in Mahon (2007), Li et al. (2011), Page and Hughes (2010), Steinke et al. (2016), Verma et al. (2019), Tims et al. (2021) and Ghazali et al. (2022). Genetic data of these additional species which include *Denariusa*, were directly retrieved from GenBank database along with information on collection localities and museum vouchers of examined specimens (Table S1). Our ambassid dataset comprises 49 terminal taxa representing 33 species and six genera. The freshwater genus *Tetracentrum*, endemic to south New Guinea, is the only missing genus in our study (for more information on *Tetracentrum*, see Allen & Burgess, 1990).

We completed our dataset with 22 non-Ambassidae taxa selected from different actinopterygian lineages, including representatives of the Ovalentaria (Table S2). The function of these taxa is twofold: 1) to provide a test for the monophyly of the family Ambassidae, and 2) to be used for calibrating time divergence within the family Ambassidae.

We selected three mitochondrial markers, the barcode portion of the cytochrome *c* oxidase subunit I (COI) gene (654 base pairs [bp]), the
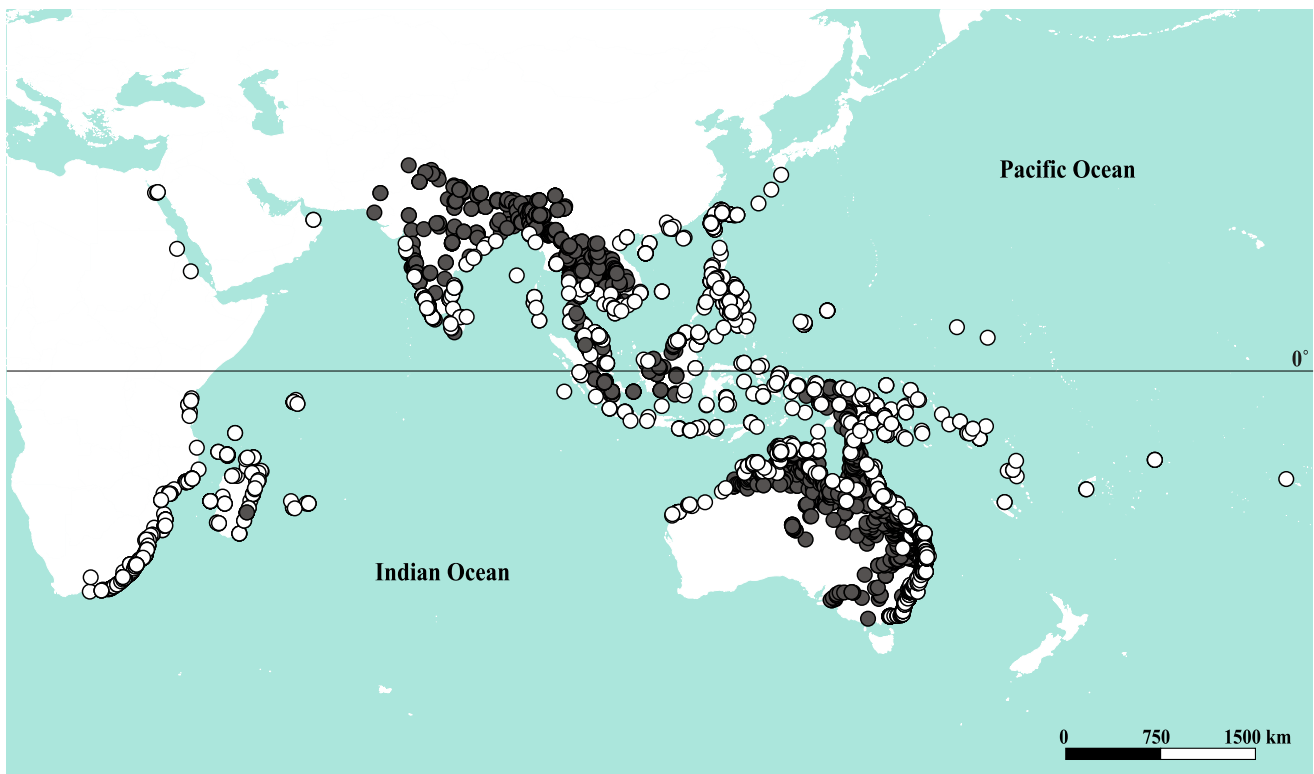

Fig. 1. Map of the Indo-West Pacific region showing the geographical distribution of the fish family Ambassidae. Raw data was extracted from GBIF database (https://www.gbif.org/). White dots represent the occurrences of marine species (from the genus *Ambassis*); grey dots represent the occurrence of freshwater species (from all genera including the Australian freshwater species of *Ambassis*, such as *Ambassis elongata* and *Ambassis agassizi*, and the Malagasy freshwater species *Ambassis fontoynonti*).

nearly complete cytochrome *b* (Cyt*b*) gene (1,108 bp), and the partial 16S ribosomal RNA (16S) gene (670 bp) along with two nuclear markers, the partial recombination activating protein 1 (RAG1) gene (1,018 bp) and the partial rhodopsin (RHO) gene (697 bp).

2.2. DNA isolation, amplification and sequence editing

DNA was extracted from fin and muscle tissues using the hexadecyltrimethylammonium bromide (CTAB) extraction method (Grewe et al., 1993). Each marker was individually amplified by Polymerase Chain Reaction (PCR). The COI gene fragment was amplified using primer pair combinations selected among the universal primers FishF1, FishR1, FishF2 and FishR2 (Ward et al., 2005); the 16S gene fragment was amplified using the primers 16Sar-L (Palumbi et al., 1991) and 16Srev2 (Agorreta et al., 2013); the RAG1 gene fragment was amplified using primer pair combinations selected among the universal primers RAG_F1, RAG_R1, RAG_F2 and RAG_R2 (Mat Jaafar, 2014). We amplified the Cyt*b* and RHO genes fragments using two newly designed Ambassidae-specific primer pairs: AmbCytb_F (5′-CTT GAA AAA CCA CCG TTG TYA TTC AAC-3′) and AmbCytb_R (5′-AGA ATC YTA GCT TTG GGA GTT A-3′), and AmbRhod_F (5′-TTC CCC ATC AAC TTC CTC ACT-3′) and AmbRhod R (5′-AGC TCT TGG CAA AGA AGG CTG-3′).

PCR mixture was performed in a final volume of 25 µL containing 15.75 µL ddH2O, 5.5 µL 5x MyTaq Red reaction buffer (Bioline Reagents Ltd, UK), 0.5 µL of each primer, 0.25 µL iTaq DNA polymerase (iNtRON Biotechnology, Inc., Gyeonggi-do, Korea) and 2.5 µL template DNA. PCR amplification was performed using a Major Cycler, CYCLER-25 (Major Science, USA), with annealing temperature set to 47.9 °C for COI, Cyt*b* and 16S, 48.9 °C for RAG1 and 50 °C for RHO. Unpurified PCR products were sent to Apical Scientific Sdn. Bhd. (Selangor, Malaysia) for sequencing analysis by an ABI3730XL capillary sequencer (Applied Biosystems, USA), using the same primer-pair as used for PCR.

For each individual gene, the chromatograms of forward and reverse sequences were separately edited prior to assembly. Each nuclear gene (RAG1 and RHO) sequence was carefully checked by eye to identify allelic polymorphism (double peaks), and the IUPAC nucleotide ambiguity code was used to code heterozygous sites when present. The protein-coding gene sequences (i.e., COI, Cyt*b*, RAG1 and RHO) were translated into proteins to verify the absence of stop codons. The separate alignments of these genes were straightforward and done by eye. The 16S gene fragment was aligned using MAFFT v.7 web server (https://mafft.cbrc.jp/alignment/server/) (Katoh et al. 2019) under default parameters. All sequences generated for this study are deposited in GenBank database (GenBank accession numbers provided in Table S1).

2.3. Phylogenetic analyses

Phylogenetic relationships among ambassid taxa were inferred using Maximum Likelihood (ML) and Bayesian methods of reconstruction. We first checked for conflicting phylogenetic signal among genes (and between mitochondrial and nuclear datasets) by reconstructing individual gene ML trees, using RaxML-HPC2 (Stamatakis, 2014) and a GTR + G model of sequence evolution (Fig. S1-S3). Nodal supports were estimated with 1000 standard bootstrap (BS) replicates. In absence of significant phylogenetic conflict in these trees, we combined all markers together (total aligned positions = 4,147 bp) and simultaneously analysed them. The best partitioned scheme and models of sequence evolution were estimated with PartitionFinder2 (Lanfear et al., 2017). Our dataset was divided into seven partitions (each with its own model of sequence evolution): first (GTR + I + G), second (GTR + I + G) and third (TIM + I + G) codon positions of the concatenated COI and Cyt*b* genes, 16S gene (GTR + I + G) and first (TVM + G), second (TVM + I + G) and third (TVM + I + G) codon positions of the concatenated RAG1 and RHO genes. We then conducted a partitioned ML phylogenetic analysis of our dataset using RaxML-HPC2. The tree was visualised and edited using the

software FigTree v1.4.4 (Rambaut, 2018).

We then estimated a time-calibrated phylogeny using a partitioned Bayesian method incorporating a relaxed molecular clock method using BEAST2 v.2.6.7 (Bouckaert et al., 2019). Only one specimen per species was retained. Substitution and clock models were unlinked across all partitions whereas tree models were linked. A lognormal relaxed-clock model was applied to each partition under a Yule tree prior. The fossil record of the family Ambassidae is not informative in providing reliable time calibration information within the family (Text S1). Therefore, we time calibrated our tree by using nine percomorph fossils selected from Hughes et al. (2018) and Ghezelayagh et al. (2022). These studies inferred large scale time trees of the Actinopterygii, using an extensive set of fossils to calibrate time divergence. The list of fossils we used are listed Table S3 and their respective paleontological justification and references for selecting these calibration points are given in Hughes et al. (2018) and/or Ghezelayagh et al. (2022). The monophylies of nine outgroup pairs, selected for fossil-based time calibration (see Table S3), were a priori constrained. For each time constraint, age followed an exponential distribution prior, with a strict minimum age set to the minimum age of the corresponding fossil and a soft maximum set as given in Hughes et al. (2018) and/or Ghezelayagh et al. (2022). The tree was oriented by forcing the monophyly of all taxa excluding *Neobythites* (Ophidiiformes). Based on available evidence, the age of the root of the tree was set to 135 million years (Hughes et al. 2018; Ghezelayagh et al. 2022).

We conducted two separate runs, each consisted of 500 million generations. Parameters and trees were sampled every 50 thousand generations. After checking for convergence (Effective Sampling Size > 100) with Tracer v1.7.1 (Rambaut et al., 2018), the first 25 % cycles of each run were discarded as burn-in, and the remaining cycles were combined using LogCombiner v.2.6.7 (Bouckaert et al., 2019). The maximum credibility consensus tree was calculated using TreeAnnotator v.2.6.7 (Bouckaert et al., 2019). The posterior probabilities of inferred clades and the mean and 95 % Highest Posterior Density (HPD) interval of divergence time estimates were visualised and edited using FigTree v.1.4.4. All phylogenetic analyses were conducted on XSEDE in the CIPRES Science Gateway (Miller et al., 2010).

Finally, a Bayesian time-calibrated species tree of Ambassidae was also reconstructed using a multispecies coalescent-based approach as implemented in StarBeast3 v.1.0.5 (Douglas et al., 2022), a package available in BEAST2 suite. The RHO gene dataset (697 positions, 40 taxa), RAG1 gene dataset (1,018 positions, 40 taxa) and the concatenated mitochondrial genes dataset (2,432 positions, 55 taxa) were unlinked and assigned their own model of molecular evolution, a Species Relaxed Clock model, and a Yule tree model. The fossil-based calibration procedure was the same as described above. Two independent MCMC analyses were carried out, each for 100 million generations, and were sampled every 50 thousand generations. We discarded 25 % of the trees of each run as burn-in before combining the remaining into a single file and calculating the maximum credibility consensus tree, mean node ages and their 95 % HPD as described above.

2.4. Character evolution reconstruction

Habitat preference evolution relative to salinity was inferred using the likelihood-based symmetric one-rate (Mk1) model of ancestral character reconstruction implemented in Mesquite v.3.7.0 (Maddison and Maddison, 2019) onto our time-calibrated Bayesian phylogenetic trees previously inferred with BEAST2 and StarBeast3. We categorised our taxa as either marine/brackish or freshwater following Fraser-Brunner (1955), Allen and Burgess (1990), Roberts (1995), Froese and Pauly (2023) and our field observations.

The habitat preference of the most recent common ancestor (MRCA) of the family Ambassidae was also specifically reconstructed using a Bayesian approach implemented in BayesTraits v.4 (Pagel & Meades, 2006). The habitat preferences of terminal taxa were categorized as

described in the previous paragraph. The ADDNODE command was used to specify the node of interest (i.e., the MRCA of the family Ambassidae). Two independent MCMC runs were conducted, each of 40 million generations, sampling every 1,000 generations and discarding the first 25 % as burn-in. The average of the mean values of the proportional likelihoods was calculated with Tracer.

2.5. Biogeographical analysis

Ancestral geographical range evolution was inferred on our time-calibrated Bayesian trees using four different models: dispersal-extinction cladogenesis (DEC) (Matzke, 2014) and DEC + J (Ree et al., 2008); ML version of Dispersal-Vicariance analysis (DIVAlike) (Ronquist and Sanmartín, 2011) and DIVAlike + J (Matzke, 2014) implemented in the R software BioGeoBEARS v1.1 (Matzke, 2018). Although Ree and Sanmartín (2018) critiqued the parameter "+J" (jump dispersal at speciation), Matzke (2022) answered to all concerns raised by these authors. The BAYAREAlike model was not considered because it excludes the vicariance process (Landis et al., 2013; Matzke, 2013). The log-likelihood (LnL) and corrected Akaike Information Criterion (AICc) were calculated for each model-based reconstruction using Bio-GeoBEARS and then compared to choose the best fitting one. The distribution of each terminal species (regardless of its salinity preference) was assigned to four areas: Australia (including Papua New-Guinea) (A), South Asia (B), Southeast Asia (C) and Africa (including Madagascar) (D). The maximum number of areas that ancestors at the nodes can occupy was not constrained and set to four, and no area combinations were excluded. Prior to analysis, the outgroups were removed from the tree file to ensure the area range estimation analysis is only for the family Ambassidae.

3. Results

3.1. Phylogenetic relationships

The concatenated alignment of the five genes consists of 4,147 characters, of which 1,394 are variable, and 1,269 parsimony-informative. The ML and Bayesian phylogenetic analyses of the combined dataset resulted in well resolved and overall congruent trees (ML and Bayesian trees shown in Fig. 2 and Fig. S4, respectively). In these trees, the family Ambassidae is monophyletic (supported by a bootstrap support [BS] of 100 % and a posterior probability [PP] of 1). The two largest ambassid genera, *Ambassis* and *Parambassis*, are each not monophyletic. Instead, our analyses identified four main lineages among Ambassidae which are labelled 1 to 4 in Fig. 2. The first lineage corresponds to the Australian freshwater genus *Denariusa*. Lineage 2 is composed of five Australian species of *Ambassis* (*A. jacksoniensis, A. agassizii, A. agrammus, A. elongata* and *A. macleayi*) (BS = 95 %, PP = 1). Lineage 3 comprises seven species of *Ambassis* (*A. ambassis, A. natalensis, A. interrupta, A. macracanthus, A. dussumieri, A. marianus* and *A. vachellii*) (BS = 100 %, PP = 1). Lineage 4, the largest one, comprises two species of *Ambassis* (*A. nalua* and *A. kopsii*) along with all species currently classified into the genera *Parambassis, Chanda, Gymnochanda* and *Paradoxodacna* (BS = 99 %, PP = 1). The relationships among these four lineages (1 to 4) are well supported, with *Denariusa* being the sister group to the other lineages (BS = 93 %, PP = 1) and Lineage 4 the sister group of Lineages 2 + 3 (BS = 82 %, PP = 1).

The multispecies coalescent tree inferred with StarBeast3 (Fig S5) is topologically similar to the ML and BEAST2-Bayesian phylogenetic trees built from our concatenated dataset (Fig. 2 and Fig. S4) in inferring: 1) the monophyly of Ambassidae, 2) the same phylogenetic position of *Denariusa* (sister group to the rest of the Ambassidae), and 3) the respective monophyly of Lineages 2, 3 and 4. However this tree differs from the others by its overall poor node-support, with most PP within the Ambassidae below 0.6 (and as low as 0.3) and by some topological incongruence within Lineages 2 and 4.

3.2. Divergence time estimation

Our concatenated Bayesian time-calibrated phylogenetic tree is shown in Fig. S4. The estimated age of the crown group family Ambassidae is 55.7 Ma (95 % Highest Posterior Density [HPD] = 63.4–48.7 Ma). The early diversification of the family (leading to the four main lineages) took place within about 11 million years, between 55.7 and 44.3 Ma; with 44.3 Ma (95 % HPD = 50.5–37.7 Ma) corresponding to the divergence of Lineages 2 and 3. The analysis further reveals that Lineages 2 and 3 started to diversify at 32.0 Ma (95 % HPD = 39.6–24.5 Ma) and 30.1 Ma (95 % HPD = 39.1–21.1 Ma), respectively. Lineage 4 starts to diversify at 39.0 Ma (95 % HPD = 44.9–33.7 Ma).

The time divergences inferred from the multispecies coalescent tree (Fig S5) were overall about 10 % younger. For example, the age of the crown group Ambassidae was estimated to only 51.7 Ma (instead of 55.7 Ma previously).

3.3. Ancestral habitat reconstruction

The reconstructions of the ancestral habitat preference relative to salinity is shown in Fig. 3 (using BEAST2 tree) and Fig S6 (using Star-Beast3 tree). Each reconstruction indicates that the most recent common ancestor (MRCA) of the family was most likely freshwater adapted ($p = 0.66$ and 0.88). This result is consistent with the BayesTraits analysis ($p = 0.89$ for freshwater, using BEAST2 tree). Then, we identified three and four independent transition events from a freshwater to a marine environment when using the BEAST2 tree and the StarBeast3 tree, respectively. Using the BEAST2 tree, the first habitat shift occurred between mid-Eocene to early Oligocene (between 44.3 and 32.0 Ma) in the Lineage 3, leading to several marine species of *Ambassis* ($p = 0.56$). The second event led to the Australian marine *A. jacksoniensis* and occurred during mid-Oligocene, after 30.1 Ma ($p = 0.65$). The last one led to *A. nalua* and *A. kopsii* between late Eocene to mid-Oligocene (between 39.0 and 29.6 Ma) ($p = 0.65$). The analysis based on the StarBeast3 inferred the same first two events. However, because *A. nalua* and *A. kopsii* do not form a monophyletic group in this tree, two independent freshwater-to-marine transitions were needed, one leading to *A. nalua* and the other one leading to *A. kopsii* (Fig S6).

3.4. Ancestral range estimation

Using the BEAST2 tree, all four model-based reconstructions of ancestral areas resulted in almost similar scenarios (Fig. 4 and Fig. S7). Based on the corrected Akaike information criterion (AICc), the DEC + J model is favoured (AICc = 140.3) (Table S4; Fig. 4). The key findings in this reconstruction (Fig. 4) can be summarised as follows: 1) The distribution of the most recent common ancestor of Ambassidae was Australia; 2) the endemic Oriental (South Asia + Southeast Asia) taxa were reconstructed as a single long-distance range extension event from Australian ancestors, dated between 55.7 and 49 Ma (event "1" in Fig. 4); 3) the marine species of Lineage 3 were reconstructed as a long-distance range extension from Australia to Africa dated between 44.3 and 32 Ma (event "2" in Fig. 4) before seven long-distance range extensions among regions; 4) the ancestral area reconstruction among freshwater species is statistically better supported than that among marine species; 5) ambassids colonised South Asia from Southeast Asia about 34.4 Ma through a single founder event (event "3" in Fig. 4); 6) subsequently, two independent lineages of Southeast Asian ambassids originated through a founder event from South Asia (leading to *Parambassis siamensis* and *Gymnochanda filamentosa*); and finally 7) *Parambassis gulliveri* colonised the Australian region from Southeast Asia after 20.4 Ma.

Using the StarBeast3 tree, the ancestral area reconstruction during the early evolution of the Ambassidae provided a similar scenario, with Australia as the region of origin, and two early trans-marine dispersal
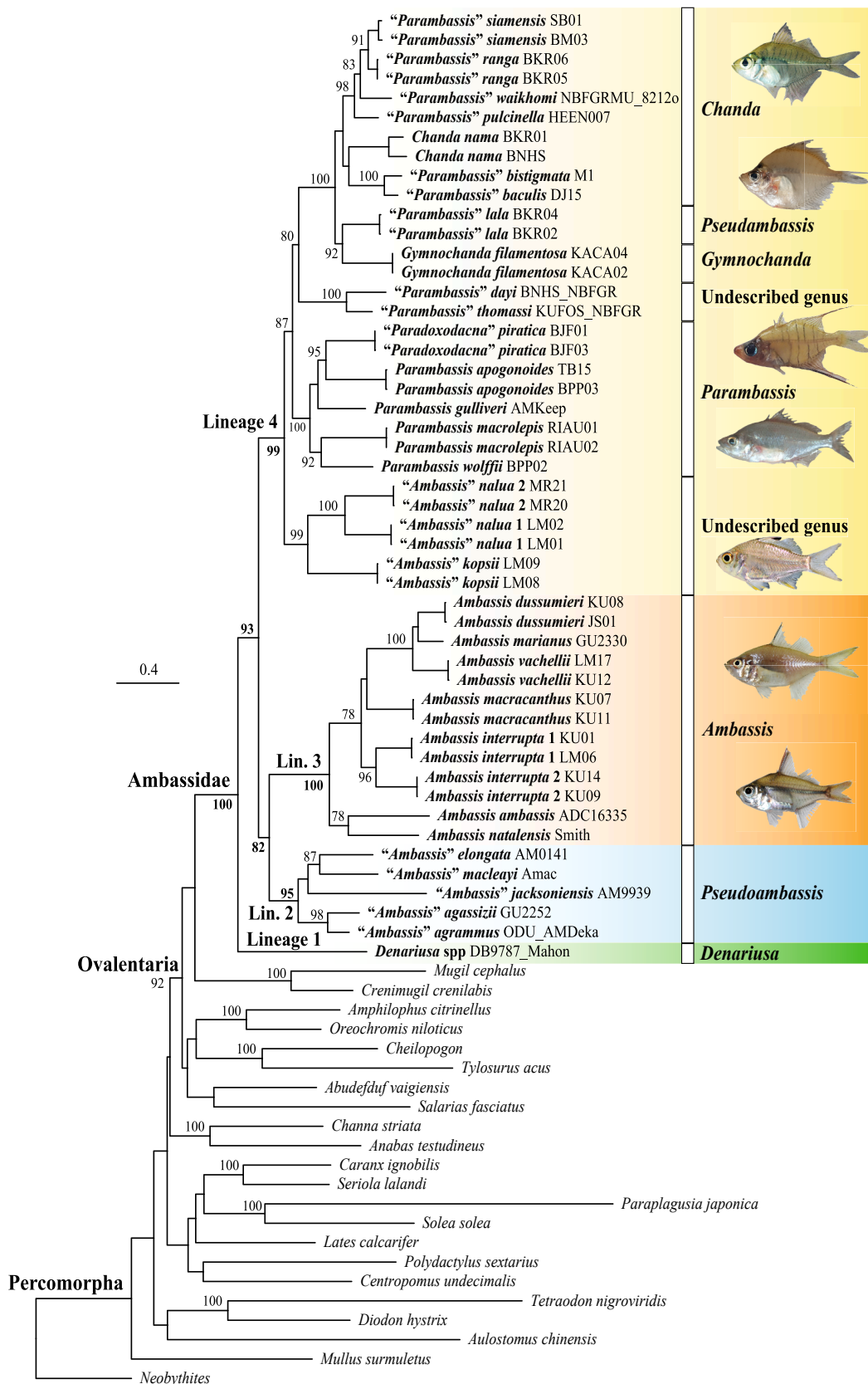
Fig. 2. Maximum Likelihood (ML) phylogenetic tree of the family Ambassidae based on concatenated information from five molecular markers (COI, Cyt*b*, 16S, RAG1 and RHO genes; total: 4,124 nucleotide positions). The software RaxML-HPC2 was used to build this tree. Tree rooted with *Neobythites* (Ophidiiformes). Branch lengths proportional to number of substitutions. Bootstrap proportions indicated at nodes when >75 %. The generic names of ambassid species needing their generic positions to be re-investigated are indicated with quotation marks. A tentative revised generic classification of Ambassidae is shown at the right of the figure.
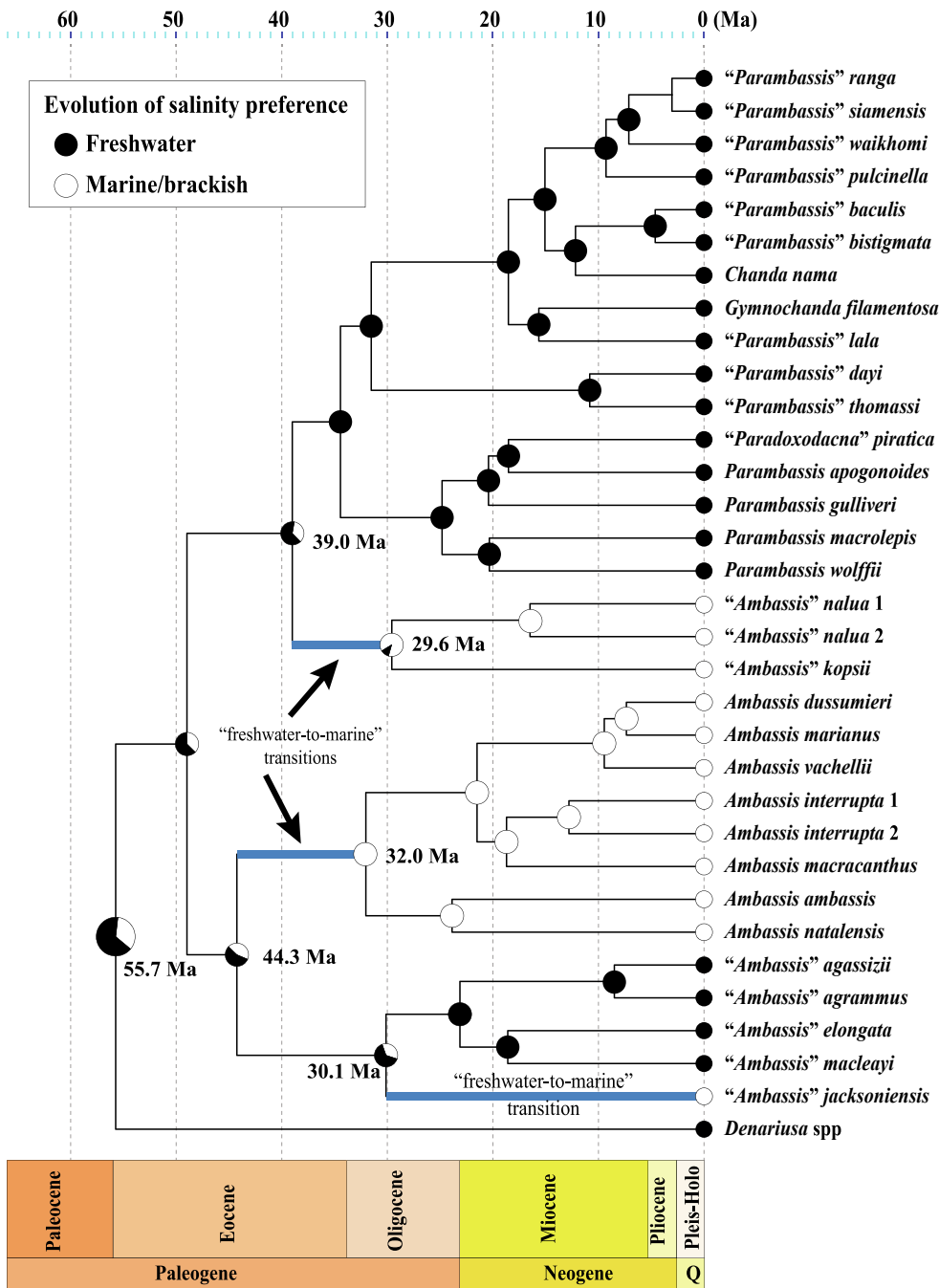
Fig. 3. Maximum-likelihood reconstruction of the evolution of salinity preference within the family Ambassidae inferred on the multi-locus time-calibrated Bayesian phylogeny (inferred with BEAST2). The "Mk1" model of character evolution was used. Salinity preference is classified into two states: "marine/brackish" indicated in white and "freshwater" in black. The relative probabilities of the alternative states (sum = 1) are drawn using pie charts at each node. Freshwater-to-marine transitions are highlighted with blue branches. Salinity preferences were extracted from Fraser-Brunner (1955), Allen and Burgess (1990), Robert (1995), Froese and Pauly (2023) and our own field observations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

events. But the direction and timing of these two events are slightly different (Fig S8).

4. Discussion

The wide distribution of Ambassidae in coastal and freshwater ecosystems throughout the IWP region makes this group a good model to further examine the biogeography of this region. We first briefly examine the phylogenetic implications of our results on the current classification of the family Ambassidae and call for some taxonomic changes. Then, we discuss the distribution of ambassids as the result of post-Gondwanan long-distance dispersal (range extension events) from Australia to other regions during the early-mid Cenozoic. Finally, we highlight a paradox observed in Ambassidae and some other circum-Indian Ocean groups of freshwater fishes: the absence of extinct or living marine taxa attesting their past marine dispersals.

4.1. Taxonomic implications

Our phylogenetic results (Fig. 2; Figs. S4-S5) call for a thorough taxonomic revision of the two most speciose genera, *Ambassis* and *Parambassis*, because each of these two genera is not monophyletic. *Ambassis* is polyphyletic and comprises at least three independent lineages (Fig. 2). *Parambassis* is paraphyletic relative to *Chanda, Paradoxodacna* and *Gymnochanda*.

For rendering the classification of the Asian freshwater ambassids natural in minimising the number of modifications, we recommend the following actions: 1) to restrict the composition of the genus *Parambassis* to the following freshwater species: *P. apogonoides* (type species of *Parambassis*), *P. wolffii, P. macrolepis, P. gulliveri*, along with
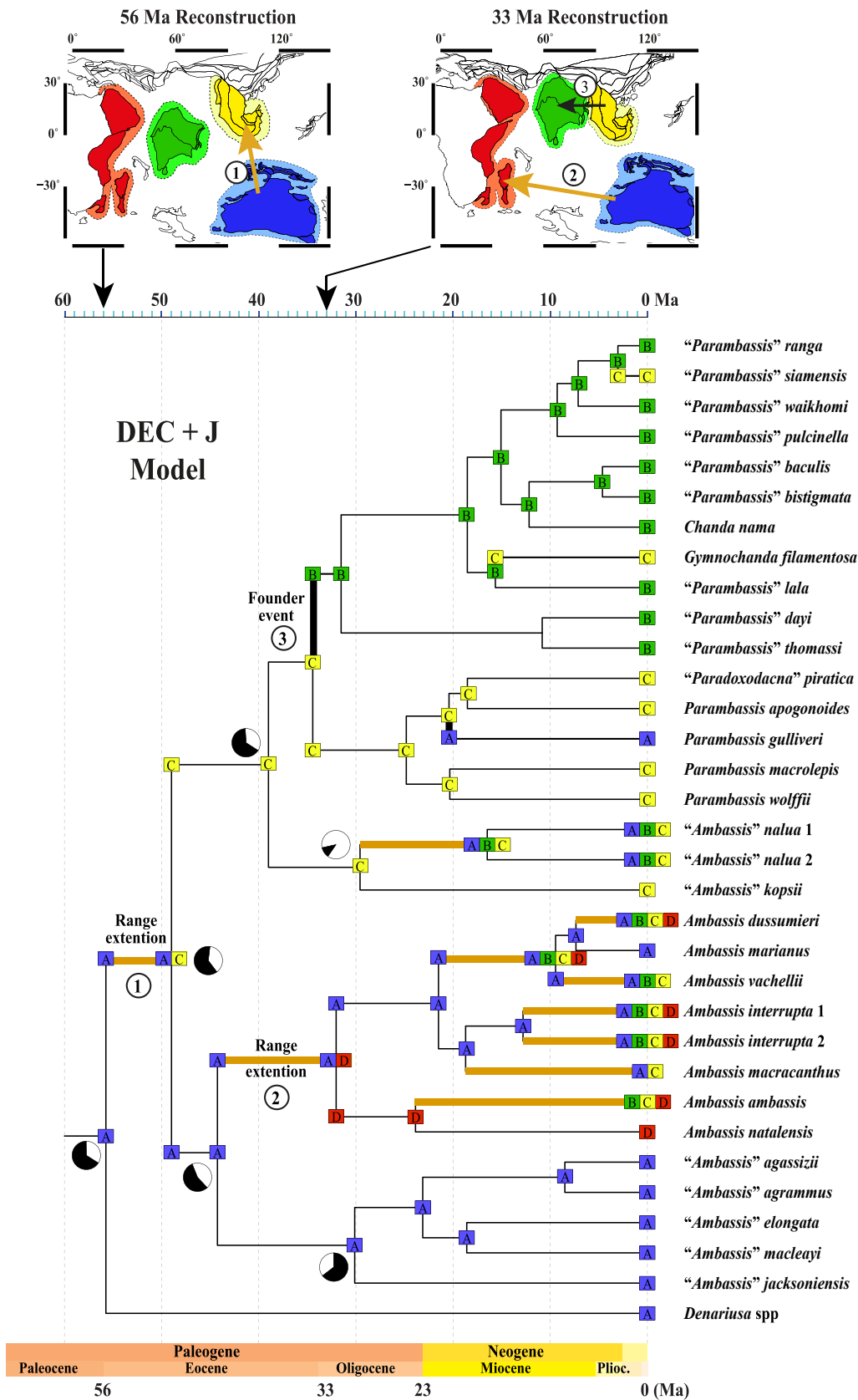
Fig. 4. Ancestral area estimates of the family Ambassidae under the unconstrained reconstruction model DEC + J, using the multi-locus time-calibrated Bayesian phylogeny (inferred with BEAST2). Letters from A to D represent the freshwater + coastal marine regions (see the inserted two palaeomaps at the top) used for the reconstruction: A, Australia (including Papua-New Guinea) (blue); B, South Asia (=Indian plate) (green); C, Southeast Asia (yellow); D, East Africa (including Madagascar) (red). The geographical distributions of extant taxa are indicated in the tree, immediately left to their corresponding names. The ancestral ranges at nodes indicate the inferred ancestral distributions before speciation and selected ancestral ranges at corner positions represent geographical ranges immediately after speciation. Founder events are highlighted with thick black branches whereas range extensions are indicated with thick orange branches. Pie charts nearby selected ancestral area reconstructions show the probability (white) of the corresponding reconstruction (the sum of probabilities of all other reconstructions is shown in black). Time scale is in millions of years ago (Ma). Insert palaeomaps represent reconstructions of the Indian Ocean region, at 56 and 33 Ma, in which are indicated with arrows the direction of the main dispersal events. Regional code-colour corresponds to that used in the phylogenetic ancestral area analysis. The palaeomaps were created at www.odsn.de with data from Hay et al. (1999). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Paradoxodacna piratica; 2) to revalidate the monospecific genus *Pseudambassis* for *P. lala*; 3) to transfer all other small-scale species of *Parambassis* of the Oriental region (i.e. *P. siamensis, P. ranga, P. baculis, P. pulcinella*, etc.), into the genus *Chanda* (type species *Chanda nama*); 4) to describe a new genus to accommodate *P. dayi* and *P. thomassi*. The two species of *Parambassis* from Papua-New Guinea not examined in the current study, herein (i.e., *P. alleni* and *P. confinis*) are left in this genus until further evidence.

Regarding the situation of the genus *Ambassis*, our results support its dismantlement into three smaller monophyletic genera: 1) a restricted *Ambassis* (type species *A. ambassis*) including most of the marine/brackish species, 2) a revalidated genus *Pseudoambassis* (type species *P. macleayi*) for the Australian species classified into *Ambassis* by Fraser-Brunner (1955) (i.e., *P. macleayi, P. elongata, P. agassizii, P. agrammus* and *P. jacksoniensis*) and 3) an undescribed genus comprising *A. nalua* and *A. kopsii*.

This revised generic taxonomic classification recognises 10 genera within the family Ambassidae, including two undescribed genera and the genus *Tetracentrum* which was not available for our study. There is one more species of particular interest that was not examined in this study, the Malagasy *Ambassis fontoynonti* (Pellegrin, 1932). Knowing the phylogenetic position of this species is interesting for two reasons: This species exhibits several "primitive" characteristics for this family, such as a weak development of head serration, questioning its current generic placement (Fraser-Brunner, 1955; Anderson and Heemstra, 2003) and this is the only species of freshwater *Ambassis* outside the Australian region (Pellegrin, 1932).

4.2. Ambassidae originated and early diversified in freshwaters

The large majority of the 400+ fish families have a narrow range of tolerance relative to salinity, being either freshwater or marine adapted, indicating that habitat transitions relative to salinity were rare evolutionary events at this taxonomic level (Bloom and Lovejoy, 2012; Kirchhoff et al., 2017). The Ambassidae is one of the mere 4 % fish families which comprises both marine/brackish and freshwater species (Vega and Wiens, 2012). Several species of the genus *Ambassis* preferentially live in coastal marine and brackish environments (Martin, 1988), whereas species in all other genera along with a few species of *Ambassis*, are restricted to freshwater environments. Our habitat preference reconstruction supports a freshwater origin for the family. Although the current statistical evidence for this hypothesis is not overwhelming, we note that future phylogenetic analysis of the freshwater genus *Tetracentrum* along with the freshwater *Ambassis fontoynonti* may increase the support for this hypothesis. In contrast, the inclusion of the five missing marine species of *Ambassis* will likely have a limited impact on the habitat preference inference of Ambassidae because our taxonomic sampling already captured most of the morphological diversity of *Ambassis*.

We infer only three freshwater-to-marine transitions and no marine-to-freshwater transition in the course of the Ambassidae evolution. The absence of marine-to-freshwater transition has important biogeographical implications because it means that all common ancestors shared by the freshwater ambassids (i.e., *Denariusa*, the Australian freshwater species of *Ambassis* and the Oriental freshwater species of Ambassidae) were also freshwater adapted; there is no phylogeny-based evidence of any intermediate marine ancestors among these lineages (but see below). The Ambassidae is, therefore, considered as a freshwater-originating family with living marine taxa being secondarily derived.

The oldest (about 66 Ma) fossils of Ambassidae (†*Dapalis erici* and †*Ambassidarum cappetai*) were likely freshwater (Nolf et al., 2008; Kapur and Khosla, 2019). Their phylogenetic position is not resolved relative to living forms, but our time-diversification scale suggests these fossils are stem Ambassidae because they are older than the most recent common ancestor of the living ambassids. Nonetheless, these fossils indicate that ambassids were already adapted to freshwater environments in the early

Cenozoic; this agrees with our finding that the most recent common ancestor of the Ambassidae was freshwater adapted. This finding has biogeographical implications.

4.3. Early Cenozoic long-distance marine dispersal across the Indian Ocean

The vicariant hypothesis, linked to the Gondwana fragmentation followed by the India drifting ("Indian ferry" hypothesis), is rejected for the Ambassidae because molecule-based dating studies support an early Cenozoic, post-Gondwanan age for the family (e.g., Hughes et al., 2018; Ghezelayagh et al., 2022; this study). The fossil record of Ambassidae is meagre but its stratigraphy agrees with that of the general fossil record of the Teleostei (see Benton et al., 2000 and Text S1). It also rejects the "Indian ferry" hypothesis because the oldest fossils are younger than the late fragmentation of Gondwana. The oldest fossil records assigned to Ambassidae are the 66 My-old otolith fossils of †*Dapalis erici* from a freshwater paleoecosystem in the Rangapur intertrappean succession in India and †*Ambassidarum cappetai* from a mixed assemblage of freshwater and marine elements in the Nagpur intertrappean bed in India too (Nolf et al., 2008; Kapur and Khosla, 2019).

Our ancestral area reconstruction combined with the ancestral habitat preference revealed that the most recent common ancestor of Ambassidae originated in the Australian freshwater region, 55.7 Ma. Two independent long-distance dispersal events from the Australian region explain the ambassid colonisation of the Indian Ocean region, both marine and freshwater (Fig. 4).

An Australian origin for this family agrees with the observation that the morphologically most generalised ambassid genera, *Denariusa* and *Tetracentrum*, are endemic to this region (Fraser-Brunner, 1955). Both genera are freshwater as well. We note that the oldest (66 Ma) ambassid fossils have been found in India, not in Australia, and there is no ambassid fossil of early Cenozoic age found in Australia. However, there is no evidence that the Indian otolith-based fossils belong to the crown group Ambassidae (Nolf et al., 2008); therefore, they do not contradict our hypothesis. We also predict the presence of ambassid fossils of early Cenozoic age in Australia.

The distribution of the Oriental (South Asia + Southeast Asia) clade of Ambassidae (Lineage 4 in Fig. 2) is explained by a single range extension event from Australia to Southeast Asia that occurred between 55.7 and 49 Ma (Fig. 4). At that time, the Australian and Southeast Asia regions were separated by several hundred kilometres of marine environment (Ali and Aitchison, 2008; Salles et al., 2023). We note that this dispersal event is not corroborated by our ancestral habitat preference reconstruction because the most recent common ancestor of Lineage 4 is suggested to have been freshwater adapted. In the absence of possible alternative hypotheses explaining a freshwater dispersal (such as the drifting of a raft comprising a freshwater pond or the transport of eggs by birds), we hypothesize that the intermediate marine forms needed to accomplish such marine dispersal must have been selectively extinct. The origin of the freshwater *Parambassis gulliveri* indirectly supports this hypothesis (Fig. 4). *Parambassis gulliveri* is a freshwater species distributed in the Australian region. In our study, this species is nested within a freshwater clade of Southeast Asian species. We inferred that *P. gulliveri* colonised the Australian region after a founder event (around 20 Ma) from Southeast Asia; this event did not leave any marine intermediates. In this case, the phylogenetic signal supporting this founder event is stronger and led to the hypothesis that marine intermediates were selectively extinct.

A second marine long-distance dispersal from Australia to Africa region is needed to explain the wide distribution of the Lineage 3 that comprises only marine species of *Ambassis*. This event, broadly dated between 44.3 and 32 Ma, is linked with a freshwater-to-marine transition event as shown Fig. 3. Further discussion on the biogeography of this Lineage is currently difficult because 1) there are several species which are not included in our analysis and 2) the precise distribution of

many species in this lineage are still unknown. As a consequence, all inferred range contractions and expansions in this lineage are statistically only weakly supported.

4.4. Southeast Asia-South Asia interchanges

Our analysis shows that all South Asian freshwater species originated after a founder event from Southeast Asia that we estimate occurred during the late Eocene, 34.4 Ma. At that time, India was already landlocked to Eurasia, offering ample possibilities of biotic interchanges between Southeast Asia and India (=South Asia) (Klaus et al., 2016; Salles et al., 2023). Klaus et al. (2016) estimated that dispersal between India and Southeast Asia started as early as 50 Ma; but it is only after 35 Ma that the number of dispersals significantly increased as a direct consequence of the contact between these two tectonic regions. Later in the evolution of Lineage 4, freshwater ambassids re-entered Southeast Asia from South Asia, leading to the genus *Gymnochanda* (after 18.5 Ma) and *Parambassis siamensis* (after 3 Ma).

4.5. Another case of long-distance marine dispersal across the Indian Ocean

The fossil record along with recent molecule-based fossil-calibrated dating studies show that there are several trans-oceanic freshwater taxa in the Indian Ocean region for which their distribution can only be explained by early–mid Cenozoic marine long-distance dispersal because the age of these taxa strictly post-dated the tectonic fragmentation of the Gondwana (Ali and Aitchison, 2008; Salles et al., 2023). These trans-oceanic freshwater taxa include Cichlidae distributed in Africa-Madagascar-South Asia (along with South and Central America) and early diverged 60 Ma (Friedman et al., 2013; Matschiner et al., 2020), Melanotaeniidae-Bedotiidae distributed in the Australian region and Madagascar and early diverged 44 Ma or younger (Campanella et al., 2015; Ghezelayagh et al., 2022), Aplocheilidae distributed in Madagascar, Seychelles, South Asia and Southeast Asia and early diverged 40 Ma (Cui et al., 2021), *Scleropages* distributed in Australian region and Southeast Asia diverged 60 Ma (Lavoué, 2015), Notopteridae distributed in Africa, South Asia and Southeast Asia, diverged 60 Ma (Lavoué et al., 2012), cave gobiids distributed in Madagascar and Australia (Chakrabarty et al., 2012) diverged 40 Ma as re-estimated by de Bruyn et al. (2013), cave shrimps distributed in Madagascar and Australia (Jurado-Rivera et al., 2017) and Ambassidae (this study). There are several possibilities to explain the absence of marine-related forms in these groups: 1) the selective extinction of the ancestral marine forms; 2) freshwater taxa retain tolerance to high salinity; 3) unrecognised early-mid Cenozoic freshwater routes connecting different regions. If the first possibility is correct, it is predicted to find marine fossils nested within these taxa. As far as we know, there are no marine fossils found in Cichlidae, Melanotaeniidae, Aplocheilidae and Notopteridae. Marine fossils assigned to *Scleropages* and Ambassidae are of unresolved phylogenetic positions or they represent stem taxa. Regarding the second possibility, Cui et al. (2021) revealed that one species of freshwater aplocheilid found in Seychelles can develop in saltwater and this does not affect its survivability. These authors concluded that the physiological tolerance of this species to marine water explain its marine dispersal. There is no evidence for the third possibility.

5. Conclusion

The most recent common ancestor of the glassperch family Ambassidae was freshwater adapted and lived in the Australian region in the early Cenozoic. Range extension and founder events from the Australian region and across the Indian Ocean, coupled with salinity adaptation, explain the current distribution of Ambassidae. The fossil record does not contradict this scenario but it is only partially informative. We suggest further studies to complete our findings on the biogeography of the family Ambassidae, such as examining the phylogenetic positions of the genus *Tetracentrum*, native of New Guinea, and *Ambassis fontoynonti* from Madagascar or determining the range of the salinity tolerance of freshwater ambassids. Finally, more studies on the biogeography of other freshwater taxa distributed surrounding the Indian Ocean, are needed to provide an overview on the post-Gondwanan breakup dispersal process in this region.

Submission declaration and verification:

The work described has not been published previously, it is not under consideration for publication elsewhere, its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

CRediT authorship contribution statement

Siti Zafirah Ghazali: Investigation, Methodology, Data curation, Writing – original draft. **Sébastien Lavoué:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition. **Tedjo Sukmono:** Resources, Data curation, Writing – review & editing. **Ahasan Habib:** Resources, Data curation, Writing – review & editing. **Min Pau Tan:** Writing – review & editing, Supervision. **Siti Azizah Mohd Nor:** Conceptualization, Resources, Writing – review & editing, Supervision, 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

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2023.107832.

References

Agorreta, A., San Mauro, D., Schliewen, U., Van Tassell, J.L., Kovačić, M., Zardoya, R., Rüber, L., 2013. Molecular phylogenetics of Gobioidei and phylogenetic placement of European gobies. Mol. Phylogenet. Evol. 69, 619–633. https://doi.org/10.1016/j.ympev.2013.07.017.

Aitchison, J.C., Ali, J.R., Davis, A.M., 2007. When and where did India and Asia collide? J. Geophys. Res. Solid Earth 112 (B5).

Ali, J.R., Aitchison, J.C., 2008. Gondwana to Asia: Plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through latest Eocene (166–35 Ma). Earth Sci. Rev. 88, 145–166. https://doi.org/10.1016/j.earscirev.2008.01.007.

Allen, R.G., Burgess, E.W., 1990. A review of the glassfishes (Chandidae) of Australia and New Guinea. Rec. West. Aust. Mus. 34, 139–206.

Anderson, M.E., Heemstra, P.C., 2003. Review of the glassfishes (Perciformes: Ambassidae) of the western Indian Ocean. Cybium. 27, 199–209.

Benton, M., Wills, M., Hitchin, R., 2000. Quality of the fossil record through time. Nature. 403, 534–537. https://doi.org/10.1038/35000558.

Blessing, J.J., Marshall, J.C., Balcombe, S.R., 2010. Humane killing of fishes for scientific research: a comparison of two methods. J. Fish Biol. 76, 2571–2577. https://doi.org/10.1111/j.1095-8649.2010.02633.x.

Bloom, D.D., Lovejoy, N.R., 2012. Molecular phylogenetics reveals a pattern of biome conservatism in New World anchovies (family Engraulidae). J. Evol. Biol. 25, 701–715. https://doi.org/10.1111/j.1420-9101.2012.02464.x.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F.K., Müller, N.F., Ogilvie, H.A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M.A., Wu, C.-H., Xie, D., Zhang, C., Stadler, T., Drummond, A.J., Pertea, M., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15 (4), e1006650.

Campanella, D., Hughes, L.C., Unmack, P.J., Bloom, D.D., Piller, K.R., Ortí, G., 2015. Multi-locus fossil-calibrated phylogeny of Atheriniformes (Teleostei, Ovalentaria). Mol. Phylogenet. Evol. 86, 8–23. https://doi.org/10.1016/j.ympev.2015.03.001.

Chakrabarty, P., Davis, M.P., Sparks, J.S., Schubert, M., 2012. The first record of a trans-oceanic sister-group relationship between obligate vertebrate Troglobites. PLoS ONE 7 (8), e44083.

Cui, R., Tyers, A.M., Malubhoy, Z.J., Wisotsky, S., Valdesalici, S., Henriette, E., Kosakovsky Pond, S.L., Valenzano, D.R., 2021. Ancestral transoceanic colonisation and recent population reduction in a nonannual killifish from the Seychelles archipelago. Mol. Ecol. 30, 3610–3623. https://doi.org/10.1111/mec.15982.

de Bruyn, M., Stelbrink, B., Page, T., Phillips, M., Lohman, D., Albrecht, C., Hall, R., von Rintelen, K., Ng, P.K.L., Shih, H.-T., Carvalho, G., von Rintelen, T., 2013. Time and space in biogeography: response to Parenti and Ebach (2013). J. Biogeogr. 40, 2204–2206. https://doi.org/10.1111/jbi.12166.

Douglas, J., Jiménez-Silva, C.L., Bouckaert, R., 2022. StarBeast3: Adaptive parallelized Bayesian inference under the multispecies coalescent. Syst. Biol. 71, 901–916. https://doi.org/10.1093/sysbio/syac010.

Feng, Y.J., Blackburn, D.C., Liang, D., Hillis, D.M., Wake, D.B., Cannatella, D.C., Zhang, P., 2017. Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous-Paleogene boundary. PNAS USA 114, E5864–E5870. https://doi.org/10.1073/pnas.1704632114.

Fraser-Brunner, A., 1955. A synopsis of the centropomid fishes of the subfamily Chandinae, with descriptions of a new genus and two new species. Raffles Bull. Zool. 25, 185–213.

Fricke, R., Eschmeyer, W.N., van der Laan, R. (Eds), 2023. Eschmeyer's Catalog of Fishes: Genera, Species, References. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp. Electronic version accessed 1 March 2023.

Friedman, M., Keck, B.P., Dornburg, A., Eytan, R.I., Martin, C.H., Darrin Hulsey, C., Wainwright, P.C., Near, T.J., 2013. Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. Proc. Royal Soc. B. 280, 2131733. https://doi.org/10.1098/rspb.2013.1733.

Froese, R., Pauly, D. (Eds), 2023. FishBase. World Wide Web electronic publication. http://www.fishbase.org, version (02/2023). Electronic version accessed February 2023.

Ghazali, S.Z., Lavoué, S., Khaironizam, M.Z., Jamaluddin, J.A.F., Ahmad, A., Ratmuangkhwang, S., Mohd Nor, S.A., 2022. Low morphological and genetic variation within the glass-perchlet *Parambassis siamensis* (Teleostei: Ambassidae) in Peninsular Malaysia. Zootaxa. 5128, 486–502. https://doi.org/10.11646/zootaxa.5128.4.2.

Ghezelayagh, A., Harrington, R.C., Burress, E.D., Campbell, M.A., Buckner, J.C., Chakrabarty, P., Glass, J.R., McCraney, W.T., Unmack, P.J., Thacker, C.E., Alfaro, M.E., Friedman, S.T., Ludt, W.B., Cowman, P.F., Friedman, M., Price, S.A., Dornburg, A., Faircloth, B.C., Wainwright, P.C., Near, T.J., 2022. Prolonged morphological expansion of spiny-rayed fishes following the end-Cretaceous. Nat. Ecol. Evol. 6, 1211–1220. https://doi.org/10.1038/s41559-022-01801-3.

Grewe, P.M., Krueger, C.C., Aquadro, C.F., Bermingham, E., Kincaid, H.L., May, B., 1993. Mitochondrial DNA variation among lake trout (*Salvelinus namaycush*) strains stocked into Lake Ontario. Can. J. Fish. Aquat. 50, 2397–2403. https://doi.org/10.1139/f93-264.

Hay, W.W., DeConto, R.M., Wold, C.N., Wilson, K.M., Voigt, S., Schulz, M., Rossby Wold, A., Dullo, W.-C., Ronov, A.B., Balukhovsky, A.N., Söding, E., 1999. Alternative global Cretaceous paleogeography. In: Barrera, E., Johnson, C.C. (Eds.), Evolution of the Cretaceous ocean–climate system. Geological Society of America, Boulder, pp. 1–48.

Hughes, L.C., Ortí, G., Huang, Y.u., Sun, Y., Baldwin, C.C., Thompson, A.W., Arcila, D., Betancur-R., R., Li, C., Becker, L., Bellora, N., Zhao, X., Li, X., Wang, M., Fang, C., Xie, B., Zhou, Z., Huang, H., Chen, S., Venkatesh, B., Shi, Q., 2018. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. PNAS USA 115 (24), 6249–6254.

Jurado-Rivera, J.A., Pons, J., Alvarez, F., Botello, A., Humphreys, W.F., Page, T.J., Iliffe, T.M., Willassen, E., Meland, K., Juan, C., Jaume, D., 2017. Phylogenetic evidence that both ancient vicariance and dispersal have contributed to the biogeographic patterns of anchialine cave shrimps. Sci. Rep. 7, 2852. https://doi.org/10.1038/s41598-017-03107-y.

Kapur, V.V., Khosla, A., 2019. Faunal elements from the Deccan volcano-sedimentary sequences of India: A reappraisal of biostratigraphic, palaeoecologic, and palaeobiogeographic aspects. Geol. J. 54, 2797–2828. https://doi.org/10.1002/gj.3379.

Karanth, K.P., 2021. Dispersal vs. vicariance: the origin of India's extant tetrapod fauna. Front. Biogeogr. 13, 1–13. https://doi.org/10.21425/F5FBG48678.

Katoh, K., Rozewicki, J., Yamada, K.D., 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. 20 (4), 1160–1166. https://doi.org/10.1093/bib/bbx108.

Kirchhoff, K.N., Hauffe, T., Stelbrink, B., Albrecht, C., Wilke, T., 2017. Evolutionary bottlenecks in brackish water habitats drive the colonisation of fresh water by stingrays. J. Evol. Biol. 30, 1576–1591. https://doi.org/10.1111/jeb.13128.

Klaus, S., Morley, R., Plath, M., Zhang, Y.-P., Li, J.-T., 2016. Biotic interchange between the Indian subcontinent and mainland Asia through time. Nat. Commun. 7, 12132. https://doi.org/10.1038/ncomms12132.

Landis, M.J., Matzke, N.J., Moore, B.R., Huelsenbeck, J.P., 2013. Bayesian analysis of biogeography when the number of areas is large. Syst. Biol. 62, 789–804. https://doi.org/10.1093/sysbio/syt040.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2017. Partitionfinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773. https://doi.org/10.1093/molbev/msw260.

Lavoué, S., 2015. Testing a time hypothesis in the biogeography of the arowana genus *Scleropages* (Osteoglossidae). J. Biogeogr. 42, 2427–2439. https://doi.org/10.1111/jbi.12585.

Lavoué, S., Miya, M., Arnegard, M.E., Sullivan, J.P., Hopkins, C.D., Nishida, M., Murphy, W.J., 2012. Comparable ages for the independent origins of electrogenesis in African and South American weakly electric fishes. PLoS ONE. 7 (5), e36287.

Li, C., Betancur-R, R., Smith, W.L., Ortí, G., 2011. Monophyly and interrelationships of Snook and Barramundi (Centropomidae sensu Greenwood) and five new markers for fish phylogenetics. Mol. Phylogenet. Evol. 60, 463–471. https://doi.org/10.1016/j.ympev.2011.05.004.

Maddison, W.P., Maddison, D.R., 2019. Documentation for Mesquite: a modular system for evolutionary analysis (Version 3.61). http://www.mesquiteproject.org/.

Mahon, A.R., 2007. Molecular phylogenetics of perciform fishes using the nuclear recombination activating gene 1. Unpublished D. Phil. Thesis, Old Dominion University.

Martin, T.J., 1988. Interaction of salinity and temperature as a mechanism for spatial separation of three co-existing species of Ambassidae (Cuvier)(Teleostei) in estuaries on the south-east coast of Africa. J. Fish Biol. 33 (sa), 9–15.

Mat Jaafar, T.N.A., 2014. DNA Barcoding and Population Genetic Structure of Malaysian Marine Fishes. Unpublished D. Phil. Thesis, Bangor University.

Matschiner, M., Böhne, A., Ronco, F., Salzburger, W., 2020. The genomic timeline of cichlid fish diversification across continents. Nat. Commun. 11 (1), 5895. https://doi.org/10.1038/s41467-020-17827-9.

Matzke, N.J., 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Front. Biogeogr. 5, 242–248. https://doi.org/10.21425/f5fbg19694.

Matzke, N.J., 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol. 63, 951–970. https://doi.org/10.1093/sysbio/syu056.

Matzke, N.J., 2022. Statistical comparison of DEC and DEC+J is identical to comparison of two ClaSSE submodels, and is therefore valid. J. Biogeogr. 49, 1805–1824. https://doi.org/10.1111/jbi.14346.

Matzke, N.J.. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis with R scripts (Version 1.1.1). https://zenodo.org/record/1478250.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. 2010 Gateway computing environments workshop, GCE 2010.

Nolf, D., Rana, R.S., Prasad, G.V.R., 2008. Late Cretaceous (Maastrichtian) fish otoliths from the Deccan Intertrappean Beds, India: A revision. Bull. Inst. r. Sci. Nat. Belg Sci. de la Terre. 78, 239–259.

Page, T.J., Hughes, J.M., 2010. Comparing the performance of multiple mitochondrial genes in the analysis of Australian freshwater fishes. J. Fish Biol. 77, 2093–2122.

Pagel, M., Meade, A., 2006. Bayesian analysis of correlated evolution of discrete characters by reversible–jump Markov chain Monte Carlo. Am. Nat. 167 (6), 808–825.

Palumbi, S., Martin, A., Romano, S., McMillan, W., Stice, L., Grabowski, G., 1991. The simple fool's guide to PCR. Dept. of Zoology and Kewalo Marine Laboratory University of Hawaii, Honolulu.

Pellegrin, J., 1932. Poissons nouveaux de Madagascar recueillis par M. Catala. Bull. Soc. Zool. Fr. 50, 425–426.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904. https://doi.org/10.1093/sysbio/syy032.

Rambaut, A., 2018. FigTree v1.4.4. https://github.com/rambaut/figtree/releases.

Ree, R.H., Sanmartín, I., 2018. Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. J. Biogeogr. 45, 741–749. https://doi.org/10.1111/jbi.13173.

Ree, R.H., Smith, S.A., Baker, A., 2008. Maximum Likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57 (1), 4–14.

Roberts, T.R., 1995. Systematic revision of tropical Asian freshwater glassperch (Ambassidae), with descriptions of three new species. Nat. Hist. Bull. Siam Soc. 42, 263–290.

Roberts, T.R., Jumnongthai, J., 1999. Miocene fishes from Lake Phetchabun in north-central Thailand, with descriptions of new taxa of Cyprinidae, Pangasiidae, and Chandidae. Nat. Hist. Bull. Siam Soc. 47, 153–189.

Ronquist, F., Sanmartín, I., 2011. Phylogenetic methods in biogeography. Annu. Rev. Ecol. Evol. Syst. 42 (1), 441–464.

Salles, T., Husson, L., Rey, P., Mallard, C., Zahirovic, S., Boggiani, B.H., Coltice, N., Arnould, M., 2023. Hundred million years of landscape dynamics from catchment to global scale. Science 379, 918–923. https://doi.org/10.1126/science.add2541.

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

Steinke, D., Connell, A.D., Hebert, P.D.N., Valdez-Moreno, M., 2016. Linking adults and immatures of South African marine fishes. Genome. 59 (11), 959–967.

Thieme, P., Schnell, N.K., Parkinson, K., Moritz, T., 2022. Morphological characters in light of new molecular phylogenies: the caudal-fin skeleton of Ovalentaria. R. Soc. Open Sci. 9, 211605 https://doi.org/10.1098/rsos.211605.

Tims, A.R., Unmack, P.J., Ho, S.Y.W., 2021. A fossil-calibrated time-tree of all Australian freshwater fishes. Mol. Phylogenet. Evol. 161, 107180 https://doi.org/10.1016/j.ympev.2021.107180.

Vega, G.C., Wiens, J.J., 2012. Why are there so few fish in the sea? Proc. Royal Soc. B. 279, 2323–2329. https://doi.org/10.1098/rspb.2012.0075.

Verma, C.R., Kumkar, P., Raghavan, R., Katwate, U., Paingankar, M.S., Dahanukar, N., 2019. Glass in the water: Molecular phylogenetics and evolution of Indian glassy perchlets (Teleostei: Ambassidae). J. Zoolog. Syst. Evol. 57, 623–631. https://doi.org/10.1111/jzs.12273.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D.N., 2005. DNA barcoding Australia's fish species. J. Zoolog. Syst. Evol. 360, 1847–1857. https://doi.org/10.1098/rstb.2005.1716.