

African lates perches (Teleostei, Latidae, Lates): Paraphyly of Nile perch and recent colonization of Lake Tanganyika

Peer-reviewed author version

Koblmüller, Stephan; Schöggel, Christian A.; Lorber, Clemens J.; VAN STEENBERGE, Maarten; KMENTOVA, Nikol; VANHOVE, Maarten & Zangl, Lukas (2021) African lates perches (Teleostei, Latidae, Lates): Paraphyly of Nile perch and recent colonization of Lake Tanganyika. In: MOLECULAR PHYLOGENETICS AND EVOLUTION, 160 (Art N° 107141).

DOI: 10.1016/j.ympev.2021.107141

Handle: <http://hdl.handle.net/1942/33797>

1 African lates perches (Teleostei, Latidae, *Lates*): paraphyly of Nile perch and recent colonization of
2 Lake Tanganyika

3

4 Stephan Koblmüller¹, Christian A. Schögggl¹, Clemens J. Lorber¹, Maarten Van Steenberge², Nikol
5 Kmentová^{3,4}, Maarten P. M. Vanhove^{3,4,5}, Lukas Zangl^{1,6,7}

6

7 ¹ Institute of Biology, University of Graz, Universitätsplatz 2, 8010 Graz, Austria

8 ² Operational Directorate Taxonomy and Phylogeny, Royal Belgian Institute for Natural Sciences,
9 Vautierstraat 29, 1000 Brussels, Belgium.

10 ³ Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, 611 37,
11 Brno, Czech Republic

12 ⁴ Hasselt University, Centre for Environmental Sciences, Research Group Zoology: Biodiversity &
13 Toxicology, Agoralaan Gebouw D, B-3590 Diepenbeek, Belgium

14 ⁵ Zoology Unit, Finnish Museum of Natural History, University of Helsinki, P.O. Box 17, Helsinki FI-
15 00014, Finland

16 ⁶ Universalmuseum Joanneum, Studienzentrum Naturkunde, Weinzöttlstraße 16, 8045 Graz, Austria

17 ⁷ ÖKOTEAM - Institute for Animal Ecology and Landscape Planning, Bergmannngasse 22, 8010 Graz,
18 Austria

19

20 Abstract

21 Lates perches of the genus *Lates* (Latidae) are large piscivorous fishes, with a strikingly disjunct
22 distribution range in coastal areas and estuaries of the Indo-Pacific region and in some large African
23 freshwater systems. Previous phylogenetic hypotheses based on osteological and ontogenetic data
24 suggested paraphyly of the African representatives, or even the small Lake Tanganyika species
25 assemblage, with respect to the remaining *Lates* species. Based on a multilocus phylogeny, however,
26 we show that extant African lates perches are monophyletic. The Nile perch, *L. niloticus*, which is
27 widely distributed in the Nilo-Sudan region and Central Africa, comprises three distinct lineages and
28 is paraphyletic with respect to the four endemic Lake Tanganyika species. We find that diversification
29 of extant African *Lates* happened only as recently as the Pliocene. With the extensive, in part much
30 older fossil record, this suggests repeated extinction and (re-)colonization of hydrological systems.
31 We further find that *Lates* started to diversify in Lake Tanganyika only in the Pleistocene, which is
32 much more recent than other fish radiations endemic to Lake Tanganyika, implying that they
33 radiated in the presence of other top predators already in this ecosystem.

34

35 Keywords: Africa, divergence time estimation, freshwater fish, multilocus phylogeny, radiation

36

37 Introduction

38 Lates perches (Latidae) are a family of large piscivorous fishes that occur in marine, brackish and
39 freshwater habitats in the Indo-Pacific region and in African freshwater systems. Previously, lates
40 perches were together with snooks (*Centropomus*) in Centropomidae, but Mooi and Gill (1995)
41 recommended the recognition of two distinct families, Centropomidae and Latidae. Thus, currently
42 Latidae is considered a distinct family of three extant genera and fourteen species (Van der Laan et
43 al., 2014; Iwatsuki et al., 2018; Fricke et al., 2021; but see Betancur-Rr et al., 2017 for an alternative
44 view). *Hypopterus* and *Psammoperca* include one and two species, respectively, and inhabit marine
45 and brackish environments in the Indo-Pacific (Iwatsuki et al., 2018). *Lates* includes marine and
46 freshwater species with a strikingly disjunct distribution range (Fig. 1a). Four species of *Lates*, *L.*
47 *calcarifer* (Bloch, 1790), *L. japonicus* Katayama & Taki, 1984, *L. lakdiva* Pethiyagoda & Gill, 2012 and
48 *L. uwisara* Pethiyagoda & Gill, 2012, occur in coastal areas and estuaries of the Indo-Pacific region,
49 but not along the African coast or Madagascar. The remaining seven *Lates* species inhabit large
50 African freshwater systems. Several species of *Lates* can grow large (>1m) and are thus highly
51 valuable food and game fish in both commercial and artisanal fisheries, such as the Barramundi (*L.*
52 *calcarifer*) or the Nile perch (*L. niloticus* (Linnaeus, 1758)) (Fig. 1b,c).

53 Among the seven African species, the Nile perch has by far the largest distribution range, being
54 native to numerous large river systems and lakes in the Nilo-Sudanic zone and Central Africa. In Lake
55 Albert and Lake Turkana, it occurs sympatrically with *L. macrophthalmus* Worthington, 1929 and *L.*
56 *longispinis* Worthington, 1932, respectively. It was also introduced in several rivers and lakes,
57 including Lake Victoria, where it caused a well-documented ecological disaster (Goldschmidt 1996).
58 Four species, *L. angustifrons* Boulenger, 1906, *L. mariae* Steindachner, 1909, *L. microlepis*
59 Boulenger, 1898 and *L. stappersii* (Boulenger, 1914) are endemic to Lake Tanganyika. Lates perches
60 are among the top predators in their ecosystems, with spatial segregation according to habitat
61 preference when different *Lates* species occur sympatrically (Harrison, 1991; Coulter, 1991).

62 Greenwood (1976) and Otero (2004), mainly based on osteological characteristics, suggested
63 paraphyly of the African species of *Lates* and Otero et al. (2004) even found some weak support for a
64 paraphyletic Lake Tanganyika assemblage (Fig. 1d). Non-monophyly of *Lates* in Lake Tanganyika was
65 also supported by ontogenetic data (Kinoshita and Tshibangu, 1997). A recent study on the molecular
66 phylogenetic relationships among snooks and lates perches had insufficient taxon coverage to test
67 either of these two scenarios (Li et al., 2011). Considering the uncertain intrageneric relationships, it
68 is still not clear whether inhabiting freshwater or marine systems is the ancestral state in *Lates* and
69 how often marine-freshwater transitions occurred during latid evolution.

70 We used multilocus sequence data to test whether i) living in freshwater or marine systems is the
71 ancestral state within *Lates*, and ii) African *Lates* is monophyletic and iii) the Lake Tanganyika
72 endemics constitute a monophylum. Furthermore, we provide a temporal framework for latid
73 diversification.

74

75 Material and Methods

76 Our study comprises 31 newly sequenced samples of African *Lates* spp. and some previously
77 published latid and centropomid data (Li et al., 2011). Samples (ethanol-preserved fin clips) were
78 obtained from museum collections, local fish markets, or the aquarium trade (Suppl. Table 1).

79 Unfortunately, no unambiguously identified material of the Lake Albert and Lake Turkana endemics
80 *L. macrophthalmus* and *L. longispinis*, respectively, could be obtained. These species, however,
81 appear to be closely related to the local *L. niloticus* (Hauser et al., 1998), such that the general
82 biogeographic patterns won't change.

83 DNA was extracted from fin clips using a rapid Chelex protocol (Richlen and Barber, 2005) or the
84 DNeasy Blood and Tissue Kit (Qiagen). One mitochondrial (16S rDNA) and 11 nuclear genes (FIC
85 domain-containing protein (ficd), kelch repeat and BTB (POZ) domain containing 4 (KBTBD4),
86 LOC562320 (KIAA1239), cardiac muscle myosin heavy chain 6 alpha (myh6), pleiomorphic adenoma

87 protein-like 2 (plagl2), receptor-interacting serine–threonine kinase 4 (RIPK4), si:dkey-174m14.3
88 (sidkey), zgc:85947 (SLC10A3), brain super conserved receptor 2 gene (sreb2), zic family member 1
89 (zic1), zinc finger 536 protein fragment (znf536)) were amplified and sequenced following Li et al.
90 (2011). Sequencing products were purified with Sephadex™ G-50 (GE Healthcare) and visualized on
91 an ABI 3130xl capillary sequencer (Applied Biosystems). All newly generated sequences have been
92 deposited on GenBank (see Suppl. Table 1).

93 Sequences were aligned using MUSCLE (Edgar, 2004) as implemented in MEGA 7.0 (Kumar et al.,
94 2016). Three different datasets were subjected to phylogenetic analysis: (i) mtDNA, (ii) all nuclear
95 genes concatenated, and (iii) nuclear plus mitochondrial genes concatenated. Representatives of
96 *Psammoperca* and *Centropomus* were used as outgroups. (Li et al., 2011) Standard phylogenetic
97 analyses were conducted via the PhyloSuite platform (Zhang et al., 2020). Maximum likelihood (ML)
98 and Bayesian inference (BI) phylogenetic analyses employed the partitioning schemes and
99 substitution models determined under the “greedy search” algorithm in PartitionFinder 2 (Lanfear et
100 al., 2017) with the Bayesian information criterion (BIC), with a starting scheme that partitioned by
101 locus and codon position within each locus (for best partitioning schemes and substitution models
102 see Suppl. Table 2). A standard maximum likelihood tree search was conducted using IQ-TREE
103 (Nguyen et al., 2015). Node support was assessed with 1000 standard bootstrap replicates. Bayesian
104 phylogenetic inference was performed in MrBayes v.3.2.6 (Ronquist et al., 2012). Posterior
105 probabilities were obtained from Metropolis-coupled Markov chain Monte Carlo simulations (2
106 independent runs; 6 chains; 10 million (dataset i) or 20 million (datasets ii & iii) generations, sampling
107 frequency 1,000; 25% burn-in). Chain stationarity and parameter convergence were assessed in
108 Tracer 1.7 (Rambaut et al., 2018) and via the average standard deviation of split frequencies (well
109 below 0.01 in all datasets), and post burn-in trees were summarized in a 50% majority rule consensus
110 tree. Visualization of ML and BI trees was done in FigTree 1.4.2 (available at
111 <http://tree.bio.ed.ac.uk/software/figtree>).

112 Additionally, a multispecies coalescent analysis based on all loci was conducted using the StarBeast2
113 package (Heled and Drummond, 2010) implemented in BEAST v.2.5.2 (Bouckaert et al., 2014). Here,
114 following Li et al. (2011) and Betancur-R et al. (2013), the tree was rooted with representatives of
115 *Centropomus* (Supplementary Table 1). Substitution, clock, and tree models were unlinked across all
116 loci. For each locus we employed the best fitting substitution model inferred based on the BIC in
117 ModelFinder (Kalyaanamoorthy et al., 2017) (Supplementary Table 2). As likelihood ratio tests in
118 TREE-PUZZLE 5.3.rc16 (Schmidt et al., 2002) rejected a clock-like evolution only for the ficd and
119 KIAA1239 data, we applied an uncorrelated lognormal relaxed clock model for these two loci and a
120 strict clock model for all other loci under a birth-death tree prior and analytical population size
121 integration. Absolute divergence times were obtained by employing two calibration schemes. The
122 first calibration scheme employed a lognormal prior (mean = 0.67, s.d. = 0.8, offset = 50 MYA) for the
123 Centropomidae-Latidae divergence, a secondary calibration following Betancur-R et al. (2013) and
124 references therein, and a lognormal prior (mean = 1.9, s.d. = 0.7, offset = 29.5 MYA) for the
125 *Psammoperca-Lates* split, based on the oldest *Lates* fossil (*L. qatraniensis* from the Jebel Qatrani
126 Formation, Fayum Depression, Egypt, 29.5-30.2 MYA; Murray and Attia, 2004) and the assumed
127 minimum root age of 50 mya (Betancur-R et al., 2013). The second calibration scheme used the same
128 two calibration points, but in addition employed a lognormal prior (mean = 1.8, s.d. = 0.5, offset =
129 15.5) for the split of the African *Lates* from the Indo-Pacific species, based on the oldest *L. niloticus*
130 fossils (*L. cf. niloticus* from the Ghaba deposits in Oman, 15.5-17.5 MYA & *L. niloticus* from the As-
131 Sarrar deposits in Saudi Arabia, 15.0-17.0 MYA; Otero and Gayet, 2001). Two independent replicates
132 were run with random starting seeds and chain lengths of 5×10^8 generations and parameters were
133 sampled every 5,000 steps. Chain convergence and stationarity of model parameter posterior
134 distributions were assessed in Tracer 1.7, runs were combined with LogCombiner (part of the BEAST2
135 package), and species trees were visualized in DensiTree2 (part of the BEAST2 package).

136

137 Results

138 Phylogenetic inference based on BI and ML produced highly consistent tree topologies and was also
139 largely congruent among all datasets (Fig. 1 and Supplementary Figs. 3, 4). Inconsistencies concerned
140 nodes that received low statistical support in at least one of two tree building algorithms. In
141 particular the branching order among *L. calcarifer*, *L. japonicus* and the African *Lates* differed among
142 datasets and tree building algorithms. Within African *Lates*, four main lineages were resolved: *L.*
143 *niloticus* from western Africa, *L. niloticus* from the Nile system (Lakes Albert and Turkana), *L. niloticus*
144 from the Congo basin, and the Lake Tanganyika species flock. *Lates niloticus* always resulted as
145 paraphyletic. In all but the mitochondrial trees, western and Nilotic *L. niloticus* were resolved as
146 sister groups in all but the mitochondrial trees, but nodal support for the alternative mitochondrial
147 relationships was low. A sister group relationship between Lake Tanganyika *Lates* and *L. niloticus*
148 from the Congo basin was recovered in all analyses. Within the Lake Tanganyika lineage, all analyses
149 but those based on mtDNA, resolved *L. microlepis* as sister group to the rest, followed by *L.*
150 *stappersii*, *L. mariae* and *L. angustifrons*. A sister group relationship between the latter two species
151 was well supported by the mitochondrial and full datasets, whereas the nuclear data resolved *L.*
152 *stappersii* and *L. angustifrons* as sister species, albeit with low statistical support.

153 A congruent topology was also recovered by the multispecies coalescent tree (Fig. 2), but *L. calcarifer*
154 and *L. japonicus* were resolved as sister species, albeit with weak statistical support. Divergence time
155 estimates suggest that, depending on the calibration scheme, the African and Indo-Pacific
156 representatives diverged 11.61 (95% HPD, 9.85 - 13.46) or 17.76 (95% HPD, 16.41 – 19.32) MYA and
157 that extant African lineages of *Lates* diverged only very recently, with a most recent common
158 ancestor (MRCA) 2.62 (95% HPD, 2.01 - 3.26) or 3.69 (95% HPD, 2.85 – 4.59) MYA and the onset of
159 the radiation in Lake Tanganyika 1.27 (95% HPD, 0.87 – 1.70) or 1.76 (95% HPD, 1.20 – 2.35) MYA.

160

161 Discussion

162 We investigated the phylogenetic relationships of lates perches, with emphasis on African
163 representatives. We found that African *Lates* constitutes a monophylum, with *L. niloticus* being
164 composed of three distinct lineages and paraphyletic with respect to the monophyletic, endemic
165 Lake Tanganyika species flock, the ancestors of which colonized the lake only well after its formation
166 and much later than the ancestors of most other endemic fish radiations.

167 Our inferred phylogenetic relationships are in stark contrast to previous morphology-based
168 phylogenetic hypotheses that suggested either a sister group relationship between the Lake
169 Tanganyika *Lates* species flock and all other African and Indo-Pacific species (Greenwood, 1976) or a
170 sister group relationship between three of the Lake Tanganyika species (*L. mariae*, *L. microlepis*, *L.*
171 *stappersii*) and all other *Lates* species (Otero, 2004). *Lates* spp. have a conserved general morphology
172 with considerable intraspecific variation, but – with the exception of the Lake Tanganyika species –
173 often little interspecific differences. Indeed, even detailed analysis of osteological characters did not
174 fully resolve the intrageneric relationships in *Lates*. The hypothesis that the Lake Tanganyika species
175 and all other *Lates* species are reciprocally monophyletic entities was mainly based on an elongation
176 of the ethmovomerine region in the Lake Tanganyika species (Greenwood, 1976). A non-
177 monophyletic Lake Tanganyika assemblage on the other hand, was supported by the presence of a
178 well-developed pleurosphenoid pedicle in *L. angustifrons* and all non-Lake Tanganyika species, and
179 its absence in *L. mariae*, *L. microlepis* and *L. stappersii* (Otero, 2004). We show that the Lake
180 Tanganyika species did not diverge earliest but occupy a derived position in the *Lates* tree. This
181 implies that the elongation of the ethmovomerine region happened only late in the evolution of
182 *Lates*. Also, the well-developed pleurosphenoid pedicle evolved independently in *L. angustifrons* and
183 the non-Lake Tanganyika species, or, alternatively, was lost in the other Lake Tanganyika species.

184 *Lates niloticus* comprises three lineages: in western Africa, in the Nile basin (including Lake Turkana),
185 and the Congo basin. The Congo lineage is sister to the Lake Tanganyika flock, rendering *L. niloticus*
186 paraphyletic. Analysis of microsatellite data (Basiita et al., 2018) and mitochondrial control region
187 sequences (Mwanja et al., 2013) already showed a deep east-west divergence (Lakes Victoria,

188 Nabugabo, Albert, Kyoga, Turkana & Chamo versus Senegal and Niger Rivers) in *L. niloticus*. Yet, the
189 third lineage and the paraphyly are unexpected. Similar to the pattern we found in *L. niloticus*, what
190 has been assumed to be a single species *L. calcarifer*, turned out to include two to three deeply
191 divergent genetic lineages (Lin et al., 2006; Vij et al., 2014). Only recently, *L. lakdiva* and *L. uwisara*
192 were described upon critical morphological examination of museum specimens originally assigned to
193 *L. calcarifer* (Pethiyagoda and Gill, 2012). Therefore, we suggest that the rather conserved
194 morphology throughout *Lates* and especially within the allegedly widespread *L. niloticus* (in Africa)
195 and *L. calcarifer* (in the Indo-Pacific) might have caused an underestimation of species numbers. Non-
196 monophyly of a widespread African freshwater fish species with respect to a lacustrine species flock
197 was also found for the African sharptooth catfish, *Clarias gariepinus* (Burchell, 1822) and the species
198 assemblage of *Bathyclarias* endemic to Lake Malawi (Van Steenberge et al., 2020).

199 The fossil record suggests that a marine or estuarine lifestyle is the ancestral state within *Lates*
200 (Otero et al., 2017). Several fossil *Lates* species are known from freshwater environments in Africa,
201 Europe, Asia Minor and the Arabian Peninsula, potentially pointing to repeated colonization of
202 continental freshwater systems (Otero et al., 2017). Regarding the extant species only, colonization
203 of freshwaters happened once in the ancestor of the African *Lates* species. *Lates* appeared in the
204 fossil record along the African Tethyan coast in the Oligocene and started to diversify in coastal
205 waters prior to the closure of the Tethys (Otero, 2018). Considering the long and extensive fossil
206 record (15.5 MYA; Otero and Gayet, 2001) and paraphyly of *L. niloticus*, it seems plausible that the
207 Indo-Pacific *Lates* diverged from *L. niloticus* well after fossils assigned to the latter species appeared.
208 Therefore, we inferred divergence times both with and without *L. niloticus* as calibration point.
209 Including the *L. niloticus* calibration point produced divergence time estimates for the split between
210 the African and the Indo-Pacific lineages consistent with a proposed origin of *L. calcarifer* shortly
211 before the closure of the Tethys in the Lower Miocene (Otero et al., 2017). Alternatively, a
212 colonization of the Indian Ocean from the Mediterranean Sea via potential connections during
213 periods of high sea level in the early Middle Miocene may have given rise to the Indo-Pacific lineages

214 of *Lates* (Otero et al., 2017). Divergence times inferred without the *L. niloticus* calibration point
215 suggested an even more recent origin of the Indo-Pacific lineages in the Middle to Upper Miocene.
216 Whether these represent a monophylum or two distinct lineages that independently colonized the
217 Indo-Pacific still remains unclear. We emphasize, however, that absolute divergence times need to
218 be regarded with caution here, as only very few and comparatively old calibrations points were
219 available for inferring rather recent events. Hence, if anything, the inferred age estimates for recent
220 events should have been overestimated and actual divergence times might be even younger (Zheng
221 et al. 2011).

222 Regardless of the calibration scheme used, there is a considerable time lag between the divergence
223 of the African and Indo-Pacific lineages and that of extant African species of *Lates*, that started only
224 in the Pliocene, indicating repeated extinction and (re-)colonization of hydrological systems. The
225 large-scale phylogeographic pattern of extant African *Lates* is somewhat reminiscent of the patterns
226 observed in tigerfish, *Hydrocynus* spp., another large top predator (Goodier et al., 2011), suggesting
227 that in both genera the present-day patterns were mainly shaped by geomorphic evolution that
228 fragmented and reconfigured drainage basins (Goodier et al., 2011; and references herein). Yet,
229 unlike in *L. niloticus*, some haplotype sharing and closely related haplotypes were found in
230 *Hydrocynus* spp. from distinct drainage basins (Goodier et al., 2011), indicating that headwater
231 captures and extreme flooding events in geographically close swampy and/or headwater regions
232 seem to have had a more profound effect on the phylogeographic structure of *Hydrocynus* spp. than
233 *L. niloticus*. *Clarias gariepinus*, another large fish occurring across large parts of Africa, shows more
234 recent intraspecific divergence than Nile perch and tigerfish (Van Steenberge et al., 2020). Being able
235 to survive conditions that cause other taxa to go extinct, *C. gariepinus* does not have the (at least in
236 part) discontinuous distribution ranges that, in the other two taxa, presumably result from local
237 extinctions (Cotteril & de Wit, 2011).

238 Our divergence time analyses suggest, regardless of the calibrations scheme employed, that
239 colonization of Lake Tanganyika by *L. niloticus* and radiation into the four extant species *L.*

240 *angustifrons*, *L. mariae*, *L. microlepis* and *L. stappersii*, is much more recent than for other endemic
241 fish radiations (spiny eels: Brown et al., 2010; squeaker catfishes: Day et al., 2013; claroteid catfishes:
242 Pear et al., 2014; cichlids: Irisarri et al., 2018; Ronco et al., 2021) which all colonized and started to
243 radiate in the lake right with the onset of its formation 9-12 MYA and/or the establishment of a truly
244 lacustrine habitat with deepwater conditions 5-6 MYA (Cohen et al., 1993). Notably, also
245 morphological and genetic data on monogenean and digenean parasites of *Lates* spp. are consistent
246 with our results (Kmentová et al., 2020a, b). The four Lake Tanganyika endemics are clearly the top
247 predators in the lake's benthopelagic and eupelagic zone and show some clear niche differentiation
248 with respect to their preferred habitat. As juveniles, *L. mariae* and *L. microlepis* are found in the
249 shallow littoral habitat, whereas as adults they are pelagic top predators, with *L. mariae* typically
250 found at greater depth (Coulter, 1991). *Lates angustifrons* prefers, throughout its life, inshore rocky
251 habitat and *L. stappersii* is a pelagic species throughout its life, forming large shoals that follow and
252 prey upon Lake Tanganyika's endemic clupeids (Coulter, 1991).

253 Throughout the Pleistocene, the time frame of *Lates* diversification in Lake Tanganyika, Lake
254 Tanganyika experienced recurrent drastic lake level fluctuations. Although for most of this period
255 detailed information on these fluctuations is lacking, extrapolations from Lake Malawi, for which a
256 good lake level record is available for the last 1.2 MY (Ivory et al., 2016), indicate that there were
257 times when Lake Tanganyika's sub-basins might have been separated. During these periods of
258 separation allopatric divergence might have taken place, followed by character displacement (i.e.
259 specialization to particular niches) upon secondary contact once lake levels rose. *Lates* spp. are highly
260 mobile species that can easily cover large distances in short time. They share their niche with large
261 piscivorous benthopelagic and eupelagic cichlid species of the genera *Boulengerochromis*,
262 *Bathybates* and *Hemibates*. These cichlids show no to very little phylogeographic structure on a lake-
263 wide scale (Koblmüller et al., 2015, 2019), and there is no reason to assume this would be different in
264 *Lates* spp..

265 The propensity of *Lates* to diversify in lakes becomes further evident from *L. macrophthalmus* and *L.*
266 *longispinis*, endemics of Lakes Albert and Turkana, respectively, where they probably diverged from,
267 and occur in sympatry with, *L. niloticus*. As in Lake Tanganyika, the *Lates* species in these two smaller
268 and younger rift lakes diverged in their habitat preferences. Whereas in both lakes *L. niloticus*
269 appears to be most common in inshore waters, *L. macrophthalmus* and *L. longispinis* are smaller
270 offshore species (Harrison, 1991). Even if morphologically intermediate individuals seem to exist in
271 these lakes, the little genetic data available on *L. macrophthalmus* and *L. longispinis* suggest that they
272 are indeed distinct, recently diverged species (Hauser et al., 1998).

273 To conclude, our multilocus phylogenetic analysis recovered phylogenetic relationships among extant
274 *Lates* species that are entirely different to what has been previously suggested based on
275 morphological data. Divergence time estimates indicate recent divergence of African taxa and a very
276 recent colonization of and radiation in Lake Tanganyika. The observed deep divergence within and
277 paraphyly of *L. niloticus* points to the urgent need of more detailed phylogeographic studies and
278 thorough morphological investigation, of especially museum specimens, to get an idea about the
279 natural distribution limits of the distinct *L. niloticus* lineages. This is particularly important in the light
280 of the species' introduction to various African water bodies as important food and sport fish.

281

282 Acknowledgements

283 This research was supported by the Systematics Research Fund (SRF; to SK), by the Czech Science
284 Foundation (GA19-13573S, to MVS, NK and MPMV; and P505/12/G112 (ECIP)), the Belspo-Brain
285 (Belgian Science Policy) project HIPE, the King Leopold III fund for nature exploration and
286 conservation, the Research Foundation - Flanders (FWO), the Special Research Fund of Hasselt
287 University (BOF20TT06, to MPMV) and by a joint program between the Austrian Agency for
288 International Cooperation in Education and Research (OEAD; project number CZ 08/2018, to SK) and
289 the Czech Ministry of Education, Youth and Sports (project number 8J18AT007, to MPMV). Many

290 thanks to T. Moelants for providing the samples from the Congo Basin, the collection of which was
291 supported by an Action 2 program of the Belgian Federal Science Policy Office and J. Snoeks, M.
292 Mulongaibalu and N. Vranken for aid in sampling at the Lake Albert.

293

294 References

295 Basiita, R.K., Zenger, K.R., Mwanja, M.T., Jerry, D.R., 2018. Gene flow and genetic structure in Nile
296 perch, *Lates niloticus*, from African freshwater rivers and lakes. PLoS ONE 13, e0200001.

297 <https://doi.org/10.1371/journal.pone.0200001>

298 Betancur-R, R., Broughton, R.E., Wiley, E.O., Carpenter, K., López, J.A., Li C, Holcroft, N.I., Arcila, D.,
299 Sanciangco, M., Cureton li, J.C., Zhang, F., Buser, T., Campbell, M.A., Ballesteros, J.A., Roa-Varon, A.,
300 Willis, S., Borden, W.C., Rowley, T., Reneau, P.C., Hough, D.J., Lu, G., Grande, T., Arratia, G., Ortí, G.,
301 2013. The tree of life and a new classification of bony fishes. PLoS Curr. 5,
302 ecurrents.tol.53ba26640df0ccaee75bb165c8c26288.

303 <https://doi.org/10.1371/currents.tol.53ba26640df0ccaee75bb165c8c26288>

304 Betancur-R, R., Wiley, E.O., Arratia, G. Acero, A., Bailly, N., Miya, M., Lecointre, G., Orti, G., 2017.

305 Phylogenetic classification of bony fishes. BMC Evo. Biol. 17, 162. [https://doi.org/10.1186/s12862-](https://doi.org/10.1186/s12862-017-0958-3)
306 [017-0958-3](https://doi.org/10.1186/s12862-017-0958-3)

307 Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A.,
308 Drummond, A.J., 2014. BEAST2: A software platform for Bayesian evolutionary analysis. PLoS
309 Comput. Biol. 10, e1003537. <https://doi.org/10.1371/journal.pcbi.1003537>

310 Brown, K.J., Rüber, L., Bills, R., Day, J.J., 2010. Mastacembelid eels support Lake Tanganyika as an
311 evolutionary hotspot of diversification. BMC Evol. Biol. 10, 188. [https://doi.org/10.1186/1471-2148-](https://doi.org/10.1186/1471-2148-10-188)
312 [10-188](https://doi.org/10.1186/1471-2148-10-188)

313 Cohen, A.S., Soreghan, M.J., Scholz, C.A., 1993. Estimating the age of formation of lakes: An example
314 from Lake Tanganyika, East African Rift system. *Geology* 21, 511-514.

315 Cotterill, F.P.D., de Wit, M., 2011. Geocodynamics and the Kalahari epeirogeny: Linking its genomic
316 record, tree of life and palimpsest into a unified narrative of landscape evolution. *S. Afr. J. Geol.* 114,
317 489– 514. <https://doi.org/10.2113/gssaig.114.3-4.489>

318 Coulter, G. W., 1991. *Lake Tanganyika and Its Life*. Oxford University Press, Oxford.

319 Day, J.J., Peart, C.R., Brown, K.J., Friel, J.P., Bills, R., Moritz, T., 2013. Continental diversification of an
320 African catfish radiation (Mochokidae: *Synodontis*). *Syst. Biol.* 62, 351-365.
321 <https://doi.org/10.1093/sysbio/syt001>

322 Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
323 *Nucleic Acids Res.* 32, 1792-1797. <https://doi.org/10.1093/nar/gkh340>

324 Fricke, R., Eschmeyer, W.N., van der Laan, R., 2021. Eschmeyer's catalog of fishes: genera, species,
325 references. (<http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>) –
326 Electronic version accessed 02 February 2021.

327 Goldschmidt, T., 1996. *Darwin's Dreampond. Drama in Lake Victoria*. MIT Press, Cambridge, MA.

328 Goodier, S.A.M., Cotterill, F.P.D., O’Ryan, C., Skelton, P.H., de Wit, M.J., 2011. Cryptic diversity of
329 African tigerfish (genus *Hydrocynus*) reveals palaeogeographic signature of linked Neogene
330 geotectonic events. *PLoS ONE* 6, e28775. <https://doi.org/10.1371/journal.pone.0028775>

331 Greenwood, P.H., 1976. A review of the family Centropomidae (Pisces, Perciformes). *Bull. Br. Mus.*
332 *Nat. Hist. (Zool.)* 29, 1-81.

333 Harrison, K., 1991. The taxonomy of East African Nile Perch, *Lates* spp. (Perciformes,
334 Centropomidae). *J. Fish Biol.* 38, 175-186.

335 Harrison, K., 1991. The taxonomy of East African Nile perch, *Lates* spp. (Perciformes,
336 Centropomidae). J. Fish Biol. 38, 175-186.

337 Hauser, L., Carvalho, G.R., Pitcher, T.J., Ogutu-Ohwayo, R., 1998. Genetic affinities of an introduced
338 predator: Nile perch in Lake Victoria, East Africa. Mol. Ecol. 7, 849-857.
339 <https://doi.org/10.1046/j.1365-294x.1998.00399.x>

340 Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biol.
341 Evol. 27, 570-580. <https://doi.org/10.1093/molbev/msp274>

342 Irisarri, J., Singh, P., Koblmüller, S., Torres-Dowdall, J., Henning, F., Franchini, P., Fischer, C., Lemmon,
343 A.R., Moriarty Lemmon, E., Thallinger, G.G., Sturmbauer, C., Meyer, A., 2018. Phylogenomics
344 uncovers early hybridization and adaptive loci shaping the radiation of Lake Tanganyika cichlid fishes.
345 Nature Comm. 9, 3159. <https://doi.org/10.1038/s41467-018-05479-9>

346 Ivory, S.J., Blome, M.W., King, J.W., McGlue, M.M., Cole, J.E., Cohen, A.S., 2016. Environmental
347 change explains cichlid adaptive radiation at Lake Malawi over the past 1.2 million years. Proc. Natl.
348 Acad. Sci. 113, 11895-11900. <https://doi.org/10.1073/pnas.1611028113>

349 Iwatsuki, Y., Newman, S.J., Tanaka, F., Russell, B.C., 2018. Validity of *Psammoperca datnioides*
350 Richardson 1848 and redescrptions of *P. waigiensis* Cuvier in Cuvier Valenciennes 1828 and
351 *Hypopterus macropterus* (Günther 1859) in the family Latidae (Perciformes) from the Indo-West
352 Pacific. Zootaxa 4402, 467-486. <https://doi.org/10.11646/zootaxa.4402.3.3>

353 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermini, L.S., 2017. ModelFinder:
354 fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587-589.
355 <https://doi.org/10.1038/nmeth.4285>

356 Kinoshita, I., Tshibangu, K.K., 1997. Larvae of four *Lates* from Lake Tanganyika. Bull. Mar. Sci. 60, 89-
357 99.

358 Kmentová, N., Bray, R.A., Koblmüller, S., Artois, T., De Keyzer, E.L.R., Gelnar, M., Vanhove, M.P.M.,
359 Georgieva, S., 2020a. Uncharted digenean diversity in Lake Tanganyika: cryptogonimids (Digenea:
360 Cryptogonimidae) infecting endemic lates perches (Actinopterygii: Latidae). *Parasit. Vectors* 13, 221.
361 <https://doi.org/10.1186/s13071-020-3913-x>

362 Kmentová, N., Koblmüller, S., Van Steenberge, M., Artois, T., Muterezi Bukinga, F., Mulimbwa
363 N'sibula, T., Muzumani Risasi, D., Masilya Mulungula, P., Gelnar, M., Vanhove, M.P.M., 2020b. Failure
364 to diverge in African Great Lakes: The case of *Dolicirroplectanum lacustre* gen. nov. comb. nov.
365 (Monogenea, Diplectanidae) infecting latid hosts. *J. Great Lakes Res.* 46, 1113-1130.
366 <https://doi.org/10.1016/j.jglr.2019.09.022>

367 Koblmüller, S., Odhiambo, E.A., Sinyinza, D., Sturmbauer, C., Sefc, K.M., 2015. Big fish, little
368 divergence: phylogeography of Lake Tanganyika's giant cichlid, *Boulengerochromis microlepis*.
369 *Hydrobiologia* 748, 29–38. <https://doi.org/10.1007/s10750-014-1863-z>

370 Koblmüller, S., Zangl, L., Börger, C., Daill, D., Vanhove, M.P.M., Sturmbauer, C., Sefc, K.M., 2019. Only
371 true pelagics mix: comparative phylogeography of deepwater bathybatine cichlids from Lake
372 Tanganyika. *Hydrobiologia* 832, 93–103. <https://doi.org/10.1007/s10750-018-3752-3>

373 Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version
374 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870-1875. <https://doi.org/10.1093/molbev/msw054>

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

378 Li, C., Betancur-R., R., Smith, W.L., Ortí, G., 2011. Monophyly and interrelationships of snook and
379 barramundi (Centropomidae *sensu* Greenwood) and five new markers for fish phylogenetics. *Mol.*
380 *Phylogenet. Evol.* 60, 463-471. <https://doi.org/10.1016/j.ympev.2011.05.004>

381 Lin, G., Lo, L.C., Zhu, Z.Y., Feng, F., Chou, R., Yue, G.H., 2006. The complete mitochondrial genome
382 sequence and characterization of single-nucleotide polymorphisms in control region of the Asian
383 seabass (*Lates calcarifer*). Mar. Biotechnol. 8, 71-89. <https://doi.org/10.1007/s10126-005-5051-z>

384 Mooi, R.D., Gill, A.C., 1995. Association of epaxial musculature with dorsal-fin pterygophores in
385 acanthomorph fishes, and its phylogenetic significance. Bull. Br. Mus. Nat. Hist. (Zool.) 61, 121-137.

386 Murray, A.M., Attia, Y.S., 2004. A new species of *Lates* (Teleostei: Perciformes) from the Lower
387 Oligocene of Egypt. J. Vertebr. Paleontol. 24, 299-308. <https://doi.org/10.1671/1948>

388 Mwanja, M.T., Muwanika, V., Masembe, C., Nyakaana, S., Mwanja, W.W., 2013. Evolutionary history
389 of Nile perch *Lates* sp. inferred from mitochondrial DNA variation analyses. Zool. Stud. 52, 59.
390 <https://doi.org/10.1186/1810-522X-52-59>

391 Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.W., 2015. IQ-TREE: a fast and effective
392 stochastic algorithm for estimating maximum likelihood phylogenies. Mol. Biol. Evol. 32, 268-274.
393 <https://doi.org/10.1093/molbev/msu300>

394 Otero, O., Gayet, M., 2001. Palaeoichthyofaunas from the Lower Oligocene and Miocene of the
395 Arabian Plate: palaeoecological and palaeobiogeographical implications. Palaeogeogr.
396 Palaeoclimatol. Palaeoecol. 165, 141-169. [https://doi.org/10.1016/S0031-0182\(00\)00158-9](https://doi.org/10.1016/S0031-0182(00)00158-9)

397 Otero, O., 2004. Anatomy, systematics and phylogeny of both recent and fossil latid fishes (Teleostei,
398 Perciformes, Latidae). Zool. J. Linn. Soc. 141, 81-133. [https://doi.org/10.1111/j.1096-
399 3642.2004.00111.x](https://doi.org/10.1111/j.1096-3642.2004.00111.x)

400 Otero, O., Murray, A., Cavin, L., Clément, G., Pinton, A., Stewart, K., 2017. African fossil fish, in:
401 Paugy, D., Levêque, C., Otero, O. (Eds.), The inland water fishes of Africa: Diversity, Evolution and
402 Human Use. IRD Editions, Marseille, France. Available on the Internet:
403 <http://books.openedition.org/irdetitions/25130>. pp. 54-86.

404 Peart, C.R., Bills, R., Wilkinson, M., Day, J.J., 2014. Nocturnal claroteine catfishes reveal dual
405 colonisation but a single radiation in Lake Tanganyika. *Mol. Phylogenet. Evol.* 73, 119-128.
406 <https://doi.org/10.1016/j.ympev.2014.01.013>

407 Pethiyagoda, R., Gill, A.C., 2012. Description of two new species of sea bass (Teleostei: Latidae: *Lates*)
408 from Myanmar and Sri Lanka. *Zootaxa* 3314, 1-16.

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

412 Richlen, M.L., Barber, P.H., 2005. A technique for the rapid extraction of microalgal DNA from single
413 live and preserved cells. *Mol. Ecol. Notes* 5, 688-691. [https://doi.org/10.1111/j.1471-](https://doi.org/10.1111/j.1471-8286.2005.01032.x)
414 [8286.2005.01032.x](https://doi.org/10.1111/j.1471-8286.2005.01032.x)

415 Ronco, F., Matschiner, M., Böhne, A., Boila, A., Büscher, H.H., El Taher, A., Indermaur, A., Malinsky,
416 M., Ricci, V., Kahmen, A., Jentoft, S., Salzburger, W., 2021. Drivers and dynamics of a massive
417 adaptive radiation in cichlid fishes. *Nature* 589, 76-81. <https://doi.org/10.1038/s41586-020-2930-4>

418 Ronquist, F., Teslenko, M., Van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L.,
419 Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and
420 model choice across a large model space. *Syst. Biol.* 61, 539-542.
421 <https://doi.org/10.1093/sysbio/sys029>

422 Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood
423 phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18, 502-504.
424 <https://doi.org/10.1093/bioinformatics/18.3.502>

425 Van der Laan, R., Eschmeyer, W.N., Fricke, R., 2014. Family-group names of recent fishes. *Zootaxa*
426 3882, 1-230. <http://dx.doi.org/10.11646/zootaxa.3882.1.1>

427 Van Steenberge, M.W., Vanhove, M.P.M., Chocha Manda, A., Larmuseau, M.H.D., Swart, B.L.,
428 Khang'Mate, F., Arndt, A., Hellemans, B., Van Houdt, J., Micha, J.-C., Koblmüller, S., Roodt-Wilding, R.,
429 Volckaert, F.A.M., 2020. Unravelling the evolution of Africa's drainage basins through a widespread
430 freshwater fish, the African sharptooth catfish *Clarias gariepinus*. J. Biogeogr. 47, 1739-1754.
431 <https://doi.org/10.1111/jbi.13858>

432 Vij, S., Purushothaman, K., Gopkrishna, G., Lau, D., Saju, J.M., Shamsudheen, K.V., Kumar, K.V.,
433 Basheer, V.S., Gopalakrishnan, A., Hossain, M.S., Sivasubbu, S., Scaria, V., Jena, J.K., Ponniah, A.G.,
434 Orbán, L., 2014. Barcoding of Asian seabass across its geographic range provide evidence for its
435 bifurcation into two distinct species. Front. Mar. Sci. 1, 30.
436 <https://doi.org/10.3389/fmars.2014.00030>

437 Zhang, F., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W.X., Wang, G.T., 2020. PhyloSuite: An integrated
438 and scalable desktop platform for streamlined molecular sequence data management and
439 evolutionary phylogenetic studies. Mol. Ecol. Res. 20, 348-355. [https://doi.org/10.1111/1755-](https://doi.org/10.1111/1755-0998.13096)
440 [0998.13096](https://doi.org/10.1111/1755-0998.13096)

441 Zheng, Y., Peng, R., Kuro-o, M., Zeng, X., 2011. Exploring patterns and extent of bias in estimating
442 divergence time from mitochondrial DNA sequence data in a particular lineage: a case study of
443 salamanders (Order Caudata). Mol. Biol. Evol. 28, 2521–2535.
444 <https://doi.org/10.1093/molbev/msr072>

445

446

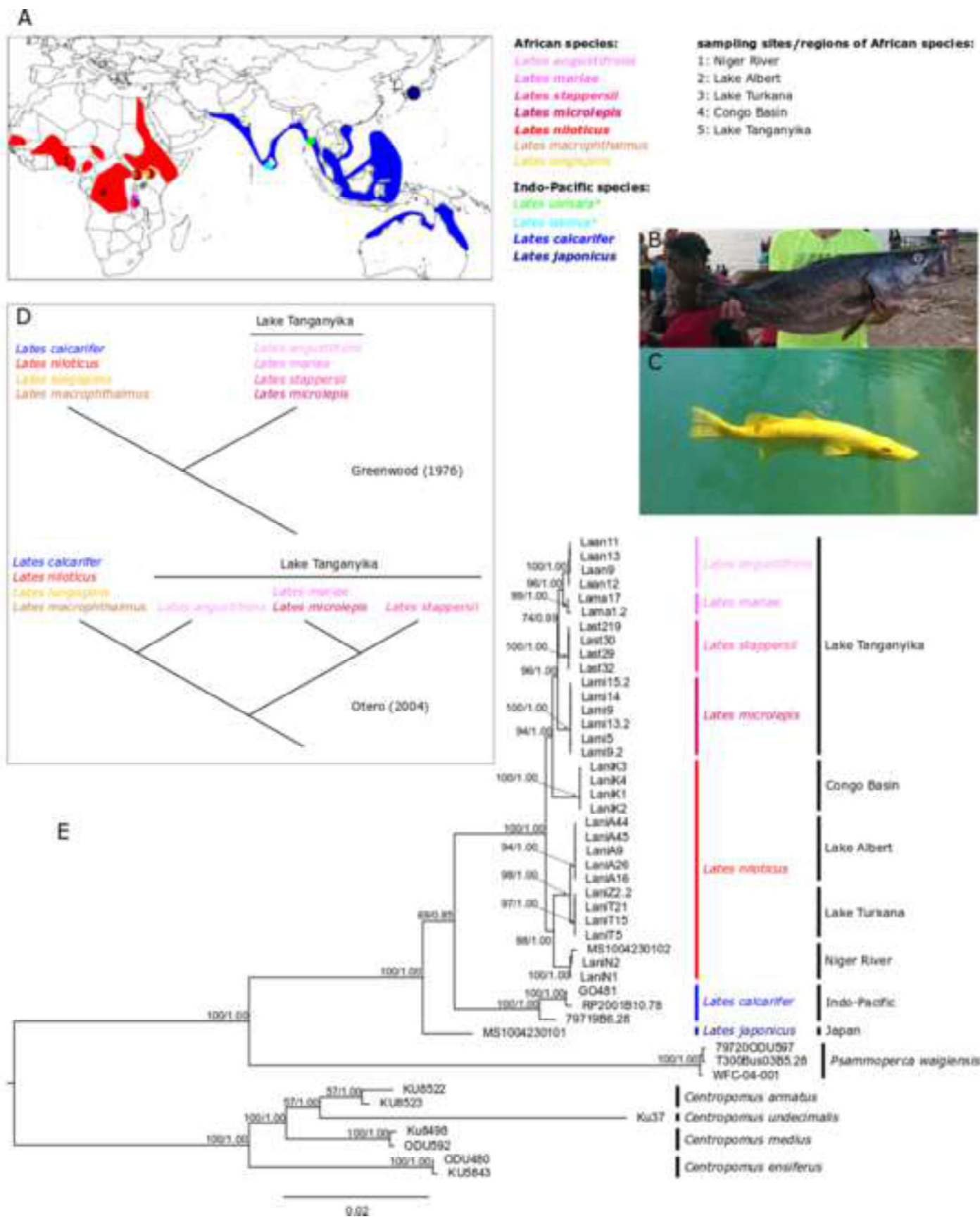
447 **Figure 1.** Diversity of *Lates* spp. (A) Natural distribution ranges of extant *Lates* species (after Otero et
448 al., 2017) and sampling sites/regions for the newly sequenced African *Lates* individuals. Regions
449 where *L. niloticus* has been introduced are not considered here. Taxa in bold are included in this
450 study. *, These two species have been described only recently; their exact distribution range is
451 unknown, only the type localities are indicated for these two species. (B) *Lates* perches are highly
452 valued food and game fish in both commercial and artisanal fisheries. Several of the species, like *L.*
453 *microlepis*, grow to considerable size (>1m). (C) The mythical golden perch is a very rare golden
454 morph of *L. angustifrons*. (D) Alternative phylogenetic hypotheses for the genus *Lates* based on
455 osteological characters (Greenwood, 1976; Otero, 2004). (E) Maximum Likelihood (ML) tree showing
456 the phylogenetic relationships among *Lates* spp. based on one mitochondrial and eleven nuclear
457 markers. As measures of nodal support bootstrap support values (for ML; only values >50 are shown)
458 and posterior probabilities (from Bayesian inference; only values >0.7 are shown) are depicted.

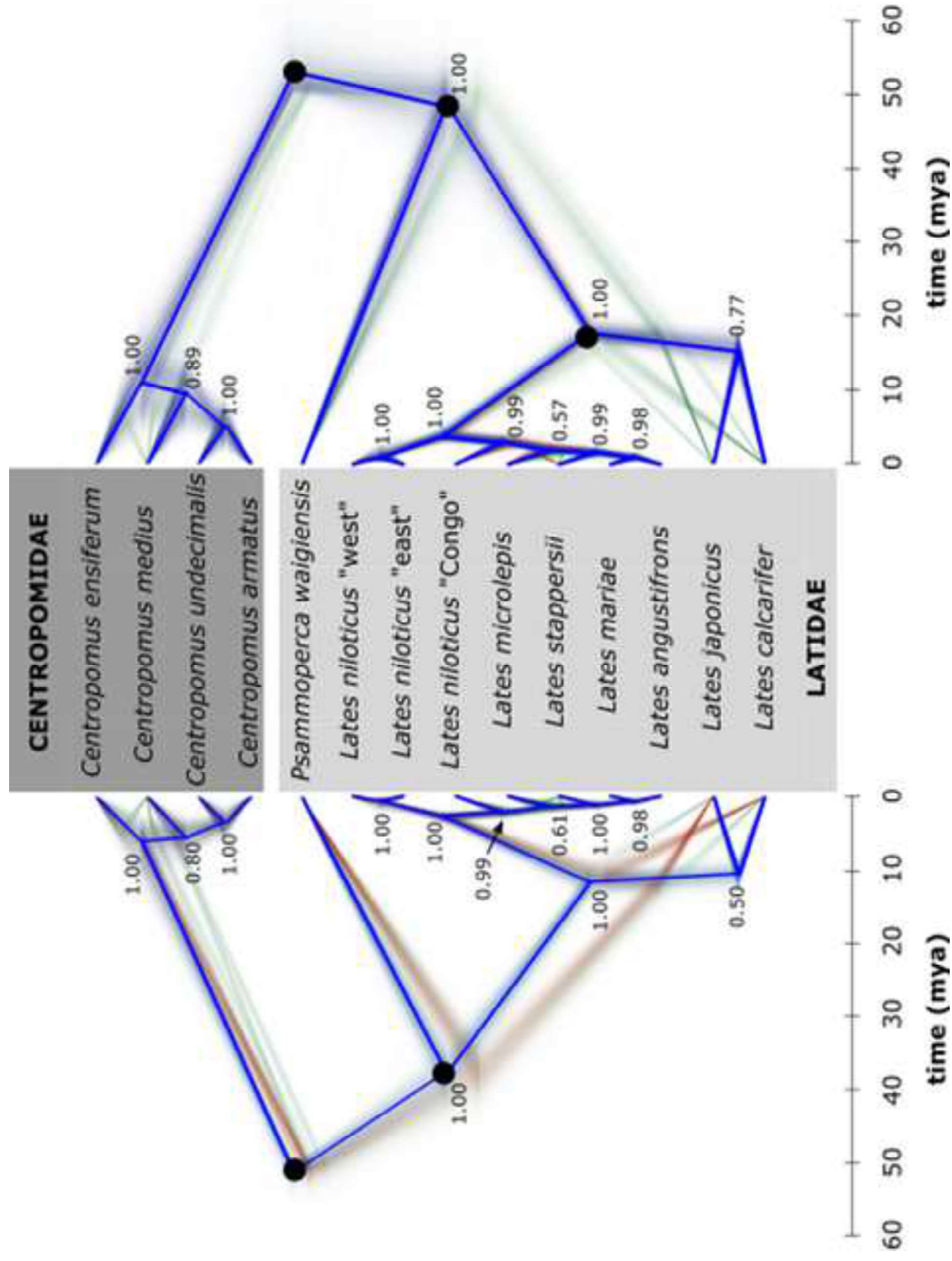
459

460 **Figure 2.** Time calibrated multispecies coalescent trees of *Lates* perches with representatives of
461 *Centropomus* used as outgroup. The consensus phylogenies from the two alternative calibration
462 schemes are superimposed on DensiTree cloudograms of alternative sampled trees, with contrasting
463 topologies highlighted by different colors. Nodal support is given as posterior probabilities (only
464 values >0.7 are shown). Black dots indicate calibration points.

465

466





time (mya)

time (mya)

Authorship statements:

Stephan Koblmüller: conceptualization, sampling, data generation, data analysis, writing – original draft, funding acquisition

Christian A. Schöggel: sampling, data generation, data analysis, writing – review

Clemens J. Lorber: sampling, data generation, writing – review

Maarten Van Steenberge: sampling, data generation, data analysis, writing – review, funding acquisition

Nikol Kmentová: conceptualization, sampling, writing – review, funding acquisition

Maarten P. M. Vanhove: conceptualization, sampling, writing – review, funding acquisition

Lukas Zangl: sampling, data generation, data analysis, writing – review