Molecular Phylogenetics and Evolution 182 (2023) 107754

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Exceptional levels of species discovery ameliorate inferences of the biogeography and diversification of an Afrotropical catfish family

Julia J. Day [a,*], Elizabeth M. Steell [a,1], Thomas R. Vigliotta [b], Lewis A. Withey [a], Roger Bills [c], John P. Friel [d], Martin J. Genner [e], Melanie L.J. Stiassny [b]

[a] Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, UK
[b] Department of Ichthyology, American Museum of Natural History, New York, NY, USA
[c] South African Institute for Aquatic Biodiversity, Private Bag, 1015, 6140 Grahamstown, South Africa
[d] Alabama Museum of Natural History, The University of Alabama, Box 870340, 35487-0340 Tuscaloosa, AL, USA
[e] School of Biological Sciences, University of Bristol, Life Sciences Building, 24, Tyndall Avenue, Bristol BS8 1TQ, UK

ARTICLE INFO

Keywords:
Biogeography
Diversification
Chiloglanis
Mochokidae
Species delimitation
Synodontis

ABSTRACT

Endeavours in species discovery, particularly the characterisation of cryptic species, have been greatly aided by the application of DNA molecular sequence data to phylogenetic reconstruction and inference of evolutionary and biogeographic processes. However, the extent of cryptic and undescribed diversity remains unclear in tropical freshwaters, where biodiversity is declining at alarming rates. To investigate how data on previously undiscovered biodiversity impacts inferences of biogeography and diversification dynamics, we generated a densely sampled species-level family tree of Afrotropical Mochokidae catfishes (220 valid species) that was *ca.* 70 % complete. This was achieved through extensive continental sampling specifically targeting the genus *Chiloglanis* a specialist of the relatively unexplored fast-flowing lotic habitat. Applying multiple species-delimitation methods, we report exceptional levels of species discovery for a vertebrate genus, conservatively delimiting a staggering *ca.* 50 putative new *Chiloglanis* species, resulting in a near 80 % increase in species richness for the genus. Biogeographic reconstructions of the family identified the Congo Basin as a critical region in the generation of mochokid diversity, and further revealed complex scenarios for the build-up of continental assemblages of the two most species rich mochokid genera, *Synodontis* and *Chiloglanis*. While *Syndontis* showed most divergence events within freshwater ecoregions consistent with largely in situ diversification, *Chiloglanis* showed much less aggregation of freshwater ecoregions, suggesting dispersal as a key diversification process in this older group. Despite the significant increase in mochokid diversity identified here, diversification rates were best supported by a constant rate model consistent with patterns in many other tropical continental radiations. While our findings highlight fast-flowing lotic freshwaters as potential hotspots for undescribed and cryptic species diversity, a third of all freshwater fishes are currently threatened with extinction, signifying an urgent need to increase exploration of tropical freshwaters to better characterise and conserve its biodiversity.

1. Introduction

Determining the drivers of the exceptional levels of species richness in tropical ecosystems is a key goal in evolutionary biology (Brown, 2013), in which large densely sampled phylogenies have been increasing used to infer diversification dynamics and/or biogeographic patterns in these systems (e.g., Day et al., 2013; De-Silva et al., 2016; Derryberry et al., 2011; Hackel et al., 2022; Liedtke et al., 2016). However, the importance of characterising biodiversity by species richness has implications for inferring evolutionary and ecological processes, and is also critical for conservation planning and management (Hortal et al., 2015), since missing species may have a dramatic effect on inferring such patterns leading to erroneous interpretation on the processes shaping biodiversity (Cusimano et al., 2012; Pybus and Harvey, 2000). This is particularly important for groups with hidden diversity where underestimation of species richness can greatly hinder understanding.

Species discovery has been transformed through the application of DNA molecular sequence data, which has greatly facilitated the

* Corresponding author.
E-mail address: j.day@ucl.ac.uk (J.J. Day).
[1] Currently: Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK.

https://doi.org/10.1016/j.ympev.2023.107754
Received 7 November 2022; Received in revised form 24 February 2023; Accepted 5 March 2023
Available online 9 March 2023

identification of cryptic species (Bickford et al., 2007), and shown many taxonomic groups to possess a proliferation of cryptic species (e.g., Feulner et al., 2006; Funk et al., 2021; Hebert et al., 2004; Martins et al., 2020). Tropical freshwaters present potential hotspots for cryptic species discovery, since freshwaters harbour high levels of endemism due to the insular nature of catchments (Strayer and Dudgeon, 2010), with over half of all freshwater fish species confined to single ecoregions (Abell et al., 2008). This is particularly pertinent for fast flowing lotic environments (rapids, cascades, and headwaters), which remain relatively underexplored, despite high endemism and genetic structuring identified in various fish clades inhabiting these high-energy systems (Alter et al., 2017; Hrbek et al., 2018; Kurata et al., 2022; Morris et al., 2016). This is likely a consequence of fast-flowing water creating potential ecological 'islands' surrounded by slower-moving waters (Hrbek et al., 2018). Furthermore, environmental constraints acting on phenotypes have led to convergent evolution in rheophilic fishes regarding body and fin shape, mouth form and orientation, among other features (Lujan and Conway, 2015). High levels of phenetic similarity is frequently observed among fish (Alter et al., 2017; Hrbek et al., 2018; Schmidt et al., 2016) and invertebrate clades (Zheng et al., 2020) likely the result of stabilising selection in these systems.

Although the Afrotropics is one of the most diverse regions on Earth (Collen et al., 2014) it suffers from a lack of comprehensive sampling, and its freshwater biodiversity remains relatively poorly known (Darwall et al., 2011; Decru et al., 2015; Skelton and Swartz, 2011), particularly within riverine habitats. Regional biodiversity studies are beginning to address this deficit (e.g., Arroyave et al., 2019; Bragança et al., 2021; Chakona et al., 2018; Decru et al., 2015; Feulner et al., 2006; Mbimbi Mayi Munene et al., 2021; Schmidt et al., 2016; Swartz et al., 2009; van der Merwe et al., 2021; Van Ginneken et al., 2017), but the extent of undiscovered diversity is unclear without more comprehensive sampling across the continent, although broader scale studies (e.g., Arroyave et al., 2020; Bragança and Costa, 2019; Daniels et al., 2015; Day et al., 2017; Day et al., 2013; Ford et al., 2019; Goodier et al., 2011; Liyandja et al., 2022; Mahulu et al., 2021) are beginning to shed light on continental-scale events that have shaped extant diversity of Afrotropical freshwater animal clades.

To investigate how data on previously undiscovered biodiversity impacts inferences of biogeography and diversification dynamics, we generated a densely sampled dated species-level family tree of the Afrotropical Mochokidae catfishes that was *ca.* 70 % complete. This was achieved through generation of novel, and harnessing published, mitochondrial and nuclear DNA sequence data. We tested if diversification rates were best supported by the null model of a constant rate, as opposed to density dependence, and by combining the timetree with geographical data, it specifically allowed us to investigate the geographic centre of origin of key clades, and if these clades possessed a common biogeographic signal. Furthermore, to explore the assumption that species richness has been underestimated in tropical lotic freshwaters considerable efforts were dedicated to extensive sampling across the African continent of the mochokid genus *Chiloglanis,* which remains largely unknown at the continental scale. To evaluate to what extent species diversity in this genus was underestimated, we applied species delimitation methods to this group prior to construction of the mochokid tree. Our biodiversity surveys included underexplored regions in East Africa, and the Congo Basin, which contains the highest diversity of freshwater fishes in the Afrotropics (Decru et al., 2015; Snoeks et al., 2011).

2. Materials and methods

2.1. Study system

Comprising *ca.* 220 valid species, the endemic African family Mochokidae is by far the most species rich of the eleven Siluriformes families present on the continent (Fricke et al., 2022). Relationships within the family have recently been evaluated based on all protein-coding mt-DNA genes (Schedel et al., 2022), with results supporting the two sister subfamilies (Chiloglanidinae and Mochokinae), with the Chiloglanidinae containing *Atopodontus*, *Aptopochilus*, *Chiloglanis*, *Euchilichthys*, and the Mochokinae containing *Acanthocleithron*, *Microsynodontis*, *Mochokiella*, *Mochokus,* and *Synodontis*. Species richness is highly unbalanced across the family and the basis of this disparity is unknown. By far the greatest diversity is within the larger sized and often distinctively marked *Synodontis* (squeaker catfish) *ca.* 130 valid species (Fricke et al., 2022) that principally occur in large rivers, but are also found in rift lakes, having radiated in Lake Tanganyika (Day and Wilkinson, 2006). In contrast, *Chiloglanis* (African suckermouth catfish) contains approximately half this diversity, with 62 valid species (Fricke et al., 2022). *Chiloglanis* are rheophilic specialists, occurring in fast-flowing lotic waters (Friel and Vigliotta, 2011). They are mostly found in upper catchment tributary streams, but do occur in large lower catchment river systems where there are suitable habitats, with a single taxon known from the rocky surge zone of Lake Malawi (Seegers, 2008). Both *Synodontis* and *Chiloglanis* have fully continental distributions, while the remaining genera are relatively depauperate with far more restricted ranges (Froese and Pauly, 2022).

Despite resolution of mochokid generic-level interrelationships (Schedel et al., 2022) a large-scale continent-wide phylogenetic analysis has not yet been conducted. To date several studies have focused on *Synodontis* (Day et al., 2013; Pinton et al., 2013) across the generic range, while only regional studies have been conducted for *Chiloglanis* (Chakona et al., 2018; Morris et al., 2016; Schmidt et al., 2016; Schmidt et al., 2014). Thus, there remains a substantial knowledge gap for this genus that hinders a family-wide interrogation of diversification dynamics and biogeography. Moreover, extensive hidden diversity within *Chiloglanis* is highly likely as previous studies have demonstrated high levels of (cryptic) species diversity at regional scales (Chakona et al., 2018; Morris et al., 2016; Schmidt et al., 2016; Schmidt et al., 2014). The genus provides an attractive study system to examine cryptic diversity and endemism in the tropical riverine environment as species are: 1) camouflaged in colouration and generally phenotypically similar, and all are 2) small in size (<100 mm SL, (Friel and Vigliotta, 2011)). Despite *Chiloglanis* species exhibiting highly similar overall appearance, a number of morphological characters have been used to distinguish valid species, including the number and distribution of maxillary and mandibular teeth, barbel lengths, oral disc shapes and sizes, body proportions, fin spines, and male caudal shapes (e.g., Friel and Vigliotta, 2011; Schmidt et al., 2017).

2.2. Specimen sampling

Approximately 350 mochokid samples were included in this study, including 227 *Chiloglanis* samples (representing 38 of 62 valid species, Fricke et al. (2022)), as well as considerable potential undescribed diversity. Multiple individuals of *Chiloglanis* were included for most species and sampling locations to test species validity (Appendix A, Table S1). Samples of *Chiloglanis* were collected from across the generic range (Fig. 1), including from all ichthyofaunal provinces encompassed in their range. Valid *Chiloglanis* species not included here are not region specific, with most (>70 %) from regions that were otherwise well-sampled in this study, including the Congo Basin, Lower Guinea Forest, and East Africa. For the majority of mochokid species, with the exception of *Microsynodontis* and Atopochilini, we included a single exemplar taxon based on previous findings (Day et al., 2013). For a list of all included mochokid samples, outgroups, voucher numbers, locality information, and Genbank numbers, see Appendix A, Table S1.

2.3. Sequence data

We sequenced two mitochondrial (mt)DNA genes: CO1 (676 bp) and Cyt *b* (1138 bp) and the nuclear (nu)DNA loci: recombination activating

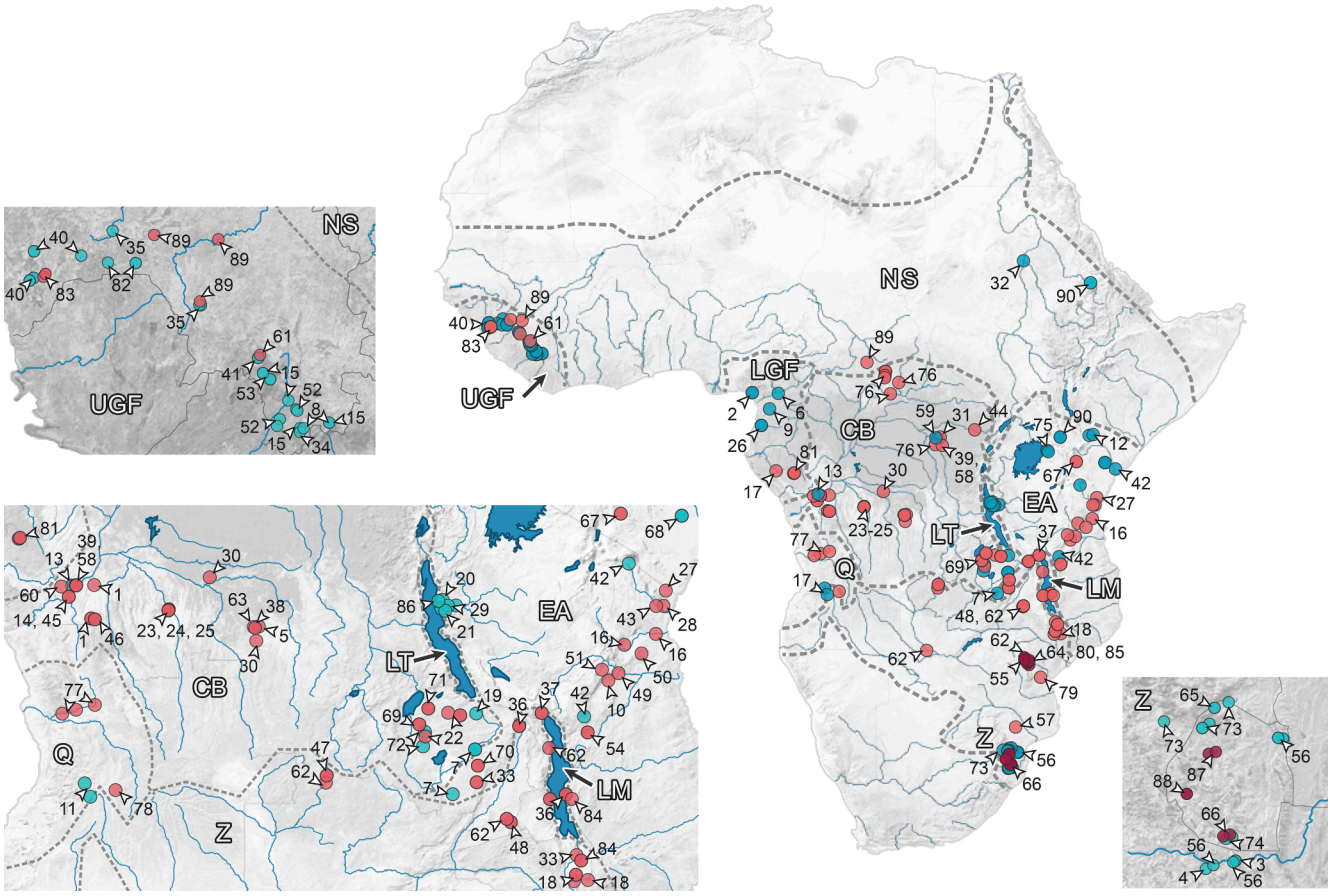

Fig. 1. Map of Africa displaying all *Chiloglanis* samples used in species delimitation analyses, along with insets of heavily sampled regions. Blue circles = described valid species; red circles = putative novel candidate species; dark red circles = previously identified potential candidate species. Numbers refer to species/candidate names (see Fig. 2a,b; Appendix A, Table S4, column H). Ichthyo-provinces are included, which are used for subsequent biogeographic analyses: Congo Basin (CB), East Africa (EA), Quanza (Q), Zambezi (Z), Nilo-Sudan (N-S), Upper Guinea Forest (UGF), Lower Guinea Forest (LGF), and Lake Tanganyika (LT) and Lake Malawi (LM)). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

gene-2 (Rag2) (921 bp) enabling these data to be combined with published data for other mochokids. The number of *Chiloglanis* individuals included for each locus was as follows: CO1 n = 188, Cyt b n = 126, Rag2 n = 108 (for other mochokids see Appendix A, Table S1). DNA was extracted from ethanol preserved fin clips or muscle tissue using QIAGEN DNeasy Blood and Tissue kit. We followed Day et al. (2013) for PCR conditions, and primers (see Appendix A, Table S2 for published and novel primers). Samples were cleaned and cycle sequenced using standard protocols and analysed using an ABI13730xl sequencer.

2.4. Phylogenetic inference

Sequences were cleaned, contiged, and aligned in Geneious v. 10.0.8 (Kearse et al., 2012) using ClustalW (Larkin et al., 2007). Alignments were further checked manually, and tRNAs were removed from Cyt *b* sequences. Analyses of sequence data included the following matrices: 1) concatenated mtDNA *Chiloglanis* matrix (1817 bp), 2) Rag 2 *Chiloglanis* matrix (921 bp), and 3) concatenated mt- and nuDNA mochokid matrix (2737 bp). All matrices were partitioned by gene and codon position to evaluate substitution models. Bayesian inference was applied to all matrices. As the concatenated mt- and nuDNA mochokid matrix (matrix 3) included some distant outgroups and not all outgroup species had full coverage of sequence data, we also ran a further analysis applying Maximum Likelihood (ML) to matrix 3 to check concordance between methods. For Bayesian analyses we applied PartitionFinder2 v2.1 (Lanfear et al., 2016) to identify the best substitution models (Appendix A, Table S3), implementing the Bayesian Information Criterion (BIC), unlinking branch lengths, and selected the greedy search algorithm. Matrices were subsequently analysed using MrBayes v3.2.6 (Ronquist et al., 2012), with two independent MrBayes analyses run for 10 million generations, sampled every 100 generations (chain temperature 0.2, 4 chains) with branch support evaluated using Bayesian posterior probabilities (BPPs). Run convergence was assessed in Tracer v1.7.1 (Rambaut et al., 2018) and to ensure effective sample size (ESS) values were > 200, with burn-in set to 10 %. Bayesian analyses were performed using CIPRES (Miller et al., 2010). For the ML analysis IQ-TREE 2.2.0 (Nguyen et al., 2015) was implemented, applying ModelFinder (Kalyaanamoorthy et al., 2017) to identify optimal partition models, with branch support evaluated using an ultra fast bootstrap (bs) approximation (Minh et al., 2013) (1000 replicates). All trees were visualised in FigTree v1.4.3 (Rambaut, 2016).

2.5. Time calibrated trees

Several time calibrated trees were generated using BEAST 2 v2.6.6 (Bouckaert et al., 2014). These included 1) concatenated mtDNA tree (matrix 1) containing all *Chiloglanis* individuals (intra-species sampling, Appendix A, Table S1), required for general mixed Yule coalescent (GMYC) species delimitation models (*Chiloglanis* n = 227, mochokid outgroups n = 12); and 2) concatenated mt- and nuDNA mochokid species tree (matrix 3), including one exemplar individual per species (Mochokidae n = 195, outgroups "Big Africa"(*sensu* Sullivan et al. (2006)) n = 16, non-"Big-Africa" n = 1) (Appendix A, Table S1, see Table S4 for *Chiloglanis* individuals included) for downstream

biogeographic and diversification analyses. The BEAST 2 package bModelTest v1.2.1 (Bouckaert and Drummond, 2017) was used to estimate substitution models, and tree and birth rate priors were set to Yule and uniform respectively for both analyses. For matrix 1, a strict clock was applied to the mtDNA dataset based on low coefficient of variation (CV) scores, and substitution rates were estimated. These data were partitioned only by loci regarding the site models (to reduce the parameters), in which TN93 + G + I and GTR + G + I were best fit to the COI and Cytb partitions respectively. For matrix 3 (the concatenated mt- and nuDNA data set), loci were partitioned by genome, and clock models were unlinked (independent clock rates automatically mean substitution rates are not estimated). For this matrix, a strict clock was selected for the mtDNA data (based on an low CV 0.274 and overall convergence from an initial run) and a relaxed log-normal clock (Drummond et al., 2006) applied to the nuDNA data (CV 0.945). To reduce parameterisation of site models, by genome partitions were also implemented, with the models TIM + G + I and TN93 + G + I selected as best fit to the mtDNA, and nuDNA partitions respectively. The oldest known fossil of the Mochokidae, a *Synodontis* dating to 34 Ma, was selected to calibrate this clade (following Day et al., 2013) using a log-normal prior (34 Ma with a 0.8 % tail probability) with 'user originate' option selected which is recommended for fossil calibrations. All other priors were default. For the mochokid dataset, a further calibration was applied to the node including all "Big-Africa" taxa using a normal prior (71.11 Ma, M = 0.01, S = 0.5) based on a secondary calibration from a mitogenomic siluriform phylogeny (Kappas et al., 2016). Analyses were run three times for 50 million generations for matrix 1 (concatenated mtDNA data), and four times for 100 million generations for matrix 3 (concatenated mt- and nuDNA data), sampling every 5000 generations, with run convergence assessed using Tracer v1.7.1 (Rambaut et al., 2018). All ESS values were > 200, except for the concatenated tree where the Yule Model prior was 177. Tree files were subsequently combined in LogCombiner v2.5.1 and trees summarised in TreeAnnotator v2.5.1 (Bouckaert et al., 2014) with burn-in was set to 10 %. Trees were again visualised in FigTree v1.4.3 (Rambaut, 2016).

2.6. Quantitative species delimitation

Species delimitation methods, particularly harnessing single-locus sequencing, provide a useful approach for determining potential species limits in clades containing large numbers of undescribed species (Talavera et al., 2013). Given this, and since broad agreement in delimitation across methods provides reasonable confidence in correct assignment of candidate species (Dellicour and Flot, 2018), we apply multiple approaches to our single-locus (mtDNA) dataset (matrix 1). Although we were unable to sequence as many samples at the Rag 2 gene, phylogenetic trees generated from these data (matrix 2) were largely congruent with those generated from mtDNA data (Appendix A, Fig. S1).

Tree-based methods were applied to the concatenated mtDNA dataset (matrix 1) since these methods have been shown to perform well over a wide range of assumptions (Fujisawa and Barraclough, 2013; Reid and Carstens, 2012; Talavera et al., 2013). These included the following: 1) ML general mixed Yule coalescent (GMYC) (Fujisawa and Barraclough, 2013) single-model using 'splits' v1.0 (Ezard et al., 2017); 2) Bayesian 'bGMYC' v1.0 (Reid and Carstens, 2012) in R v3.5.2 (R Development Core Team, 2018); 3) ML Poisson tree process (PTP) v1.0 and 4) Bayesian PTP v1.0 (Zhang et al., 2013) analyses using the Exelixis Server (Zhang et al., 2013).

The bGMYC and GMYC analysis used the BEAST2 mtDNA consensus tree (from 10,000 trees), while the ML and Bayesian PTP analyses used the MrBayes mtDNA consensus tree. For all analyses, outgroups were removed as they have been shown to skew threshold values (Reid and Carstens, 2012) leaving 227 *Chiloglanis* individuals. For bGMYC the log (coalescence rate/Yule rate) ratio was checked to ensure no sub-zero values. Values above zero show model appropriateness as coalescence rate is higher than speciation rate (Reid et al., 2019). User-defined thresholds in bGMYC included 0.1 (highly conservative/lumping) – 0.9 (high splitting). Bayesian PTP analysis were run for 100 thousand generations, sampling every 100 trees with burn-in at 10 %, with run convergence assessed from a resulting trace file.

Based on the results from the four species delimitation analyses we produced a table of candidate species (Appendix A, Table S4). Valid species were used as a baseline for the correct splitting of OTUs by bGMYC thresholds. As our analyses included *ca.* 60 % of valid species diversity and many unidentified taxa, we further evaluated undescribed diversity based on the type locality and range data from Fricke et al., (2022), for species not included in our analyses. This allowed us to assess whether any potential candidate species were in fact collected from within the ranges of described species, and therefore should best be discounted in our evaluation of undescribed species diversity.

2.7. Ancestral range estimation

Geographic range evolution was estimated using the dated mochokid tree (generated from matrix 3), comparing the models: Dispersal–Extinction–Cladogenesis (DEC), DIVALIKE (Dispersal-Vicariance Analysis) and BAYAREALIKE (Bayesian Analysis of Biogeography), with the latter two models likelihood implementations of the original DIVA and BayArea models. In modelling a continental system we did not implement modifications of these models using the + J parameter, which adds the process of jump dispersal at speciation, and is more applicable to island systems (Matzke, 2014). The three models were unconstrained (allowing dispersal to and from each region), selecting max areas = 4, and were run using the R package BioGeoBEARS (Matzke, 2013). The delta Akaike Information Criterion (ΔAIC) was used to select the best model from the resulting model set.

Mochokids occur in seven of the nine ichthyo-provinces that have been identified for continental Africa (Roberts, 1975) covering all but the Maghreb (Northwest Africa) and Southern African, and include: Congo Basin (CB), East Africa (EA), Quanza (Q), Zambezi (Z), Nilo-Sudan (N-S), Upper Guinea Forest (UGF), Lower Guinea Forest (LGF). Lake Tanganyika and Lake Malawi were added as separate 'areas' due to high levels of lacustrine endemicity following Day et al. (2013) (Fig. 1). Only *Chiloglanis* and *Synodontis* are present in all seven areas, while the remaining genera have more restricted ranges (Fricke et al., 2022; Froese and Pauly, 2022). Distributional data was taken from FishBase (Froese and Pauly, 2022), except for the potential candidate species where distribution was based only on specimens included in this study due to their high levels of endemicity (Appendix A, Table S5).

2.8. Diversification rates

The dated mochokid tree (generated from matrix 3) was used for downstream diversification analyses, with all non-mochokid outgroups pruned from the tree using the R package ape 3.4 (Paradis et al., 2004) resulting in n = 195 taxa. BAMM 2.5 (Bayesian Analysis of Macroevolutionary Mixtures) (Rabosky et al., 2013) was used to investigate if the family diversified under a null hypothesis of constant rate, and if any rate shifts were identified across the tree. The following priors were generated (expected number of shifts = 1; λ Init.Prior = 2.828; λShiftPrior = 0.018; μ Init.Prior = 2.829; λIsTimeVariablePrior = 1) and a global sampling fraction of 0.68 applied. Three independent MCMC runs were each run for 10 million generations sampling every 1,000th. BAMMtools 2.0 (Rabosky et al., 2014) was used to test if MCMC simulations converged (effective sizes were > 200 (log-likelihoods > 3400; number of shift events present > 4900)) and to analyse and plot the BAMM output. The gamma (γ) statistic, and Monte Carlo constant rates (MCCR) test (Pybus & Harvey, 2000) were also computed to test if there were deviations from the pure birth model and the effect of missing species, using ape 3.4 (Paradis et al., 2004) and phytools 1.2 (Revell, 2012).

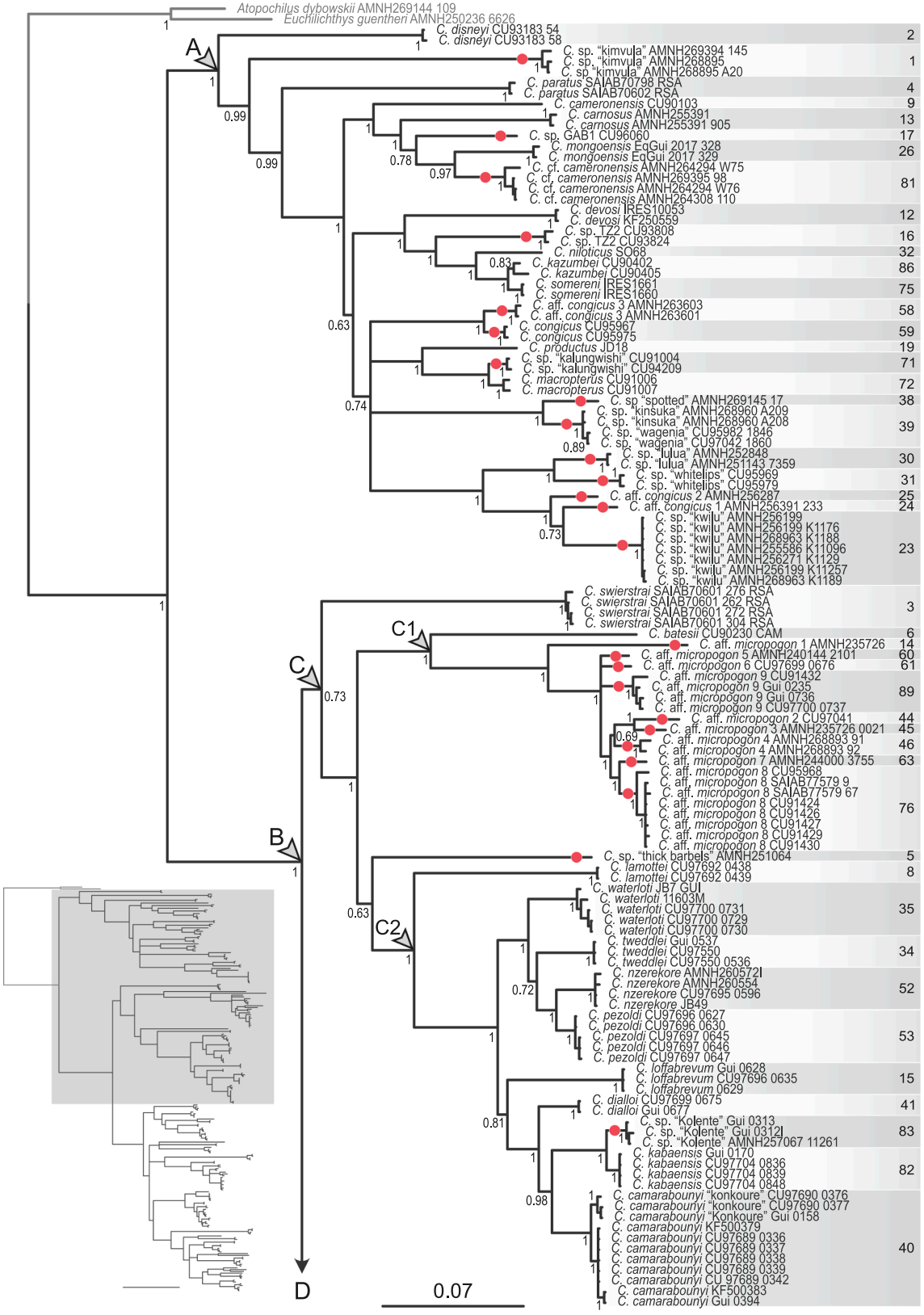

Fig. 2. Phylogeny of *Chiloglanis* based on mtDNA (Cyt b and CO1) data, highlighting putative novel species (red) and previously identified candidates (dark red), recovered across all four methods detailing: a) clades A and C; b) clade D. Insets for each figure shows the entire tree, with the pale grey box highlighting the clade of interest. Support values BPP are given below the branches, BS above the branches. Outgroups (grey branches), apart from the sister taxa to *Chiloglanis*, have been removed for clarity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

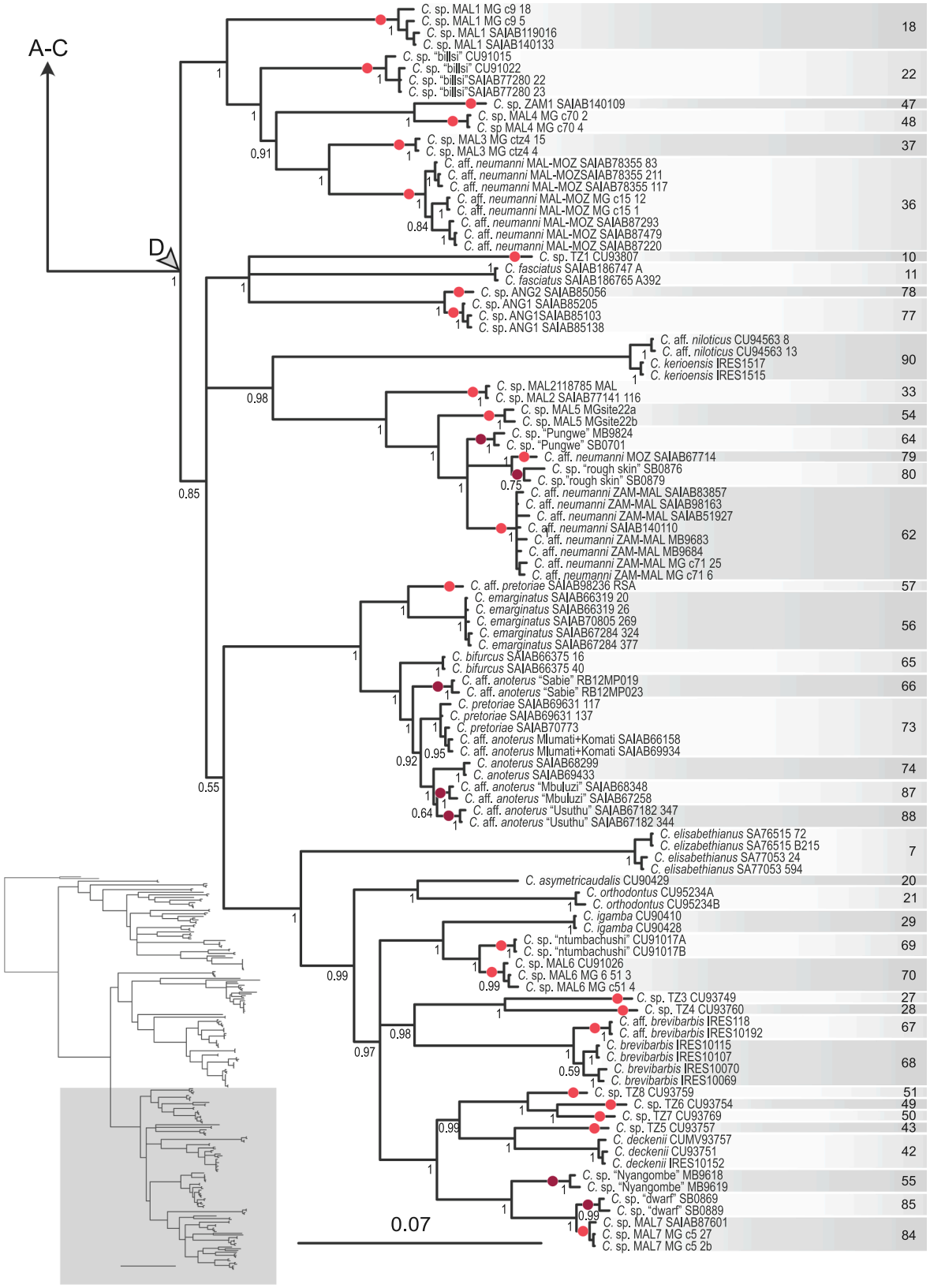

Fig. 2. (*continued*).

3. Results

3.1. Species delimitation

Overall, tree-based species delimitation methods gave similar results, identifying 92–99 operational taxonomic units (OTUs) (Appendix A, Table S4). The GMYC (single) model yielded 99 OTUs, the bGMYC model applying an intermediate 0.5 threshold (a default in many studies) estimated 92 OTUs (relative thresholds of 0.1–0.9 yielded 84–99 OTUs respectively, data not shown), and PTP methods identified 95 (ML) and 96 (Bayesian) OTUs.

Samples assigned to described species (38 of 63 species (Fricke et al., 2022)) were mostly delineated across all four tree-based methods, with no species being lumped, and only seven species identified as warranting splitting (Fig. 2a, b, Appendix A, Table S4). Two of these were previously investigated, *C. anoterus* (Morris et al., 2016) *and C. neumanni* (Chakona et al., 2018), with our results largely supporting these studies as suggestive of further cryptic species. Of the remaining taxa, *C. aff. micropogon* and *C. brevibarbis* were resolved as monophyletic, with the latter split into two OTUs, although it has not previously been considered polyspecific (Schmidt et al., 2014). Samples assigned to *Chiloglanis aff. micropogon* from the Congo Basin (no specimens from the type locality Nzokwe River, Democratic Republic of Congo were available) was delimited into nine OTUs across all methods (only five supported by bGYMC 0.1–0.9 thresholds, data not shown), indicating potentially high levels of hidden diversity within that taxon. Samples assigned to *Chiloglanis congicus, C. niloticus, C. cameronensis,* and *C. pretoriae* were resolved as non-monophyletic (Fig. 2a, b), a finding supported by the nuDNA tree (for the former two taxa common to both trees, Appendix A, Fig. S1). Further assessment is needed of all four taxa, but particularly *C. niloticus* and *C. pretoriae* since one of the OTUs of both taxa cluster with a described and candidate species respectively. The Sudanian vs. Ethiopian *C. niloticus* samples nested within clades A and D respectively (Fig. 2a, b). However, the latter samples were only supported as a distinct OTUs in 2/4 methods, and otherwise clustered with *C. kerioensis*, despite the geographic distance separating them (see Fig. 1). We also found samples of *C. pretoriae* clustered as a distinct OTU with *C. aff. anoterus* from the Mlumati + Komati Rivers, while a further *C. pretoriae* sample (Olifants River) is sister to *C. emarginatus*. (Fig. 2b).

Despite these few exceptions, the recovery of most described species as distinct OTUs (>80 %) provides a reasonably high level of confidence in the delimitation of unidentified/undescribed (i.e., those given informal names) taxa included in our dataset. Based on a conservative estimation across species-delimitation methods, where we delimit taxa based on congruence across all four implemented tree-based methods, we suggest 90 OTUs are represented in our phylogeny (Fig. 2a, b, Appendix A, Table S4). The localities of putative novel species recognized here were subsequently checked against type localities of all described species to ensure that they had not been misdiagnosed. Based on this information we did not find any match regarding the type localities of described species with unassigned taxa, and as *Chiloglanis* generally have highly restricted ranges (Morris et al., 2016; Schmidt et al., 2016; Schmidt et al., 2017) it seems unlikely that any of these novel candidates represent currently recognized species. If correct, we estimate a staggering *ca.* 50 putative novel species (56 putative species including previously identified candidates split using delimitation methods or morphometric data), representing an 80–90 % increase in species diversity for the genus. While our study includes several taxa with distinctive phenotypes as indicated by their informal names (e.g., “spotted”, “Kinsuka” blind, “thick-barbels” “white-lips”, Fig. 2a, b), the majority display limited phenotypic differences among them, highlighting that while morphologically conservative, the genus contains considerable hidden diversity.

3.2. Mochokid phylogenetic relationships and divergence estimates

Monophyly of the Mochokidae was well supported in both BI and ML analyses (0.99 BPP and 97 % BS, Appendix A, Fig. S2), with an estimated origin 64.8 Ma (95 % HPD: 60.69–68.09 Ma) based on the BEAST tree (Appendix A, Fig. S3). Two main clades are resolved, strongly supporting the subfamilies, Chiloglanidinae and Mochokinae (Appendix A, Fig. S2), with the former clade estimated to be older 52.50 Ma (HPD: 48.91–55.90 Ma) than the latter 39.72 Ma (95 % HPD: 37.18–42.14 Ma) (Appendix A, Fig. S3). Disparity in divergence estimates for these subfamilies is mirrored in the ages of the most species rich genera, with the MRCA of *Synodontis* estimated at 25.63 Ma (95 % HPD: 23.87–27.50 Ma), making it a much younger clade than *Chiloglanis* estimated at 47.53 Ma (95 % HPD: 44.65–50.8 Ma), with other mochokid clades also much younger than *Chiloglanis* (Appendix A, Fig. S3). As single fossil calibrations on internal nodes may impact divergence estimates stemward and to node(s) of interest (Duchêne et al., 2014), we also investigated if the age disparity between these clades is an artifact of a single fossil calibration. An additional analysis applying only the secondary calibration to the root of the ‘Big Africa’ clade was performed and supported this result, with similar age estimates generated (data not shown).

Chiloglanidinae included a monophyletic Atopochilini (Vigliotta, 2008), composed of *Atopodontus*, *Aptopochilus* and *Euchilichthys*, which is the sister to *Chiloglanis* (Appendix A, Fig. S2). Within the tribe, *Atopodontus* is supported as sister to *Aptopochilus* and *Euchilichthys*, however the latter two genera are resolved as non-monophyletic (Appendix A, Fig. S2). Mochokinae is composed of *Microsynodontis*, *Mochokus*, *Mochokiella*, and *Synodontis*, with each genus the successive sister in the taxon order listed. Support for the major mochokid relationships is high, typically receiving maximum support (Appendix A, Fig. S2), and relationships broadly supporting those of the mitogenome tree of Schedel et al. (2022). A single conflict between the resulting trees is the position of *Acanthocleithron chapini*. This monotypic taxon is either resolved as sister to Mochokinae in the Bayesian (0.56 BPP) and BEAST trees (latter tree depicted in Appendix A, Fig. S3), supporting (Schedel et al., 2022), or as the sister to Chiloglanidinae in the ML tree (61 % bs, Appendix A, Fig. S2), however neither hypothesis is well supported, possibly a consequence of a lack of nuclear data and/or its long branch. Irrespective of its position, *Acanthocleithron* is considerably older than the two mochokid subfamilies, and in our study, it is estimated to have diverged from the Mochokinae at 59.10 Ma (95 % HPD: 53.58–62.60 Ma, Appendix A, Fig. S3).

The species rich genera, *Chiloglanis* and *Synodontis,* are resolved into distinct subclades, with relationships generally well supported (Appendix A, Fig. S2). For *Chiloglanis* two main clades were resolved (Fig. 1, Appendix A, Fig. S2 clades A and B), with clade B further divided into two subclades C and D, all with maximum support, except clade C (0.85 BPP and 96 % BS). Overall, only some internal relationships (e.g., several nodes along the backbone of clade D) are not well resolved. The relationships within *Synodontis* are generally congruent to those documented in Day et al. (2013).

3.3. Biogeography of the Mochokidae

Results from our ancestral range reconstruction analyses find the DEC model (Fig. 3a, b) a significantly better fit than the other models, with the delta AIC > 2 units lower than the DIVALIKE model (the second-best model), while the BAYESAREALIKE model did not perform well (Table 1). The DEC and DIVALIKE models are very similar regarding ancestral area reconstructions for internal nodes differing only by a few nodes for the major mochokid lineages (see Table 2). The DEC model reconstructed the origin of the Mochokidae combining the Congo Basin (CB) and the super-region of West Africa (Nilo-Sudan (N-S), Lower Guinea Forest (LGF), and Upper Guinea Forest (UGF)), whereas the DIVALIKE model supported only the Congo Basin at the ancestral

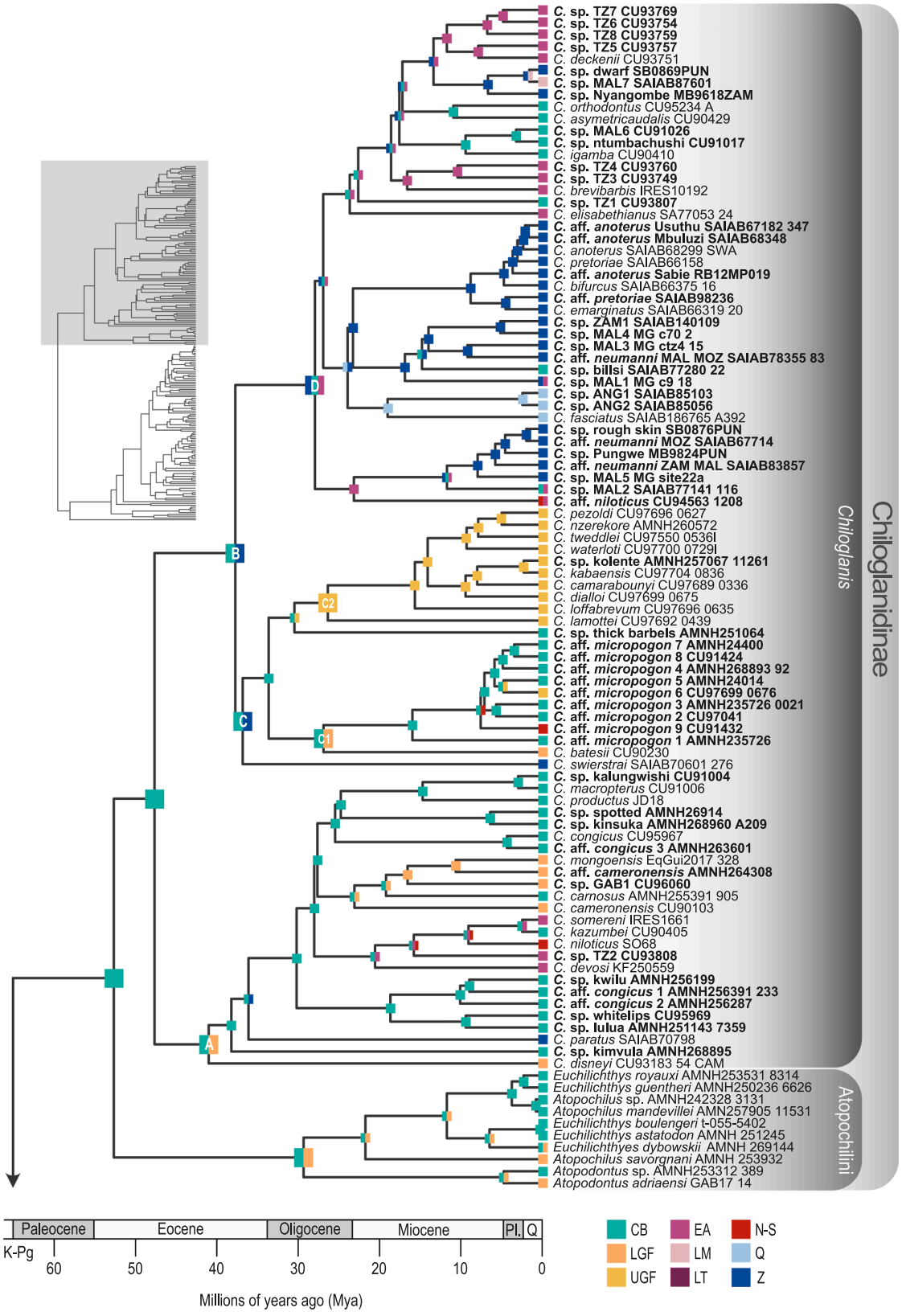

Fig. 3. Biogeographic range inheritance based on the DEC unconstrained model for the family Mochokidae using the BEAST dated concatenated mt- and nuDNA gene tree detailing: a) Chiloglanidinae, and b) Mochokinae. Insets for each figure shows the entire tree, with the pale grey box highlighting the clade of interest. Putative novel candidate species in bold. Ichthyo-provinces are as follows: Congo Basin (CB), East Africa (EA), Quanza (Q), Zambezi (Z), Nilo-Sudan (N-S), Upper Guinea Forest (UGF), Lower Guinea Forest (LGF), Lake Tanganyika (LT), Lake Malawi (LM). See Fig. S3 for 95 % confidence intervals (HPD).

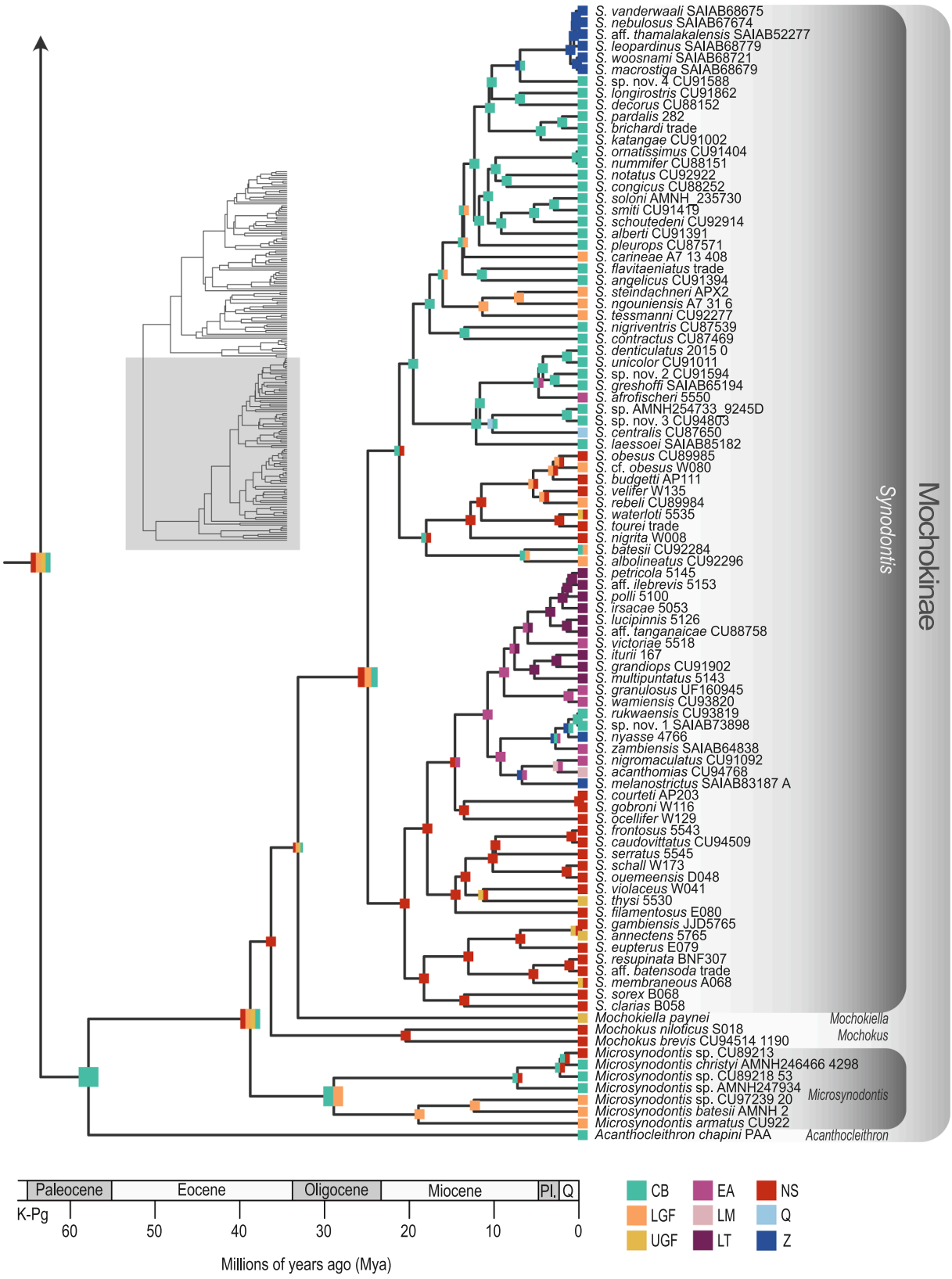

Fig. 3. (*continued*).

node (Table 2). The Congo Basin is reconstructed as ancestral for the subfamily Chiloglanidinae, and its constituent major clades, *Chiloglanis* and the Atopochilini and is also reconstructed as ancestral for the Mochokinae + *Acanthocleithron* indicating, if the DEC model is correct, a range contraction early in the history of the family. Both models reconstructed the origin of the subfamily Mochokinae as comprising several areas including the Congo Basin and a combination of other West African ichthyo-provinces (Table 2). Constituent Mochokinae genera

Table 1
Comparisons of unconstrained models (dispersal to and from an area) for the three models implemented in BioGeoBEARS: DEC, DIVALIKE, and BAYAREALIKE. Parameters are as follows: dispersal (d); extinction (e); number of parameters (k); Akaike Weights (AW). The optimal model is denoted in bold.

Model	Ln *L*	Parameters			AIC analysis		
		k	*d*	*e*	AIC	ΔAIC	AW
DEC	**−327.39**	**2**	**0.0020**	**0.0013**	**658.78**	**0.00**	**0.913**
DIVALIKE	−329.74	2	0.0027	0.0016	663.48	4.70	0.087
BAYAREALIKE	−399.15	2	0.0015	0.0365	802.30	143.52	<0.001

Table 2
Ancestral area reconstruction of the major Mochokidae lineages comparing the best fit model DEC vs. the second best DIVALIKE model.

Node		Model	
Family/Subfamily	Genus/Tribe	DEC	DIVALIKE
Mochokidae		NS UGF LGF CB	CB
	Acanthocleithron	CB	CB
Chiloglanidinae		CB	CB
	Chiloglanis	CB	CB
	Atopochilini	CB	CB
Mochokinae		NS UGF LGF CB	NS LGF CB
	Microsynodontis	CB	CB
	Mochokus	NS	NS
	Mochokiella	NS	NS UGF
	Synodontis	NS LGF CB	NS

Mochokiella and *Synodontis* are hypothesised to originate from a similar set of areas, as for the subfamily, for the DEC model, which differ to the single area origins reconstructed by the DIVALIKE model (Table 2). *Mochokus and Microsynodontis* are hypothesized to originate from single areas (N-S and CB respectively) under both models indicating range contractions.

Both models support contrasting biogeographic patterns for the two species rich genera *Synodontis* and *Chiloglanis*, which both have continental wide distributions (Fig. 3a, b). The ichthyo-provinces for *Synodontis* are largely aggregated into distinct biogeographic clades, with fewer dispersal events (Fig. 3b), supporting the findings of Day et al. (2013) and are not further described here. Conversely, a more complex biogeographic signal is inferred for *Chiloglanis*, in which clades have larger biogeographic distributions (Fig. 3a). Most of the areas occur repeatedly across the main clades (A, C, D), with many subsequent recolonisation's of the Congo Basin. Clade (A) is dominated by taxa from the CB, and to a lesser extent LGF. The latter area (LGF) includes the sister taxon (*C. disneyi*) to all other members of this clade, and a distinct geographic subclade nesting deeply within this clade, indicating that the CB has seeded the LGF in both cases (Fig. 3a). Several instances of independent dispersals from CB into other regions are inferred, including a dispersal into the Zambezi early in the clade's history, several independent colonisations into East Africa (EA), with further dispersal from this region into Nilo-Sudan (N-S) occurring much later. Clade C comprises taxa principally from the UGF and the Congo Basin, supporting a disjunct distribution between these subclades (C1, C2, Fig. 3a). In clade C there is a single dispersal event into the N-S, mirroring the timeframe of dispersal into this region observed in clade A, and also a single dispersal to the Zambezi early in the history of the clade. Clade D comprises taxa mainly from EA, Zambezi (Z) and CB ichthyo-provinces (which are reconstructed as ancestral), with diversification dominated within the former areas, with several dispersal events from EA to the Zambezi, but also from the Congo Basin to the Zambezi and subsequent recolonisation back into the Congo (Fig. 3a). There is also a single instance of dispersal from the Congo Basin to the Quanza (Q), and from the Zambezi to Lake Malawi, as identified for *Synodontis*, although colonisation of the Quanza is indicated to have occurred much earlier in *Chiloglanis*.

3.4. Diversification rates

The BAMM analysis supported a near constant diversification rate for the Mochokidae, with no rate shifts identified (a possible rate shift on the node leading to *Synodontis* had a Bayes Factor of only 2.67 after removing all outgroups) (Fig. 4a, b). The gamma (γ) statistic (Pybus and Harvey 2000) computed for the Mochokidae was slightly negative −1.184, but this value was non-significant ($p = 0.237$, two-tailed test) further supporting a constant diversification rate for the family. The impact of missing species was also investigated by applying the MMCR test (Pybus and Harvey, 2000), simulating 5000 trees for 285 taxa (including valid and possible candidate species across the family) was again non-significant ($p = 0.644$, two-tailed test). Inspection of the internal clade *Synodontis* appeared to show a declining rate (Fig. 4c) compared to the constant rate of *Chiloglanis* (Fig. 4d), which is supported by a significantly negative γ-statistic (−2.08 $p = 0.038$ vs. −0.282 $p = 0.778$ respectively), however, the MCCR test is non-significant ($p = 0.216$ vs. $p = 0.788$ respectively). Mean speciation rate (λ) for the family was estimated at 0.107 (0.091–0.128, upper/lower 95 % highest probability density [HPD], and where genera comprise > 10 species, λ was marginally lower for *Chiloglanis* 0.092 (0.071–0.119, upper/lower 95 % HPD) and *Microsynodontis* 0.093 (0.071–0.123, upper/lower 95 % HPD), but slightly higher for *Synodontis* 0.142 (0.094–0.213 upper/lower 95 % HPD).

4. Discussion

4.1. Exceptional levels of species discovery among rheophilic fishes

We identified exceptional levels of potential species discovery and endemicity in tropical fast flowing freshwaters based on our analysis of spatially extensive *Chiloglanis sampling*, supporting previous regional studies (Chakona et al., 2018; Morris et al., 2016; Schmidt et al., 2016; Schmidt et al., 2014; Schmidt et al., 2017). Although some of the undescribed diversity reported here was expected given the distinctive phenotypes of some specimens (e.g., "spotted", "Kinsuka blind", "thick barbels", "white-lips"), the majority was surprising, and likely reflects widespread and extensive cryptic diversity. Species richness estimates greatly expands current diversity to a putative *ca.* 119 species, representing > 80 % increase in species richness. The staggering increase in potential novel candidate species is, to our knowledge, one of the most substantial findings of undescribed diversity from molecular analysis in any vertebrate genus reported to date, although see other studies (e.g., Gehara et al., 2014; Van Ginneken et al., 2017).

While our estimates of proposed species diversity could be inflated due to over-splitting of taxa, the delimitation of described species in our dataset are broadly congruent across tree-based methods, which successfully recovered > 80 % of described species. That described species are generally consistently recovered (with no lumping) provides a robust baseline for unidentified candidates, with results broadly congruent across methods. Complete congruence across species-delimitation methods is rare (Dellicour and Flot, 2018), and although we have not applied allele-sharing or distance methods here, our findings provide an overall high level of confidence in biodiversity assessment, which is reassuring given that high degrees of discordance have been reported in some studies (Blair and Bryson, 2017).

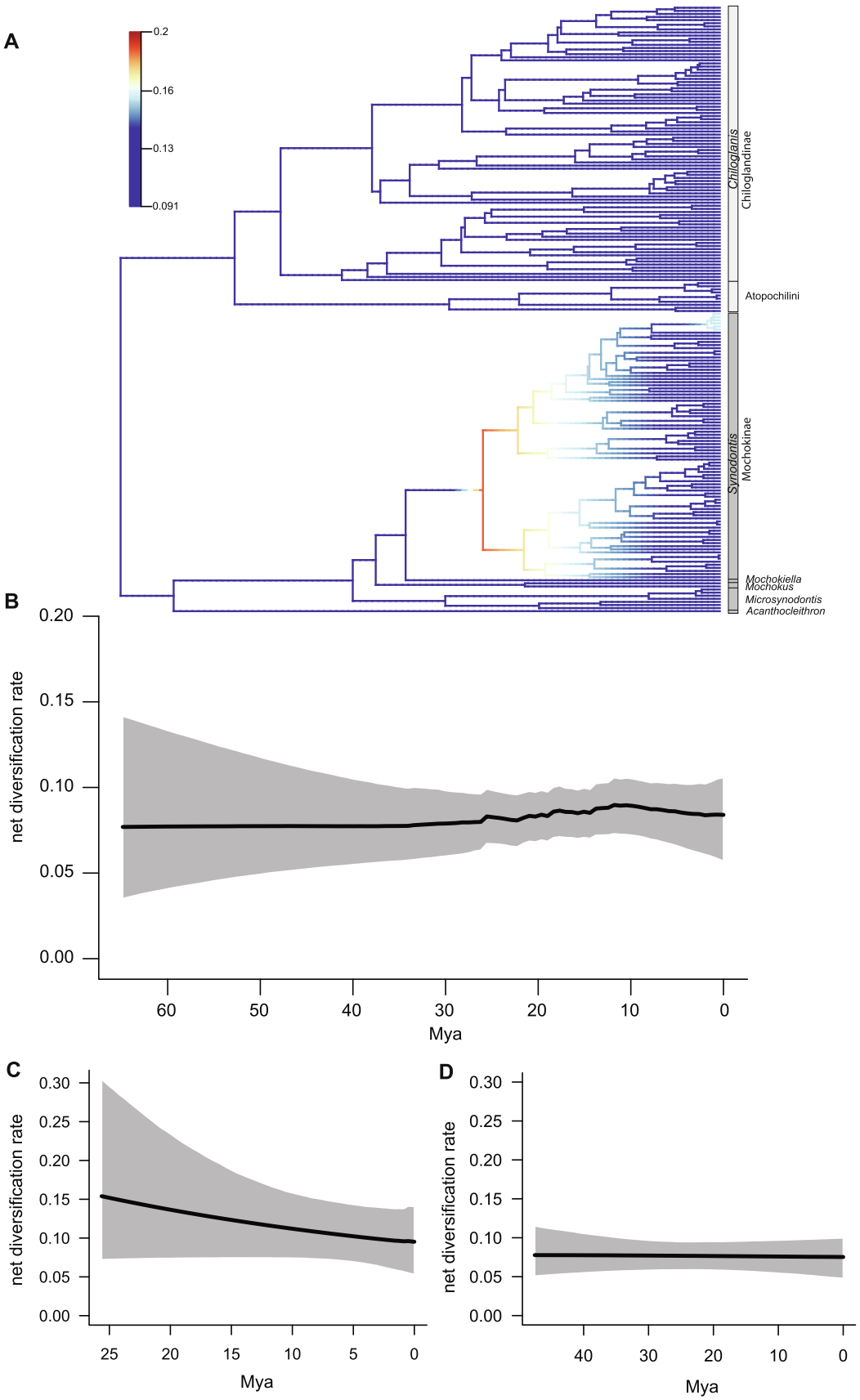

Fig. 4. A: Phylorate plot for the Mochokidae (warmer colours indicate faster rates, but note colours alone cannot be interpretted as evidence of rate heterogeneity, scale bar depicting speciation rate per Ma); Diversification rate through time plots for B: the Mochokidae; C: *Synodontis*; and D: *Chiloglanis* generated using BAMM.

Described species that species-delimitation methods suggested may actually contain multiple species typically had large biogeographic ranges (Froese and Pauly, 2022). Four of these were identified as non-monophyletic, and although non-monophyly is a potential issue with single-locus data since it could indicate ancient lineage sorting or introgression (Talavera et al., 2013), our findings were supported by our nuclear data, and therefore are unlikely to be the result of mito-nuclear discordance. Non-monophyly was further supported by both the mt- and nuDNA for many unidentified candidate *Chiloglanis* species from drainages in Tanzania and Malawi, although a number of these candidates (notably from Tanzania) are based on single individuals since we opted for spatial coverage across rivers. Despite methods such as GMYC being shown to be stable with high singleton presence (Talavera et al., 2013), further sampling of these candidates will be needed to confirm their status.

Where valid taxa were resolved as monophyletic, but were split into separate OTUs (*C. anoterus*, *C. brevibarbus*, *C. aff. micropogon*) these likely represent cases of true cryptic species as defined in Bickford et al. (2007) (i.e., two or more phenotypically similar species incorrectly classified as a single species). Indeed, the suggested delimitation of *C. anoterus* into five candidates, each with narrow ranges, was suggested by a previous study based on three independent datasets (male caudal fin morphology, mtDNA and AFLP data) (Morris et al., 2016). We note that while our identification of potential species-complexes are made solely with genetic data, they are likely valid given the ages of their divergences (Appendix A, Fig. S3). However, a recent study using a large exon genomic dataset suggested that the increasing levels of cryptic diversity (in frogs) based on distance and tree based methods could be due to gene flow as opposed to species divergence leading to inflated estimates (Chan et al., 2022). Certainly, the seven candidates identified within *C. aff. micropogon* warrant further attention regarding additional sampling and morphological investigation to confirm our hypothesis.

While we endeavoured to include dense sampling across the generic range, the high endemicity of *Chiloglanis* species in single river systems, indicates that this study may only be scratching the surface of their true diversity. Our results clearly underscore the need for more intensive sampling in poorly explored tropical freshwaters and highlights lotic habitats as potential hotspots for cryptic species discovery. Certainly, the environmental constraints acting on the phenotype in these habitats, has led to the generation of morphological conservatism and/or convergence across the clade resulting in widespread cryptic diversity in these diminutive, rapids-adapted fishes. Cryptic taxa are an increasingly common phenomenon reported in recent molecular studies of African riverine fish groups e.g., mormyrids (Feulner et al., 2006; Sullivan et al., 2002); *Pseudobarbus* (Chakona and Skelton, 2017), *Mastacembelus* spiny-eels (Day et al., 2017); *Enteromius,* and labeonine cyprinids (Van Ginneken et al., 2017), *Hydrocynus* African tigerfish (Goodier et al., 2011), tetras (Arroyave et al., 2019) and procatopodid cyprinodontiforms (Bragança et al., 2021) with some examples even in lacustrine fish radiations e.g., claroteine catfishes (Peart et al., 2018; Peart et al., 2014) and *Mastacembelus* spiny-eels (Brown et al., 2010) - where adaptive radiation via disruptive selection is a more typical mechanism, highlighting a deficiency in our taxonomic knowledge of the continent. In particular, the concept that the Congo Basin and East Africa are highly under-sampled regions of Africa that likely harbour substantial undescribed freshwater diversity (Darwall et al., 2011; Stiassny et al., 2011), is supported here.

4.2. The importance of the Congo Basin in the generation of mochokid diversity

The Congo Basin, occupying 10 % of the African continent (Kadima et al., 2011), is the most biodiverse African ichthyo-province (Harrison et al., 2016; Snoeks et al., 2011), although it is unclear to what extent this region is a centre of origin for freshwater fish clades since there are a limited number of studies attempting to quantitatively reconstruct the historical biogeography of pan-African fish clades (but see, Arroyave et al., 2020; Day et al., 2017; Day et al., 2013; Pinton et al., 2013). Our phylogeny, and the inclusion of previously unrecognised diversity, highlights the importance of the Congo Basin in the generation of diversity for a major African fish family. Certainly, drainage evolution of the proto-Congo since the end of the Cretaceous (proposed age of origin of the family, see Section 4.4) has been complex, with major river capture events occurring during the Oligocene-Eocene 30–40 Ma (Stankiewicz and de Wit, 2006), which may have facilitated diversification events. Irrespective as to whether the Congo Basin is combined with the 'West African' super-region (including N-S, LGF and UGF), as the sole centre of origin for the Mochokidae, there has been extensive diversification in this region early in the clade's history, and it is here hypothesised to be the centre of origin for the subfamily Chiloglanidinae and its constituent genera, along with subsequent repeated colonisations and diversifications events throughout the family's history. In contrast, ancestral ranges for the subfamily Mochokinae are more varied (within the 'West African' and Congo Basin ichthyo-provinces), with the Nilo-Sudan playing a prominent role. The Congo Basin has also been hypothesised as a centre of origin for several other fish clades including the characiform genera *Distichodus* (Arroyave et al. 2020) and *Hydrocynus* (Goodier et al., 2011), with considerable subsequent diversification within the region, and has been shown to harbour higher genetic diversity than other regions for continentally distributed species (Van Steenberge et al., 2020). Conversely, West Africa has been hypothesised to be the ancestral area for *Synodontis* (Day et al., 2013), African members of the synbranchiform genus *Mastacembelus* (Day et al. 2017), and a broad West African region (combined from the Upper Guinea and Western Nilo-Sudan) for the African potamonautine crabs (Daniels et al., 2015). Despite West Africa being inferred as a centre of origin for these other freshwater groups, the Congo Basin has clearly played a pivotal role in the generation of their diversity, with evidence of repeated colonisations and diversification within the basin for *Mastacembelus* spiny eels (Day et al., 2017), and several invertebrate groups, e.g., *Lanistes* gastropods (Mahulu et al., 2021) and potamonautine crabs (Daniels et al., 2015), and *Synodontis* catfish have diversified in situ in the Congo Basin to a significant degree (Fig. 3b). Certainly, our family-level study mirrors other pan-African freshwater clades in suggesting a complex scenario regarding the source of geographic diversity of the continent's freshwater communities, but further supports a central role for the Congo Basin.

4.3. Contrasting biogeographic histories of continental-wide African clades, and the role of dispersal

Our study qualitatively revealed contrasting biogeographic histories for the principal mochokid genera in the degree of geographic constraint of their constituent clades. While *Synodontis* has remarkably constrained biogeographic clades (Day et al. 2013, and this study) indicating relatively limited dispersal, *Chiloglanis* has a far more complex biogeographic history. In *Chiloglanis* most constituent clades were typically less geographically constrained, indicating dispersal as the major driver of diversification. The contrasting biogeographical signatures of these two genera, which have both colonised most of tropical Africa and diversified to a similar degree, likely reflect their differing ecologies that facilitated successful exploitation of the different freshwater habitats that they inhabit (*Synodontis* mainly large river/lake species vs. *Chiloglanis* almost exclusively rheophilic species) but may also be a consequence of their age differential. Interestingly, *Mastacembelus* (broadly co-occurring with *Synodontis*) revealed a more mixed biogeographic pattern, as while there have been multiple colonisation into the Congo Basin, all other regions were likely colonised once, indicating geographic constraint (Day et al., 2017). *Chiloglanis* occur almost exclusively in rapid flowing, high energy habitats and are characterised by high levels of endemicity; nonetheless the genus appears to have enhanced ability for dispersal, likely facilitated by their small size,

adhesive suckermouths, and broadly expanded pectoral/pelvic fins. Utilizing this anatomy some species have been observed to travel upstream across rapids and waterfalls (Morris et al., 2016; Roberts, 1975; Schmidt et al., 2016), which are typically effective barriers to dispersal to most fish clades. Enhanced dispersal ability alongside geomorphological processes such as river capture, and palaeoclimatic events, may have facilitated repeated colonisation of regions thus promoting allopatric speciation in this group (Schmidt et al., 2016).

The discovery of previously unrecognised *Chiloglanis* diversity, particularly in East African and the Zambezian ichthyo-provinces, suggest that these regions have been subjected to multiple colonisation events often with subsequent diversification, occurring late in the history of the group. At least five independent colonisations of the Zambezian region occurring at different times are inferred across all major *Chiloglanis* clades, although our constrained models (data not shown) could not determine if dispersal was both into and out of the Zambezi, or if the region has acted as a sink (i.e., dispersal has only occurred into this region), regardless the region has clearly not acted as a source. This contrasts with just a few recent colonisations of the region by *Synodontis,* indicating the Zambezi has played a more important role in the generation of diversity in *Chiloglanis*. Given the ecology of these two taxonomic groups, this finding is not surprising, as we would expect to see more colonisation events by species that typically occur in small tributary streams since river capture events of these environments are far more frequent than large river course changes. Multiple faunal transfers of the Zambezian ichthyo-province also appears to have occurred late in the history of various other freshwater clades (Arroyave et al., 2020; Daniels et al., 2015; Day et al., 2017; Mahulu et al., 2021; Ortiz-Sepulveda et al., 2020) either from East Africa and/or the Congo Basin, with varying degrees of subsequent diversification. Certainly, *Chiloglanis*, has repeatedly been able to colonise this region throughout the history of the clade, from the adjoining Congo Basin, and/or East Africa. Further evidence of multiple regional colonisations has also occurred with respect to the Nilo-Sudanic province. However, these events do not appear to have led to subsequent diversification, which contrasts to the colonisation of this region by other fish clades (Day et al., 2017; Day et al., 2013).

In contrast to other regions there is a single diversification event in the Upper Guinea Forest region, which may be due to it being separated by a savanna corridor (the Dahomey Gap) from the Lower Guinea Forest. However, this clade is estimated to have begun diversification in the Miocene (26.3 Ma [95 %HPD: 23.3–29.0 Ma]) long before the emergence of this landscape feature in the Holocene (Salzmann and Hoelzmann, 2005). An epicontinental sea has been suggested to have been a causative factor in the separation of killifish clades from east and west of the Dahomey Gap (Murphy and Collier, 1997), but this feature lasted only until the Early Eocene. While the disjunct distribution of clade C (Fig. 3a) may indicate a vicariant event, it is also plausible that there are missing species and/or multiple species that have since gone extinct from the Nilo-Sudan that subsequently seeded the Upper Guinea Forest. Certainly, rivers of the West African coast need further exploration to determine if this pattern represents a sampling gap rather than a reflection of biogeographic history.

4.4. Age estimates of the Mochokidae and constituent genera

We estimated the age of origin of the Mochokidae as *ca.* 66 Ma placing the emergence of the family around, or just after, the K-Pg boundary, while mochokid genera typically originated much more recently, except the monotypic *Acanthocleithron*. Cladogenesis of these principal mochokid lineages appears to have occurred asynchronously, with key divergence events estimated from the Early Eocene through the early Miocene, although much of the diversification within the constituent genera is suggested to have occurred during middle Miocene and onwards. The Miocene epoch is considered a key episode in the diversification of African fishes (Arroyave et al., 2020; Day et al., 2017; Day et al., 2013; Schwarzer et al., 2009) and freshwater invertebrates (Daniels et al., 2015; Mahulu et al., 2021; Ortiz-Sepulveda et al., 2020). This is thought to be due to widespread geotectonic uplift on the continent, along with major climatic shifts, such as the Middle Miocene climatic optimum, leading to increased river discharge, as well as repeated aridification events (Zachos et al., 2001). Such dynamic hydrological landscapes during this period likely facilitated widespread allopatric speciation across these diverse groups (Daniels et al., 2015). Of note is the disparity in ages of the two transcontinental genera, with *Chiloglanis* estimated to have diverged considerably earlier than *Synodontis*. The relatively recent diversification of *Synodontis*, along with biotic and abiotic factors, may explain their differing biogeographic histories. Despite the placement of a single calibration for the MRCA of *Synodontis*, the disparity in ages of these clades, and similar divergence times across the tree, were estimated when data was reanalysed enforcing only a secondary calibration on the 'Big Africa' node. However, as with many studies investigating lower taxonomic groups, dating may be more contentious since single calibrations and/or secondary calibrations can strongly influence estimated ages (Rutschmann et al., 2007), and we acknowledge that our study will require refinement before definitive conclusions may be reached.

4.5. Unexceptional diversification rates despite high species richness

The potential increase in mochokid species richness, specifically within the Chiloglanidinae, provides novel information to test evolutionary and biogeographic hypotheses more accurately. Here we demonstrated that missing taxa can have a significant effect on biogeographic reconstruction. Yet despite the discovery of numerous potential candidate species greatly augmenting species richness within *Chiloglanis,* diversification rates across the family remained unchanged (Fig. 4) and are similar to those generated for other riverine fishes (Miller, 2021). A near constant rate was previously identified for the mochokid genus *Synodontis* (Day et al. 2013), and unsurprisingly we find this model best supports the entire family. This pattern contrasts with the general pattern in larger species-level clades that typically show an early burst followed by a slow-down (Phillimore and Price, 2008). The early burst model is hypothesized to have been a result of the rapid saturation of available niches, often interpreted in terms of adaptive radiation and is typically reported for insular systems (e.g., Day et al., 2008; Reddy et al., 2012; Seehausen, 2004), although this signal may occur in the absence of ecology (Pigot et al., 2010). However, a constant / near-constant rate has been identified across other tropical continental faunas, including several other African terrestrial and aquatic clades (Liedtke et al., 2016; Day et al. 2017), and has been interpretated as a consequence of these clades not yet reaching their ecological carrying capacity, and/or tropical continental diversification not being as limited by ecological opportunities.

The precise factors that have promoted speciation in *Chiloglanis* and *Synodontis* have yet to be investigated. However, *Chiloglanis* have clearly specialised to live in extreme, high-flow habitats, likely leading to reduced competition allowing for exploitation of this extreme niche, while other mochokids are typically found in large, slower flowing rivers and lacustrine conditions (Froese and Pauly, 2022). However, in the case of *Synodontis* prior occupancy by other mochokid lineages appears not to have inhibited subsequent diversification within the genus. Clearly, future studies investigating details of habitat occupancy, trophic biology, and morphology, as well as genomic interrogation will help to clarify the noteworthy disparities in species richness across this large, endemic African family.

5. Conclusions

We highlight exceptional levels of undiscovered biological diversity in the form of cryptic species from African freshwaters, specifically in high-energy, lotic habitats, and show how such species discovery can

impact our understanding of evolutionary patterns and processes within lineages. Including sampling from across the range of described and newly discovered taxa, our study suggests a near 80 % increase in species diversity in the diminutive catfish genus *Chiloglanis*. Our results highlight the need for intensified exploration of African rivers, which remain woefully under surveyed (Snoeks et al., 2011), including a focus on different freshwater regimes, such as headwaters and rapids as potential centres of species richness and endemism. Our study, along with previous regional studies, highlights exceptionally high levels of endemicity within *Chiloglanis*, with species/candidate species often narrowly restricted to fast flowing sections of single river systems. Such a significant increase in endemic diversity also has implications for species conservation, particularly in the light of expected range contractions due to global climate change, and dam construction (Winemiller et al., 2016), indicating higher losses of cryptic evolutionary lineages as opposed to described morphospecies (Bálint et al., 2011). Currently, *ca.* 65 % of *Chiloglanis* species listed by IUCN have a conservation status of 'Least concern' or 'Data deficient' (IUCN, 2022). However through more detailed sampling of described taxa, we and others (Chakona et al., 2018; Morris et al., 2016; Schmidt et al., 2016) show that many species considered of 'Least concern' with large ranges likely represent species-complexes in which actual range sizes are considerably reduced, which would likely alter their risk of endangerment and conservation status. Greater sampling effort of tropical freshwaters is especially pertinent given that around a third of all freshwater fishes are now threatened with extinction (Hughes, 2021). Given these predictions and the undiscovered diversity based on our, and other, studies, it is likely that the biodiversity of this major ecological vertebrate grouping is disappearing before species diversity, as well as population diversity or connectivity, has been fully characterised. Furthermore, our results provide a much-needed phylogenetic framework for ongoing and future taxonomic studies. Detailed morphological study will be a necessary augmentation of our molecular data for the characterization and formal taxonomic description of many of these previously unrecognized *Chiloglanis* lineages and species.

Funding

Funding for MLJS was provided by the US National Science Foundation (DEB-05423540 and DEB-1655227) and the AMNH Axelrod Research curatorship (AA100023, AA130107). This work was further supported by a Leverhulme Trust-Africa Award; NSF Grant (DEB-0315963): PBI: All Catfish Species (Siluriformes) – Phase I of an Inventory of the Otophysi; NSF Grant (DEB-1023403) PBI: All Cypriniformes Species – Phase II of an Inventory of the Otophysi. A Genetics Society Summer Studentship provided support for Elizabeth Steell.

CRediT authorship contribution statement

Julia J. Day: Conceptualization, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Elizabeth M. Steell:** Investigation, Formal analysis, Writing – review & editing, Visualization. **Thomas R. Vigliotta:** Resources, Writing – review & editing. **Lewis A. Withey:** Investigation. **Roger Bills:** Resources, Writing – review & editing. **John P. Friel:** Resources. **Martin J. Genner:** Resources, Writing – review & editing. **Melanie L.J. Stiassny:** Resources, Writing – review & editing.

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.

Data availability

Sequence data are available from NCBI. Nexus files of aligned sequence data: Matrix 1 (mtDNA *Chiloglanis* data); Matrix 2 (Rag2 *Chiloglanis* data); Matrix 3 (concatenated mt and nuDNA Mochokid data) are available from Dryad https://doi.org/10.5061/dryad.tb2rbp04t.

Acknowledgements

For assistance with collection and exportation permitting MLJS is grateful to the Ministère de la Recherché Scientifique et de l'Innovation Technique, Institut National de Recherche en Sciences Exactes et Naturelles (IRSEN) and Groupe d'Etude et de Recherche sur la Diversité Biologique (GERBID) in the Republic of Congo, and to the Ministère de l'Agriculture et du Développement Rural, Sécretariat Général de l'Agriculture, Pêche et Elevage, Direction des Pêches and the Université de Kinshasa, Cabinet du Recteur, in the Democratic Republic of Congo. We thank Benjamin Ngatunga, Cyprian Katongo, Alan Smit, George Turner and Jennifer Swanstrom for their contributions to sampling in Tanzania and Zambia. We thank the Tanzania Commission for Science and Technology, Zambia Department of Fisheries and Malawi Department of Fisheries for sampling permissions. We thank Wendy Hart, Farrell McKenzie, Fraser Simpson and Keziah Conroy for helping with DNA extractions, PCRs, and for sequencing. We thank the reviewers for their comments.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2023.107754.

References

Abell, R., Thieme, M.L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., Coad, B., Mandrak, N., Balderas, S.C., Bussing, W., 2008. Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. Bioscience 58, 403–414.

Alter, S.E., Munshi-South, J., Stiassny, M.L.J., 2017. Genomewide SNP data reveal cryptic phylogeographic structure and microallopatric divergence in a rapids-adapted clade of cichlids from the Congo River. Mol. Ecol. 26, 1401–1419.

Arroyave, J., Martinez, C.M., Stiassny, M.L.J., 2019. DNA barcoding uncovers extensive cryptic diversity in the African long-finned tetra Bryconalestes longipinnis (Günther 1864) (Alestidae: Characiformes). J. Fish Biol. 95, 379–392.

Arroyave, J., Denton, J.S.S., Stiassny, M.L.J., 2020. Pattern and timing of diversification in the African freshwater fish genus Distichodus (Characiformes: Distichodontidae). BMC Evol. Biol. 20, 48.

Bálint, M., Domisch, S., Engelhardt, C.H.M., Haase, P., Lehrian, S., Sauer, J., Theissinger, K., Pauls, S.U., Nowak, C., 2011. Cryptic biodiversity loss linked to global climate change. Nat. Clim. Chang. 1, 313–318.

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K., Das, I., 2007. Cryptic species as a window on diversity and conservation. Trends Ecol. Evol. 22, 148–155.

Blair, C., Bryson, R.W.J., 2017. Cryptic diversity and discordance in single-locus species delimitation methods within horned lizards (Phrynosomatidae: Phrynosoma). Mol. Ecol. Resour. 17, 1168–1182.

Bouckaert, R., Drummond, A., 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. BMC Evol. Biol. 17.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537.

Bragança, P.H.N., Costa, W.J.E.M., 2019. Multigene fossil-calibrated analysis of the African lampeyes (Cyprinodontoidei: Procatopodidae) reveals an early Oligocene origin and Neogene diversification driven by palaeogeographic and palaeoclimatic events. Org. Divers. Evol. 19, 303–320.

Bragança, P.H.N., Van der Zee, J., Chakona, A., Schmidt, R.C., Stiassny, M.L.J., 2021. Following the Mangroves: diversification in the banded lampeye Aplocheilichthys spilauchen (Duméril, 1861) (Cyprinodontiformes: Procatopodidae) along the Atlantic coast of Africa. Hydrobiologia 848, 1433–1453.

Brown, J.H., 2013. Why are there so many species in the tropics? J. Biogeogr. 41, 8–22.

Brown, K.J., Rüber, L., Bills, R., Day, J.J., 2010. Mastacemblid eels support Lake Tanganyika as an evolutionary hotspot of diversification. BMC Evol. Biol. 10.

Chakona, A., Kadye, W.T., Bere, T., Mazungula, D.N., Vreven, E., 2018. Evidence of hidden diversity and taxonomic conflicts in five stream fishes from the Eastern Zimbabwe Highlands freshwater ecoregion. Zookeys Jun 19(768), 69–95.

Chakona, A.H., Skelton, P., 2017. A review of the Pseudobarbus after (Peters, 1864) species complex (Teleostei, Cyprinidae) in the eastern Cape Fold Ecoregion of South Africa. ZooKeys 657, 109–140.
Chan, K.O., Hutter, C.R., Wood Jr., P.L., Su, Y.-C., Brown, R.M., 2022. Gene flow increases phylogenetic structure and inflates cryptic species estimations: A case study on widespread Philippine puddle frogs (Occidozyga laevis). Syst. Biol. 71, 40–57.
Collen, B., Whitton, F., Dyer, E.E., Baillie, J.E., Cumberlidge, N., Darwall, W.R., Pollock, C., Richman, N.I., Soulsby, A.M., Böhm, M., 2014. Global patterns of freshwater species diversity, threat and endemism. Glob. Ecol. Biogeogr. 23, 40–51.
Cusimano, N., Stadler, T., Susanne, S., Renner, A., 2012. A new method for handling missing species in diversification analysis applicable to randomly or nonrandomly sampled phylogenies. Syst. Biol. 61, 785–792.
Daniels, S.R., Phiri, E.E., Klaus, S., Albrecht, C., Cumberlidge, N., 2015. Multilocus phylogeny of the Afrotropical freshwater crab fauna reveals historical drainage connectivity and transoceanic dispersal since the Eocene. Syst. Biol. 64, 549–567.
Darwall, W.R.T., Smith, K.G., Allen, D.J., Holland, R.A., Harrison, I.J., Brooks, E.G.E., 2011. The diversity of life in African freshwaters: Under water, under threat. An analysis of the status and distribution of freshwater species throughout mainland Africa. IUCN, Cambridge.
Day, J.J., Cotton, J.A., Barraclough, T.G., 2008. Tempo and mode of diversification of Lake Tanganyika cichlid fishes. PLoS ONE 3, e1730.
Day, J.J., Peart, C.R., Brown, K.J., Friel, J.P., Bills, R., Moritz, T., 2013. Continental diversification of an African catfish radiation (Mochokidae: *Synodontis*). Syst. Biol. 62, 351–365.
Day, J.J., Fages, A., Brown, K.J., Vreven, E.J., Stiassny, M.L.J., Bills, R., Friel, J.P., Rüber, L., 2017. Multiple independent colonizations into the Congo Basin during the continental radiation of African Mastacembelus spiny eels. J. Biogeogr. 44, 2308–2318.
Day, J.J., Wilkinson, M., 2006. On the origin of the Synodontis catfish flock from Lake Tanganyika. Biol. Lett. 2, 548–552.
Decru, E., Moelants, T., De Gelas, K., Vreven, E., Verheyen, E., Snoeks, J., 2015. Taxonomic challenges in freshwater fishes: a mismatch between morphology and DNA barcoding in fish of the north-eastern part of the Congo Basin. Mol. Ecol. Resour. 16, 342–352.
Dellicour, S., Flot, J.-F., 2018. The hitchhiker's guide to single-locus species delimitation. Mol. Ecol. Resour. 18, 1234–1246.
Derryberry, E.P., Claramunt, S., Derryberry, G., Chesser, R.T., Cracraft, J., Aleixo, A., Pérez-Emán, J., Remsen, J.V.J., Brumfield, R.T., 2011. Lineage diversification and morphological evolution in a large-scale continental radiation: the neotropical ovenbirds and woodcreepers (Aves: Furnariidae). Evolution 65, 2973–2986.
De-Silva, D.L., Elias, M., Willmott, K., Mallet, J., Day, J.J., 2016. Diversification of clearwing butterflies with the rise of the Andes. J. Biogeogr. 43, 44–58.
Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.
Duchêne, S., Lanfear, R., Ho, S.Y.W., 2014. The impact of calibration and clock-model choice on molecular estimates of divergence times. Mol. Phylogenet. Evol. 78, 277–289.
Ezard, T., Fujisawa, T., Barraclough, T.G., 2017. splits: SPecies' LImits by Threshold Statistics. R package version 1.0-19/r52. https://R-Forge.
Feulner, P.G.D., Kirschbaum, F., Schugardt, C., Ketmaier, V., Tiedemann, R., 2006. Electrophysiological and molecular genetic evidence for sympatrically occuring cryptic species in African weakly electric fishes (Teleostei: Mormyridae: Campylomormyrus). Mol. Phylogenet. Evol. 39, 198–208.
Ford, A.G.P., Bullen, T.R., Pang, L., Genner, M.J., Bills, R., Flouri, T., Ngatunga, B.P., Rüber, L., Schliewen, U.K., Seehausen, O., Shechonge, A., Stiassny, M.L.J., Turner, G. F., Day, J.J., 2019. Molecular phylogeny of Oreochromis (Cichlidae: Oreochromini) reveals mito-nuclear discordance and multiple colonisation of adverse aquatic environments. Mol. Phylogenet. Evol. 136, 215–226.
Fricke, R., Eschmeyer, W.N., van der Laan, R., 2022. Eschmeyer's catalog of fishes: genera, species, references. https://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp. Accessed 22/04/2022.
Friel, J.P., Vigliotta, T.R., 2011. Three new species of African suckermouth catfishes, genus Chiloglanis (Siluriformes: Mochokidae), from the lower Malagarsi and Luiche rivers of western Tanzania. Zootaxa 3063, 1–21.
Froese, R., Pauly, D., 2022. FishBase. www.fishbase.org. Accessed 22/04/2022.
Fujisawa, T., Barraclough, T.G., 2013. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. Syst. Biol. 62, 707–724.
Funk, W.C., Marcel, C., Santiago, R.R., 2021. High levels of cryptic species diversity uncovered in Amazonian frogs. Proc. R. Soc. B Biol. Sci. 279, 1806–1814.
Gehara, M., Crawford, A.J., Orrico, V.G.D., Rodríguez, A., Lötters, S., Fouquet, A., 2014. High levels of diversity uncovered in a widespread nominal taxon: Continental Phylogeography of the Neotropical Tree Frog Dendropsophus minutus. PLoS ONE 9, e103958.
Goodier, S.A.M., Cotterill, F.P.D., O'Ryan, C., Skelton, P.H., de Wit, M.J., 2011. Cryptic diversity of African Tigerfish (Genus Hydrocynus) reveals palaeogeographic signatures of linked Neogene geotectonic events. PLoS ONE 6, e28775.
Hackel, J., Henkel, T.W., Moreau, P.A., De Crop, E., Verbeken, A., Sà, M., Buyck, B., Neves, M.A., Vasco-Palacios, A., Wartchow, F., Schimann, H., Carriconde, F., Garnica, S., Courtecuisse, R., Gardes, M., Manzi, S., Louisanna, E., Roy, M., 2022. Biogeographic history of a large clade of ectomycorrhizal fungi, the Russulaceae, in the Neotropics and adjacent regions. New Phytol. 236, 698–713.
Harrison, I.J., Brummett, R.E., Stiassny, M.L.J., 2016. Congo River Basin. In: Finlayson, M., Milton, G.R., Prentice, R.C., Davidson, N.C. (Eds.), The Wetland Book: Distribution, Description and Conservation. Springer, Dordrecht.
Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. Proceedings of the National Academy of Sciences, U.S.A. 101, 14812–14817.
Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M., Ladle, R.J., 2015. Seven shortfalls that beset large-scale knowledge of biodiversity. Annu. Rev. Ecol. Evol. Syst. 46, 523–549.
Hrbek, T., Meliciano, N.V., Zuanon, J., Farias, I.P., 2018. Remarkable geographic structuring of rheophilic fishes of the Lower Araguaia River. Front. Genet. 9, 295.
Hughes, K., 2021. The World's Forgotten Fishes. WWF International, pp. 1–48.
IUCN, 2022. The IUCN Red List of threatened species. Version 2022-2. https://www.iucnredlist.org/ . Accessed 4 April 2022.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649.
Kadima, E., Delvaux, D., Sebagenzi, S.N., Tack, L., Kabeya, S.M., 2011. Structure and geological history of the Congo Basin: an integrated interpretation of gravity, magnetic and reflection seismic data. Basin Res. 23, 499–527.
Kalyaanamoorthy, S., Minh, B., Wong, T., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.
Kappas, I., Vittas, S., Pantzartzi, C.N., Drosopoulou, E., Scouras, Z.G., 2016. A time-calibrated mitogenome phylogeny of catfish (Teleostei: Siluriformes). PLoS ONE 11, e0166988.
Kurata, N.P., Hickerson, M.J., Hoffberg, S.L., Gardiner, N., Stiassny, M.L.J., Alter, S.E., 2022. Riverscape genomics of cichlid fishes in the lower Congo: Uncovering mechanisms of diversification in an extreme hydrological regime. Mol. Ecol. 31, 3516–3532.
Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 772–773.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.
Liedtke, H.C., Müller, H., Rödel, M.O., Menegon, M., Gonwouo, L.N., Barej, M.F., Gvoždík, V., Schmitz, A., Channing, A., Nagel, P., Loader, S.P., 2016. No ecological opportunity signal on a continental scale? Diversification and life-history evolution of African true toads (Anura: Bufonidae). Evolution 70, 1717–1733.
Liyandja, T.L.D., Armbruster, W.A., Poopola, M.O., Stiassny, M.L.J., 2022. Evolutionary convergence in body shape obscures taxonomic diversity in species of the African forskalii-group: case study of Labeo parvus Boulenger 1902 and L. ogunensis Boulenger 1910. J. Fish Biol. 101, 898–913.
Lujan, N.K., Conway, K.W., 2015. Life in the Fast Lane: A Review of Rheophily in Freshwater Fishes. In: Riesch, R., Tobler, M., Plath, M. (Eds.), Extremophile Fishes. Springer, Switerland, pp. 107–136.
Mahulu, A., Stelbrink, B., Van Bocxlaer, B., Riedel, F., Albrecht, C., 2021. Going with the flow? Diversification of gastropods reflects drainage evolution in Africa. J. Biogeogr. 48, 1579–1593.
Martins, F.C., Cox, S.C., Irestedt, M., Pŕýs-Jones, R.P., Day, J.J., 2020. A comprehensive molecular phylogeny of Afrotropical white-eyes (Aves: Zosteropidae) highlights prior underestimation of mainland diversity and complex colonization history. Mol. Phylogenet. Evol. 149, 106843.
Matzke, N.J., 2013. BioGeoBEARS: BioGeography with Bayesian (and likelihood) Evolutionary Analysis in R Scripts. http://cran.r-project.org/web/packages/BioGeoBEARS/.
Matzke, NJ., 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol. 63 (6), 951–970. https://doi.org/10.1093/sysbio/syu056.
Miller, E.C., 2021. Comparing diversification rates in lakes, rivers, and the sea. Evolution 75, 2055–2073.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), pp. 1–8.
Minh, B., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195.
Morris, J., Ford, A.G.P., Ali, J., Peart, C.P., Bills, R., Day, J.J., 2016. High levels of genetic structure and striking phenotypic variability in a sexually dimorphic suckermouth catfish from the African Highveld. Biol. J. Linn. Soc. 117, 528–546.
Mbimbi Mayi Munene, J.J., Stiassny, M.L.J., Monsembula Iyaba, R.J.C., Liyandja, T.L.D., 2021. Fishes of the Lulua River (Kasai basin, central Africa): a continental hotspot of ichthyofaunal diversity under threat. Diversity 13, 341.
Murphy, W.J., Collier, G.E., 1997. A molecular phylogeny for Aplocheiloid fishes (Atherinomorpha, Cyprinodontiformes): the role of vicariance and origins of annualism. Mol. Phylongenet. Evol. 14, 790–799.
Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating Maximum-Likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.
Ortiz-Sepulveda, C.M., Stelbrink, B., Vekemans, X., Albrecht, C., Riedel, F., Todd, J.A., Van Bocxlaer, B., 2020. Diversification dynamics of freshwater bivalves (Unionidae: Parreysiinae: Coelaturini) indicate historic hydrographic connections throughout the East African Rift System. Mol. Phylogenet. Evol. 148, 106816.
Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20, 289–290.

Peart, C.R., Bills, R., Wilkinson, M., Day, J.J., 2014. Nocturnal claroteine catfishes reveal dual colonisation but a single radiation in Lake Tanganyika. Mol. Phylogenet. Evol. 73, 119–128.
Peart, C.P., Dasmahapatra, K.K., Day, J.J., 2018. Contrasting geographic structure in evolutionarily divergent Lake Tanganyika catfishes. Ecol. Evol. 8, 2688–2697.
Phillimore, A.B., Price, T.D., 2008. Density-dependent cladogenesis in birds. PloS Biol. 6, e71.
Pigot, A.L., Phillimore, A.B., Owens, I.P.F., Orme, C.D.L., 2010. The shape and temporal dynamics of phylogenetic trees arising from geographic speciation. Syst. Biol. 59, 660–673.
Pinton, A., Agnèse, J.-F., Paugy, D., Otero, O., 2013. A large-scale phylogeny of Synodontis (Mochokidae, Siluriformes) reveals the influence of geological events on continental diversity during the Cenozoic. Mol. Phylogenet. Evol. 66, 1027–1040.
Pybus, O.G., Harvey, P.H., 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. Proc. R. Soc. Lond. B 267, 2267–2272.
R Development Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, http://www.R-project.org/.
Rabosky, D.L., Santini, F., Eastman, J., Smith, S.A., Sidlauskas, B., Chang, J., Alfaro, M.E., 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. Nat. Commun. 4, 1958.
Rabosky, D.L., Grundler, M., Anderson, C., Shi, J.J., Brown, J.W., Huang, H., Larson, J.G., 2014. BAMMtools: An R package for the analysisof evolutionary dynamics on phylogenetic trees. Methods Ecol. Evol. 5, 701–707.
Rambaut, A., Drummond, A., Xie, D., Baele, G., Suchard, M., 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904.
Rambaut, A., 2016. FigTree: a graphical viewer for phylogenetics trees. http://tree.bio.ed.ac.uk/software/figtree/.
Reddy, S., Driskell, A., Rabosky, D.L., Hackett, S.J., Schulenberg, T.S., 2012. Diversification and the adaptive radiation of the vangas of Madagascar. Proc. R. Soc., Ser. B 279, 2062–2071.
Reid, A.-A.-O., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T.J., Kidd, K.A., MacCormack, T.J., Olden, J.D., Ormerod, S.J., Smol, J.P., Taylor, W.W., Tockner, K., Vermaire, J.C., Dudgeon, D., Cooke, S.J., 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. Biol. Rev. Camb. Philos. Soc. 94, 849–873.
Reid, N.M., Carstens, B.C., 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evol. Biol. 12, 196.
Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3, 217–223.
Roberts, T.R., 1975. Geographical distribution of African freshwater fishes. Zool. J. Linn. Soc. 57, 249–319.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Rutschmann, F., Eriksson, T., Salim, K.A., Conti, E., 2007. Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points. Syst. Biol. 56, 591–608.
Salzmann, U., Hoelzmann, P., 2005. The Dahomey Gap: an abrupt climatically induced rain forest fragmentation in West Africa during the late Holocene. The Holocene 15, 1–10.
Schedel, F.D.B., Chakona, A., Sidlauskas, B.L., Popoola, M.O., Wingi, N.U., Neumann, D., Vreven, E.J.W.M.N., Schliewen, U.K., 2022. New phylogenetic insights into the African catfish families Mochokidae and Austroglanididae. J. Fish Biol. 100, 1171–1186.
Schmidt, R.C., Bart, H.L., Nyingi, D.W., Gichuki, N.N., 2014. Phylogeny of suckermouth catfishes (Mochokidae: Chiloglanis) from Kenya: The utility of Growth Hormone introns in species level phylogenies. Mol. Phylogenet. Evol. 79, 415–421.
Schmidt, R.C., Bart Jr., H.L., Pezold, F., 2016. High levels of endemism in suckermouth catfishes (Mochokidae: Chiloglanis) from the Upper Guinean forests of West Africa. Mol. Phylogenet. Evol. 100, 199–205.
Schmidt, R.C., Bart, J., Pezold, F., Friel, J.P., 2017. A biodiversity hotspot heats up: Nine new species of suckermouth catfishes (Mochokidae: Chiloglanis) from Upper Guinean forest streams in West Africa. Copeia 105, 301–338.
Schwarzer, J., Misof, B., Tautz, D., Schliewen, U., 2009. The root of the East African cichlid radiations. BMC Evol. Biol. 9, 186.
Seegers, L., 2008. The catfishes of Africa. A handbook for identification and maintenance. Aquaristik - Consulting & Service GmbH, Rodgau.
Seehausen, O., 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19, 198–207.
Skelton, P.H., Swartz, E.R., 2011. Walking the tightrope: trends in African freshwater systematic ichthyology. J. Fish Biol. 79, 1413–1435.
Snoeks, J., Harrison, I.J., Stiassny, M.L.J., 2011. The status and distribution of freshwater fishes. In: Darwall, W.R.T., Smith, K.G., Allen, D.J., Holland, R.A., Harrison, I.J., Brooks, E.G.E. (Eds.), The diversity of life in African freshwaters: underwater, under threat. IUCN, Cambridge, U.K. and Gland, Switzerland, pp. 42–73.
Stankiewicz, J., de Wit, M.J., 2006. A proposed drainage evolution model for Central Africa—did the Congo flow east? J. Afr. Earth Sc. 44, 75–84.
Stiassny, M.L.J., Brummett, R.E., Harrison, I.J., Monsembula, R., Mamonekene, V., 2011. The status and distribution of the freshwater fishes of Central Africa. In: Brooks, E.G.E., Allen, D.J., Darwall, W.T. (Eds.), The Status and Distribution of Freshwater Biodiversity in Central Africa. IUCN Red List of Threatened Species, Regional Assessment, IUCN, Gland, Switzerland, pp. 27–46.
Strayer, D.L., Dudgeon, D., 2010. Freshwater biodiversity conservation: Recent progress and future challenges. J. N. Am. Benthol. Soc. 29, 344–358.
Sullivan, J.P., Lavoué, S., Hopkins, C.D., 2002. Discovery and phylogenetic analysis of a riverine species flock of African electric fishes (Mormyridae: Teleostei). Evolution 56, 597–616.
Sullivan, J.P., Lundberg, J.G., Hardman, M., 2006. A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences. Mol. Phylogenet. Evol. 41, 636–662.
Swartz, E.R., Skelton Ph Fau - Bloomer, P., Bloomer, P., 2009. Phylogeny and biogeography of the genus Pseudobarbus (Cyprinidae): shedding light on the drainage history of rivers associated with the Cape Floristic Region. Mol. Biol. Evol. 51, 75–84.
Vigliotta, T.R., 2008. A phylogenetic study of the African catfish family Mochokidae (Osteichthyes, Ostariophysi, Siluriformes), with a key to genera. Proc. Acad. Nat. Sci. Phila. 157, 73–136.
Talavera, G., Dincă, V., Vila, R., 2013. Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. Methods Ecol. Evol. 4, 1101–1110.
van der Merwe, P.D.W., Cotterill, F.P.D., Kandziora, M., Watters, B.R., Nagy, B., Genade, T., Flügel, T.J., Svendsen, D.S., Bellstedt, D.U., 2021. Genomic fingerprints of palaeogeographic history: The tempo and mode of rift tectonics across tropical Africa has shaped the diversification of the killifish genus Nothobranchius (Teleostei: Cyprinodontiformes). Mol. Phylogenet. Evol. 158, 106988.
Van Ginneken, M., Decru, E., Verheyen, E., Snoeks, J., 2017. Morphometry and DNA barcoding reveal cryptic diversity in the genus Enteromius (Cypriniformes: Cyprinidae) from the Congo basin, Africa. Eur. J. Taxon. 310.
Van Steenberge, M.W., Vanhove, M.P.M., Chocha Manda, A., Larmuseau, M.H.D., Swart, B.L., Khang'Mate, F., Arndt, A., Hellemans, B., Van Houdt, J., Micha, J.C., Koblmüller, S., Roodt-Wilding, R., Volckaert, F.A.M., 2020. Unravelling the evolution of Africa's drainage basins through a widespread freshwater fish, the African sharptooth catfish *Clarias gariepinus*. J. Biogeogr. 47, 1739–1754.
Winemiller, K.O., McIntyre, P.B., Castello, L., Fluet-Chouinard, E., Giarrizzo, T., Nam, S., Baird, I.G., Darwall, W., Lujan, N.K., Harrison, I., Stiassny, M.L.J., Silvano, R.A.M., Fitzgerald, D.B., Pelicice, F.M., Agostinho, A.A., Gomes, L.C., Albert, J.S., Baran, E., Petrere, M., Zarfl, C., Mulligan, M., Sullivan, J.P., Arantes, C.C., Sousa, L.M., Koning, A.A., Hoeinghaus, D.J., Sabaj, M., Lundberg, J.G., Armbruster, J., Thieme, M.L., Petry, P., Zuanon, J., Vilara, G.T., Snoeks, J., Ou, C., Rainboth, W., Pavanelli, C.S., Akama, A., Soesbergen, A.v., Sáenz, L., 2016. Balancing hydropower and biodiversity in the Amazon, Congo, and Mekong. Science 351, 128–129.
Zachos, J.C., Pagani, M., Sloan, L., Thomas, E., 2001. Trends, rhythms, and aberrations in the global climate 65 Ma to present. Science 292, 686–693.
Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29, 2869–2876.
Zheng, C., Ye, Z., Zhu, X., Zhang, H., Dong, X., Chen, P., Bu, W., 2020. Integrative taxonomy uncovers hidden species diversity in the rheophilic genus Potamometra (Hemiptera: Gerridae). Zool. Scr. 49, 174–186.